Estatística descritiva (1ed)

Apostilas de aula com exemplos em R

Djuri Vieira John Fabio Aguilar Sánchez
Luis Francisco Gómez López

2023-12-19

Índice

Bem-vindos 1

Prefácio 3

I Estatística e dados 5

1 Visão geral 7

2 Dados 9

II Visualização de dados 11

3 Tabelas 13

4 Gráficos 2D 15

III Medidas-resumo 17

5 Medidas de tendência central 19

6 Medidas de posição 21

7 Medidas de dispersão 23

8 Medidas de forma 25

IV Probabilidade 27

9 Experimento aleatório e espaço de probabilidade 29

10 Interpretações da Probabilidade 31

11 Consequências dos axiomas de probabilidade **33**

12 Independência e probabilidade condicional **35**

13 Regras de contagem **37**

V Variáveis aleatórias **39**

14 Distribuições de probabilidade discretas **41**

15 Distribuições de probabilidade contínuas **43**

Referências **45**

Apêndices **47**

A Introdução ao R **47**

B Teoria ingênua dos conjuntos **49**
B.1 Conjuntos 49
B.2 Operações com conjuntos 52
B.3 O conjunto vazio e o conjunto potência 55

Bem-vindos

Prefácio

Parte I

Estatística e dados

Capítulo 1

Visão geral

Capítulo 2

Dados

Parte II

Visualização de dados

Capítulo 3

Tabelas

Capítulo 4

Gráficos 2D

Parte III

Medidas-resumo

Capítulo 5

Medidas de tendência central

Capítulo 6

Medidas de posição

Capítulo 7

Medidas de dispersão

Capítulo 8

Medidas de forma

Parte IV

Probabilidade

Capítulo 9

Experimento aleatório e espaço de probabilidade

Capítulo 10

Interpretações da Probabilidade

Capítulo 11

Consequências dos axiomas de probabilidade

Capítulo 12

Independência e probabilidade condicional

Capítulo 13

Regras de contagem

Parte V

Variáveis aleatórias

Capítulo 14

Distribuições de probabilidade discretas

Capítulo 15

Distribuições de probabilidade contínuas

Referências

Halmos, Paul R. 2001. *Teoria ingenua dos conjuntos.* Rio de Janeiro: Editora Ciencia Moderna.

Apêndice A

Introdução ao R

Apêndice B

Teoria ingênua dos conjuntos

A teoria dos conjuntos é um ramo da matemática que estuda coleções denominadas conjuntos. Compreender a teoria dos conjuntos é fundamental, pois ela serve como base para a teoria da probabilidade, que por sua vez é crucial para o estudo da estatística. No entanto, um conhecimento básico da teoria dos conjuntos é suficiente para entender os princípios fundamentais da probabilidade e da estatística, sem a necessidade de um formalismo excessivo [1].

B.1 Conjuntos

Definição B.1 (Conjunto)**.** Um **conjunto** é uma coleção não ordenada de elementos únicos, ou pode ser uma coleção vazia, sem nenhum elemento.

Podemos denotar um conjunto usando uma letra arbitrária como A e descrevê-lo listando seus elementos entre chaves. Por exemplo, $A = \{1, 2\}$ é o conjunto cujos elementos são os números 1 e 2. Com base em Definição B.1 e na notação anterior, é importante fazer as seguintes observações:

- $A = \{1, 2\}$ e $B = \{2, 1\}$ são o mesmo conjunto porque conjuntos são coleções não ordenadas onde a ordem não é definida.
- $C = \{1, 1, 2, 2\}$ não está bem definido porque um conjunto contém elementos únicos, onde a especificação correta seria $C = \{1, 2\}$.
- Existe um conjunto, denotado por $\emptyset = \{\}$, chamado **conjunto vazio**, que não possui elementos.

[1] Para uma apresentação detalhada e clara da teoria dos conjuntos usando um sistema de axiomas, você pode consultar (Halmos 2001)

- É possível que os elementos de um conjunto sejam eles próprios conjuntos. Por exemplo, $D = \{\{1, 2\}, 3\}$ é um conjunto que contém o conjunto $\{1, 2\}$ e o número 3

O pacote R `sets` pode ser usado para ilustrar as ideias mencionadas acima para entender o conceito de conjunto. Primeiramente, podemos criar dois conjuntos e verificar se os dois conjuntos são iguais:

```
library(sets)
A <- set(1, 2)
A
#> {1, 2}
B <- set(2, 1)
B
#> {1, 2}
A == B
#> [1] TRUE
```

Também podemos verificar a propriedade de elementos únicos em um conjunto:

```
C <- set(1, 1, 2, 2)
C
#> {1, 2}
```

Além disso, podemos criar um conjunto vazio:

```
vazio <- set()
vazio
#> {}
```

Por último, podemos definir um conjunto cujos elementos podem ser conjuntos:

```
D <- set(A, 3)
D
#> {3, {1, 2}}
```

Definição B.2 (Relação de pertença)**.** Se a é um elemento de A, expressamos essa condição como $a \in A$. Caso contrário, expressamos que a não é um elemento de A com $a \notin A$. Na teoria dos conjuntos, $\in$ é conhecido como a relação *"é um elemento de"*.

Por exemplo, se $A = \{1, 2\}$ então $1 \in A$ e $3 \notin A$. Em R, podemos verificar se um elemento pertence a um conjunto usando o operador `%e%` do pacote `sets`, que representa a relação $\in$:

```
1 %e% A
#> [1] TRUE
3 %e% A
#> [1] FALSE
```

Às vezes, não é possível listar os elementos de um conjunto porque os elementos são infinitos ou porque não sabemos exatamente quais são. No entanto, se sabemos a propriedade que cada elemento deve ter, podemos usar uma notação matemática conhecida como **notação construtor de conjuntos** para descrever o conjunto. Essa notação é especificada como $\{x \in \Omega : P(x)\}$, onde x é um elemento genérico que pertence a Ω com a propriedade $P(x)$. Ω[2] é conhecido como **universo de discurso** e se refere ao conjunto que contém todos os elementos em consideração a partir dos quais o valor de x pode ser escolhido. Por exemplo:

- $E = \{x \in \Omega : x \text{ é um cachorro}\}$ onde E é o conjunto que contém todos os x que são cachorros. Nesse caso, $P(x)$ refere-se a ter a propriedade de ser um cachorro e Ω pode ser o conjunto dos seres vivos.
- $F = \{x \in \mathbb{N} : x > 5\}$ onde F contém todos os números maiores que 5 que são números naturais. Nesse caso, $P(x)$ refere-se a todos os x maiores que 5 e Ω é o conjunto dos números naturais.

Infelizmente, o R só pode manipular objetos que podem ser representados como números e que são finitos. Portanto, o uso de pacotes como `sets` não permite representar conjuntos como E ou F no R.

Definição B.3 (Subconjunto)**.** Um conjunto A é um **subconjunto** de um conjunto B se todos os elementos de A também sejam elementos de B. Essa condição pode ser expressa através da notação $A \subseteq B$. Por outro lado, se existir um elemento que pertença a A, mas não pertença a B, representamos essa situação como $A \nsubseteq B$. Na teoria dos conjuntos, $\subseteq$ é conhecido como a relação de *"inclusão"*.

A Figura B.1 ilustra a Definição B.3 usando um **diagrama de Venn**[3].

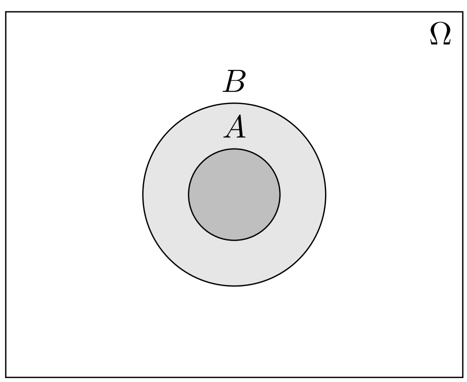

Figura B.1: $A \subseteq B$ representado por um diagrama de Venn onde Ω é o universo de discurso

[2] Ω é chamado Ômega

[3] Um diagrama de Venn é uma ferramenta visual que utiliza formas fechadas para representar conjuntos e ilustrar como seus elementos se relacionam.

Por exemplo, se $A = \{1, 2\}$ e $G = \{1, 2, 3\}$ então $A \subseteq G$ porque $1 \in A$, $1 \in G$, $2 \in A$ e $2 \in G$. No entanto, $G \nsubseteq A$ porque $3 \in G$ e $3 \notin A$. Em R, podemos verificar se um conjunto é um subconjunto de um conjunto usando o operador `<=` do pacote `sets`, que representa a relação $\subseteq$:

```
G <- set(1, 2, 3)
G
#> {1, 2, 3}
A <= G
#> [1] TRUE
G <= A
#> [1] FALSE
```

Definição B.4 (Igualdade de conjuntos)**.** Com base em Definição B.3 podemos estabelecer uma definição equivalente para a igualdade de conjuntos. Dois conjuntos, A e B, são considerados iguais, $A = B$, se e somente se $A \subseteq B$ e $B \subseteq A$. Em outras palavras, ambos os conjuntos devem conter exatamente os mesmos elementos para serem considerados iguais.

B.2 Operações com conjuntos

Definição B.5 (União de conjuntos)**.** A união de dois conjuntos A e B, denotada por $A \cup B$, é o conjunto de todos os elementos que estão em A **ou** em B. $A \cup B$ também é um conjunto e pode ser definido usando notação construtor de conjuntos como $A \cup B = \{x \in \Omega : x \in A \textbf{ ou } x \in B\}$

A Figura B.2 ilustra a Definição B.5 usando um diagrama de Venn.

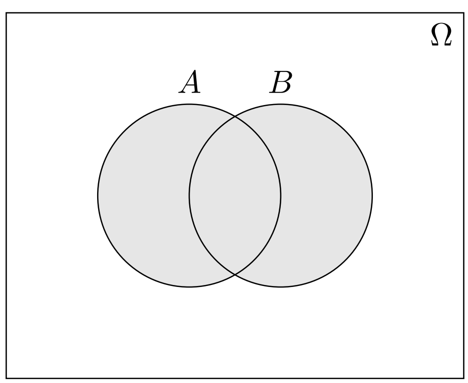

Figura B.2: $A \cup B$ representado por um diagrama de Venn onde Ω é o universo de discurso

Por exemplo, se $A = \{1, 2\}$ e $H = \{2, 3\}$, então $A \cup H = \{1, 2, 3\}$. Em R, utilizando o pacote `sets`, o operador `|` é utilizado para representar a união, $\cup$, de dois conjuntos da seguinte maneira:

```
H <- set(2, 3)
A | H
#> {1, 2, 3}
```

Definição B.6 (Interseção de conjuntos)**.** A interseção de dois conjuntos A e B, denotada por $A \cap B$, é o conjunto de todos os elementos que estão em A **e** em B. $A \cap B$ também é um conjunto e pode ser definido usando notação construtor de conjuntos como $A \cap B = \{x \in \Omega : x \in A \textbf{ e } x \in B\}$.

A Figura B.3 ilustra a Definição B.6 usando um diagrama de Venn.

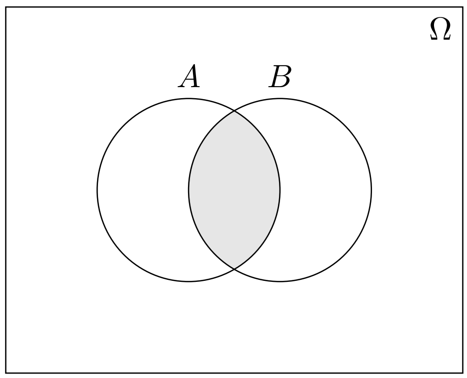

Figura B.3: $A \cap B$ representado por um diagrama de Venn onde Ω é o universo de discurso

Por exemplo, se $A = \{1, 2\}$ e $H = \{2, 3\}$, então $A \cap H = \{2\}$. Em R, utilizando o pacote `sets`, o operador `&` é utilizado para representar a interseção, $\cap$, de dois conjuntos da seguinte maneira:

```
A & H
#> {2}
```

Definição B.7 (Diferença de conjuntos)**.** A diferença de dois conjuntos A e B, denotada por $A \setminus B$, é o conjunto de todos os elementos que pertencem a A, mas não a B. $A \setminus B$ também é um conjunto e pode ser definido usando notação construtor de conjuntos como $A \setminus B = \{x \in \Omega : x \in A \text{ e } x \notin B\}$.

A Figura B.4 ilustra a Definição B.7 usando um diagrama de Venn.

Por exemplo, se $A = \{1, 2\}$ e $H = \{2, 3\}$, então $A \setminus H = \{1\}$ e $H \setminus A = \{3\}$. Em R, utilizando o pacote `sets`, o operador `-` é utilizado para representar a diferença, $\setminus$, de dois conjuntos da seguinte maneira:

```
A - H
#> {1}
H - A
#> {3}
```

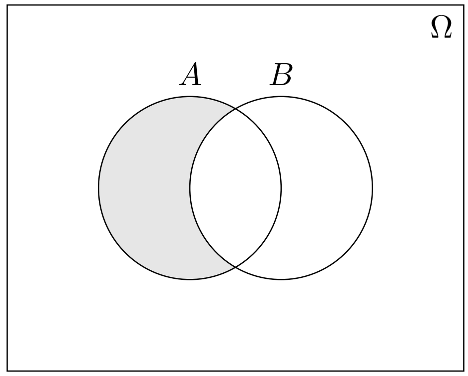

Figura B.4: $A \setminus B$ representado por um diagrama de Venn onde Ω é o universo de discurso

Definição B.8 (Complemento de um conjunto)**.** Se A e Ω são conjuntos, onde e Ω é o universo de discurso, o complemento de A, denotado por A^c, é o conjunto $\Omega \setminus A$. Ou seja, $A^c = \{x \in \Omega : x \notin A \text{ e } A \subseteq \Omega\}$.

A Figura B.5 ilustra a Definição B.8 usando um diagrama de Venn.

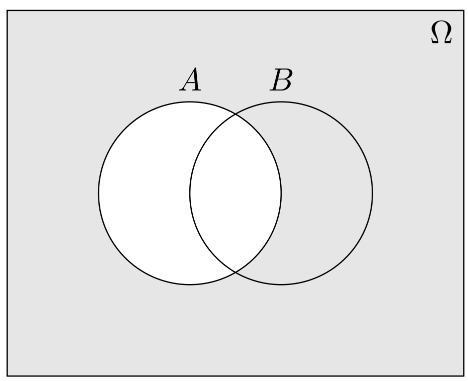

Figura B.5: A^c representado por um diagrama de Venn onde Ω é o universo de discurso

Por exemplo, se $A = \{1, 2\}$ e $\Omega = \{1, 2, 3, 4\}$, então $A^c = \{3, 4\}$. Em R, utilizando o pacote `sets`, podemos determinar A^c da seguinte maneira:

```
omega <- set(1, 2, 3, 4)
omega - A
#> {3, 4}
```

B.3 O conjunto vazio e o conjunto potência

O conjunto vazio, $\emptyset$, tem certas características que podem parecer contraintuitivas. Em primeiro lugar, só pode haver um conjunto vazio porque quaisquer dois conjuntos que não contenham nenhum elemento são idênticos. Conforme declarado na Definição B.4, conjuntos são considerados iguais se eles têm os mesmos elementos. Como ambos os conjuntos vazios não contêm elementos, eles são considerados o mesmo conjunto.

Outra propriedade aparentemente contraintuitiva é que o conjunto vazio é um subconjunto de todos os conjuntos. Se temos um conjunto A, e afirmamos que o conjunto vazio não é um subconjunto de A, denotado por $\emptyset \subseteq A$, então, de acordo com a Definição B.3, deve existir um elemento que pertença a $\emptyset$, mas não a A. No entanto, como o conjunto vazio não tem nenhum elemento, é impossível que um elemento pertença a $\emptyset$. Portanto, a única maneira de evitar uma contradição é aceitar que o conjunto vazio é um subconjunto de todos os conjuntos, denotado por $\emptyset \subseteq A$.

Podemos resumir os resultados acima da seguinte maneira:

Teorema B.1 (Unicidade do conjunto vazio)**.** *Só existe um conjunto vazio. Em outras palavras, se $\emptyset$ e $\emptyset'$ são ambos conjuntos vazios, então $\emptyset$ é igual a $\emptyset'$, $\emptyset = \emptyset'$.*

Teorema B.2 (Propriedade de subconjunto do conjunto vazio)**.** *O conjunto vazio é um subconjunto de todos os conjuntos. Para qualquer conjunto A, o conjunto vazio, $\emptyset$, é um subconjunto de A, $\emptyset \subseteq A$.*

Há também um conjunto chamado **conjunto das potências**. O conjunto das potências de um conjunto A, denotado como $\mathcal{P}(A)$, é um conjunto que contém todos os subconjuntos de A. Podemos defini-lo como:

Definição B.9 (Conjunto das potências)**.** Se A é um conjunto, então o conjunto que contém todos os subconjuntos de A, denotado como $\mathcal{P}(A)$, é definido como $\mathcal{P}(A) = \{B : B \subseteq A\}$.

Por exemplo, se $A = \{1, 2\}$ então $\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, A\} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$ porque $\emptyset \subseteq A$ pelo Teorema B.2, $\{1\} \subseteq A$, $\{2\} \subseteq A$ e $A = \{1, 2\} \subseteq A$. Em R, podemos construir o conjunto das potências de um conjunto A, $\mathcal{P}(A)$, como `2^A` utilizando o pacote `sets` da seguinte maneira:

```
potencia_A <- 2^A
potencia_A
#> {{}, {1}, {2}, {1, 2}}
```