

Estatística Descritiva II

Felipe Figueiredo

## Estatística Descritiva II

Medidas sumárias

Felipe Figueiredo

Instituto Nacional de Traumatologia e Ortopedia

## Sumário



- Medidas de Tendência Central
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
- Boxplot
  - Resumo

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência

Medidas de Dispersão

> ledidas de osição

Boxplot

esumo

4日ト 4周ト 4 三ト 4 三 ・ 夕久へ



Felipe Figueiredo

 Medidas sumárias resumem a informação contida nos dados em um pequeno conjunto de números.

- Medidas sumárias de populações se chamam parâmetros, e são representadas por letras gregas (μ, σ, etc).
- Medidas sumárias de amostras se chamam estatísticas e são representadas por letras comuns  $(\bar{x}, s, \text{ etc})$ .
- Geralmente trabalhamos com estatísticas descritivas.

Estatística Descritiva II Felipe

Medidas de Tendência Central

Medidas de Dispersão

Medidas de

Royplot



Estatística Descritiva II

Felipe Figueiredo

 Medidas sumárias resumem a informação contida nos dados em um pequeno conjunto de números.

- Medidas sumárias de populações se chamam parâmetros, e são representadas por letras gregas ( $\mu$ ,  $\sigma$ , etc).
- Medidas sumárias de amostras se chamam estatísticas
- Geralmente trabalhamos com estatísticas descritivas.



Estatística Descritiva II

Felipe Figueiredo

 Medidas sumárias resumem a informação contida nos dados em um pequeno conjunto de números.

- Medidas sumárias de populações se chamam parâmetros, e são representadas por letras gregas ( $\mu$ ,  $\sigma$ , etc).
- Medidas sumárias de amostras se chamam estatísticas e são representadas por letras comuns ( $\bar{x}$ , s, etc).
- Geralmente trabalhamos com estatísticas descritivas.



Estatística Descritiva II

Felipe Figueiredo

 Medidas sumárias resumem a informação contida nos dados em um pequeno conjunto de números.

- Medidas sumárias de populações se chamam parâmetros, e são representadas por letras gregas ( $\mu$ ,  $\sigma$ , etc).
- Medidas sumárias de amostras se chamam estatísticas e são representadas por letras comuns ( $\bar{x}$ , s, etc).
- Geralmente trabalhamos com estatísticas descritivas.

## Sumário



- Medidas de Tendência Central
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
- Boxplo

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Média

Mediana Moda Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



 A média (aritmética) leva em conta todos os dados disponíveis, e indica (em muitas situações) o ponto de maior acumulação de dados.

Notação: média populacional (μ)

$$\mu = \sum_{j=1}^{N} \frac{x_j}{N}$$

Notação: média amostral (x̄)

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$$

Nem sempre pertence ao dataset.

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Média Mediana Moda

Moda Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



 A média (aritmética) leva em conta todos os dados disponíveis, e indica (em muitas situações) o ponto de maior acumulação de dados.

Notação: média populacional (μ)

$$\mu = \sum_{j=1}^{N} \frac{x_j}{N}$$

Notação: média amostral (x̄)

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$$

Nem sempre pertence ao dataset.

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Média Mediana Moda

Comparação Medidas de

Medidas de

Posição

Boxplot



 A média (aritmética) leva em conta todos os dados disponíveis, e indica (em muitas situações) o ponto de maior acumulação de dados.

Notação: média populacional (μ)

$$\mu = \sum_{j=1}^{N} \frac{x_j}{N}$$

Notação: média amostral (x̄)

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$$

Nem sempre pertence ao dataset.

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Média Mediana Moda Comparação

Medidas de

Medidas de Posicão

Boxplot



 A média (aritmética) leva em conta todos os dados disponíveis, e indica (em muitas situações) o ponto de maior acumulação de dados.

Notação: média populacional (μ)

$$\mu = \sum_{j=1}^{N} \frac{x_j}{N}$$

Notação: média amostral (x̄)

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$$

Nem sempre pertence ao dataset.

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Média Mediana Moda Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Foram observados os seguintes níveis de colesterol de uma amostra de pacientes. Qual é o nível médio de colesterol nestes pacientes?

$$x_1 = 142$$

$$x_2 = 144$$

$$x_3 = 176$$

$$x_4 = 203$$

$$x_5 = 134$$

$$x_6 = 191$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{990}{6} = 165$$

#### Estatística Descritiva II

#### Felipe Figueiredo

Medidas de Tendência Central

Média Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot

## Sumário



- Medidas de Tendência Central
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
- 4 Boxplo

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Média Mediana

Moda Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Média

Mediana

Moda Comparação

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot

Resumo

#### 

## Definition

- Notação: M<sub>d</sub>
- Divide o dataset ao meio
- Costuma pertencer ao dataset



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Média Mediana

Moda

Comparação

Medidas de

Medidas de

Boxplot

Resumo

Definition

- Notação: M<sub>d</sub>
- Divide o dataset ao meio
- Costuma pertencer ao dataset



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Média Mediana

Moda

Comparação

Medidas de

Medidas de

Boxplot

Resumo

4 D > 4 A > 4 B > 4 B > 9 Q Q

## Definition

- Notação: M<sub>d</sub>
- Divide o dataset ao meio
- Costuma pertencer ao dataset



Estatística Descritiva II

Felipe Figueiredo

endência Central

Média Mediana

Moda

Comparação

Medidas de Dispersão

Medidas de

Boxplot

Resumo

### Definition

- Notação: M<sub>d</sub>
- Divide o dataset ao meio
- Costuma pertencer ao dataset



- Para se calcular a mediana, deve-se ordenar os dados.
- Encontrar o valor do meio se *n* for ímpar.
- Encontrar a média dos dois valores do meio se n for par.

### Example

Conforme no exemplo anterior

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



- Para se calcular a mediana, deve-se ordenar os dados.
- Encontrar o valor do meio se *n* for ímpar.
- Encontrar a média dos dois valores do meio se n for par.

### Example

Conforme no exemplo anterior

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparaçã

Medidas de

Medidas de

Boxplot



- Para se calcular a mediana, deve-se ordenar os dados.
- Encontrar o valor do meio se *n* for ímpar.
- Encontrar a média dos dois valores do meio se n for par.

### Example

Conforme no exemplo anterior

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparaçã

Medidas de Dispersão

Medidas de Posição

Boxplot



- Para se calcular a mediana, deve-se ordenar os dados.
- Encontrar o valor do meio se *n* for ímpar.
- Encontrar a média dos dois valores do meio se n for par.

### Example

Conforme no exemplo anterior

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central Média Mediana

Moda Comparação

Medidas de

Medidas de Posição

Boxplot



- Para se calcular a mediana, deve-se ordenar os dados.
- Encontrar o valor do meio se *n* for ímpar.
- Encontrar a média dos dois valores do meio se n for par.

### Example

### Conforme no exemplo anterior

$$x_1 = 142$$
  
 $x_2 = 144$ 

$$x_3 = 170$$

$$x_4 = 203$$

$$M_d = \frac{144 + 176}{2} = 160$$

#### Estatística Descritiva II

#### Felipe Figueiredo

ledidas de endência entral

Mediana Moda

Modidae (

Medidas de

Posição

Boxplot



- Para se calcular a mediana, deve-se ordenar os dados.
- Encontrar o valor do meio se n for ímpar.
- Encontrar a média dos dois valores do meio se n for par.

### Example

### Conforme no exemplo anterior

 $x_5 = 134$ 

 $x_1 = 142$ 

 $x_2 = 144$ 

 $x_3 = 176$ 

 $x_6 = 191$ 

 $x_4 = 203$ 

$$M_d = \frac{144 + 176}{2} = 160$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central <sup>Média</sup> Mediana

Moda Comparação

Medidas d

Dispersão

Posição

Boxplot



- Para se calcular a mediana, deve-se ordenar os dados.
- Encontrar o valor do meio se *n* for ímpar.
- Encontrar a média dos dois valores do meio se n for par.

### Example

### Conforme no exemplo anterior

$$x_5 = 134$$

$$x_1 = 142$$

$$x_2 = 144$$

$$x_3 = 176$$

$$x_6 = 191$$

$$x_4 = 203$$

$$M_d = \frac{144 + 176}{2} = 160$$

#### Estatística Descritiva II

#### Felipe Figueiredo

Medidas de Fendência Central Média Mediana

Moda Comparação

Medidas de

Medidas de Posicão

Boxplot

## Sumário



- Medidas de Tendência Central
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
- Boxplo

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Comparação

Medidas de Dispersão

Medidas de

Boxplot

Resumo

### 4 D > 4 A > 4 E > 4 E > 9 Q P

### Definition

- Notação: Mo
- Sempre pertence ao dataset.
- Não é necessariamente única: o dataset pode ser bimodal, ou mesmo multimodal.
- Não necessariamente existe: amodal



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Moda Comparação

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot

Resumo

4 D > 4 A > 4 B > 4 B > B 9 Q Q

### Definition

- Notação: M<sub>o</sub>
- Sempre pertence ao dataset.
- Não é necessariamente única: o dataset pode ser bimodal, ou mesmo multimodal.
- Não necessariamente existe: amodal



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

**woda** Comparação

Medidae d

Medidas de Dispersão

Medidas de Posição

Boxplot

Resumo

## Definition

- Notação: M<sub>o</sub>
- Sempre pertence ao dataset.
- Não é necessariamente única: o dataset pode se bimodal, ou mesmo multimodal.
- Não necessariamente existe: amodal



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

**woda** Comparação

Medidas di

Medidas de Dispersão

Medidas de Posição

Boxplot

Resumo

4 D > 4 A > 4 B > 4 B > B = 900

### Definition

- Notação: M<sub>o</sub>
- Sempre pertence ao dataset.
- Não é necessariamente única: o dataset pode ser bimodal, ou mesmo multimodal.
- Não necessariamente existe: amodal



Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Mediana Moda

Comparaçã

Medidas d

Medidas de Dispersão

Medidas de Posição

Boxplot

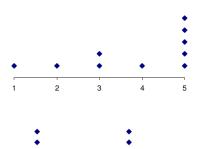
Resumo

4D > 4A > 4B > 4B > 4 B > 900

### Definition

- Notação: M<sub>o</sub>
- Sempre pertence ao dataset.
- Não é necessariamente única: o dataset pode ser bimodal, ou mesmo multimodal.
- Não necessariamente existe: amodal





Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central Média Mediana

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot

Resumo

Figura: Diagrama de pontos para dados (a) unimodal, (b) bimodal (Fonte: Reis, Reis, 2002)

5

7

3

## Sumário



- Medidas de Tendência Central
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
- Boxplot

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda Comparação

Comparação

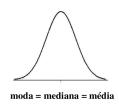
Medidas de Dispersão

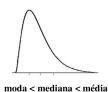
> Medidas de Posição

Boxplot

# Comparação entre as Medidas Centrais







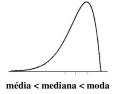


Figura: (a) Simétrica, (b) Assimétrica à esquerda, (c) Assimétrica à direita (Fonte: Reis, Reis 2002)

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central Média Mediana

Comparação

Medidas de Dispersão

> Medidas de Posição

Boxplot

## Robustez da Média



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Medidas de

Medidas de

Boxplot

Resumo

nesumo

- A média é mais usada, mas não é robusta.
- É distorcida na presença de outliers (valores discrepantes, extremos)

# Comparação entre as Medidas Centrais



### Example

### Considere o seguinte dataset

$$\{1,1,2,4,7\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

$$\bullet \ \mu = \frac{1+1+2+4+7}{5} = \frac{15}{5} = 3$$

- *M<sub>d</sub>* = 2
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Média Mediana

Moda Comparação

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot

# Comparação entre as Medidas Centrais



### Example

Considere o seguinte dataset

$$\{1, 1, 2, 4, 7\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

$$\bullet \ \mu = \frac{1+1+2+4+7}{5} = \frac{15}{5} = 3$$

- $M_d = 2$
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Média Mediana

Moda Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Considere o seguinte dataset

$$\{1,1,2,4,7\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

$$\bullet \ \mu = \frac{1+1+2+4+7}{5} = \frac{15}{5} = 3$$

- $M_d = 2$
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Medidas de

Medidas de Posição

Boxplot



## Example

Considere o seguinte dataset

$$\{1, 1, 2, 4, 7\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

$$\bullet \ \mu = \frac{1+1+2+4+7}{5} = \frac{15}{5} = 3$$

- $M_d = 2$
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Considere o seguinte dataset

$$\{1, 1, 2, 4, 7\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

- M<sub>d</sub> = 2
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Mediana

Moda Comparação

Comparagao

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Considere o seguinte dataset

$$\{1, 1, 2, 4, 7\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

$$\bullet \ \mu = \frac{1+1+2+4+7}{5} = \frac{15}{5} = 3$$

- M<sub>d</sub> = 2
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

edidas de endência entral

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Considere agora este outro dataset

$$\{1, 1, 2, 4, 32\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

- $M_d = 2$
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Média Mediana

Moda Comparação

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Considere agora este outro dataset

$$\{1, 1, 2, 4, 32\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

- $M_d = 2$
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

ledidas de endência entral

Média Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Considere agora este outro dataset

$$\{1,1,2,4,32\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

- $M_d = 2$
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Considere agora este outro dataset

$$\{1, 1, 2, 4, 32\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

$$\bullet \ \mu = \frac{1+1+2+4+32}{5} = \frac{40}{5} = 8$$

- M<sub>d</sub> = 2
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

edidas de endência entral

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Considere agora este outro dataset

$$\{1, 1, 2, 4, 32\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

$$\bullet \ \mu = \frac{1+1+2+4+32}{5} = \frac{40}{5} = 8$$

- M<sub>d</sub> = 2
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

edidas de endência entral

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



## Example

Considere agora este outro dataset

$$\{1, 1, 2, 4, 32\}$$

- N = 5
- As medidas descritivas centrais para estes dados são:

$$\bullet \ \mu = \frac{1+1+2+4+32}{5} = \frac{40}{5} = 8$$

- $M_d = 2$
- $M_0 = 1$

Estatística Descritiva II

Felipe Figueiredo

edidas de endência entral

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



#### Exercício

Um pesquisador observou as seguintes idades (anos) para uma amostra: 35, 33, 37, 33, 34.

#### Determine:

- A média amostral  $(\bar{x})$
- ② A mediana  $(M_d)$
- 3 A moda  $(M_o)$

## Solução

- $M_d = 34$
- $M_0 = 33$

Estatística Descritiva II

Felipe Figueiredo

Medidas de endência

Média Mediana

Moda Comparação

Comparação

Medidas de Dispersão

> Medidas de Posição

Boxplot



#### Exercício

Um pesquisador observou as seguintes idades (anos) para uma amostra: 35, 33, 37, 33, 34.

Determine:

- A média amostral  $(\bar{x})$
- ② A mediana  $(M_d)$
- $\odot$  A moda ( $M_o$ )

# Solução

$$\bar{x} = \frac{35 + 33 + 37 + 33 + 34}{5} = 34.4$$

$$M_d = 34$$

$$M_0 = 33$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



#### Exercício

Um pesquisador observou as seguintes idades (anos) para uma amostra: 35, 33, 37, 33, 34.

Determine:

- A média amostral  $(\bar{x})$
- 2 A mediana  $(M_d)$
- 3 A moda  $(M_o)$

# Solução

$$\bar{x} = \frac{35 + 33 + 37 + 33 + 34}{5} = 34.4$$

$$M_d = 34$$

$$M_o = 33$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot



#### Exercício

Um pesquisador observou as seguintes idades (anos) para uma amostra: 35, 33, 37, 33, 34.

Determine:

- A média amostral  $(\bar{x})$
- 2 A mediana  $(M_d)$
- 3 A moda  $(M_o)$

# Solução

$$\bar{x} = \frac{35 + 33 + 37 + 33 + 34}{5} = 34.4$$

**2** 
$$M_d = 34$$

$$M_0 = 33$$

Estatística Descritiva II

Felipe Figueiredo

ledidas de endência entral

Mediana Moda

Comparação

Medidas de Dispersão

> Medidas de Posição

Boxplot



#### Exercício

Um pesquisador observou as seguintes idades (anos) para uma amostra: 35, 33, 37, 33, 34.

Determine:

- A média amostral  $(\bar{x})$
- 2 A mediana  $(M_d)$
- $\odot$  A moda ( $M_o$ )

# Solução

$$\bar{x} = \frac{35 + 33 + 37 + 33 + 34}{5} = 34.4$$

**2** 
$$M_d = 34$$

$$M_0 = 33$$

Estatística Descritiva II

Felipe Figueiredo

ledidas de endência entral

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de Posição

Boxplot

## Resumo



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Medidas de

Medidas de

tolaxo

Resumo

#### Média mais usual

- Mediana na presença de outliers
- Moda quando a distribuição das frequências for bimodal ou multimodal.

## Resumo



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Medidas de

Medidas de

Boxplot

Resumo

4 D > 4 A > 4 B > 4 B > 9 Q P

- Média mais usual
- Mediana na presença de outliers
- Moda quando a distribuição das frequências fo bimodal ou multimodal.

## Resumo



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Mediana Moda

Comparação

Medidas de Dispersão

Medidas de

Boxplot

Resumo

#### 4D > 4A > 4B > 4B > B 990

Média mais usual

- Mediana na presença de outliers
- Moda quando a distribuição das frequências for bimodal ou multimodal.

# Variabilidade em Medições



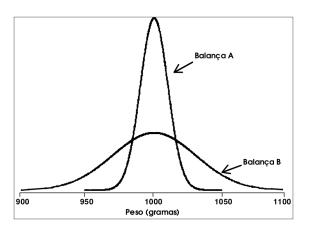


Figura: Variabilidade da medição de uma esfera metálica de 1000g. Balança A, "imprecisão" de 50g, balança B, "imprecisão" de 100g (Fonte: Reis, Reis, 2002)

Estatística Descritiva II

Felipe Figueiredo

ledidas de endência entral

#### Medidas de Dispersão

Amplitude
Desvios em relação
à media
Variância
Desvio Padrão

Exercícios Coeficiente de Variação

viedidas de Posição

Boxplot



# Sumário



- - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude

  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- - Quartis
  - Percentis e Decis

Estatística Descritiva II

Felipe Figueiredo

Amplitude



# **Amplitude**



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

#### Amplitude

Desvios em relaç à media

à media Variância

Desvio Padrão Exercícios

Variação

Posição

Boxplot

Resumo

A amplitude dos dados identifica o intervalo de ocorrência de todos os dados observados

•  $A = x_{max} - x_{min}$ 

Example

Seja o dataset {21, 12, 20, 4, 75, 40, 39, 63} Então, a amplitude é: A = 75 - 4 = 71

# Amplitude



Estatística Descritiva II Felipe Figueiredo

Amplitude

A amplitude dos dados identifica o intervalo de ocorrência de todos os dados observados

 $\bullet$   $A = X_{max} - X_{min}$ 

## Example

Seja o dataset {21, 12, 20, 4, 75, 40, 39, 63} Então, a amplitude é: A = 75 - 4 = 71

# Sumário



- Medidas de Tendência Centra
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- 2 Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
- Boxplo

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência

Medidas de Dispersão Amplitude

Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot





Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude
Desvios em relação
à media

Variância Desvio Padrão Exercícios

Medidas de Posição

Boxplot

- Uma maneira de entender a variabilidade do dataset é analisar os desvios em relação à média.
- Cada desvio é a diferença entre o valor do dado e a média.



Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão

Amplitude

Desvios em relação

Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot

- Uma maneira de entender a variabilidade do dataset é analisar os desvios em relação à média.
- Cada desvio é a diferença entre o valor do dado e a média.



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude

Desvios em relaç

Desvios em relação à media

Variância Desvio Padrão Exercícios Coeficiente de

Medidas de Posicão

Boxplot

Resumo

#### Mas os desvios...

- 1 são tão numerosos quanto os dados
- 2 têm sinal (direção do desvio)
- 3 têm soma nula



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude Desvios em relação

Desvios em relaçã à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot

Resumo

#### Mas os desvios...

- 1 são tão numerosos quanto os dados
- 2 têm sinal (direção do desvio)
- têm soma nula



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude

Desvios em relação

Desvios em relaçã à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posicão

Boxplot

Resumo

#### Mas os desvios...

- 1 são tão numerosos quanto os dados
- 2 têm sinal (direção do desvio)
- têm soma nula



## Example

 $\{1, 2, 3, 4, 5\}$ 

Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão Amplitude

Desvios em relação à media Variância

Desvio Padrão Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot



# Example $\{1, 2, 3, 4, 5\}$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão Amplitude Desvios em relação

à media
Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot



## Example

 $\{1, 2, 3, 4, 5\}$ 

- N = 5
- $\bar{x} = 3$

 $D_1 = 1 - 3 = -2$   $D_2 = 2 - 3 = -1$   $D_3 = 3 - 3 = 0$ 

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência

Medidas de Dispersão Amplitude Desvios em relação

à media
Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot



## Example

 $\{1, 2, 3, 4, 5\}$ 

- N = 5
- $\mathbf{o} \ \bar{x} = 3$

 $D_1 = 1 - 3 = -2$   $D_2 = 2 - 3 = -1$   $D_3 = 3 - 3 = 0$ 

①  $D_4 = 4 - 3 = 1$ 

 $D_5 = 5 - 3 = 2$ 

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência

Medidas de Dispersão <sub>Amplitude</sub>

Desvios em relação à media Variância Desvio Padrão

Desvio Padrão Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot



## Example

 $\{1, 2, 3, 4, 5\}$ 

- N = 5
- $\bar{x} = 3$

 $D_1 = 1 - 3 = -2$   $D_2 = 2 - 3 = -1$   $D_2 = 3 - 3 = 0$ 

 $\mathbf{O}$   $D_4 = 4 - 3 = 1$ 

 $D_5 = 5 - 3 = 2$ 

Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão Amplitude Desvios em relação

à media
Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot



## Example

$$\{1,2,3,4,5\}$$

• 
$$\bar{x} = 3$$

$$D_1 = 1 - 3 = -2$$

2 
$$D_2 = 2 - 3 = -$$

3 
$$D_3 = 3 - 3 = 0$$

$$\mathbf{\Phi} \ D_4 = 4 - 3 = 1$$

**6** 
$$D_5 = 5 - 3 = 2$$

#### Estatística Descritiva II

#### Felipe Figueiredo

#### Medidas de Tendência

#### Medidas de Dispersão

#### Amplitude Desvios em relação à media

# Variância Desvio Padrão Exercícios Coeficiente de Variação

#### Medidas de Posição

Boxplot



## Example

• N = 5•  $\bar{x} = 3$ 

$$\{1, 2, 3, 4, 5\}$$

$$\mathbf{0} \ D_1 = 1 - 3 = -2$$

$$D_2 = 2 - 3 = -3$$

**3** 
$$D_3 = 3 - 3 = 0$$

**4** 
$$D_4 = 4 - 3 = 1$$

**6** 
$$D_5 = 5 - 3 = 2$$

#### Estatística Descritiva II

#### Felipe Figueiredo

#### Medidas de Tendência

#### Medidas de Dispersão

# Amplitude Desvios em relação

#### Variância Desvio Padrão Exercícios

à media

# Medidas de

Boxplot



## Example

$$\{1, 2, 3, 4, 5\}$$

$$\bullet$$
  $\bar{x}=3$ 

$$O_1 = 1 - 3 = -2$$

$$Q D_2 = 2 - 3 = -1$$

3 
$$D_3 = 3 - 3 = 0$$

$$\mathbf{\Phi} \ D_4 = 4 - 3 = 1$$

**6** 
$$D_5 = 5 - 3 = 2$$

#### Estatística Descritiva II

#### Felipe Figueiredo

#### Medidas de Tendência

#### Medidas de Dispersão

# Amplitude Desvios em relação à media

#### Variância Desvio Padrão Exercícios

# Medidas de

Boxplot

# Desvios em relação à média



### Example

$$\{1, 2, 3, 4, 5\}$$

• 
$$N = 5$$

$$\bullet$$
  $\bar{x}=3$ 

$$\mathbf{0} \ D_1 = 1 - 3 = -2$$

$$Q D_2 = 2 - 3 = -1$$

3 
$$D_3 = 3 - 3 = 0$$

$$\mathbf{0}$$
  $D_4 = 4 - 3 = 1$ 

**6** 
$$D_5 = 5 - 3 = 2$$

#### Estatística Descritiva II

#### Felipe Figueiredo

Desvios em relação à media

# Desvios em relação à média



## Example

$$\{1, 2, 3, 4, 5\}$$

• 
$$N = 5$$

$$\bullet$$
  $\bar{x}=3$ 

$$D_2 = 2 - 3 = -1$$

3 
$$D_3 = 3 - 3 = 0$$

$$\mathbf{0} \ D_4 = 4 - 3 = 1$$

**6** 
$$D_5 = 5 - 3 = 2$$

#### Estatística Descritiva II

#### Felipe Figueiredo

Medidas de Tendência

Medidas de Dispersão

Amplitude
Desvios em relação

à media Variância Desvio Padrão

Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot

# Desvios em relação à média



# Example

$$\{1,2,3,4,5\}$$

$$\bar{x}=3$$

$$D_1 = 1 - 3 = -2$$

$$D_2 = 2 - 3 = -1$$

**3** 
$$D_3 = 3 - 3 = 0$$

$$\mathbf{Q} D_4 = 4 - 3 = 1$$

**6** 
$$D_5 = 5 - 3 = 2$$

#### Estatística Descritiva II

#### Felipe Figueiredo

### Medidas de Tendência

# Medidas de

#### Dispersão Amplitude

#### Desvios em relação à media

# Variância Desvio Padrão Exercícios Coeficiente de

#### Medidas de Posição

Boxplot

# Soma dos desvios



#### Estatística Descritiva II

Felipe Figueiredo

#### Medidas de Tendência Central

#### Medidas de Dispersão

Amplitude

#### Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

#### Medidas de Posição

Boxplot

Resumo

# Example

Somando tudo:

$$D_1 + D_2 + D_3 + D_4 + D_5 =$$

$$(-2) + (-1) + 0 + 1 + 2 = 0$$

# Como proceder?



 Como extrair alguma informação útil (e sumária!) dos desvios?

Problema: sinais

Pergunta

Como tirar os sinais dos desvios?

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão

Amplitude

Desvios em relação

à media
Variância
Desvio Padrão
Exercícios
Conficiente do

Medidas de

Boxplot

# Como proceder?



 Como extrair alguma informação útil (e sumária!) dos desvios?

Problema: sinais

Pergunta

Como tirar os sinais dos desvios?

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão

Amplitude

Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot

# Como proceder?



 Como extrair alguma informação útil (e sumária!) dos desvios?

Problema: sinais

### Pergunta

Como tirar os sinais dos desvios?

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão Amplitude

Desvios em relação à media

Desvio Padrão Exercícios Coeficiente de

Medidas de

Boxplot



Tomando-se o módulo dos desvios temos:

### **Definition**

Desvio médio absoluto (MAD) é a média dos desvios absolutos

- É uma medida de dispersão robusta (pouco influenciada por outliers)
- Módulo não tem boas propriedades matemáticas (analíticas e algébricas).
- Pouco usado para inferência (baixa eficiência)

Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão

Amplitude Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot



Tomando-se o módulo dos desvios temos:

### Definition

Desvio médio absoluto (MAD) é a média dos desvios absolutos

- É uma medida de dispersão robusta (pouco influenciada por outliers)
- Módulo não tem boas propriedades matemáticas (analíticas e algébricas).
- Pouco usado para inferência (baixa eficiência)

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot





Tomando-se o módulo dos desvios temos:

### Definition

Desvio médio absoluto (MAD) é a média dos desvios absolutos

- É uma medida de dispersão robusta (pouco influenciada por outliers)
- Módulo não tem boas propriedades matemáticas (analíticas e algébricas).
- Pouco usado para inferência (baixa eficiência)

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude

Desvios em relação

à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posicão

Boxplot





Tomando-se o módulo dos desvios temos:

### Definition

Desvio médio absoluto (MAD) é a média dos desvios absolutos

- É uma medida de dispersão robusta (pouco influenciada por outliers)
- Módulo não tem boas propriedades matemáticas (analíticas e algébricas).
- Pouco usado para inferência (baixa eficiência)

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude

Desvios em relação

à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot



### Example

$$\{1, 2, 3, 4, 5\}, \bar{x} = 3$$

 $|U_2| = |2 - 3| = 1$ 

 $|D_3| = |3-3| = 0$ 

 $|D_4| = |4-3| = 1$ 

 $|D_5| = |5 - 3| = 2$ 

Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão

Amplitude Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot



### Example

$$\{1, 2, 3, 4, 5\}, \bar{x} = 3$$

$$|D_1| = |1 - 3| = 2$$

$$|D_2| = |2 - 3| = 1$$

$$|D_3| = |3-3| = 0$$

$$|D_4| = |4-3| = 1$$

**6** 
$$|D_5| = |5 - 3| = 2$$

$$MAD = \frac{\sum D_i}{5} = 2$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão

Amplitude

Desvios em relação à media Variância

Desvio Padrão Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot



### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$|D_1| = |1 - 3| = 2$$

$$|D_2| = |2 - 3| =$$

$$|D_3| = |3-3| = 0$$

$$|D_4| = |4-3| = 1$$

**5** 
$$|D_5| = |5 - 3| = 2$$

$$MAD = \frac{\sum D_i}{5} = 2$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot



### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$|D_1| = |1 - 3| = 2$$

$$|D_2| = |2-3| = 1$$

$$|D_3| = |3-3| = 0$$

$$|D_4| = |4-3| = 1$$

**5** 
$$|D_5| = |5 - 3| = 2$$

$$MAD = \frac{\sum D_i}{5} = 2$$

#### Estatística Descritiva II

#### Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot



### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$|D_1| = |1 - 3| = 2$$

$$|D_2| = |2-3| = 1$$

$$|D_3| = |3-3| = 0$$

$$|D_4| = |4-3| = 1$$

**6** 
$$|D_5| = |5 - 3| = 2$$

$$MAD = \frac{\sum D_i}{5} = 2$$

#### Estatística Descritiva II

#### Felipe Figueiredo

#### Medidas de Tendência Central

#### Medidas de Dispersão

#### Desvios em relação à media

# Variância Desvio Padrão Exercícios Coeficiente de

#### Medidas de Posição

Boxplot



### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$|D_1| = |1 - 3| = 2$$

$$|D_2| = |2-3| = 1$$

$$|D_3| = |3-3| = 0$$

$$|D_4| = |4-3| = 1$$

**6** 
$$|D_5| = |5 - 3| = 2$$

$$MAD = \frac{\sum D_i}{5} = 2$$

#### Estatística Descritiva II

#### Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

#### Amplitude Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot



### Example

$$\{1, 2, 3, 4, 5\}, \bar{x} = 3$$

$$|D_1| = |1 - 3| = 2$$

$$|D_2| = |2-3| = 1$$

$$|D_3| = |3-3| = 0$$

$$|D_4| = |4-3| = 1$$

**6** 
$$|D_5| = |5-3| = 2$$

$$MAD = \frac{\sum D_i}{5} = 3$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude

Desvios em relação

à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot



Estatística Descritiva II Felipe Figueiredo

### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$|D_1| = |1 - 3| = 2$$

$$|D_2| = |2-3| = 1$$

$$|D_3| = |3-3| = 0$$

$$|D_4| = |4-3| = 1$$

**6** 
$$|D_5| = |5 - 3| = 2$$

#### Tendência Central

Medidas de Dispersão

Amplitude
Desvios em relação
à media

Variância
Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot

Resumo

 $MAD = \frac{\sum D_i}{5} = 2$ 

# Uma proposta "melhor"



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão Amplitude

Desvios em relação à media

Variância Desvio Padrão Exercícios

Medidas de

Boxplot

- Uma outra maneira de eliminar os sinais é elevar ao quadrado cada desvio.
- Preserva boas propriedades matemáticas
- Calculando a média dos quadrados dos desvios (desvios quadráticos) temos ...

# Uma proposta "melhor"



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude

Desvios em relação

à media

a media
Variância
Decuio Padrão

Desvio Padrão
Exercícios
Coeficiente de

Medidas de Posição

Boxplot

- Uma outra maneira de eliminar os sinais é elevar ao quadrado cada desvio.
- Preserva boas propriedades matemáticas
- Calculando a média dos quadrados dos desvios (desvios quadráticos) temos ...

# Uma proposta "melhor"



Estatística Descritiva II

Felipe Figueiredo

Desvios em relação

à media

- Uma outra maneira de eliminar os sinais é elevar ao quadrado cada desvio.
- Preserva boas propriedades matemáticas
- Calculando a média dos quadrados dos desvios (desvios quadráticos) temos ...

# Sumário



Estatística Descritiva II

Felipe

Figueiredo

- Medidas de Tendência Centra
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis

Boxplo



Tendência Central

Dispersão

Amplitude

Desvios em relação

Variância

Desvio Padrão Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot





### **Definition**

### A variância é a média dos desvios quadráticos.

Variância populacional

$$\sigma^2 = \frac{\sum (x_j - \mu)^2}{N}$$

Variância amostral

$$s^{2} = \frac{\sum (x_{i} - \bar{x})^{2}}{n - 1}$$

- Conveniente do ponto de vista matemático (boas propriedades algébricas e analíticas).
- Unidade quadrática, pouco intuitiva para interpretação de resultados

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão Amplitude

à media Variância

Desvio Padrão Exercícios

Coeficiente de Variação

Posição

Boxplot

Resumo

◆ロト→同ト→三ト ● 夕久で



### **Definition**

A variância é a média dos desvios quadráticos.

Variância populacional

$$\sigma^2 = \frac{\sum (x_j - \mu)^2}{N}$$

Variância amostral

$$s^{2} = \frac{\sum (x_{i} - \bar{x})^{2}}{n - 1}$$

- Conveniente do ponto de vista matemático (boas propriedades algébricas e analíticas).
- Unidade quadrática, pouco intuitiva para interpretação de resultados

Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Dispersão
Amplitude
Desvios em relação

à media Variância

Desvio Padrão Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot

Resumo

4 D > 4 B > 4 B > 4 B > 9 Q P



### **Definition**

A variância é a média dos desvios quadráticos.

Variância populacional

$$\sigma^2 = \frac{\sum (x_j - \mu)^2}{N}$$

Variância amostral

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

- Conveniente do ponto de vista matemático (boas propriedades algébricas e analíticas).
- Unidade quadrática, pouco intuitiva para interpretação de resultados

Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão Amplitude

Amplitude Desvios em relação à media

#### Variância

Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot

Resumo

◆ロト→同ト→三ト ● 夕久で



### **Definition**

A variância é a média dos desvios quadráticos.

Variância populacional

$$\sigma^2 = \frac{\sum (x_j - \mu)^2}{N}$$

Variância amostral

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

- Conveniente do ponto de vista matemático (boas propriedades algébricas e analíticas).
- Unidade quadrática, pouco intuitiva para interpretação de resultados.

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão Amplitude

Amplitude Desvios em relação a media

### Variância

Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot



### **Definition**

A variância é a média dos desvios quadráticos.

Variância populacional

$$\sigma^2 = \frac{\sum (x_j - \mu)^2}{N}$$

Variância amostral

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

- Conveniente do ponto de vista matemático (boas propriedades algébricas e analíticas).
- Unidade quadrática, pouco intuitiva para interpretação de resultados.

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão

Amplitude

Desvios em relação

à media Variância

Desvio Padrão Exercícios

Variação Medidas de

Boxplot

Roxbio

Resumo

◆ロト→同ト→三ト ● 夕久で



## Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

 $D_1^2 = (1-3)^2 = (-2)^2 = 4$   $D_2^2 = (2-3)^2 = (-1)^2 = 1$ 

 $D_2 = (2 - 3)^2 - (2 - 3)^2 = 0$ 

 $D_4^2 = (4-3)^2 = 1^2 = 1$ 

**6**  $D_{\rm f}^2 = (5-3)^2 = 2^2 = 4$ 

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão Amplitude

Amplitude Desvios em relação à media

Variância

Desvio Padrão Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot



## Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$D_1^2 = (1-3)^2 = (-2)^2 = 4$$

$$D_2^2 = (2-3)^2 = (-1)^2 = 1$$

$$D_3^2 = (3-3)^2 = 0^2 = 0$$

$$D_4^2 = (4-3)^2 = 1^2 = 1$$

**6** 
$$D_5^2 = (5-3)^2 = 2^2 = 4$$

$$s^2 = \frac{\sum D_i^2}{4} = 2.5$$

#### Estatística Descritiva II

#### Felipe Figueiredo

#### Medidas de Tendência Central

### Medidas de Dispersão

#### Amplitude Desvios em relação à media

### Variância

Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot



## Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$D_1^2 = (1-3)^2 = (-2)^2 = 4$$

3 
$$D_3^2 = (3-3)^2 = 0^2 = 0$$

$$D_4^2 = (4-3)^2 = 1^2 = 1$$

**6** 
$$D_5^2 = (5-3)^2 = 2^2 = 4$$

$$g^2 = \frac{\sum D_i^2}{1} = 2.5$$

#### Estatística Descritiva II

#### Felipe Figueiredo

#### Medidas de Tendência Central

### Medidas de Dispersão

# Amplitude Desvios em relação à media

### Variância

#### Desvio Padrão Exercícios Coeficiente de Variação

#### Medidas de Posição

#### Boxplot



Estatística Descritiva II Felipe Figueiredo

### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$D_1^2 = (1-3)^2 = (-2)^2 = 4$$

$$2 D_2^2 = (2-3)^2 = (-1)^2 = 1$$

$$D_3^2 = (3-3)^2 = 0^2 = 0$$

$$D_4^2 = (4-3)^2 = 1^2 = 1$$

**6** 
$$D_5^2 = (5-3)^2 = 2^2 = 4$$

Medidas de Dispersão

Amplitude
Desvios em relação

Desvios em relação à media Variância

### Variância

Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot



### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$D_1^2 = (1-3)^2 = (-2)^2 = 4$$

$$2 D_2^2 = (2-3)^2 = (-1)^2 = 1$$

$$D_3^2 = (3-3)^2 = 0^2 = 0$$

$$D_4^2 = (4-3)^2 = 1^2 = 1$$

**6** 
$$D_5^2 = (5-3)^2 = 2^2 = 4$$

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão

Amplitude
Desvios em relação à media

### Variância

Desvio Padrão Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot



Estatística Descritiva II Felipe

### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$D_1^2 = (1-3)^2 = (-2)^2 = 4$$

$$D_3^2 = (3-3)^2 = 0^2 = 0$$

$$D_5^2 = (5-3)^2 = 2^2 = 4$$

Variância



### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$D_1^2 = (1-3)^2 = (-2)^2 = 4$$

$$D_2^2 = (2-3)^2 = (-1)^2 = 1$$

$$D_3^2 = (3-3)^2 = 0^2 = 0$$

$$D_5^2 = (5-3)^2 = 2^2 = 4$$

#### Estatística Descritiva II Felipe

Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude
Desvios em relação
à media

### Variância

Desvio Padrão Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot



Estatística Descritiva II Felipe Figueiredo

### Example

$$\{1,2,3,4,5\}, \bar{x}=3$$

$$D_1^2 = (1-3)^2 = (-2)^2 = 4$$

$$D_2^2 = (2-3)^2 = (-1)^2 = 1$$

$$D_3^2 = (3-3)^2 = 0^2 = 0$$

$$D_4^2 = (4-3)^2 = 1^2 = 1$$

**6** 
$$D_5^2 = (5-3)^2 = 2^2 = 4$$

Variância

 $s^2 = \frac{\sum D_i^2}{4} = 2.5$ 

## Sumário



- Medidas de Tendência Centra
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
- Boxplo

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência

Dispersão

Amplitude

Desvios em relação

a media Variância Desvio Padrão

Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot





### Definition

### O desvio padrão é a raiz quadrada da variância.

Desvio padrão populacional

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

Desvio padrão amostral

$$s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Amplitude
Desvios em relação
à media

Variância

Desvio Padrão Exercícios

Coeficiente de Variação

Medidas de Posição

Boxplot



### Definition

O desvio padrão é a raiz quadrada da variância.

Desvio padrão populacional

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

Desvio padrão amostral

$$s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão

Amplitude

à media Variância

Variância Desvio Padrão

Exercícios Coeficiente de

Medidas de

Boxplot



### Definition

O desvio padrão é a raiz quadrada da variância.

Desvio padrão populacional

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

Desvio padrão amostral

$$s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão
Amplitude
Desvios em relação

à media Variância

Desvio Padrão Exercícios

Coeficiente de Variação

Medidas de Posição

Boxplot



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão
Amplitude
Desvios em relação

à media
Variância
Desvio Padrão

Exercícios Coeficiente de

Medidas de Posição

Boxplot

Resumo

### É a medida mais usada, por estar na mesma escala (unidade) dos dados.

- Boas propriedades matemáticas
- Boas propriedades como estimador (Inferência



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão
Amplitude
Desvios em relação

à media Variância

Desvio Padrão Exercícios

Coeficiente de Variação

Medidas de Posição

Boxplot

- É a medida mais usada, por estar na mesma escala (unidade) dos dados.
- Boas propriedades matemáticas
- Boas propriedades como estimador (Inferência



Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Dispersão

Amplitude

Desvios em relação

à media

Variância

Desvio Padrão

Exercícios
Coeficiente de

Medidas de Posição

Boxplot

- É a medida mais usada, por estar na mesma escala (unidade) dos dados.
- Boas propriedades matemáticas
- Boas propriedades como estimador (Inferência)



### Estatística Descritiva II

Felipe Figueiredo

### Desvio Padrão

### Example

$$\{1, 2, 3, 4, 5\}, \bar{x} = 3$$
  
 $s^2 = 2.5$   
 $s = \sqrt{s^2} = \sqrt{2.5} = 1.58$ 

## Sumário



- Medidas de Tendência Centra
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
- Boxplo

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência

Medidas de Dispersão Amplitude Desvios em relacão

Variância Desvio Padrão Exercícios

Coeficiente de Variação

Medidas de Posição

Boxplot





### Exercício

Um pesquisador observou as seguintes idades (anos) para uma amostra: 35, 33, 37, 33, 34.

### Determine:

- A variância amostral (s²)
- O desvio padrão amostral (s)

### Solução

Lembrando que  $\bar{x} = 34.4$ , temos:

 $(35-34.4)^2+(33-34.4)^2$ 

 $=\frac{0.36+1.96+6.76+1.96+0.16}{4}=2.8$ 

 $s = \sqrt{2.8} = 1.67$ 

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência

Medidas de Dispersão Amplitude Desvios em relação

> à media Variância Desvio Padrão

Exercícios Coeficiente de

Medidas de

Boxplot



### Exercício

Um pesquisador observou as seguintes idades (anos) para uma amostra: 35, 33, 37, 33, 34.

Determine:

- A variância amostral (s²)
- O desvio padrão amostral (s)

### Solução

Lembrando que  $\bar{x} = 34.4$ , temos:

$$s^{2} = \frac{(35 - 34.4)^{2} + (33 - 34.4)^{2} + \dots}{5 - 1}$$
$$= \frac{0.36 + 1.96 + 6.76 + 1.96 + 0.16}{4} = 2.8$$

2 
$$s = \sqrt{2.8} = 1.67$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

> esvios em relação nedia riância

Desvio Padrão Exercícios

Coeficiente de Variação

Medidas de Posição

Boxplo



### Exercício

Um pesquisador observou as seguintes idades (anos) para uma amostra: 35, 33, 37, 33, 34.

Determine:

- A variância amostral (s²)
- O desvio padrão amostral (s)

### Solução

Lembrando que  $\bar{x} = 34.4$ , temos:

$$s^2 = \frac{(35 - 34.4)^2 + (33 - 34.4)^2 + \dots}{5 - 1}$$

$$= \frac{0.36 + 1.96 + 6.76 + 1.96 + 0.16}{4} = 2.8$$

2  $s = \sqrt{2.8} = 1.67$ 

Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência

Medidas de Dispersão

> mplitude lesvios em relação media ariância

Exercícios
Coeficiente de

Coeficiente de Variação

Medidas de Posição

Boxplo



### Exercício

Um pesquisador observou as seguintes idades (anos) para uma amostra: 35, 33, 37, 33, 34.

Determine:

- 1 A variância amostral (s²)
- O desvio padrão amostral (s)

### Solução

Lembrando que  $\bar{x} = 34.4$ , temos:

$$s^2 = \frac{(35 - 34.4)^2 + (33 - 34.4)^2 + \dots}{5 - 1}$$

$$= \frac{0.36 + 1.96 + 6.76 + 1.96 + 0.16}{4} = 2.8$$

2 
$$s = \sqrt{2.8} = 1.67$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência

Medidas de Dispersão Amplitude

> svios em relação nedia riância svio Padrão

Desvio Padrão Exercícios

Coeficiente de Variação

Medidas de Posição

Boxplot

## Sumário



- Medidas de Tendência Centra
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
  - Boxplot



Felipe Figueiredo

Medidas de Tendência

Dispersão
Amplitude
Desvios em relação
à media

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot





## Definition

$$CV = \frac{\sigma}{\mu}$$

- Normaliza a variabilidade em relação à média
- Permite comparar a variabilidade de datasets não relacionados (mesmo que não usem a mesma unidade)
- Só deve ser usado para grandezas em escala (i.e. possui um "zero" não arbitrário, ou "zero absoluto")

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão Amplitude Desvios em relaçã à media Variância Desvio Padrão

Medidas de Posição

Coeficiente de Variação

Boxplot



### Definition

$$CV = \frac{\sigma}{\mu}$$

- Normaliza a variabilidade em relação à média
- Permite comparar a variabilidade de datasets não relacionados (mesmo que não usem a mesma unidade)
- Só deve ser usado para grandezas em escala (i.e. possui um "zero" não arbitrário, ou "zero absoluto")

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão Amplitude Desvios em relar à media Variância Desvio Padrão Exercícios

> Medidas de Posição

Boxplot

Variação



### Definition

$$CV = \frac{\sigma}{\mu}$$

- Normaliza a variabilidade em relação à média
- Permite comparar a variabilidade de datasets não relacionados (mesmo que não usem a mesma unidade)
- Só deve ser usado para grandezas em escala (i.e. possui um "zero" não arbitrário, ou "zero absoluto")

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão
Amplitude
Desvios em relação à media
Variância

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot



### Definition

$$CV = \frac{\sigma}{\mu}$$

- Normaliza a variabilidade em relação à média
- Permite comparar a variabilidade de datasets não relacionados (mesmo que não usem a mesma unidade)
- Só deve ser usado para grandezas em escala (i.e. possui um "zero" não arbitrário, ou "zero absoluto")

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão Amplitude Desvios em relaç à media Variância

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot



### Example

x =Estatura e y =Perímetro abdominal.

x =	y =
181.2	76.3
173.7	66.7
169.0	73.3
184.1	74.8
174.4	82.7

79.6

172.6

# maior variabilidade?

Qual das duas amostras tem

- Calcular a média x
- 2 Calcular a variância  $s^2 = \frac{\sum (x_i \bar{x})^2}{2}$
- 3 Calcular o desvio padrão  $s = \sqrt{s^2}$
- $CV = \frac{s}{\bar{v}}$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão
Amplitude
Desvios em relação
à media

Variância Desvio Padrão Exercícios Coeficiente de Variação

Medidas de Posição

Boxplot



### Example

x =Estatura e y =Perímetro abdominal.

X =	y =
181.2	76.3
173.7	66.7
169.0	73.3
184.1	74.8
174.4	82.7
172.6	79.6

Qual das duas amostras tem maior variabilidade?

**1** Calcular a média  $\bar{x}$ 

2 Calcular a variância 
$$s^2 = \frac{\sum (x_i - \bar{x})^2}{2}$$

3 Calcular o desvio padrão  $s = \sqrt{s^2}$ 

$$CV = \frac{s}{\bar{y}}$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão
Amplitude
Desvios em relação
à media

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot



### Example

x =Estatura e y =Perímetro abdominal.

x =	y =
181.2	76.3
173.7	66.7
169.0	73.3
184.1	74.8

174.4 82.7 172.6 79.6 Qual das duas amostras tem maior variabilidade?

- Calcular a média  $\bar{x}$
- 2 Calcular a variância  $s^2 = \frac{\sum (x_i \bar{x})^2}{n 1}$ 
  - 3 Calcular o desvio padrão  $s = \sqrt{s^2}$
- **4**  $CV = \frac{5}{5}$

### Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão
Amplitude
Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot



### Example

x =Estatura e y =Perímetro abdominal.

x =	y =
181.2	76.3
173.7	66.7
169.0	73.3
184.1	74.8
174.4	82.7
172.6	79.6

Qual das duas amostras tem maior variabilidade?

- Calcular a média  $\bar{x}$
- 2 Calcular a variância  $s^2 = \frac{\sum (x_i \bar{x})^2}{n 1}$
- 3 Calcular o desvio padrão  $s = \sqrt{s^2}$

$$CV = \frac{S}{V}$$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão
Amplitude
Desvios em relação
à media

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot



### Example

x =Estatura e y =Perímetro abdominal.

x =	y =
181.2	76.3
173.7	66.7
169.0	73.3
184.1	74.8
174.4	82.7
172.6	79.6

Qual das duas amostras tem maior variabilidade?

- Calcular a média  $\bar{x}$
- 2 Calcular a variância  $s^2 = \frac{\sum (x_i \bar{x})^2}{n 1}$
- 3 Calcular o desvio padrão  $s = \sqrt{s^2}$
- $OV = \frac{s}{\bar{v}}$

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão
Amplitude
Desvios em relação à media
Variância

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplo



### Example

*x* = Estatura e *y* = Perímetro abdominal

abaomma.		
x =	y =	
181.2	76.3	
173.7	66.7	
169.0	73.3	
184.1	74.8	
174.4	82.7	

172.6 79.6

$$\bar{x}=175.8~s_x=5.7$$
  
 $\bar{y}=75.7~s_y=5.5$   
 $CV_x=3.24\%$   
 $CV_y=7.27\%$   
Resposta: O perímetro  
abdominal tem maior  
variabilidade que a altura.

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Dispersão

Amplitude

Desvios em relação à media

Variância
Desvio Padrão
Exercícios
Coeficiente de
Variação

Medidas de Posição

Boxplot

## Medidas de Posição



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas d Dispersão

Medidas de Posição Quartis

Percentis e Decis

oxplot

Resumo

 Permitem estabelecer informações quantitativas relativas à ordem dos dados

## Medidas de Posição



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas d Dispersão

Medidas de Posição Quartis

Quartis Percentis e Decis

Boxplot

Resumo

 Permitem estabelecer informações quantitativas relativas à ordem dos dados

## Sumário



- Medidas de Tendência Centra
  - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- Medidas de Dispersão
  - Amplitude
  - Desvios em relação à media
  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis
- 4 Boxplot

Estatística Descritiva II

> Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de Posição

Quartis

ercentis e Dec

oxplot

resumo



### Definition

Dividem o dataset em quatro partes, cada uma com 25% dos dados

- Q<sub>1</sub>, primeiro quartil, representa os primeiros 25% dos dados
- Q<sub>2</sub>, segundo quartil, representa os primeiros 50% dos dados
- Q<sub>3</sub>, terceiro quartil, representa os primeiros 75% dos dados

## Pergunta

O que podemos dizer sobre o segundo quartil  $(Q_2)$ ?

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de

Quartis

ercentis e De

Boxplot



### Definition

Dividem o dataset em quatro partes, cada uma com 25% dos dados

- Q<sub>1</sub>, primeiro quartil, representa os primeiros 25% dos dados
- $Q_2$ , segundo quartil, representa os primeiros 50% dos
- $Q_3$ , terceiro quartil, representa os primeiros 75% dos

Estatística Descritiva II

Felipe Figueiredo

Quartie



### Definition

Dividem o dataset em quatro partes, cada uma com 25% dos dados

- Q<sub>1</sub>, primeiro quartil, representa os primeiros 25% dos dados
- Q<sub>2</sub>, segundo quartil, representa os primeiros 50% dos dados
- Q<sub>3</sub>, terceiro quartil, representa os primeiros 75% dos dados

Pergunta

O que podemos dizer sobre o segundo quartil  $(Q_2)$ ?

Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão

Medidas de

Quartis

ercentis e De

oxplot



### Definition

Dividem o dataset em quatro partes, cada uma com 25% dos dados

- Q<sub>1</sub>, primeiro quartil, representa os primeiros 25% dos dados
- Q<sub>2</sub>, segundo quartil, representa os primeiros 50% dos dados
- Q<sub>3</sub>, terceiro quartil, representa os primeiros 75% dos dados

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de

Quartis

ercentis e De

oxplot

lesumo

### Pergunta

O que podemos dizer sobre o segundo quartil  $(Q_2)$ ?



### Definition

Dividem o dataset em quatro partes, cada uma com 25% dos dados

- Q<sub>1</sub>, primeiro quartil, representa os primeiros 25% dos dados
- Q<sub>2</sub>, segundo quartil, representa os primeiros 50% dos dados
- Q<sub>3</sub>, terceiro quartil, representa os primeiros 75% dos dados

## Pergunta

O que podemos dizer sobre o segundo quartil  $(Q_2)$ ?

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de

Quartis

rcentis e Decis

oxplot



Estatística Descritiva II

Felipe Figueiredo

Quartie

### Example

Os pesos de 102 bebês nascidos em uma certa maternidade ao longo de um ano foram anotados e ordenados. Um certo bebê ocupa o  $Q_3$  deste dataset.

Isto significa que aproximadamente 75% dos bebês



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de

Quartis

Percentis e De

Boxplot

Resumo

### Example

Os pesos de 102 bebês nascidos em uma certa maternidade ao longo de um ano foram anotados e ordenados. Um certo bebê ocupa o  $Q_3$  deste dataset.

 Isto significa que aproximadamente 75% dos bebês nascidos nesta maternidade tem peso menor ou igual a ele (Mazel Tov!).

## Sumário



- - Média
  - Mediana
  - Moda
  - Comparação entre as Medidas Centrais
- - Amplitude

  - Variância
  - Desvio Padrão
  - Exercícios
  - Coeficiente de Variação
- Medidas de Posição
  - Quartis
  - Percentis e Decis

Estatística Descritiva II

Felipe Figueiredo

Percentis e Decis

### Percentis e Decis



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

> Posição Quartis

Percentis e Decis

Boxplot

Resumo

### Definition

O percentil de ordem k (onde k é qualquer valor entre 0 e 100), denotado por  $P_k$ , é o valor tal que k% dos valores do dataset são menores ou iguais a ele.

- Generalizam a idéia dos quartis
- Dividem o dataset em 100 partes
- Maior granularidade na ordem
- Decis: dividem o dataset em 10 partes



#### Estatística Descritiva II Felipe Figueiredo

Medidas de Tendência

Medidas de Dispersão

> Posição Quartis

Percentis e Decis

Boxplot

Resumo

### **Definition**

- Generalizam a idéia dos quartis
- Dividem o dataset em 100 partes
- Maior granularidade na ordem
- Decis: dividem o dataset em 10 partes



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Posição

Quartis

Percentis e Decis

Percentis e Dec

Boxplot

Resumo

### **Definition**

- Generalizam a idéia dos quartis
- Dividem o dataset em 100 partes
- Maior granularidade na ordem
- Decis: dividem o dataset em 10 partes



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

COSIÇÃO

Quartis

Percentis e Decis

Percentis e Dec

Boxplot

Resumo

### Definition

- Generalizam a idéia dos quartis
- Dividem o dataset em 100 partes
- Maior granularidade na ordem
- Decis: dividem o dataset em 10 partes



#### Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Posição

Quartis

Percentis e Decis

Povolot

Resumo

### Definition

- Generalizam a idéia dos quartis
- Dividem o dataset em 100 partes
- Maior granularidade na ordem
- Decis: dividem o dataset em 10 partes



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

> Medidas de Posição

Boxplot

Resumo

Resumo

### Gráfico que ilustra a dispersão dos dados pelos quartis

- Retângulo que representa a Distância Interquartílica (DQ = Q<sub>3</sub> - Q<sub>1</sub>)
- Barra interna que representa a mediana ( $Q_2$ )
- Limite superior (barra vertical): Q<sub>3</sub> + 1.5 · DQ
- Limite inferior (barra vertical):  $Q_1 1.5 \cdot DQ$
- Outliers como pontos, círculos ou estrelas
- Conveniente para comparar vários grupos ou amostras



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

> Medidas de Posição

Boxplot

Resumo

### 4 □ → 4 個 → 4 目 → 4 目 → 9 Q (~)

- Gráfico que ilustra a dispersão dos dados pelos quartis
- Retângulo que representa a Distância Interquartílica (DQ = Q<sub>3</sub> - Q<sub>1</sub>)
- Barra interna que representa a mediana ( $Q_2$ )
- Limite superior (barra vertical): Q<sub>3</sub> + 1.5 · DQ
- Limite inferior (barra vertical):  $Q_1 1.5 \cdot DQ$
- Outliers como pontos, círculos ou estrelas
- Conveniente para comparar vários grupos ou amostras



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas d Dispersão

> Medidas de Posição

Boxplot

Resumo

nesumo

- Gráfico que ilustra a dispersão dos dados pelos guartis
- Retângulo que representa a Distância Interquartílica  $(DQ = Q_3 Q_1)$
- Barra interna que representa a mediana (Q2)
- Limite superior (barra vertical):  $Q_3 + 1.5 \cdot DQ$
- Limite inferior (barra vertical):  $Q_1 1.5 \cdot DQ$
- Outliers como pontos, círculos ou estrelas
- Conveniente para comparar vários grupos ou amostras



Estatística Descritiva II

Felipe Figueiredo

**Boxplot** 

- Gráfico que ilustra a dispersão dos dados pelos guartis
- Retângulo que representa a Distância Interguartílica  $(DQ = Q_3 - Q_1)$
- Barra interna que representa a mediana  $(Q_2)$
- Limite superior (barra vertical): Q<sub>3</sub> + 1.5 ⋅ DQ
- Limite inferior (barra vertical):  $Q_1 1.5 \cdot DQ$
- Conveniente para comparar vários grupos ou amostras



#### Estatística Descritiva II

Felipe Figueiredo

**Boxplot** 

- Gráfico que ilustra a dispersão dos dados pelos guartis
- Retângulo que representa a Distância Interguartílica  $(DQ = Q_3 - Q_1)$
- Barra interna que representa a mediana  $(Q_2)$
- Limite superior (barra vertical): Q<sub>3</sub> + 1.5 ⋅ DQ
- Limite inferior (barra vertical): Q₁ 1.5 · DQ
- Conveniente para comparar vários grupos ou amostras



Estatística Descritiva II

Felipe Figueiredo

**Boxplot** 

Gráfico que ilustra a dispersão dos dados pelos guartis

- Retângulo que representa a Distância Interguartílica  $(DQ = Q_3 - Q_1)$
- Barra interna que representa a mediana  $(Q_2)$
- Limite superior (barra vertical): Q<sub>3</sub> + 1.5 ⋅ DQ
- Limite inferior (barra vertical):  $Q_1 1.5 \cdot DQ$
- Outliers como pontos, círculos ou estrelas
- Conveniente para comparar vários grupos ou amostras



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de Posição

Boxplot

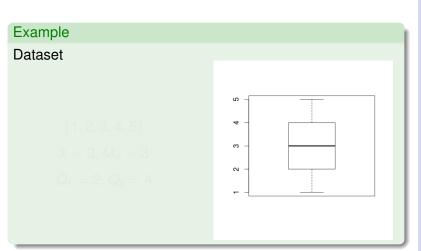
.

Resumo

Gráfico que ilustra a dispersão dos dados pelos quartis

- Retângulo que representa a Distância Interquartílica  $(DQ = Q_3 Q_1)$
- Barra interna que representa a mediana (Q2)
- Limite superior (barra vertical): Q<sub>3</sub> + 1.5 ⋅ DQ
- Limite inferior (barra vertical):  $Q_1 1.5 \cdot DQ$
- Outliers como pontos, círculos ou estrelas
- Conveniente para comparar vários grupos ou amostras





Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão

Medidas de

Boxplot

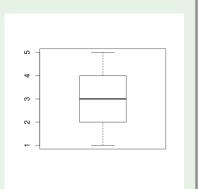
Boxbiot



### Example

### **Dataset**

 $\{1, 2, 3, 4, 5\}$  $\bar{x} = 3, M_d = 3$  $Q_1 = 2, Q_3 = 4$ 



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de

Boxplot



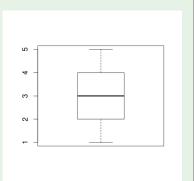
## Example

### Dataset

$$\{1, 2, 3, 4, 5\}$$

$$\bar{x} = 3, M_d = 3$$

$$Q_1 = 2, Q_3 = 4$$



#### Estatística Descritiva II

#### Felipe Figueiredo

#### Medidas de Tendência Central

## Medidas de

## Dispersão

### Posição

#### Boxplot



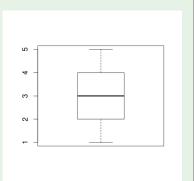
## Example

### Dataset

$$\{1, 2, 3, 4, 5\}$$

$$\bar{x} = 3, M_d = 3$$

$$Q_1 = 2, Q_3 = 4$$



#### Estatística Descritiva II

#### Felipe Figueiredo

#### Medidas de Tendência Central

## Medidas de

## Dispersão

### Posição

#### Boxplot





Exemplo do colesterol

 $X_1 = 142$ 

 $x_2 = 144$ 

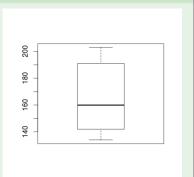
 $x_3 = 176$ 

 $x_4 = 203$ 

 $x_5 = 134$ 

 $x_{\rm s} = 191$ 

 $\bar{x} = 165 \ M_{2} = 160$ 



Estatística Descritiva II

Felipe Figueiredo

Medidas de endência Central

Medidas de Dispersão

Medidas de

Boxplot



## Example

### Exemplo do colesterol

 $x_1 = 142$ 

 $x_2 = 144$ 

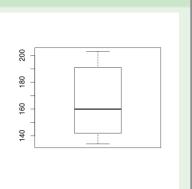
 $x_3 = 176$ 

×3 = 170

 $x_5 = 134$ 

 $x_6 = 191$ 

 $\bar{x} = 165, M_d = 160$ 



Estatística Descritiva II

Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão

Medidas de

Boxplot



## Example

### Exemplo do colesterol

 $x_1 = 142$ 

 $x_2 = 144$ 

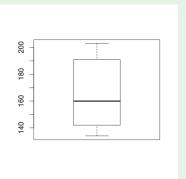
 $x_3 = 176$ 

 $x_4 = 203$ 

 $x_5 = 134$ 

 $x_6 = 191$ 

 $\bar{x} = 165, M_d = 160$ 



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de

Boxplot



## Example

### Exemplo do colesterol

 $x_1 = 142$ 

 $x_2 = 144$ 

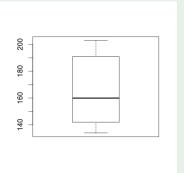
 $x_3 = 176$ 

 $x_4 = 203$ 

 $x_5 = 134$ 

 $x_6 = 191$ 

 $\bar{x} = 165, M_d = 160$ 



#### Estatística Descritiva II

#### Felipe Figueiredo

Medidas de Tendência Central

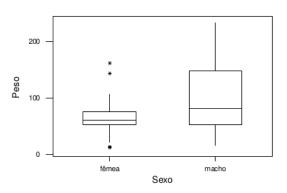
Medidas de Dispersão

Medidas de

Boxplot

## Boxplot: duas amostras





Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de Posição

Boxplot

Resumo

Figura: Boxplots para dois grupos de dados (Fonte: Reis, Reis, 2002)



Estatística Descritiva II

Felipe

Figueiredo

Ao iniciar sua Análise Exploratória de Dados (EDA), você pode visualizar sua amostra com:

- Resumo dos cinco números
  - Valor mínimo
  - Primeiro quartil Q<sub>1</sub>
  - Mediana (e/ou média)
  - Terceiro quartil Q<sub>3</sub>
  - Valor máximo
- 2 Boxplo



Medidas de

Dispersão

osição

oxplot





Estatística Descritiva II

Felipe

Figueiredo

Ao iniciar sua Análise Exploratória de Dados (EDA), você pode visualizar sua amostra com:

- Resumo dos cinco números
  - Valor mínimo
  - Primeiro quartil Q<sub>1</sub>

  - Valor máximo







Ao iniciar sua Análise Exploratória de Dados (EDA), você pode visualizar sua amostra com:

- Resumo dos cinco números
  - Valor mínimo
  - Primeiro quartil Q1
  - Mediana (e/ou média)
  - Terceiro quartil Q
  - Valor máximo
- 2 Boxplo



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

osição

loxplot



Ao iniciar sua Análise Exploratória de Dados (EDA), você pode visualizar sua amostra com:

- Resumo dos cinco números
  - Valor mínimo
  - Primeiro quartil Q<sub>1</sub>
  - Mediana (e/ou média)
  - Terceiro quartil Q
  - Valor máximo
- 2 Boxplo



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

osição

loxplot



Ao iniciar sua Análise Exploratória de Dados (EDA), você pode visualizar sua amostra com:

- Resumo dos cinco números
  - Valor mínimo
  - Primeiro quartil Q<sub>1</sub>
  - Mediana (e/ou média)
  - Terceiro quartil Q3
  - Valor máximo
- 2 Boxplo



Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de

Boxplot



Ao iniciar sua Análise Exploratória de Dados (EDA), você pode visualizar sua amostra com:

- Resumo dos cinco números
  - Valor mínimo
  - Primeiro quartil Q<sub>1</sub>
  - Mediana (e/ou média)
  - Terceiro quartil Q<sub>3</sub>
  - Valor máximo
- Boxplot



Felipe Figueiredo

Medidas de Fendência Central

Medidas de Dispersão

Medidas de

Boxplot





Ao iniciar sua Análise Exploratória de Dados (EDA), você pode visualizar sua amostra com:

- Resumo dos cinco números
  - Valor mínimo
  - Primeiro quartil Q<sub>1</sub>
  - Mediana (e/ou média)
  - Terceiro quartil Q<sub>3</sub>
  - Valor máximo
- Boxplot

Estatística Descritiva II

Felipe Figueiredo

Medidas de Tendência Central

Medidas de Dispersão

Medidas de

Boxplot

