

Medidas resumo

Fernando de Pol Mayer

Laboratório de Estatística e Geoinformação (LEG)
Departamento de Estatística (DEST)
Universidade Federal do Paraná (UFPR)



Este conteúdo está disponível por meio da Licença Creative Commons 4.0
(Atribuição/NãoComercial/PartilhaIgual)

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Características importantes de qualquer conjunto de dados

- Centro
- Variação
- Distribuição
- Valores atípicos

Medidas
resumo

Introdução

Medidas de
centro

Moda
Mediana
Média

Medidas de
variação
Amplitude
Desvio médio
Variância
Desvio-
padrão
Coeficiente
de Variação
Exercícios

Medidas de
posição
Percentis
Quartis
Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Medidas resumo

Introdução

Medidas de centro

Moda Mediana Média

Medidas de variação

Amplitude Desvio médio Variância Desvio- padrão Coeficiente de Variação Exercícios

Medidas de posição

Percentis Quartis

Exercícios

Definição

É um valor no centro, ou meio, do conjunto de dados

Ferramentas para **resumo** e **análise** de dados

- Média
- Mediana
- Moda

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana Média

Medidas de variação

Amplitude Desvio médio

Variância

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

A **moda** é o valor que ocorre com **maior frequência** em um conjunto de dados

Dependendo do conjunto de dados, ele pode ser

- **Sem moda** quando nenhum valor se repete
- **Unimodal** quando existe apenas um valor repetido com maior frequência
- **Bimodal** quando existem dois valores com a mesma maior frequência
- **Multimodal** quando mais de dois valores se repetem com a mesma frequência

Medidas
resumo

Introdução

Medidas de
centro

Moda

Mediana
Média

Medidas de
variação

Amplitude
Desvio médio

Variancia

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de
posição

Percentis
Quartis

Exercícios

Qual é a moda?

A) 2 5 7 9 13 15 22

B) 16 19 19 21 21 21 23 27

C) 2 7 7 13 15 15 22

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana Média

Medidas de variação

Amplitude Desvio médio

Variancia

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Qual é a moda?

ótimo	bom	bom	péssimo	bom	bom	ótimo
ótimo	bom	ótimo	bom	ótimo	bom	bom
ótimo	bom	péssimo	bom	péssimo	bom	péssimo
bom	bom	bom	bom	ótimo	bom	péssimo
ótimo	ótimo	bom	péssimo			

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana Média

Medidas de variação

Amplitude Desvio médio

Variancia

Desvio- padrão

Coefficiente de Variação

Exercícios

Medidas de posição

Percentis Quartis

Exercícios

Vantagens

- **Resistente** à valores extremos
- É a única medida de centro que pode ser usada para dados **qualitativos**

Desvantagens

- É uma medida **viesada**

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

1 Introdução

2 Medidas de tendência central

- Moda
- **Mediana**
- Média

3 Medidas de variação

- Amplitude
- Desvio médio
- Variância
- Desvio-padrão
- Coeficiente de Variação
- Exercícios

4 Medidas de posição relativa

- Percentis
- Quartis

5 Exercícios

A **mediana** é uma medida de centro que é o **valor do meio**, quando os dados são arranjados de maneira **ordenada**

É o valor cuja posição separa o conjunto de dados em duas partes iguais

Quando as observações são ordenadas em ordem crescente, vamos denotar a menor observação por $x_{(1)}$, a segunda por $x_{(2)}$, e assim por diante, obtendo-se

$$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n-1)} \leq x_{(n)}$$

Estas observações ordenadas são chamadas de **estatísticas de ordem**.

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Por exemplo, se cinco observações de uma variável forem $x_1 = 8$, $x_2 = 4$, $x_3 = 3$, $x_4 = 8$, $x_5 = 7$, então

$$3 \leq 4 \leq 7 \leq 8 \leq 8$$

E as estatísticas de ordem são: $x_{(1)} = 3$, $x_{(2)} = 4$, $x_{(3)} = 7$, $x_{(4)} = 8$, $x_{(5)} = 8$.

Nesse exemplo, a mediana (Md) é 7, pois é o valor que separa o conjunto de dados em duas partes iguais.

Mas note que o número de observações é par. Caso fosse ímpar, a mediana seria a média aritmética das duas observações centrais.

De maneira geral, a mediana de uma variável X pode ser definida por:

$$Md(X) = \begin{cases} x_{(\frac{n+1}{2})} & \text{se } n \text{ ímpar} \\ \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} & \text{se } n \text{ par} \end{cases}$$

Portanto, no exemplo anterior, se tivéssemos

$$3 \leq 4 \leq 7 \leq 8 \leq 8 \leq 9$$

Então

$$Md = \frac{x_{(3)} + x_{(4)}}{2} = \frac{7 + 8}{2} = 7,5$$

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio- padrão

Coefficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Número ímpar de elementos

2 4 6 7 11

Medidas
resumo

Introdução

Medidas de
centro

Moda

Mediana

Média

Medidas de
variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de
posição

Percentis

Quartis

Exercícios

Número par de elementos

2 4 7 9 11 13

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Vantagens

- Medida **resistente**
- Não é influenciada pela presença de valores extremos

Desvantagens

- É uma medida **viesada**

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - **Média**
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

A **média aritmética** de um conjunto de dados é a medida de tendência central encontrada pela soma de todos os valores, dividida pelo número total de elementos, ou seja,

$$\bar{x} = \frac{1}{n} \cdot (x_1 + x_2 + \cdots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i$$

No exemplo anterior, temos então que a média de 3, 4, 7, 8, 8 é

$$\begin{aligned}\bar{x} &= \frac{1}{5} \cdot (3 + 4 + 7 + 8 + 8) \\ &= \frac{1}{5} \cdot (30) \\ &= 6\end{aligned}$$

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio- padrão

Coefficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Considere a nota das provas de 5 alunos em uma sala com 30 alunos

7,0 3,0 5,5 6,5 8,0

Note que a média é o **ponto de equilíbrio de massa** dos dados

Medidas
resumo

Introdução

Medidas de
centro

Moda
Mediana
Média

Medidas de
variação

Amplitude
Desvio médio

Variancia

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de
posição

Percentis
Quartis

Exercícios

Considere o valor dos salários de todos os 6 empregados de uma pequena empresa

860,00 750,00 980,00 1.200,00 790,00 950,00

Calcule a média populacional

Agora, se tivermos n observações da variável X , das quais f_1 são iguais a x_1 , f_2 são iguais a x_2 , \dots , f_k são iguais a x_k , então a média pode ser definida por:

$$\bar{x} = \frac{1}{n} \cdot (x_1 f_1 + x_2 f_2 + \dots + x_k f_k) = \frac{1}{n} \sum_{i=1}^k x_i f_i$$

Note que, se as frequências relativas são $fr_i = f_i/n$, então a equação acima também pode ser escrita como

$$\bar{x} = x_1 fr_1 + x_2 fr_2 + \dots + x_k fr_k = \sum_{i=1}^k x_i fr_i$$

Como exemplo, considere a tabela de frequência abaixo:

Número	f_i	fr_i
0	4	0,20
1	5	0,25
2	7	0,35
3	3	0,15
5	1	0,05
Total	20	1

A média é calculada por:

$$\begin{aligned}
 \bar{x} &= \frac{1}{20} \cdot (0 \cdot 4 + 1 \cdot 5 + \dots + 5 \cdot 1) \\
 &= \frac{1}{20} \cdot (33) \\
 &= 1,65
 \end{aligned}$$

No caso de variáveis contínuas resumidas em tabelas de frequência com intervalos de classe, a média pode ser aproximada, calculando-se o **ponto médio** de cada classe

$$PM = \frac{\lim_{inf} + \lim_{sup}}{2}$$

e supor que os valores dentro de cada classe sejam iguais ao ponto médio. Nesse caso, ficamos com a mesma situação para o caso discreto, onde a média é calculada com pares (x_i, f_i) ou (x_i, fr_i) .

Claramente isso é uma aproximação, pois estamos perdendo informação ao assumir que todos os valores de uma classe sejam iguais. Portanto, deverá haver alguma diferença entre esta média aproximada e a média que seria calculada com os valores originais.

Considere a seguinte tabela de distribuição de frequência:

Classe	f_i	fr_i
[4, 8)	10	0,278
[8, 12)	12	0,333
[12, 16)	8	0,222
[16, 20)	5	0,139
[20, 24)	1	0,028
Total	36	1

Considerando os pontos médios de cada classe, a média é calculada por

$$\begin{aligned}
 \bar{x} &= \frac{1}{36} \cdot (6 \cdot 10 + 10 \cdot 12 + \dots + 22 \cdot 1) \\
 &= \frac{1}{36} \cdot (404) \\
 &= 11,22
 \end{aligned}$$

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Vantagens

- Medida **não viesada**
- A média tende a ser mais **consistente** do que outras medidas de centro

Desvantagens

- Sensível à valores extremos
- Medida não **resistente**

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio- padrão

Coefficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Você está procurando um estágio nas empresas A e B. Cada empresa oferece remuneração por 20 horas semanais com as seguintes característica (em salários mínimos)

	A	B
média	2,5	2,0
mediana	1,7	1,9
moda	1,5	1,9

Qual você escolheria?

Para notar como a média é influenciada pela presença de valores extremos

$$5 \ 7 \ 10 \ 13 \ 15 \Rightarrow \bar{x} = 10 \text{ e } Me = 10$$

$$5 \ 7 \ 10 \ 13 \ 65 \Rightarrow \bar{x} = 20 \text{ e } Me = 10$$

Nos casos onde se deseja comparar bases de dados diferentes, normalmente a mediana é mais indicada, por ser uma medida mais **robusta**, *não influenciada por valores extremos*

Média, Mediana, e Moda

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio- padrão

Coefficiente de Variação

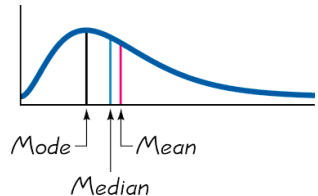
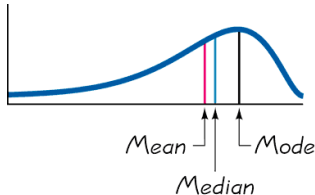
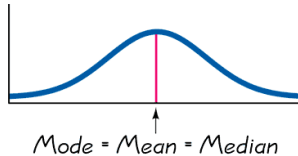
Exercícios

Medidas de posição

Percentis

Quartis

Exercícios



Exemplo: Os dados abaixo se referem ao percentual de cobertura de vegetação em duas áreas de uma floresta.

Área A: 43 47 48 51 51 55 55 57 59

Área B: 20 22 45 46 53 54 56 57

- Calcule a média, a mediana e a moda para a área A. Qual a medida de tendência central melhor representa esse conjunto de dados? Por quê?
- Calcule a média, a mediana e a moda para a área B. Qual a medida de tendência central melhor representa esse conjunto de dados? Por quê?

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

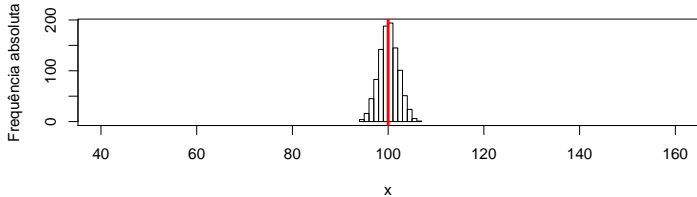
Exercícios

O resumo de um conjunto de dados exclusivamente por uma medida de centro, **esconde** toda a informação sobre a variabilidade do conjunto de observações

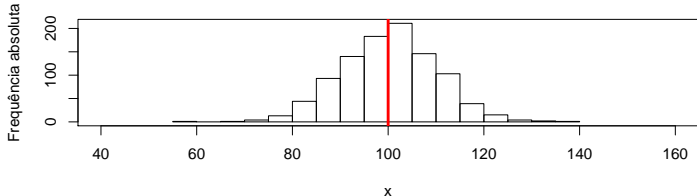
Não é possível analisar um conjunto de dados apenas através de uma medida de tendência central

Por isso precisamos de medidas que resumam a **variabilidade** dos dados em relação à um valor central

$N(100, 4)$



$N(100, 100)$



Cinco grupos de alunos se submeteram a um teste, obtendo as seguintes notas

Grupo	Notas	\bar{x}
A	3, 4, 5, 6, 7	5
B	1, 3, 5, 7, 9	5
C	5, 5, 5, 5, 5	5
D	3, 5, 5, 7	5
E	3, 5, 5, 6, 6	5

O que a média diz a respeito das notas quando comparamos os grupos?

Definição

São medidas estatísticas que caracterizam o quanto um conjunto de dados está disperso em torno de sua tendência central

Ferramentas para **resumo** e **análise** de dados

- Amplitude
- Desvio-médio
- Variância
- Desvio-padrão
- Coeficiente de Variação

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

A **amplitude** de um conjunto de dados é a diferença entre o maior e o menor valor.

$$AMP = \max - \min$$

Como a amplitude usa **apenas** os valores máximo e mínimo, é muito **sensível** a valores extremos

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Calcule a média e a amplitude do número de acertos em uma prova com 50 questões

31 27 42 35 47 28 7 45 15 20

Calcule a média e a amplitude para a idade de um grupo de pessoas

4 3 4 3 4 3 21

Para melhorar a medida de variabilidade, devemos considerar **todos os dados disponíveis**

A melhor forma de se fazer isso é considerar o **desvio** de cada valor em relação à média

Como queremos um **resumo** da variabilidade, devemos fazer a **soma** dos desvios

Considere as notas do grupo A do exemplo acima ($\bar{x} = 5$)

Grupo A	$x_i - \bar{x}$
3	-2
4	-1
5	0
6	1
7	2
Soma	0

Como a soma dos desvios é **sempre** zero, temos duas alternativas

- Considerar o total dos desvios absolutos (em módulo)

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

- Considerar o total dos quadrados dos desvios

$$\sum_{i=1}^n (x_i - \bar{x})^2$$

O uso destes totais pode causar dificuldades quando comparamos conjuntos de dados de tamanhos diferentes. Desse modo é mais conveniente exprimir estas medidas como **médias** (dividindo as somas por n). Assim teremos:

- Desvio médio
- Variância

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

O **desvio médio** é definido como a média aritmética dos desvios em módulo (valor absoluto)

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

No exemplo anterior

Grupo A	$x_i - \bar{x}$	$ x_i - \bar{x} $
3	-2	2
4	-1	1
5	0	0
6	1	1
7	2	2
Soma	0	6

$$DM = \frac{6}{5} = 1,2$$

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Mas, o desvio médio é baseado em uma operação **não algébrica** (módulo), o que cria dificuldades em análises posteriores

Além disso, é uma medida **viesada**

Uma alternativa melhor é a **soma dos quadrados dos desvios**

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - **Variância**
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

A **variância** é definida como a *média aritmética* da soma dos quadrados dos desvios.

Variância amostral

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

Uma fórmula alternativa da variância pode ser obtida desenvolvendo-se o quadrado no numerador da expressão anterior

$$s^2 = \frac{1}{n} \left[\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \right]$$

No exemplo anterior

Grupo A	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
3	-2	2	4
4	-1	1	1
5	0	0	0
6	1	1	1
7	2	2	4
Soma	0	6	10

$$s^2 = \frac{10}{5} = 2$$

Assim como no caso da média, se tivermos n observações da variável X , das quais f_1 são iguais a x_1 , f_2 são iguais a x_2 , \dots , f_k são iguais a x_k , então a variância pode ser definida por:

$$s^2 = \frac{1}{n} \sum_{i=1}^n f_i (x_i - \bar{x})^2 = \sum_{i=1}^n fr_i (x_i - \bar{x})^2$$

Ou, pela fórmula alternativa

$$\begin{aligned} s^2 &= \frac{1}{n} \left[\sum_{i=1}^n x_i^2 \cdot f_i - \frac{(\sum_{i=1}^n x_i \cdot f_i)^2}{n} \right] \\ &= \sum_{i=1}^n x_i^2 \cdot fr_i - \frac{(\sum_{i=1}^n x_i \cdot fr_i)^2}{n} \end{aligned}$$

Como exemplo, considere a tabela de frequência abaixo ($\bar{x} = 1,65$):

Número	f_i	fr_i	$x_i - \bar{x}$
0	4	0,20	-1,65
1	5	0,25	-0,65
2	7	0,35	0,35
3	3	0,15	1,35
5	1	0,05	3,35
Total	20	1	

A variância pode ser calculada por:

$$\begin{aligned}
 s^2 &= \frac{1}{20} \cdot [4 \cdot (-1,65)^2 + 5 \cdot (-0,65)^2 + \dots + 1 \cdot (3,35)^2] \\
 &= \frac{1}{20} \cdot (30,55) \\
 &= 1,528
 \end{aligned}$$

Considere a seguinte tabela de distribuição de frequência ($\bar{x} = 11,22$):

Classe	f_i	fr_i	$x_i - \bar{x}$
[4, 8)	10	0,278	-5,222
[8, 12)	12	0,333	-1,222
[12, 16)	8	0,222	2,778
[16, 20)	5	0,139	6,778
[20, 24)	1	0,028	10,778
Total	36	1	

Considerando os pontos médios de cada classe, a variância pode ser calculada por

$$\begin{aligned}
 \bar{x} &= \frac{1}{36} \cdot [6 \cdot (-5,222)^2 + 10 \cdot (-1,222)^2 + \dots + 22 \cdot (10,778)^2] \\
 &= \frac{1}{36} \cdot (698,22) \\
 &= 19,395
 \end{aligned}$$

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

A variância amostral s^2 é considerada um estimador **não viesado** da variância populacional σ^2

É utilizada em diversos métodos estatísticos e caracteriza todas as distribuições de probabilidade

No entanto, as *unidades da variância são diferentes das unidades dos dados originais* (são medidas ao quadrado, como notas ao quadrado ou cm^2)

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

**Desvio-
padrão**

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

O **desvio-padrão** é a raiz quadrada da variância

Desvio-padrão amostral

$$s = \sqrt{s^2}$$

Sendo que s^2 é calculada de qualquer uma das formas anteriores.

Propriedades do desvio-padrão

- É uma medida de variação de todos os dados em relação à **média**
- É sempre positivo ou nulo
 - Valores mais distantes da média tem desvio-padrão maior
 - Valores mais próximos da média tem desvio-padrão menor
- A unidade do desvio-padrão é a mesma dos dados originais (por exemplo notas ou cm)
- A inclusão de valores **extremos** pode afetar drasticamente o valor do desvio-padrão

Medidas resumo

Introdução

Medidas de
centro

Moda

Mediana

Média

Medidas de
variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de
posição

Percentis

Quartis

Exercícios

Exemplo: Os dados abaixo se referem ao percentual de cobertura de vegetação em duas áreas de uma floresta.

Área A: 43 47 48 51 51 55 55 57 59

Área B: 20 22 45 46 53 54 56 57

- Calcule o desvio-padrão para as duas áreas.
- Podemos comparar essas duas medidas? O que podemos concluir?

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio-
padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

O **Coeficiente de Variação** (CV) mede a dispersão dos dados em relação à média (medido em %)

Coeficiente de variação amostral

$$CV = \frac{s}{\bar{x}} \cdot 100\%$$

É utilizado para se comparar a variação de um ou mais conjuntos de dados

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio-
padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Qual o Coeficiente de Variação para as duas áreas do exemplo anterior:

Área A: 43 47 48 51 51 55 55 57 59

Área B: 20 22 45 46 53 54 56 57

O que podemos concluir?

O Coeficiente de Variação é muito útil também para se comparar dados medidos em escalas diferentes. Por exemplo

	Média	Desvio-padrão
Altura	174 cm	7 cm
Peso	78 kg	12 kg

Sópodemos comparar o desvio-padrão com unidades diferentes através do CV

$$CV_A = \frac{7}{174} \cdot 100\% = 4\% \quad CV_P = \frac{12}{78} \cdot 100\% = 15,4\%$$

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Considere a tabela de frequência abaixo:

Classe	f_i
1,0 ┤ 2,5	3
2,5 ┤ 4,0	5
4,0 ┤ 5,5	3
5,5 ┤ 7,0	7
7,0 ┤ 8,5	9
8,5 ┤ 10,0	13

Calcule a média, a variância, o desvio-padrão, e o CV para este conjunto de dados.

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Definição

Percentis são medidas de posição, denotados por P_1, P_2, \dots, P_{99} que dividem os dados em 100 grupos, com cerca de 1% cada grupo

Por exemplo

- O 50º percentil, P_{50} , tem cerca de 50% dos valores abaixo dele, e 50% de valores acima dele
 - Nesse caso, $P_{50} = \text{Mediana}$

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Para determinar um percentil:

- Encontre a posição

$$\text{Pos}P_i = \frac{i(n+1)}{100}, \quad i = 1, \dots, 99$$

- Se o valor for fracionário calcule o valor intermediário

Calcule o P_{30} e o P_{65} para os dados abaixo

15 21 28 25 30 11 17 12 25 20 16 23 12 10

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Definição

Quartis são medidas de posição, denotadas por Q_1 , Q_2 , Q_3 que dividem um conjunto de dados em 4 grupos, com cerca de 25% dos valores em cada grupo

Q_1 (**Primeiro quartil**): Separa os 25% inferiores dos 75% superiores dos valores ordenados

Q_2 (**Segundo quartil**): O mesmo que a mediana. Separa os 50% valores ordenados inferiores dos 50% superiores

Q_3 (**terceiro quartil**): Separa os 75% valores ordenados inferiores dos 25% superiores

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Para determinar um quartil:

- Encontre a posição

$$\text{Pos}Q_i = \frac{i(n+1)}{4}, \quad i = 1, \dots, 3$$

- Se o valor for fracionário calcule o valor intermediário

Calcule os quartis para os dados abaixo

15 21 28 25 30 11 17 12 25 20 16 23 12 10

Medidas resumo

Introdução

Medidas de
centro

Moda

Mediana
Média

Medidas de
variação

Amplitude
Desvio médio

Variancia

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de
posição

Percentis

Quartis

Exercícios

Gráfico de caixa ou resumo dos cinco números

Resumo dos 5 números

O resumo dos cinco números consiste no valor mínimo, primeiro quartil, segundo quartil (mediana), terceiro quartil, e no valor máximo

Gráfico de caixa

O gráfico de caixa, ou *boxplot*, é uma representação gráfica do resumo dos cinco números

Para os valores

15 21 28 25 30 11 17 12 25 20 16 23 12 10

o **gráfico de caixa** correspondente é

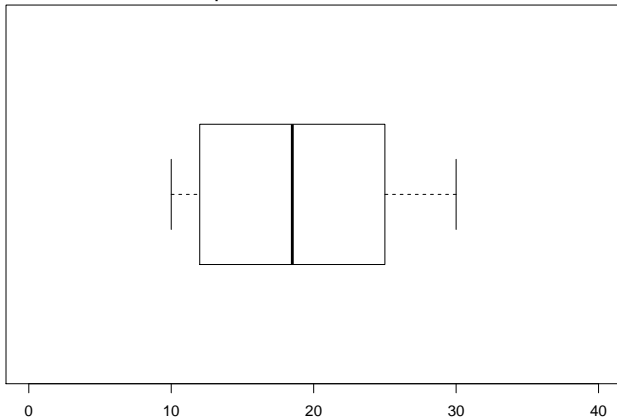
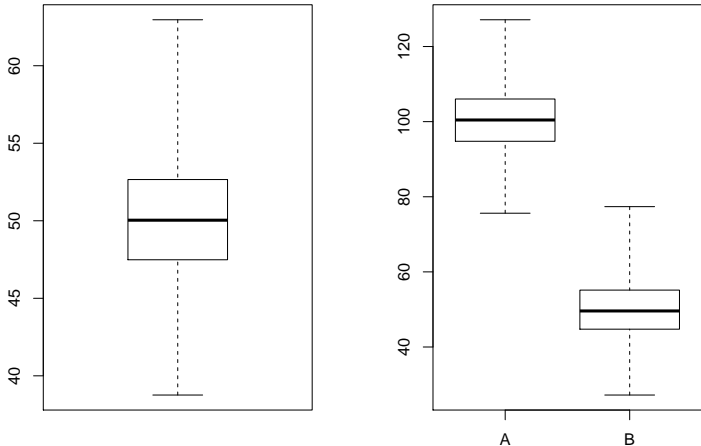


Gráfico de caixa ou resumo dos cinco números



Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio- padrão

Coeficiente de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

Exemplo: o tempo de espera, em minutos, para o atendimento em uma central telefônica, para homens e mulheres, foi registrado como abaixo

Homens: 5 2 7 9 3 4 3 1 3 8
Mulheres: 3 5 7 4 5 6 7 6 5 4

- Monte o resumo dos cinco números e o gráfico de caixa para homens e mulheres juntos
- Monte o resumo dos cinco números e o gráfico de caixa para homens e mulheres separados

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variancia

Desvio-
padrão

Coefficiente
de Variação

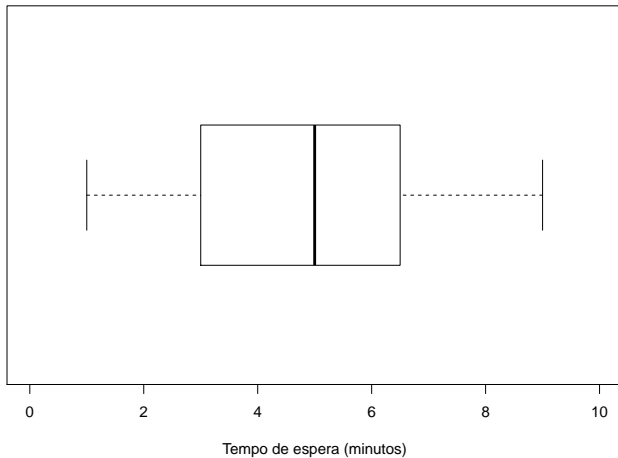
Exercícios

Medidas de posição

Percentis

Quartis

Exercícios



Medidas resumo

Introdução

Medidas de
centro

Moda

Mediana

Média

Medidas de
variação

Amplitude

Desvio médio

Variancia

Desvio-
padrão

Coeficiente
de Variação

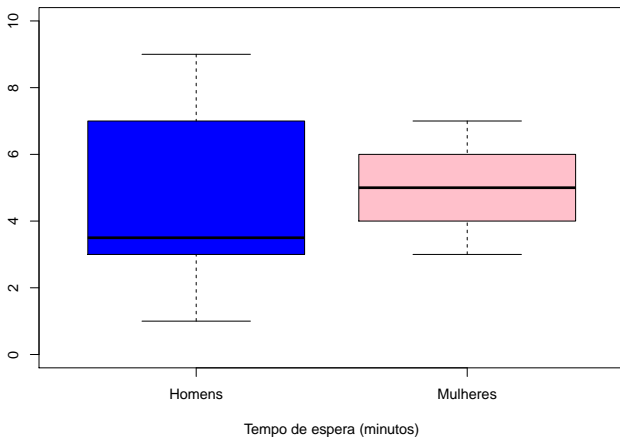
Exercícios

Medidas de
posição

Percentis

Quartis

Exercícios



Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- 1 Introdução
- 2 Medidas de tendência central
 - Moda
 - Mediana
 - Média
- 3 Medidas de variação
 - Amplitude
 - Desvio médio
 - Variância
 - Desvio-padrão
 - Coeficiente de Variação
 - Exercícios
- 4 Medidas de posição relativa
 - Percentis
 - Quartis
- 5 Exercícios

Medidas
resumo

Introdução

Medidas de
centro

Moda

Mediana
Média

Medidas de
variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coefficiente
de Variação

Exercícios

Medidas de
posição

Percentis

Quartis

Exercícios

Exercícios 2–7, 9–15 do capítulo 3 do livro (pgs. 94–96).

Pinto, SS; Silva, CS. **Estatística, Vol I**. Rio Grande: Editora da FURG, 2010. [Cap. 3]

Medidas resumo

Introdução

Medidas de centro

Moda

Mediana

Média

Medidas de variação

Amplitude

Desvio médio

Variância

Desvio-
padrão

Coeficiente
de Variação

Exercícios

Medidas de posição

Percentis

Quartis

Exercícios

- Magalhães, MN; Lima, ACP. **Noções de Probabilidade e Estatística**. São Paulo: EDUSP, 2008. [Cap. 1]
- Bussab, WO; Morettin, PA. **Estatística básica**. São Paulo: Saraiva, 2006. 526 p. [Cap. 3]