Introdução à Probabilidade e Estatística

Teoria de Probabilidades

Variáveis Aleatórias

Departamento de Matemática
Universidade de Évora

Ano lectivo de 2016/17

Patrícia Filipe
Dulce Gomes

Teoria das Probabilidades

Experiência Aleatória e Espaço Amostra

Uma experiência diz-se aleatória sse verificar as seguintes características:

- cada vez que é efetuada desconhecemos à partida qual o resultado que vamos obter;
- é conhecido o conjunto de todos os resultados possíveis;
- a experiência pode ser repetida em condições similares e existe regularidade quando é repetida muitas vezes.

O espaço amostra ou universo, Ω, é o conjunto de todos os resultados possíveis de uma experiência aleatória.

Acontecimentos

Um acontecimento A respeitante a um determinado espaço amostra Ω e associado a uma experiência aleatória é simplesmente um conjunto de resultados possíveis. Dizemos que A se realizou, se o resultado da experiência aleatória, ω, é um elemento de A, i.e., $\omega \in A$.

Um acontecimento pode ser um conjunto formado por um só elemento de Ω, e nesse caso, diz-se acontecimento elementar chamamos aos outros acontecimentos compostos.

Ao espaço amostra, Ω, também se chama acontecimento certo.

Álgebra dos Acontecimentos

Representando os acontecimentos por A, B, C, ..., temos:

- União de acontecimentos: $A \cup B$, consiste na realização de pelo menos um dos acontecimentos. Ocorre A ou B.
- Interseção de acontecimentos: $A \cap B$ ocorre sse ocorrerem A e B, em simultâneo ou em sequência.
- Diferença de acontecimentos: $A - B$ (ou $A \backslash B$) ocorre A mas não B.
  Se $A = \Omega$, $\Omega - B = \bar{B}$ diz-se o complementar de B, e ocorre quando não ocorre B.

Propriedades das Operações — União

Propriedades	União
Comutativa	$A \cup B = B \cup A$
Associativa	$(A \cup B) \cup C = A \cup (B \cup C)$
Distributiva	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
Idempotência	$A \cup A = A$
Lei do Complementar	$A \cup \overline{A} = \Omega$
Leis de De Morgan	$\overline{A \cap B} = \overline{A} \cup \overline{B}$
Elemento Neutro	$A \cup \emptyset = A$
Elemento Absorvente	$A \cup \Omega = \Omega$

Propriedades das Operações — Intersecção

Propriedades	Intersecção
Comutativa	$A \cap B = B \cap A$
Associativa	$(A \cap B) \cap C = A \cap (B \cap C)$
Distributiva	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
Idempotência	$A \cap A = A$
Lei do Complementar	$A \cap \overline{A} = \emptyset$
Leis de De Morgan	$\overline{A \cup B} = \overline{A} \cap \overline{B}$
Elemento Neutro	$A \cap \Omega = A$
Elemento Absorvente	$A \cap \emptyset = \emptyset$

Definição Clássica (ou de Laplace) de Probabilidade

Para uma experiência aleatória em que os resultados possíveis são disjuntos e igualmente prováveis.

Probabilidade Laplaciana

$$P(A) = \frac{\text{Número de casos favoráveis a A}}{\text{Número de casos possíveis}}$$

Conceito frequencista de probabilidade

Se em N realizações de uma experiência aleatória, o acontecimento A ocorreu n vezes, diz-se que a frequência relativa de A nas N realizações é

$$f_A = \frac{n}{N}$$

a frequência relativa do acontecimento A.
Define-se probabilidade do acontecimento A, como o número para que tende a frequência relativa f_A, quando se aumenta o número de provas, ou seja,

$$P[A] = \lim_{N \to \infty} f_A$$

Exemplo: O jogo das cartas

Vamos considerar um baralho de cartas normal. Ou seja, um baralho com 52 cartas com 4 naipes: ouros, espadas, copas e paus. Existem 13 cartas de cada naipe.
Qual a probabilidade de retirar ao acaso do baralho um rei de ouros?

- Qual o acontecimento a que pretendemos atribuir uma probabilidade?
- Número de resultado possíveis associados à experiência aleatória de retirar uma carta do baralho?
- Número de resultado favoráveis associados à experiência aleatória de retirar uma carta do baralho?

Seja A="Sair um rei de ouros" $\leftrightsquigarrow$ Acontecimento Elementar

Espaço de Resultados:

$$\begin{aligned} \Omega &= \{\text{Ás}\diamondsuit,\ 2\diamondsuit,\ 3\diamondsuit \ldots, \text{Dama}\spadesuit,\ \text{Rei}\spadesuit\} \\ \sharp(\Omega) &= 52 \end{aligned}$$

$$P(A) = \frac{1}{52} = 0.01923 \ \ (1,923\%)$$

Exemplo: O jogo continua...

Suponhamos agora que voltamos a colocar a carta no baralho. Voltamos a baralhar as cartas e retiramos de novo, ao acaso, uma carta.
Qual a probabilidade de retirar agora, ao acaso, um às de espadas?

Qual a probabilidade de retirar, ao acaso, um às?

Qual a probabilidade de retirar, ao acaso, uma carta copas de número par?

Exemplo: O jogo continua ...

Suponhamos agora que retirámos uma carta do baralho (um às de copas) e não a voltamos a colocar no baralho.
Perante este cenário, qual é a probabilidade de ao retirar uma outra carta do baralho esta seja um dois de copas?
E qual a probabilidade de sair um às?

A extração faz-se Com Reposição ou Sem Reposição?

Definição Axiomática de Probabilidade

Teoremas

1. $P(\emptyset) = 0 \quad \rightsquigarrow \quad \emptyset$ é o acontecimento impossível;
2. $0 \leqslant P(A) \leqslant 1$;
3. $P(\overline{A}) = 1 - P(A)$;
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;

Probabilidade Condicionada

Suponhamos que retirámos uma carta do baralho (um às de copas) e não a voltamos a colocar no baralho. Perante este cenário, qual é a probabilidade de ao retirar uma outra carta do baralho esta seja um dois de copas? A="Às de copas"; B="Dois de copas".

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \Leftrightarrow P(B|A) = \frac{P(A \cap B)}{P(A)} \quad (P(A) > 0)$$

De onde se deduz a seguinte expressão para a probabilidade conjunta:

$$P(A \cap B) = P(B|A) \times P(A) \quad \text{ou} \quad P(A \cap B) = P(A|B) \times P(B)$$

Acontecimentos Independentes

Suponhamos que retirámos uma carta do baralho (um às de copas) e a voltamos a colocar no baralho. Qual é a probabilidade de ao retirar uma outra carta do baralho esta seja um dois de copas? A="Dois de copas"; B="Às de copas".

A e B dizem-se independentes sse

$$P(A \cap B) = P(A) \times P(B)$$

ou

$$P(A|B) = P(A) \text{ e } P(B|A) = P(B)$$

Acontecimentos Mutuamente Exclusivos

Acontecimentos disjuntos, incompatíveis ou mutuamente exclusivos são acontecimentos em que a realização de um deles implica a não realização do outro.

A e B são disjuntos $\Leftrightarrow A \cap B = \emptyset \Leftrightarrow P(A \cap B) = 0$

Acontecimentos independentes versus acontecimentos mutuamente exclusivos

Dados dois acontecimentos, A e B, tais que $P[A] > 0$ e $P[B] > 0$,

- se os acontecimentos são *mutuamente exclusivos,* então

$$A \cap B = \emptyset \text{ e } P[A \cap B] = 0,$$

  pelo que não podem ser independentes, pois para tal,

$$P[A \cap B] = P[A] . P[B] > 0;$$

- se os acontecimentos são independentes, então

$$P[A \cap B] = P[A] . P[B] > 0,$$

  não podendo ser mutuamente exclusivos, pois para tal, $P[A \cap B] = 0$.

Exemplo

Na empresa Sojoga será desenvolvida uma aplicação para Smartphones que permita decidir qual o melhor caminho para que um indivíduo se desloque de um qualquer ponto A para um outro qualquer ponto B, tendo em conta o tráfego em tempo real. Um estudo prévio revelou que 72% das pessoas afirmaram que iriam comprar a aplicação, e destas 27% afirmaram registar-se no site da empresa. Tendo em conta os resultados e admitindo que 37% das pessoas se iriam registar no site, diga qual a probabilidade de uma pessoa, seleccionada ao acaso,

(a) não comprar a aplicação?

(b) comprar a aplicação e registar-se no site?

(c) registar-se no site, sabendo que não comprou a aplicação?

Como apresentar a informação numa tabela de dupla entrada?

Exemplo

A poluição do ar de certa cidade é causada essencialmente por gases industriais (75% dos casos) e por gases dos escapes de automóveis (25% dos casos). Nos próximos 4 anos prevê-se que a possibilidade de controlar com sucesso a poluição dado que provêem dessas duas fontes de poluição sejam de 70% e de 60%, respectivamente.

(a) Qual a probabilidade de haver controlo bem sucedido da poluição do ar dessa cidade (i.e. controlar com sucesso pelo menos uma das fontes de poluição) nos próximos 4 anos?

Seja C="Controlo bem sucedido da poluição do ar", I="Gases industriais" e E="Gases de escape".

$P(I) = 0,75 \quad P(E) = 0,25 \quad P(C|I) = 0,7 \quad P(C|E) = 0,6.$ $P(C) = ?$

R: $P(C) = P(C|I)P(I) + P(C|E)P(E) = 0,7 \times 0,75 + 0,6 \times 0,25 = 0,675.$

Teorema da Probabilidade Total $\rightsquigarrow$

Exemplo

A poluição do ar de certa cidade é causada essencialmente por gases industriais (75% dos casos) e por gases dos escapes de automóveis (25% dos casos). Nos próximos 4 anos prevê-se que a possibilidade de controlar com sucesso a poluição dado que provêem dessas duas fontes de poluição sejam de 70% e de 60%, respectivamente.

(a) Qual a probabilidade de haver controlo bem sucedido da poluição do ar dessa cidade (i.e. controlar com sucesso pelo menos uma das fontes de poluição) nos próximos 4 anos?

Seja C=”Controlo bem sucedido da poluição do ar“, I=”Gases industriais“ e E=”Gases de escape“.

$P(I) = 0,75 \quad P(E) = 0,25 \quad P(C|I) = 0,7 \quad P(C|E) = 0,6$. $P(C) = ?$

R: $P(C) = P(C|I)P(I) + P(C|E)P(E) = 0,7 \times 0,75 + 0,6 \times 0,25 = 0,675.$

Teorema da Probabilidade Total $\rightsquigarrow$

Exemplo

A poluição do ar de certa cidade é causada essencialmente por gases industriais (75% dos casos) e por gases dos escapes de automóveis (25% dos casos). Nos próximos 4 anos prevê-se que a possibilidade de controlar com sucesso a poluição dado que provêem dessas duas fontes de poluição sejam de 70% e de 60%, respectivamente.

(a) Qual a probabilidade de haver controlo bem sucedido da poluição do ar dessa cidade (i.e. controlar com sucesso pelo menos uma das fontes de poluição) nos próximos 4 anos?

Seja C="Controlo bem sucedido da poluição do ar", I="Gases industriais" e E="Gases de escape".

$P(I) = 0,75 \quad P(E) = 0,25 \quad P(C|I) = 0,7 \quad P(C|E) = 0,6$. $P(C) = ?$

R: $P(C) = P(C|I)P(I) + P(C|E)P(E) = 0,7 \times 0,75 + 0,6 \times 0,25 = 0,675$.

Teorema da Probabilidade Total $\rightsquigarrow$

Partição do universo

Os acontecimentos A_1, A_2,..., A_n definem uma partição em Ω, quando se verificam simultaneamente as seguintes condições:

(i) a união de todos os acontecimentos é o próprio espaço amostra, Ω:

$$\bigcup_{i=1}^{n} A_i = \Omega;$$

(ii) os acontecimentos são mutuamente exclusivos, dois a dois.

$$A_i \cap A_j = \emptyset, \quad i \neq j, \; i, j = 1, 2, ..., n;$$

(iii) todos os acontecimentos têm probabilidade não nula:

$$P[A_i] > 0, \quad i = 1, 2, ..., n.$$

Teorema da probabilidade total

Se os acontecimentos A_1, A_2,..., A_n definem uma partição sobre Ω, então para qualquer acontecimento B definido em Ω, tem-se que

$$\begin{aligned} P[B] &= \sum_{i=1}^{n} P[B|A_i].P[A_i] = \\ &= P[B|A_1].P[A_1] + P[B|A_2].P[A_2] + ... + P[B|A_n].P[A_n]. \end{aligned}$$

$(P(C) = P(C|I)P(I) + P(C|E)P(E) = 0,7 \times 0,75 + 0,6 \times 0,25 = 0,675.)$

Teorema da probabilidade total

Se os acontecimentos A_1, A_2,..., A_n definem uma partição sobre Ω, então para qualquer acontecimento B definido em Ω, tem-se que

$$\begin{aligned} P[B] &= \sum_{i=1}^{n} P[B|A_i].P[A_i] = \\ &= P[B|A_1].P[A_1] + P[B|A_2].P[A_2] + ... + P[B|A_n].P[A_n]. \end{aligned}$$

($P(C) = P(C|I)P(I) + P(C|E)P(E) = 0,7 \times 0,75 + 0,6 \times 0,25 = 0,675$.)

Exemplo (continuação)

(b) Constatando-se que houve controlo bem sucedido de poluição do ar ao fim de 4 anos, qual a probabilidade de isso ser devido ao controlo de gases dos escapes de automóveis?

Existem efeitos na poluição do ar que são originados por duas causas possíveis (e disjuntas). Conhecem-se essas probabilidades (**probabilidades à** *priori*). O que se pretende agora é calcular uma probabilidades à *posteriori*. Ou seja, após se ter verificado a ocorrência da poluição do ar, qual a probabilidade desta ter sido determinada por uma das causas.

$P(E|C) = ?$

Exemplo (continuação)

(b) Constatando-se que houve controlo bem sucedido de poluição do ar ao fim de 4 anos, qual a probabilidade de isso ser devido ao controlo de gases dos escapes de automóveis?

Existem efeitos na poluição do ar que são originados por duas causas possíveis (e disjuntas). Conhecem-se essas probabilidades (**probabilidades à** *priori*). O que se pretende agora é calcular uma probabilidades à *posteriori*. Ou seja, após se ter verificado a ocorrência da poluição do ar, qual a probabilidade desta ter sido determinada por uma das causas.

$P(E|C) = ?$

Teorema da Bayes

Se A_1, A_2,..., A_n definem uma partição sobre Ω, então para qualquer acontecimento B definido em Ω, com $P\left[B\right] > 0$,

Fórmula de Bayes

$$P\left[A_j \left| B\right.\right] = \frac{P\left[A_j\right].P\left[B\left|A_j\right.\right]}{\sum\limits_{i=1}^{n} P\left[A_i\right]P\left[B\left|A_i\right.\right]},\ \ j = 1,\ 2, ..., n.$$

($P(E|C) = \frac{P(C|E)P(E)}{P(C)} = \frac{0{,}6\times 0{,}25}{0{,}675} = 0,222$.)

Variáveis Aleatórias

Definição de Variável Aleatória

Variável aleatória

Chama-se variável aleatória (v.a), X, a toda a função que associa um número real a cada elemento do espaço de resultados Ω, ou seja,

$$\begin{aligned} X: \quad & \Omega \rightarrow \mathbb{R} \\ & \omega \rightarrow x = X(\omega) \end{aligned}$$

1. Uma variável aleatória é discreta $\rightsquigarrow$ se toma um nº finito ou infinito numerável de valores.
2. Uma variável aleatória é contínua $\rightsquigarrow$ se toma valores num certo intervalo.

Variável Aleatória Discreta

Exemplo: Consideremos a experiência aleatória que consiste no lançamento de uma moeda equilibrada três vezes consecutivas. Sendo F = "sair cara" e C = "sair coroa" o conjunto de todos os resultados possíveis é dado por

$$\begin{aligned}\Omega &= \{(F,F,F),\ (F,F,C),\ (F,C,F),\\ &\quad (C,F,F),\ (F,C,C),\ (C,F,C),\\ &\quad (C,C,F),\ (C,C,C)\}\end{aligned}$$

Neste espaço amostra, podemos definir uma variável aleatória

$$X = \text{número de vezes que saiu caras.}$$

Esta variável pode tomar valores 0, 1, 2 ou 3.

Distribuição de Probabilidade de uma Variável Aleatória Discreta

À função que a todo o x faz corresponder a respectiva probabilidade chama-se função massa de probabilidade (abreviadamente, f.m.p.) da v.a. X

e representa-se por $P(X = x_i)$, $f(x_i)$ ou simplesmente $f(x)$.

$X = x_i$	0	1	2	Soma
$P(X = x_i)$	2/9	5/9	2/9	1

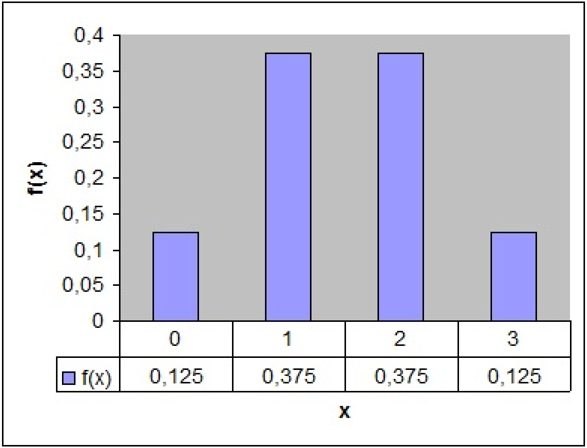
Propriedades

1. $0 \leqslant f(x_i) \leqslant 1$
2. $\sum_{i=1}^{n} f(x_i) = 1$

Exemplo (continuação):
$X =$número de vezes que saiu caras em três lançamentos de uma moeda,

- $P[X = 0] = P[(C \cap C \cap C)] = \frac{1}{8}$
- $P[X = 1] = P[(F \cap C \cap C) \cup (C \cap F \cap C) \cup (C \cap C \cap F)] = \frac{3}{8}$
- $P[X = 2] = P[(F \cap F \cap C) \cup (F \cap C \cap F) \cup (C \cap F \cap F)] = \frac{3}{8}$
- $P[X = 3] = P[(F \cap F \cap F)] = \frac{1}{8}$

x	0	1	2	3
$f(x)$	$1/8$	$3/8$	$3/8$	$1/8$


0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
f(x)
0
1
2
3
f(x)
0,125
0,375
0,375
0,125
x

Função de Distribuição

$$\boxed{F(x) = P(X \leqslant x) \Rightarrow \text{Função de Distribuição da v.a. X}}$$

Propriedades:

Para qualquer f.d. $F(x)$, dados os números reais $x, y,\ (y > x)$,

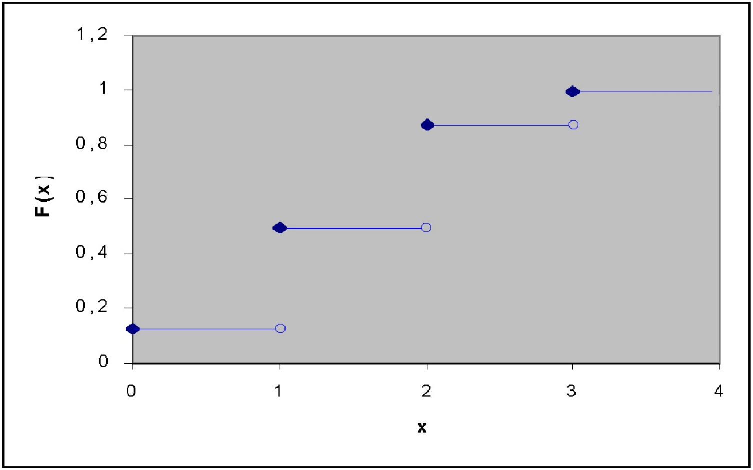
1. $0 \leqslant F(x) \leqslant 1$
2. $P(X < x) < P(X < y) \Longleftrightarrow F(x) < F(y)$
3. $P(X > x) = 1 - P(X \leqslant x) \Longleftrightarrow P(X > x) = 1 - F(x)$
4. $P(x < X \leqslant y) = P(X \leqslant y) - P(X \leqslant x) = F(y) - F(x)$.

Exemplo (continuação):

A função distribuição da variável X é dada por

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leqslant x < 1 \\ 4/8 & 1 \leqslant x < 2 \\ 7/8 & 2 \leqslant x < 3 \\ 1 & x \geqslant 3 \end{cases}$$

- $F(0) = P[X \leqslant 0] = \sum_{x \leqslant 0} f(x) = 1/8;$
- $F(1) = P[X \leqslant 1] = \sum_{x \leqslant 1} f(x) = 4/8;$
- $F(2) = P[X \leqslant 2] = \sum_{x \leqslant 2} f(x) = 7/8;$
- $F(3) = P[X \leqslant 3] = \sum_{x \leqslant 3} f(x) = 1;$


1,2
1
0,8
0,6
0,4
0,2
0
F(x)
0
1
2
3
4
x

Distribuição de Probabilidade de uma Variável Aleatória Contínua

X v.a. com f.d. $F(x)$. Se existir uma função real de variável real $f(x) > 0$, não-negativa, tal que

$$F(x) = \int_{-\infty}^{x} f(u)\, du \quad x \in \mathbb{R}$$

Neste caso, diz-se que a v.a. X é contínua.

$f(x) \Rightarrow$ Função Densidade de probabilidade (f.d.p.) da v.a. X

Propriedades

1. $f(x) \geqslant 0$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. $\int_{x}^{y} f(u)du = F(y) - F(x)$, $x < y$
4. $F'(x) = \dfrac{dF(x)}{dx} = f(x)$

Algumas situações típicas do cálculo de probabilidades envolvendo variáveis aleatórias contínuas:

1. $P[X = x] = 0;$
2. $P[X < a] = P[X \leqslant a] = F(a);$
3. $P[X > a] = 1 - P[X \leqslant a] = 1 - F(a);$
4. $P[a < X < b] = P[a \leqslant X < b] = P[a < X \leqslant b] =$
   $P[a \leqslant X \leqslant b] = F(b) - F(a), \; \forall\, a, b : a < b;$

Momentos — Variáveis Discretas

Valor Esperado (ou Média Populacional) da v.a. X discreta (que toma os valores $x_1, x_2 \ldots, x_n$):

$$\mu = E[X] = \sum_{i=1}^{n} x_i P(X = x_i)$$

$$\mu = E[X] = \sum_{i=1}^{n} x_i f(x_i)$$

Nota: A este parâmetro também se chama *momento de ordem 1 em relação à origem*.

Variância da v.a. X discreta (que toma os valores $x_1, x_2 \ldots, x_n$):

$$\sigma^2 = Var[X] = \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i)$$

$$\sigma^2 = Var[X] = E[X^2] - (E[X])^2 = \sum_{i=1}^{n} x_i^2 f(x_i) - \mu^2$$

Desvio-padrão da variável aleatória X (ou desvio-padrão populacional) $\rightsquigarrow$ $\sigma = \sqrt{\sigma^2}$

Nota: À Variância também se chama *momento de 2ª ordem em relação à média*.

Momentos — Variáveis Contínuas

Valor Esperado da variável aleatória X (que toma os valores em $\mathbb{R}$):

$$\mu = E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

Variância da v.a. X contínua (que toma os valores em $\mathbb{R}$):

$$\sigma^2 = Var[X] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

ou

$$\sigma^2 = E[X^2] - (E[X])^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2$$

Vectores Aleatórios

A $(X_1, X_2, \ldots, X_k)$ chama-se Vector Aleatório Multivariado

1. $F(x_1, x_2, \ldots, x_k) = P(X_1 \leqslant x_1, X_2 \leqslant x_2, \ldots, X_k \leqslant x_k)$ Função de Distribuição Conjunta do vector $(X_1, X_2, \ldots, X_k)$.
2. $f(x_1, x_2, \ldots, x_k)$ Função Densidade de Probabilidade Conjunta do vector $(X_1, X_2, \ldots, X_k)$.

Caso Particular $k = 2 \Rightarrow (X, Y)$ Vector Aleatório Bivariado ou Par aleatório

Função de probabilidade conjunta

A função de probabilidade conjunta do par aleatório (X, Y) é uma função $f(x, y)$ que associa a cada elemento de $\mathbb{R}^2$ uma probabilidade,

$$f_{(X,Y)}(x, y) = P[X = x, Y = y].$$

Verifica as seguintes propriedades:

1. $0 \leqslant f_{(X,Y)}(x, y) \leqslant 1, \ \forall (x, y) \in \mathbb{R}^2$;
2. $\sum_i \sum_j f_{(X,Y)}(x_i, y_j) = 1$.

Função de distribuição conjunta

Dada uma v.a. bidimensional (X, Y), discreta, a função de distribuição conjunta de (X, Y) é definida da seguinte forma:

$$F_{(X,Y)}(x, y) = P[X \leqslant x,\ Y \leqslant y] = \sum_{x_i \leqslant x} \sum_{y_j \leqslant y} f_{(X,Y)}(x_i, y_j),$$

e satisfaz as seguintes condições:

1. $\lim\limits_{x \to -\infty} F_{(X,Y)}(x, y) = 0$, com y fixo;
2. $\lim\limits_{y \to -\infty} F_{(X,Y)}(x, y) = 0$, com x fixo;
3. $\lim\limits_{x,y \to -\infty} F_{(X,Y)}(x, y) = 0$;
4. $\lim\limits_{x,y \to +\infty} F_{(X,Y)}(x, y) = 1$;
5. $0 \leqslant F_{(X,Y)}(x, y) \leqslant 1, \ \ \forall\ (x,\ y) \in \mathbb{R}^2$;
6. $F_{(X,Y)}(x_1, y_1) \leqslant F_{(X,Y)}(x_2, y_2), \ \ \forall\ x_1 < x_2,\ y_1 < y_2$.

Funções de probabilidade marginais

Dada uma v.a. bidimensional (X, Y), é possível definir:

função de probabilidade marginal de X,

$$f_X(x) = P[X = x, -\infty < Y < +\infty] = \sum_y f_{(X,Y)}(x, y)$$

função de probabilidade marginal de Y

$$f_Y(y) = P[-\infty < X < +\infty, Y = y] = \sum_x f_{(X,Y)}(x, y)$$

Funções de probabilidade condicionadas

Dada uma v.a. bidimensional discreta (X, Y), definimos

função de probabilidade de X condicionada a $Y = y$

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

função de probabilidade de Y condicionada a $X = x$

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

Valor esperado condicionado

Dada uma v.a. bidimensional discreta (X,Y), definimos

Valor esperado de X condicionado a $Y=y$

$$\mu_{X|Y=y} = E[X|Y=y] = \sum_{i=1}^{n} x_i f_{X|Y=y}(x_i)$$

Valor esperado de Y condicionado a $X=x$

$$\mu_{Y|X=x} = E[Y|X=x] = \sum_{i=1}^{n} y_i f_{Y|X=x}(y_i)$$

Variância condicionada

Dada uma v.a. bidimensional discreta (X,Y), definimos

Variância de X condicionada a $Y=y$

$$\sigma^2_{X|Y=y} = Var[X|Y=y] = \sum_{i=1}^{n} (x_i - \mu_{X|Y=y})^2 f_{X|Y=y}(x_i)$$

Variância de Y condicionada a $X=x$

$$\sigma^2_{Y|X=x} = Var[Y|X=x] = \sum_{i=1}^{n} (y_i - \mu_{Y|X=x})^2 f_{Y|X=x}(y_i)$$

Variáveis Independentes

X e Y, dizem-se **independentes** sse

$$f_{(X,Y)}\left(x,y\right)=f_X\left(x\right)\,.f_Y\left(y\right),\ \forall\ \left(x,y\right)\in\mathbb{R}^2.$$

Consequentemente:

$$f_{X|Y=y}\left(x\right)=f_X\left(x\right),,\ \forall\ x\in\mathbb{R}.$$

$$f_{Y|X=x}\left(y\right)=f_Y\left(y\right),,\ \forall\ y\in\mathbb{R}.$$

Propriedades dos Momentos

Propriedades da Média:

Sejam X e Y duas v.a.'s e k uma dada constante.

1. $E[k] = k$
2. $E[kX] = kE[X]$
3. $E[X \pm Y] = E[X] \pm E[Y]$
4. Se X e Y forem independentes então $E\left[X\,Y\right] = E\left[X\right].E\left[Y\right]$

Propriedades dos Momentos

Propriedades da Variância:

Sejam X e Y duas v.a.'s e k uma dada constante.

1. $\text{Var}[k] = 0$
2. $\text{Var}[kX] = k^2\text{Var}[X]$
3. $\text{Var}[k + X] = \text{Var}[X]$
4. $\text{Var}[X \pm Y] = \text{Var}[X] + \text{Var}[Y] \pm 2\text{Cov}(X, Y)$

Covariância

Define-se a Covariância entre X e Y por

$$\sigma_{X,Y} = \text{Cov}(X, Y) = E[(X-\mu_X)(Y-\mu_Y)] = \sum_i \sum_j (x_i - \mu_X)(y_j - \mu_Y) f_{(X\,Y)}(x_i, y_j)$$

Uma outra fórmula para calcular a covariância é

$$\sigma_{X,Y} = Cov(X,\ Y) = E[XY] - E[X]\,E[Y]\,.$$

onde

$$E[XY] = \sum_i \sum_j x_i y_j f_{(X,Y)}(x_i, y_j)$$

Avalia a influência que o afastamento de X, em relação à sua média, tem sobre o afastamento de Y à sua média.
Com base nesta medida, deduz-se ainda que:

$$E[X \times Y] = \text{Cov}(X, Y) + E[X] \times E[Y]$$

Propriedades dos Momentos

Propriedades da Covariância: $\sigma_{X,Y} = \text{Cov}(X,Y)$
Sejam X e Y duas v.a.'s e a, b, c e d duas constante.

- X e Y são variáveis independentes $\Rightarrow Cov\,(X,Y) = 0$
  (**Nota:** O recíproco pode não ser verdadeiro. O facto de $Cov\,(X,Y) = 0$ não implica a independência entre X e Y, pode existir uma relação não linear entre as variáveis.);
- $Cov\,(X,X) = Var\,[X]\,;$
- $Cov\,(aX+b, cY+d) = acCov\,(X,Y)$.

A covariância depende das unidades em que se exprimem as variáveis aleatórias X e Y. Sendo assim, é importante a introdução de um parâmetro para caracterizar a intensidade da ligação entre X e Y, mas que não dependa das unidades, como é o caso do coeficiente de correlação.

Coeficiente de correlação

O coeficiente de correlação é definido como:

$$\rho = \rho_{X,Y} = \frac{Cov\left[X, Y\right]}{\sqrt{Var\left[X\right].Var\left[Y\right]}} = \frac{\sigma_{X,Y}}{\sigma_X \, \sigma_Y}$$

Propriedades do coeficiente de correlação: Sejam X e Y duas variáveis aleatórias, e a, b, c e d constantes

- $-1 < \rho_{X,Y} < 1$;
- Se X e Y são variáveis aleatórias independentes, então $\rho_{X,Y} = 0$;
- O coeficiente de correlação não se altera quando as variáveis sofrem uma transformação linear positiva, ou seja,

$$\rho_{aX+b,cY+d} = \rho_{X,Y} \quad \text{se } ac > 0.$$