Estatística Descritiva II

Medidas sumárias

Felipe Figueiredo

Instituto Nacional de Traumatologia e Ortopedia

Sumário

1. Medidas de Tendência Central
   - Média
   - Mediana
   - Moda
   - Comparação entre as Medidas Centrais
2. Medidas de Dispersão
   - Amplitude
   - Desvios em relação à media
   - Variância
   - Desvio Padrão
   - Coeficiente de Variação
3. Medidas de Posição
   - Quartis
   - Percentis
   - Escore padrão

Medidas Sumárias

- 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, etc).
- Geralmente trabalhamos com estatísticas descritivas.

Medidas Sumárias

- 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, etc).
- Geralmente trabalhamos com estatísticas descritivas.

Medidas Sumárias

- 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, etc).
- Geralmente trabalhamos com estatísticas descritivas.

Medidas Sumárias

- 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, etc).
- Geralmente trabalhamos com estatísticas descritivas.

Sumário

1. Medidas de Tendência Central
    - Média
    - Mediana
    - Moda
    - Comparação entre as Medidas Centrais
2. Medidas de Dispersão
    - Amplitude
    - Desvios em relação à media
    - Variância
    - Desvio Padrão
    - Coeficiente de Variação
3. Medidas de Posição
    - Quartis
    - Percentis
    - Escore padrão

Média

- 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 ($\bar{x}$)

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

- Nem sempre pertence ao dataset.

Média

- 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 ($\bar{x}$)

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

- Nem sempre pertence ao dataset.

Média

- 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 ($\bar{x}$)

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

- Nem sempre pertence ao dataset.

Média

- 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 ($\bar{x}$)

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

- Nem sempre pertence ao dataset.

Média

Example

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

$$\begin{aligned} x_1 &= 142 \\ x_2 &= 144 \\ x_3 &= 176 \\ x_4 &= 203 \\ x_5 &= 134 \\ x_6 &= 191 \end{aligned}$$

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

Sumário

1. Medidas de Tendência Central
   - Média
   - Mediana
   - Moda
   - Comparação entre as Medidas Centrais
2. Medidas de Dispersão
   - Amplitude
   - Desvios em relação à media
   - Variância
   - Desvio Padrão
   - Coeficiente de Variação
3. Medidas de Posição
   - Quartis
   - Percentis
   - Escore padrão

Mediana

Definition

A mediana é o dado que ocupa a posição central nos dados ordenados.

- Notação: M_d
- Divide o dataset ao meio
- Costuma pertencer ao dataset

Mediana

Definition

A mediana é o dado que ocupa a posição central nos dados ordenados.

- Notação: M_d
- Divide o dataset ao meio
- Costuma pertencer ao dataset

Mediana

Definition

A mediana é o dado que ocupa a posição central nos dados ordenados.

- Notação: M_d
- Divide o dataset ao meio
- Costuma pertencer ao dataset

Mediana

Definition

A mediana é o dado que ocupa a posição central nos dados ordenados.

- Notação: M_d
- Divide o dataset ao meio
- Costuma pertencer ao dataset

Mediana

- 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

$$
\begin{aligned}
x_5 &= 134 \\
x_1 &= 142 \\
x_2 &= 144 \\
x_3 &= 176 \\
x_6 &= 191 \\
x_4 &= 203
\end{aligned}
\qquad M_d = \frac{144 + 176}{2} = 160
$$

Mediana

- 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

$$\begin{aligned} x_5 &= 134 \\ x_1 &= 142 \\ x_2 &= 144 \\ x_3 &= 176 \\ x_6 &= 191 \\ x_4 &= 203 \end{aligned} \qquad M_d = \frac{144 + 176}{2} = 160$$

Mediana

- 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

$$x_5 = 134$$
$$x_1 = 142$$
$$x_2 = 144$$
$$x_3 = 176$$
$$x_6 = 191$$
$$x_4 = 203$$

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

Mediana

- 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

$$\begin{aligned} X_5 &= 134 \\ X_1 &= 142 \\ X_2 &= 144 \\ X_3 &= 176 \\ X_6 &= 191 \\ X_4 &= 203 \end{aligned}$$

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

Mediana

- 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

$$\begin{aligned} x_5 &= 134 \\ x_1 &= 142 \\ x_2 &= 144 \\ x_3 &= 176 \\ x_6 &= 191 \\ x_4 &= 203 \end{aligned}$$

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

Mediana

- 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

$$\begin{aligned} x_5 &= 134 \\ x_1 &= 142 \\ x_2 &= 144 \\ x_3 &= 176 \\ x_6 &= 191 \\ x_4 &= 203 \end{aligned}$$

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

Mediana

- 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

$$\begin{aligned} x_5 &= 134 \\ x_1 &= 142 \\ x_2 &= 144 \\ x_3 &= 176 \\ x_6 &= 191 \\ x_4 &= 203 \end{aligned}$$

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

Sumário

1. Medidas de Tendência Central
   - Média
   - Mediana
   - Moda
   - Comparação entre as Medidas Centrais
2. Medidas de Dispersão
   - Amplitude
   - Desvios em relação à media
   - Variância
   - Desvio Padrão
   - Coeficiente de Variação
3. Medidas de Posição
   - Quartis
   - Percentis
   - Escore padrão

Moda

Definition

A moda é o dado que ocorre com maior frequência.

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

Moda

Definition

A moda é o dado que ocorre com maior frequência.

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

Moda

Definition

A moda é o dado que ocorre com maior frequência.

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

Moda

Definition

A moda é o dado que ocorre com maior frequência.

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

Moda

Definition

A moda é o dado que ocorre com maior frequência.

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

Moda

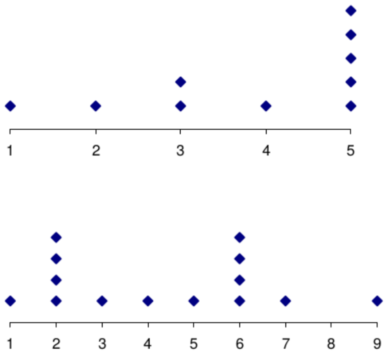

Figura: Diagrama de pontos para dados (a) unimodal, (b) bimodal (Fonte: Reis, Reis, 2002)

Sumário

1. Medidas de Tendência Central
   - Média
   - Mediana
   - Moda
   - Comparação entre as Medidas Centrais
2. Medidas de Dispersão
   - Amplitude
   - Desvios em relação à media
   - Variância
   - Desvio Padrão
   - Coeficiente de Variação
3. Medidas de Posição
   - Quartis
   - Percentis
   - Escore padrão

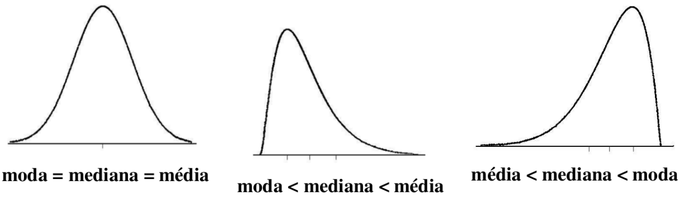

Figura: (a) Simétrica, (b) Assimétrica à esquerda, (c) Assimétrica à direita (Fonte: Reis, Reis 2002)

Robustez da Média

- 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:
- $\mu = \frac{1+1+2+4+7}{5} = \frac{15}{5} = 3$
- $M_d = 2$
- $M_o = 1$

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:
- $\mu = \frac{1+1+2+4+7}{5} = \frac{15}{5} = 3$
- $M_d = 2$
- $M_o = 1$

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:
- $\mu = \dfrac{1+1+2+4+7}{5} = \dfrac{15}{5} = 3$
- $M_d = 2$
- $M_o = 1$

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:
- $\mu = \dfrac{1+1+2+4+7}{5} = \dfrac{15}{5} = 3$
- $M_d = 2$
- $M_o = 1$

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:
- $\mu = \dfrac{1+1+2+4+7}{5} = \dfrac{15}{5} = 3$
- $M_d = 2$
- $M_o = 1$

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:
- $\mu = \dfrac{1+1+2+4+7}{5} = \dfrac{15}{5} = 3$
- $M_d = 2$
- $M_o = 1$

Comparação entre as Medidas Centrais

Example

Considere agora este outro dataset

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

- $N = 5$
- As medidas descritivas centrais para estes dados são:
- $\mu = \dfrac{1+1+2+4+32}{5} = \dfrac{40}{5} = 8$
- $M_d = 2$
- $M_o = 1$

Comparação entre as Medidas Centrais

Example

Considere agora este outro dataset

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

- $N = 5$
- As medidas descritivas centrais para estes dados são:
- $\mu = \frac{1+1+2+4+32}{5} = \frac{40}{5} = 8$
- $M_d = 2$
- $M_o = 1$

Comparação entre as Medidas Centrais

Example

Considere agora este outro dataset

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

- $N = 5$
- As medidas descritivas centrais para estes dados são:
- $\mu = \frac{1+1+2+4+32}{5} = \frac{40}{5} = 8$
- $M_d = 2$
- $M_o = 1$

Comparação entre as Medidas Centrais

Example

Considere agora este outro dataset

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

- $N = 5$
- As medidas descritivas centrais para estes dados são:
- $\mu = \dfrac{1 + 1 + 2 + 4 + 32}{5} = \dfrac{40}{5} = 8$
- $M_d = 2$
- $M_o = 1$

Comparação entre as Medidas Centrais

Example

Considere agora este outro dataset

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

- $N = 5$
- As medidas descritivas centrais para estes dados são:
- $\mu = \dfrac{1+1+2+4+32}{5} = \dfrac{40}{5} = 8$
- $M_d = 2$
- $M_o = 1$

Comparação entre as Medidas Centrais

Example

Considere agora este outro dataset

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

- $N = 5$
- As medidas descritivas centrais para estes dados são:
- $\mu = \dfrac{1 + 1 + 2 + 4 + 32}{5} = \dfrac{40}{5} = 8$
- $M_d = 2$
- $M_o = 1$

Exercícios

Exercício

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

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

Solução

1. $\bar{x} = \dfrac{35 + 33 + 37 + 33 + 34}{5} = 34.4$
2. $M_d = 34$
3. $M_o = 33$

Exercícios

Exercício

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

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

Solução

1. $$\bar{x} = \frac{35 + 33 + 37 + 33 + 34}{5} = 34.4$$
2. $M_d = 34$
3. $M_o = 33$

Exercícios

Exercício

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

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

Solução

1. $\bar{x} = \dfrac{35 + 33 + 37 + 33 + 34}{5} = 34.4$
2. $M_d = 34$
3. $M_o = 33$

Exercícios

Exercício

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

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

Solução

1. $\bar{x} = \dfrac{35 + 33 + 37 + 33 + 34}{5} = 34.4$
2. $M_d = 34$
3. $M_o = 33$

Exercícios

Exercício

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

Determine:

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

Solução

1. $\bar{x} = \dfrac{35 + 33 + 37 + 33 + 34}{5} = 34.4$
2. $M_d = 34$
3. $M_o = 33$

Resumo

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

Resumo

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

Resumo

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

Variabilidade em Medições

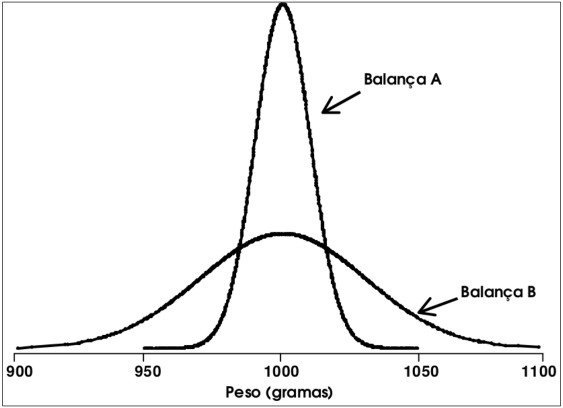

Figura: Variabilidade da medição de uma esfera metálica de 1000g. Balança A, “imprecisão” de 50g, balança B, “imprecisão” de 100g (Fonte: Reis, Reis, 2002)

Sumário

1. Medidas de Tendência Central
   - Média
   - Mediana
   - Moda
   - Comparação entre as Medidas Centrais
2. Medidas de Dispersão
   - Amplitude
   - Desvios em relação à media
   - Variância
   - Desvio Padrão
   - Coeficiente de Variação
3. Medidas de Posição
   - Quartis
   - Percentis
   - Escore padrão

Amplitude

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

- $A = x_{max} - x_{min}$

Sumário

1. Medidas de Tendência Central
   - Média
   - Mediana
   - Moda
   - Comparação entre as Medidas Centrais
2. Medidas de Dispersão
   - Amplitude
   - Desvios em relação à media
   - Variância
   - Desvio Padrão
   - Coeficiente de Variação
3. Medidas de Posição
   - Quartis
   - Percentis
   - Escore padrão

Desvios em relação à média

- Uma maneira de entender a variabilidade do dataset é analisar os desvios em relação à média.
- O desvio é a diferença entre o valor do dado
- $D = x_i - \mu$ ou $D = x_i - \bar{x}$

Desvios em relação à média

- Uma maneira de entender a variabilidade do dataset é analisar os desvios em relação à média.
- O desvio é a diferença entre o valor do dado
- $D = x_i - \mu$ ou $D = x_i - \bar{x}$

Desvios em relação à média

Medidas de Tendência Central

Medidas de Dispersão

Amplitude

Desvios em relação à media

Variância

Desvio Padrão

Coeficiente de Variação

Medidas de Posição

- Uma maneira de entender a variabilidade do dataset é analisar os desvios em relação à média.
- O desvio é a diferença entre o valor do dado
- $D = x_i - \mu$ ou $D = x_i - \bar{x}$

Sumário

1 Medidas de Tendência Central
- Média
- Mediana
- Moda
- Comparação entre as Medidas Centrais

2 Medidas de Dispersão
- Amplitude
- Desvios em relação à media
- Variância
- Desvio Padrão
- Coeficiente de Variação

3 Medidas de Posição
- Quartis
- Percentis
- Escore padrão

Variância

A variância é a média dos desvios quadráticos

- Variância populacional

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

- Variância amostral

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

- É uma medida conveniente do ponto de vista matemático (boas propriedades algébricas e analíticas).
- Como ela usa uma unidade quadrática, é pouco intuitiva do ponto de vista de interpretação para resultados.

Medidas de Tendência Central

Medidas de Dispersão

Amplitude

Desvios em relação à media

Variância

Desvio Padrão

Coeficiente de Variação

Medidas de Posição

Variância

A variância é a média dos desvios quadráticos

- Variância populacional

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

- Variância amostral

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

- É uma medida conveniente do ponto de vista matemático (boas propriedades algébricas e analíticas).
- Como ela usa uma unidade quadrática, é pouco intuitiva do ponto de vista de interpretação para resultados.

Variância

A variância é a média dos desvios quadráticos

- Variância populacional

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

- Variância amostral

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

- É uma medida conveniente do ponto de vista matemático (boas propriedades algébricas e analíticas).
- Como ela usa uma unidade quadrática, é pouco intuitiva do ponto de vista de interpretação para resultados.

Variância

A variância é a média dos desvios quadráticos

- Variância populacional

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

- Variância amostral

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

- É uma medida conveniente do ponto de vista matemático (boas propriedades algébricas e analíticas).
- Como ela usa uma unidade quadrática, é pouco intuitiva do ponto de vista de interpretação para resultados.

Sumário

1 Medidas de Tendência Central
- Média
- Mediana
- Moda
- Comparação entre as Medidas Centrais

2 Medidas de Dispersão
- Amplitude
- Desvios em relação à media
- Variância
- Desvio Padrão
- Coeficiente de Variação

3 Medidas de Posição
- Quartis
- Percentis
- Escore padrão

Desvio Padrão

O desvio padrão é a

- Desvio padrão populacional

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

- Desvio padrão amostral

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

- É a medida mais usual para mensurar a variabilidade dos dados, por estar na mesma escala (unidade) destes.

Desvio Padrão

O desvio padrão é a

- Desvio padrão populacional

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

- Desvio padrão amostral

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

- É a medida mais usual para mensurar a variabilidade dos dados, por estar na mesma escala (unidade) destes.

Desvio Padrão

O desvio padrão é a

- Desvio padrão populacional

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

- Desvio padrão amostral

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

- É a medida mais usual para mensurar a variabilidade dos dados, por estar na mesma escala (unidade) destes.

Desvio Padrão

Medidas de Tendência Central

Medidas de Dispersão

Amplitude

Desvios em relação à media

Variância

Desvio Padrão

Coeficiente de Variação

Medidas de Posição

Desvio Padrão

Exercícios

Exercício

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

1. A variância amostral (s^2)
2. O desvio padrão amostral (s)

Solução

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

1. $$s^2 = \frac{(35 - 34.4)^2 + (33 - 34.4)^2 + \ldots}{5 - 1}$$
$$= \frac{0.36 + 1.96 + 6.76 + 1.96 + 0.16}{4} = 2.8$$
2. $s = \sqrt{2.8} = 1.67$

Exercícios

Exercício

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

1. A variância amostral (s^2)
2. O desvio padrão amostral (s)

Solução

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

1. $$s^2 = \frac{(35 - 34.4)^2 + (33 - 34.4)^2 + \dots}{5 - 1}$$
   $$= \frac{0.36 + 1.96 + 6.76 + 1.96 + 0.16}{4} = 2.8$$
2. $s = \sqrt{2.8} = 1.67$

Exercícios

Exercício

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

1. A variância amostral (s^2)
2. O desvio padrão amostral (s)

Solução

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

1. $$s^2 = \frac{(35-34.4)^2 + (33-34.4)^2 + \dots}{5-1}$$
   $$= \frac{0.36 + 1.96 + 6.76 + 1.96 + 0.16}{4} = 2.8$$
2. $s = \sqrt{2.8} = 1.67$

Exercícios

Exercício

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

1. A variância amostral (s^2)
2. O desvio padrão amostral (s)

Solução

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

1. $$s^2 = \frac{(35 - 34.4)^2 + (33 - 34.4)^2 + \dots}{5 - 1} = \frac{0.36 + 1.96 + 6.76 + 1.96 + 0.16}{4} = 2.8$$
2. $s = \sqrt{2.8} = 1.67$

Sumário

1. Medidas de Tendência Central
   - Média
   - Mediana
   - Moda
   - Comparação entre as Medidas Centrais
2. Medidas de Dispersão
   - Amplitude
   - Desvios em relação à media
   - Variância
   - Desvio Padrão
   - Coeficiente de Variação
3. Medidas de Posição
   - Quartis
   - Percentis
   - Escore padrão

Coeficiente de 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)

Example

Coeficiente de 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)

Example

Coeficiente de 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)

Example

Coeficiente de 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)

Example

Medidas de Posição

- Permitem estabelecer informações quantitativas relativas à ordem dos dados
-

Medidas de Posição

- Permitem estabelecer informações quantitativas relativas à ordem dos dados
-

Medidas de Tendência Central

Medidas de Dispersão

Medidas de Posição

Quartis

Percentis

Escore padrão

Sumário

1. Medidas de Tendência Central
   - Média
   - Mediana
   - Moda
   - Comparação entre as Medidas Centrais
2. Medidas de Dispersão
   - Amplitude
   - Desvios em relação à media
   - Variância
   - Desvio Padrão
   - Coeficiente de Variação
3. Medidas de Posição
   - Quartis
   - Percentis
   - Escore padrão

Quartis

Definition

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

- Q_1, primeiro quartil, representa os primeiros 25% dos dados
- Q_2, segundo quartil, representa os primeiros 50% dos dados
- Q_3, terceiro quartil, representa os primeiros 75% dos dados

Pergunta

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

Quartis

Definition

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

- Q_1, primeiro quartil, representa os primeiros 25% dos dados
- Q_2, segundo quartil, representa os primeiros 50% dos dados
- Q_3, terceiro quartil, representa os primeiros 75% dos dados

Pergunta

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

Quartis

Definition

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

- Q_1, primeiro quartil, representa os primeiros 25% dos dados
- Q_2, segundo quartil, representa os primeiros 50% dos dados
- Q_3, terceiro quartil, representa os primeiros 75% dos dados

Pergunta

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

Quartis

Definition

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

- Q_1, primeiro quartil, representa os primeiros 25% dos dados
- Q_2, segundo quartil, representa os primeiros 50% dos dados
- Q_3, terceiro quartil, representa os primeiros 75% dos dados

Pergunta

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

Quartis

Definition

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

- Q_1, primeiro quartil, representa os primeiros 25% dos dados
- Q_2, segundo quartil, representa os primeiros 50% dos dados
- Q_3, terceiro quartil, representa os primeiros 75% dos dados

Pergunta

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

Quartis

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!).

Quartis

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

1 Medidas de Tendência Central
- Média
- Mediana
- Moda
- Comparação entre as Medidas Centrais

2 Medidas de Dispersão
- Amplitude
- Desvios em relação à media
- Variância
- Desvio Padrão
- Coeficiente de Variação

3 Medidas de Posição
- Quartis
- Percentis
- Escore padrão

Percentis

•

Sumário

1 Medidas de Tendência Central
- Média
- Mediana
- Moda
- Comparação entre as Medidas Centrais

2 Medidas de Dispersão
- Amplitude
- Desvios em relação à media
- Variância
- Desvio Padrão
- Coeficiente de Variação

3 Medidas de Posição
- Quartis
- Percentis
- Escore padrão

Escore padrão