

UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE
COMPUTAÇÃO

Fundamentos de Probabilidade e Estatística para Ciência de Dados

Resumo das aulas do Prof. Dr. Francisco Rodrigues

Bruna Zamith Santos

Agosto de 2025

Sumário

1	Teoria dos Conjuntos	2
2	Experimento Aleatório	2
3	Conceitos de Probabilidade	2
3.1	Probabilidade Frequentista	3
3.2	Probabilidade de União de Dois Eventos	3
3.3	Probabilidade Condicional	3
3.4	Partições do Espaço Amostral	4
4	Teorema de Bayes	4
5	Variáveis Aleatórias	4
6	Função de Distribuição	5
7	Esperança	6
7.1	Variável Aleatória Discreta	6
7.2	Variável Aleatória Contínua	6
7.3	Função de uma Variável Aleatória	6
7.4	Propriedades	6
8	Momento	7
8.1	Momento Estatístico	7
8.2	Momento Central	7
9	Variância	7

1 Teoria dos Conjuntos

Sejam os conjuntos:

$$A = \{1, 2, 4, 9\}, \quad B = \{3, 7, 9\}$$

- União: $A \cup B = \{1, 2, 3, 4, 7, 9\}$
- Interseção: $A \cap B = \{7, 9\}$
- Complementar de B : $B^C = \{1, 2, 4\}$
- Complementar de A : $A^C = \{3, 7\}$
- Espaço amostral (Ω): É o conjunto de todos os resultados possíveis de um experimento aleatório. Exemplo: $\Omega = \{1, 2, 3, 4, 5, 6\}$, ao lançar um dado.
- Evento (A): É um subconjunto do espaço amostral. Exemplo: $A = \{2, 4, 6\}$
- Evento impossível (\emptyset): É um evento que nunca ocorre.
- Evento certo (Ω): É o evento que sempre ocorre.
- $A \cup B$: É o evento que ocorre se A ou B (ou ambos) ocorrerem.
- $A \cap B$: É o evento que ocorre se A e B ocorrerem ao mesmo tempo.
- A^C : É o evento que ocorre se A não ocorre.
- Eventos mutuamente exclusivos: Quando $A \cap B = \emptyset$.

2 Experimento Aleatório

Um experimento aleatório é um experimento que pode ser repetido inúmeras vezes sob as mesmas condições, sendo o seu resultado incerto.

3 Conceitos de Probabilidade

Sejam Ω o espaço amostral e A um evento em Ω . Então, uma função $P(\cdot)$ é denominada probabilidade se satisfaz:

- $0 \leq P(A) \leq 1, \forall A \in \Omega$
- $P(\Omega) = 1$
- Se A_1, A_2, \dots forem eventos mutuamente exclusivos, isto é, $A_i \cap A_j = \emptyset, \forall i \neq j$, então:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Se um experimento aleatório tiver $n(\Omega)$ resultados mutuamente exclusivos e igualmente possíveis, e se um evento A conter $n(A)$ desses resultados, a probabilidade de ocorrência desse evento é definida por:

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{|A|}{|\Omega|}$$

Sejam A e B eventos em um mesmo espaço amostral, então:

- $P(\emptyset) = 0$
- $P(A) = 1 - P(A^C)$
- Se $A \subseteq B$, então $P(A) \leq P(B)$

3.1 Probabilidade Frequentista

A probabilidade de um evento é igual à sua frequência de ocorrência em um grande número de experimentos:

$$P(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n}$$

, onde n_A é o número de vezes que o evento A ocorre em n experimentos.

3.2 Probabilidade de União de Dois Eventos

Para dois eventos A e B em um mesmo espaço amostral:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

3.3 Probabilidade Condicional

Sejam dois eventos A e B em um mesmo espaço amostral Ω . A probabilidade condicional de A dado que B ocorreu é definida por:

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

Assim, A e B são eventos independentes se, e somente se:

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

Ou equivalentemente:

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

3.4 Partições do Espaço Amostral

Os eventos B_1, B_2, \dots, B_n formam uma partição do espaço amostral Ω se:

- $B_i \cap B_j = \emptyset$, para $i \neq j$, com $i, j = 1, \dots, n$
- $\bigcup_{i=1}^n B_i = \Omega$
- $P(B_i) \geq 0$, para $i = 1, \dots, n$

Seja A um evento no espaço amostral Ω e seja B_1, \dots, B_n uma partição amostral de Ω . Podemos escrever A considerando tal partição:

$$A = \bigcup_{i=1}^n (A \cap B_i)$$

$$P(A) = P\left(\bigcup_{i=1}^n A \cap B_i\right) = \sum_{i=1}^n P(A \cap B_i)$$

Sejam B_1, B_2, \dots, B_n uma partição do espaço amostral Ω . Então, qualquer evento $A \subseteq \Omega$ pode ser escrito como:

$$P(A) = \sum_{i=1}^n P(A \mid B_i) \cdot P(B_i)$$

4 Teorema de Bayes

Sejam B_1, B_2, \dots, B_n uma partição do espaço amostral Ω , e A um evento com $P(A) > 0$, então:

$$P(B_i \mid A) = \frac{P(A \mid B_i) \cdot P(B_i)}{\sum_{j=1}^n P(A \mid B_j) \cdot P(B_j)}$$

E assim podemos definir:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

5 Variáveis Aleatórias

Suponha que lancemos dois dados. O espaço amostral associado ao experimento, sendo os eventos C : “sai uma cara” e R : “sai uma coroa”, é dado por:

$$\Omega = \{CC, CR, RC, RR\}$$

Uma possível variável aleatória associada ao experimento é definida por:

$$X = \text{“número de caras obtido no experimento”}$$

- Representamos variáveis aleatórias por letras maiúsculas (X, Y, Z), enquanto usamos letras minúsculas para indicar os valores das variáveis (x, y, z).
- Se o número de valores possíveis de uma variável aleatória for finito ou infinito enumerável, dizemos que é uma variável aleatória discreta.
- Caso contrário, é uma variável aleatória contínua.

A função que atribui a cada valor da variável aleatória sua respectiva probabilidade é chamada de distribuição de probabilidade:

$$P(X = x_i) = p(x_i) = p_i, \quad i = 1, 2, 3$$

A distribuição de probabilidade também é chamada de função massa de probabilidade. E temos que:

$$\sum_{i=1}^n P(X = x_i) = 1$$

Dizemos que X é uma variável aleatória contínua se existir uma função f denominada função densidade de probabilidade (fdp) que satisfaz:

- $f(x) \geq 0, \quad \forall x \in \mathbb{R}$
- $\int_{-\infty}^{\infty} f(x) dx = 1$
- $P(a \leq X \leq b) = \int_a^b f(x) dx, \quad -\infty < a < b < \infty$
- $f(x)$ é uma função com valores positivos e área unitária.

Seja X uma variável aleatória discreta ou contínua. A probabilidade condicional de que $X \in S$ dado que $X \in V$ é:

$$P(X \in S \mid X \in V) = \frac{P(X \in S \cap V)}{P(X \in V)}$$

, onde S e V são subconjuntos do espaço da variável.

6 Função de Distribuição

A função distribuição acumulada ou simplesmente função de distribuição de uma variável aleatória X é definida por:

$$F(x) = P(X \leq x)$$

Se discreta:

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

Se contínua:

$$F(x) = \int_{-\infty}^x f(t) dt$$

Propriedades da função de distribuição:

- $0 \leq F(x) \leq 1$, $F(x)$ é não decrescente,
- $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow +\infty} F(x) = 1$
- Caso discreto: $P(a < X \leq b) = F(b) - F(a)$
- Caso contínuo: $f(x) = \frac{dF(x)}{dx}$

7 Esperança

7.1 Variável Aleatória Discreta

Seja X uma variável aleatória discreta com distribuição de probabilidade $P(X = x_i)$. O valor esperado (ou esperança matemática) é:

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

7.2 Variável Aleatória Contínua

Seja X uma variável aleatória contínua com função densidade de probabilidade $f(x)$, então:

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

7.3 Função de uma Variável Aleatória

Seja $g(X)$ uma função de uma variável aleatória discreta X . Então:

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

Seja $g(X)$ uma função de variável contínua com densidade $f(x)$. Então:

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

7.4 Propriedades

- Se $X = c$, onde c é constante, então: $E[X] = E[c] = c$
- Se c é constante: $E[cX] = c \cdot E[X]$
- Então: $E[aX + b] = a \cdot E[X] + b$

8 Momento

8.1 Momento Estatístico

Seja X uma variável aleatória discreta com valores x_1, x_2, \dots, x_k . O momento de ordem n de X é:

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

Se X for contínua:

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

8.2 Momento Central

Seja X uma variável aleatória.

- Se X é discreta, o momento central de ordem n ($n > 0$) de X é:

$$\mu_n = E[(X - E[X])^n] = \sum_{x_i} (x_i - E[X])^n \cdot P(X = x_i)$$

- Se X é contínua, então:

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

9 Variância

A variância de uma variável aleatória X é definida por:

$$V(X) = \sigma^2 = E[(X - E(X))^2]$$

O desvio padrão é igual à raiz quadrada da variância:

$$\sigma = \sqrt{V(X)}$$

Temos a propriedade de que:

$$V(X) = E[X^2] - (E[X])^2$$

Seja $g(X)$ uma função da variável aleatória X . Então,

$$V[g(X)] = E[g(X)^2] - (E[g(X)])^2$$

Seja X uma variável aleatória e a e b constantes. Então,

$$V(aX + b) = a^2 V(X)$$