Estatística Descritiva I

Bacharelado em Economia - FEA - Noturno

1º Semestre 2017

Profs. Gilberto A. Paula & Vanderlei C. Bueno

Sumário

Objetivos da Aula

Inicialmente o objetivo principal desta aula é apresentar a Estatística como uma metodologia científica para obtenção, organização e análise de dados oriundos das mais variadas áreas das Ciências Experimentais, visando auxiliar na tomada de decisões em situações de incerteza.

Objetivos da Aula

Inicialmente o objetivo principal desta aula é apresentar a Estatística como uma metodologia científica para obtenção, organização e análise de dados oriundos das mais variadas áreas das Ciências Experimentais, visando auxiliar na tomada de decisões em situações de incerteza.

Em seguida, algumas medidas de tendência central (posição ou localização) e de variabilidade (dispersão) serão introduzidas.

Sumário

Origem da Estatística

A origem da Estatística está relacionada com a coleta e construção de tabelas de dados para o governo. Há, por exemplo, registros de presos de guerra egípcios de 5000 A.C., do censo chinês de 2000 A.C. e descrição detalhada de coleta de dados em livros de Constantinopla de 310 A.C..

Evolução da Estatística

A situação evoluiu e a coleta de dados representa apenas um dos aspectos da Estatística. No século XIX, o desenvolvimento do cálculo de probabilidade e de outras metodologias matemáticas, tais como

Evolução da Estatística

A situação evoluiu e a coleta de dados representa apenas um dos aspectos da Estatística. No século XIX, o desenvolvimento do cálculo de probabilidade e de outras metodologias matemáticas, tais como

- Método de Mínimos Quadrados (Legendre, 1805)

Evolução da Estatística

A situação evoluiu e a coleta de dados representa apenas um dos aspectos da Estatística. No século XIX, o desenvolvimento do cálculo de probabilidade e de outras metodologias matemáticas, tais como

- Método de Mínimos Quadrados (Legendre, 1805)
- Distribuição Normal (Gauss, 1809)

Evolução da Estatística

A situação evoluiu e a coleta de dados representa apenas um dos aspectos da Estatística. No século XIX, o desenvolvimento do cálculo de probabilidade e de outras metodologias matemáticas, tais como

- Método de Mínimos Quadrados (Legendre, 1805)
- Distribuição Normal (Gauss, 1809)
- Teorema do Limite Central (Laplace, 1810)

Evolução da Estatística

A situação evoluiu e a coleta de dados representa apenas um dos aspectos da Estatística. No século XIX, o desenvolvimento do cálculo de probabilidade e de outras metodologias matemáticas, tais como

- Método de Mínimos Quadrados (Legendre, 1805)
- Distribuição Normal (Gauss, 1809)
- Teorema do Limite Central (Laplace, 1810)

foram fundamentais para o desenvolvimento da Estatística.

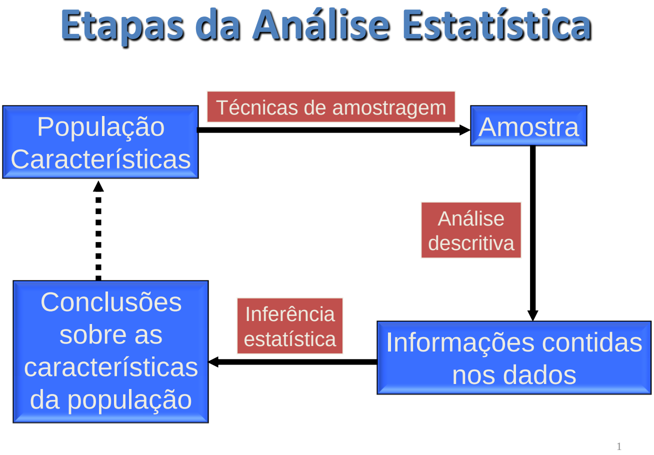
O que é Estatística

Evolução da Estatística

No século XX a Estatística evoluiu como uma área específica do conhecimento a partir do desenvolvimento da Inferência Estatística que faz uso da Teoria das Probabilidades e com ampla aplicação em ciências experimentais.

Situação Atual

A Estatística hoje consiste em uma metodologia científica para obtenção, organização e análise de dados oriundos das mais variadas áreas das ciências experimentais, cujo objetivo principal é auxiliar a tomada de decisões em situações de incerteza.


Etapas da Análise Estatística
População
Características
Técnicas de amostragem
Amostra
Análise
descritiva
Informações contidas
nos dados
Inferência
estatística
Conclusões
sobre as
características
da população
1

Amostragem

Associada à coleta de dados, a tecnologia da amostragem desenvolveu um conjunto de técnicas para obtenção de amostras convenientemente extraídas da população de interesse.

Amostragem

Associada à coleta de dados, a tecnologia da amostragem desenvolveu um conjunto de técnicas para obtenção de amostras convenientemente extraídas da população de interesse.

Exemplos de uso

- Pesquisas de mercado

Amostragem

Associada à coleta de dados, a tecnologia da amostragem desenvolveu um conjunto de técnicas para obtenção de amostras convenientemente extraídas da população de interesse.

Exemplos de uso

- Pesquisas de mercado
- Pesquisas de opinião pública

Amostragem

Associada à coleta de dados, a tecnologia da amostragem desenvolveu um conjunto de técnicas para obtenção de amostras convenientemente extraídas da população de interesse.

Exemplos de uso

- Pesquisas de mercado
- Pesquisas de opinião pública
- Ensaios clínicos

Amostragem

Associada à coleta de dados, a tecnologia da amostragem desenvolveu um conjunto de técnicas para obtenção de amostras convenientemente extraídas da população de interesse.

Exemplos de uso

- Pesquisas de mercado
- Pesquisas de opinião pública
- Ensaios clínicos
- Estudos experimentais

Estatística Descritiva

Etapa inicial da análise utilizada para descrever, organizar e resumir os dados coletados.

Estatística Descritiva

Etapa inicial da análise utilizada para descrever, organizar e resumir os dados coletados.

A disponibilidade de uma grande quantidade de dados e de métodos computacionais muito eficientes revigorou esta área da Estatística.

Probabilidade

A teoria das probabilidades auxilia na modelagem de fenômenos aleatórios, ou seja, aqueles em que está presente a incerteza.

Probabilidade

A teoria das probabilidades auxilia na modelagem de fenômenos aleatórios, ou seja, aqueles em que está presente a incerteza.

É uma ferramenta fundamental para a inferência estatística.

Inferência Estatística

Conjunto de técnicas que permite, a partir de dados amostrais, tirar conclusões sobre a população de interesse, controlando erros.

Sumário

Exemplo 1

Pesquisa de Opinião

Numa pesquisa eleitoral, um instituto de pesquisa tem como objetivo prever o resultado da eleição, utilizando uma amostra da população.

Exemplo 1

Pesquisa de Opinião

Denote por p a proporção de eleitores na população que votariam no candidato A se a eleição fosse hoje.

Exemplo 1

Pesquisa de Opinião

Denote por p a proporção de eleitores na população que votariam no candidato A se a eleição fosse hoje.

Denote por $\hat{p}$ a proporção de eleitores na amostra que votariam no candidato A se a eleição fosse hoje.

Exemplo 1

Pesquisa de Opinião

Denote por p a proporção de eleitores na população que votariam no candidato A se a eleição fosse hoje.

Denote por $\hat{p}$ a proporção de eleitores na amostra que votariam no candidato A se a eleição fosse hoje.

Podemos usar a quantidade $\hat{p}$ para estimar a proporção p.

Exemplo 1

Pesquisa de Opinião

Em vários anos de eleições, os institutos de pesquisa de opinião colhem periodicamente amostras de eleitores para obter as estimativas de intenção de voto da população.

Exemplo 1

Pesquisa de Opinião

Em vários anos de eleições, os institutos de pesquisa de opinião colhem periodicamente amostras de eleitores para obter as estimativas de intenção de voto da população.

As estimativas são fornecidas com a estimativa pontual e uma margem de erro com a qual é construída a estimativa intervalar.

Pesquisa Sensus

Os quadros apresentados a seguir referem-se à intenção de voto para presidente do Brasil para o primeiro e segundo turnos das eleições de 2010. A resposta foi estimulada e única.

Exemplo 1

Pesquisa Sensus

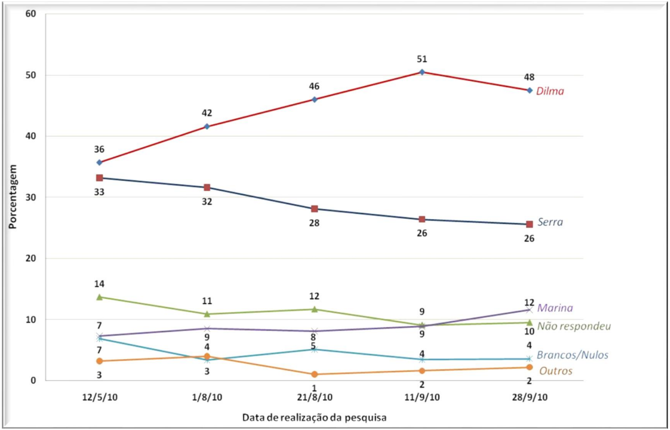
Os quadros apresentados a seguir referem-se à intenção de voto para presidente do Brasil para o primeiro e segundo turnos das eleições de 2010. A resposta foi estimulada e única.

Pergunta Realizada

Se a eleição para presidente fosse hoje e os candidatos fossem estes, em quem o(a) Sr.(Sra) votaria?

Pesquisa Sensus

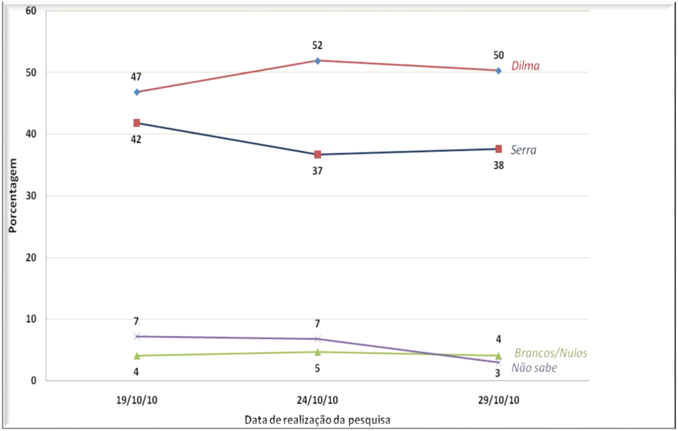
Intenção de voto para presidente do Brasil, 1º Turno – 2010



Pesquisa Sensus, em % do total de votos.
2.000 eleitores - Margem de erro de 2,2% com 95% de confiança.

1

Intenção de voto para presidente do Brasil, 2º Turno – 2010



Pesquisa Sensus, em % do total de votos.
2.000 eleitores - Margem de erro de 2,2% com 95% de confiança.

1

Sumário

Definição

Defini-se como variável aleatória qualquer característica associada a uma população. Há dois tipos de variável aleatória: Qualitativa e Quantitativa.

Variável Aleatória Qualitativa

Variável Aleatória Qualitativa

- Nominal: gênero, estado civil, cor dos olhos

Variável Aleatória Qualitativa

- Nominal: gênero, estado civil, cor dos olhos
- Ordinal: classe social, grau de instrução, faixa etária

Variável Aleatória Qualitativa

- Nominal: gênero, estado civil, cor dos olhos
- Ordinal: classe social, grau de instrução, faixa etária

Variável Aleatória Quantitativa

Variável Aleatória Qualitativa

- Nominal: gênero, estado civil, cor dos olhos
- Ordinal: classe social, grau de instrução, faixa etária

Variável Aleatória Quantitativa

- Contínua: peso, altura, salário, temperatura

Variável Aleatória Qualitativa

- Nominal: gênero, estado civil, cor dos olhos
- Ordinal: classe social, grau de instrução, faixa etária

Variável Aleatória Quantitativa

- Contínua: peso, altura, salário, temperatura
- Discreta: número de filhos, número de sinistros

Sumário

Medidas de Localização

Medidas de Localização

Algumas medidas de posição: Mínimo, Máximo, Moda, Média, Mediana, Percentis

Medidas de Localização

Algumas medidas de posição: Mínimo, Máximo, Moda, Média, Mediana, Percentis

Medidas de Dispersão

Medidas de Localização

Algumas medidas de posição: Mínimo, Máximo, Moda, Média, Mediana, Percentis

Medidas de Dispersão

Algumas medidas de dispersão: Amplitude, Intervalo Interquartil, Variância, Desvio Padrão, Coeficiente de Variação

Sumário

Medidas de Localização

Medidas de Localização

- Máximo: maior valor da amostra

Medidas de Localização

- Máximo: maior valor da amostra
- Mínimo: menor valor da amostra

Medidas de Localização

- Máximo: maior valor da amostra
- Mínimo: menor valor da amostra
- Moda: valor (ou atributo) que ocorre com maior frequência

Medidas de Localização

- Máximo: maior valor da amostra
- Mínimo: menor valor da amostra
- Moda: valor (ou atributo) que ocorre com maior frequência

Exemplo 2

Medidas de Localização

- Máximo: maior valor da amostra
- Mínimo: menor valor da amostra
- Moda: valor (ou atributo) que ocorre com maior frequência

Exemplo 2

Dados: 4, 5, 4, 6, 5, 8, 4.

Medidas de Localização

- Máximo: maior valor da amostra
- Mínimo: menor valor da amostra
- Moda: valor (ou atributo) que ocorre com maior frequência

Exemplo 2

Dados: 4, 5, 4, 6, 5, 8, 4.

max=8 min=4 mo=4

Média Amostral

Média Amostral

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média amostral (média aritmética) é definida por

Média Amostral

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média amostral (média aritmética) é definida por

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

Média Amostral

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média amostral (média aritmética) é definida por

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

Exemplo 3

Média Amostral

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média amostral (média aritmética) é definida por

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

Exemplo 3

Dados: 2, 5, 3, 7, 8.

Média Amostral

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média amostral (média aritmética) é definida por

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

Exemplo 3

Dados: 2, 5, 3, 7, 8.Portanto

$$\bar{x} = \frac{2 + 5 + 3 + 7 + 8}{5} = 5.$$

Média Geométrica

Média Geométrica

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média geométrica é definida por

Média Geométrica

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média geométrica é definida por

$$\text{Mg} = \sqrt[n]{x_1 x_2 \ldots x_n} = (x_1 x_2 \ldots x_n)^{\frac{1}{n}}$$

Média Geométrica

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média geométrica é definida por

$$\text{Mg} = \sqrt[n]{x_1 x_2 \ldots x_n} = (x_1 x_2 \ldots x_n)^{\frac{1}{n}}$$

Observações

Média Geométrica

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média geométrica é definida por

$$\text{Mg} = \sqrt[n]{x_1 x_2 \ldots x_n} = (x_1 x_2 \ldots x_n)^{\frac{1}{n}}$$

Observações

- $\log(\text{Mg}) = \frac{1}{n} \sum_{i=1}^{n} \log x_i$

Média Geométrica

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média geométrica é definida por

$$\text{Mg} = \sqrt[n]{x_1 x_2 \ldots x_n} = (x_1 x_2 \ldots x_n)^{\frac{1}{n}}$$

Observações

- $\log(\text{Mg}) = \frac{1}{n}\sum_{i=1}^{n} \log x_i$
- A média geométrica é mais apropriada que a média aritmética para descrever crescimentos proporcionais

Média Geométrica

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a média geométrica é definida por

$$\text{Mg} = \sqrt[n]{x_1 x_2 \ldots x_n} = (x_1 x_2 \ldots x_n)^{\frac{1}{n}}$$

Observações

- $\log(\text{Mg}) = \frac{1}{n} \sum_{i=1}^{n} \log x_i$
- A média geométrica é mais apropriada que a média aritmética para descrever crescimentos proporcionais
- Se X e Y são observadas nas n unidades experimentais, então $\text{Mg}(\frac{X}{Y}) = \frac{\text{Mg}(X)}{\text{Mg}(Y)}$

Exemplo

Exemplo

Dados: 2, 5, 3, 7, 8.

Exemplo

Dados: 2, 5, 3, 7, 8.Portanto

$$\begin{aligned} Mg &= \sqrt[5]{2 \times 5 \times 3 \times 7 \times 8} \\ &= \sqrt[5]{1680} \\ &= 4,416. \end{aligned}$$

Mediana

Mediana

A mediana é o valor da variável que ocupa a posição central de um conjunto de n dados ordenados.

Mediana

A mediana é o valor da variável que ocupa a posição central de um conjunto de n dados ordenados.

Posição da Mediana

Mediana

A mediana é o valor da variável que ocupa a posição central de um conjunto de n dados ordenados.

Posição da Mediana

Para um conjunto de n dados ordenados a medida corresponde ao valor que ocupa a posição

$$\frac{n+1}{2}.$$

Mediana

A mediana é o valor da variável que ocupa a posição central de um conjunto de n dados ordenados.

Posição da Mediana

Para um conjunto de n dados ordenados a medida corresponde ao valor que ocupa a posição

$$\frac{n+1}{2}.$$

Não necessariamente a mediana é um dos valores da amostra.

Exemplo 4

Exemplo 4

Dados: 2, 3, 6, 7, 8. Temos que n=5 (n ímpar).

Exemplo 4

Dados: 2, 3, 6, 7, 8. Temos que n=5 (n ímpar). A mediana é o valor que está na posição $\frac{5+1}{2} = 3$.

Exemplo 4

Dados: 2, 3, 6, 7, 8. Temos que n=5 (n ímpar). A mediana é o valor que está na posição $\frac{5+1}{2} = 3$. Portanto

$$\text{Md} = 6.$$

Mediana

Exemplo 4

Dados: 2, 3, 6, 7, 8. Temos que n=5 (n ímpar). A mediana é o valor que está na posição $\frac{5+1}{2} = 3$. Portanto

$$\text{Md} = 6.$$

Exemplo 5

Exemplo 4

Dados: 2, 3, 6, 7, 8. Temos que n=5 (n ímpar). A mediana é o valor que está na posição $\frac{5+1}{2} = 3$. Portanto

$$\text{Md} = 6.$$

Exemplo 5

Dados: 1, 2, 4, 6, 8, 9. Temos que n=6 (n par).

Exemplo 4

Dados: 2, 3, 6, 7, 8. Temos que n=5 (n ímpar). A mediana é o valor que está na posição $\frac{5+1}{2} = 3$. Portanto

$$\text{Md} = 6.$$

Exemplo 5

Dados: 1, 2, 4, 6, 8, 9. Temos que n=6 (n par). A mediana é o valor que está na posição $\frac{6+1}{2} = 3,5$.

Exemplo 4

Dados: 2, 3, 6, 7, 8. Temos que n=5 (n ímpar). A mediana é o valor que está na posição $\frac{5+1}{2} = 3$. Portanto

$$\text{Md} = 6.$$

Exemplo 5

Dados: 1, 2, 4, 6, 8, 9. Temos que n=6 (n par). A mediana é o valor que está na posição $\frac{6+1}{2} = 3,5$. Portanto

$$\text{Md} = \frac{4+6}{2} = 5.$$

Percentil

Percentil

O percentil de ordem $p \times 100$, em que $0 < p < 1$, é o valor de um conjunto de n dados ordenados que ocupa a posição

$$p \times (n + 1).$$

Percentil

O percentil de ordem $p \times 100$, em que $0 < p < 1$, é o valor de um conjunto de n dados ordenados que ocupa a posição

$$p \times (n+1).$$

Casos Particulares

Percentil

O percentil de ordem $p \times 100$, em que $0 < p < 1$, é o valor de um conjunto de n dados ordenados que ocupa a posição

$$p \times (n+1).$$

Casos Particulares

- percentil 25 = primeiro quartil (Q1)

Percentil

O percentil de ordem $p \times 100$, em que $0 < p < 1$, é o valor de um conjunto de n dados ordenados que ocupa a posição

$$p \times (n+1).$$

Casos Particulares

- percentil 25 = primeiro quartil (Q1)
- percentil 50 = mediana ou segundo quartil (Md)

Percentil

O percentil de ordem $p \times 100$, em que $0 < p < 1$, é o valor de um conjunto de n dados ordenados que ocupa a posição

$$p \times (n + 1).$$

Casos Particulares

- percentil 25 = primeiro quartil (Q1)
- percentil 50 = mediana ou segundo quartil (Md)
- percentil 75 = terceiro quartil (Q3)

Percentil

O percentil de ordem $p \times 100$, em que $0 < p < 1$, é o valor de um conjunto de n dados ordenados que ocupa a posição

$$p \times (n + 1).$$

Casos Particulares

- percentil 25 = primeiro quartil (Q1)
- percentil 50 = mediana ou segundo quartil (Md)
- percentil 75 = terceiro quartil (Q3)
- percentil 10 = primeiro decil

Quartil Q1

Quartil Q1

O primeiro quartil (percentil 25) é o valor de um conjunto de n dados ordenados que ocupa a posição

$$0,25 \times (n+1) = \frac{n+1}{4}.$$

Quartil Q1

O primeiro quartil (percentil 25) é o valor de um conjunto de n dados ordenados que ocupa a posição

$$0,25 \times (n+1) = \frac{n+1}{4}.$$

Quartil Q3

Quartil Q1

O primeiro quartil (percentil 25) é o valor de um conjunto de n dados ordenados que ocupa a posição

$$0,25 \times (n+1) = \frac{n+1}{4}.$$

Quartil Q3

O terceiro quartil (percentil 75) é o valor de um conjunto de n dados ordenados que ocupa a posição

$$0,75 \times (n+1) = \frac{3(n+1)}{4}.$$

Quartil Q1

O primeiro quartil (percentil 25) é o valor de um conjunto de n dados ordenados que ocupa a posição

$$0,25 \times (n+1) = \frac{n+1}{4}.$$

Quartil Q3

O terceiro quartil (percentil 75) é o valor de um conjunto de n dados ordenados que ocupa a posição

$$0,75 \times (n+1) = \frac{3(n+1)}{4}.$$

Não necessariamente Q1 e Q3 são valores da amostra. A mediana corresponde ao segundo quartil, Md = Q2.

Exemplo 5

Exemplo 5

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7.

Exemplo 5

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7. Temos n=10, assim a mediana fica dada por

$$\text{Md} = (3,0 + 3,1)/2 = 3,05.$$

Exemplo 5

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7. Temos n=10, assim a mediana fica dada por

$$\text{Md} = (3,0 + 3,1)/2 = 3,05.$$

Quartil Q1

Exemplo 5

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7. Temos n=10, assim a mediana fica dada por

$$\text{Md} = (3,0 + 3,1)/2 = 3,05.$$

Quartil Q1

O primeiro quartil é o valor que está na posição $\frac{10+1}{4} = 2,75$.

Exemplo 5

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7. Temos n=10, assim a mediana fica dada por

$$\text{Md} = (3,0 + 3,1)/2 = 3,05.$$

Quartil Q1

O primeiro quartil é o valor que está na posição $\frac{10+1}{4} = 2,75$. Por simplicidade tomaremos a média entre os valores nas posições 2 e 3.

Exemplo 5

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7. Temos n=10, assim a mediana fica dada por

$$\text{Md} = (3,0 + 3,1)/2 = 3,05.$$

Quartil Q1

O primeiro quartil é o valor que está na posição $\frac{10+1}{4} = 2,75$. Por simplicidade tomaremos a média entre os valores nas posições 2 e 3. Assim

$$Q1 = \frac{2,0 + 2,1}{2} = 2,05.$$

Exemplo 5

Quartil Q3

Exemplo 5

Quartil Q3

O terceiro quartil é o valor que está na posição $\frac{3(10+1)}{4} = 8,25$.

Exemplo 5

Quartil Q3

O terceiro quartil é o valor que está na posição $\frac{3(10+1)}{4} = 8,25$. Por simplicidade tomaremos a média entre os valores nas posições 8 e 9.

Quartil Q3

O terceiro quartil é o valor que está na posição $\frac{3(10+1)}{4} = 8,25$. Por simplicidade tomaremos a média entre os valores nas posições 8 e 9. Assim

$$Q3 = \frac{3,7 + 6,1}{2} = 4,9.$$

Exemplo 6

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6.

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Mediana

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Mediana

É o valor que está na posição $\frac{11+1}{2} = 6$.

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Mediana

É o valor que está na posição $\frac{11+1}{2} = 6$. Portanto, Md = 5,3.

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Mediana

É o valor que está na posição $\frac{11+1}{2} = 6$. Portanto, Md = 5,3.

Quartil Q1

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Mediana

É o valor que está na posição $\frac{11+1}{2} = 6$. Portanto, Md = 5,3.

Quartil Q1

É o valor que está na posição $\frac{(11+1)}{4} = 3$.

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Mediana

É o valor que está na posição $\frac{11+1}{2} = 6$. Portanto, Md = 5,3.

Quartil Q1

É o valor que está na posição $\frac{(11+1)}{4} = 3$. Portanto, Q1 = 1,7.

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Mediana

É o valor que está na posição $\frac{11+1}{2} = 6$. Portanto, Md = 5,3.

Quartil Q1

É o valor que está na posição $\frac{(11+1)}{4} = 3$. Portanto, Q1 = 1,7.

Quartil Q3

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Mediana

É o valor que está na posição $\frac{11+1}{2} = 6$. Portanto, Md = 5,3.

Quartil Q1

É o valor que está na posição $\frac{(11+1)}{4} = 3$. Portanto, Q1 = 1,7.

Quartil Q3

É o valor que está na posição $\frac{3(11+1)}{4} = 9$.

Exemplo 6

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6. Agora temos n=11.

Mediana

É o valor que está na posição $\frac{11+1}{2} = 6$. Portanto, Md = 5,3.

Quartil Q1

É o valor que está na posição $\frac{(11+1)}{4} = 3$. Portanto, Q1 = 1,7.

Quartil Q3

É o valor que está na posição $\frac{3(11+1)}{4} = 9$. Portanto, Q3 = 12,9.

Previsão PIB de 2016

Exercício

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2017.

Exercício

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2017.
Dados: 0,0; 0,5; 0,2; 1,0; 0,4; 0,8; 0,3; 0,5; 0,7.

Exercício

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2017.
Dados: 0,0; 0,5; 0,2; 1,0; 0,4; 0,8; 0,3; 0,5; 0,7.

Com base nessa amostra:

Exercício

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2017.
Dados: 0,0; 0,5; 0,2; 1,0; 0,4; 0,8; 0,3; 0,5; 0,7.

Com base nessa amostra:

- qual o valor médio?

Exercício

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2017.
Dados: 0,0; 0,5; 0,2; 1,0; 0,4; 0,8; 0,3; 0,5; 0,7.

Com base nessa amostra:

- qual o valor médio?
- qual o valor mediano?

Exercício

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2017.
Dados: 0,0; 0,5; 0,2; 1,0; 0,4; 0,8; 0,3; 0,5; 0,7.

Com base nessa amostra:

- qual o valor médio?
- qual o valor mediano?
- qual o percentil 90?

Sumário

Exemplo 7

Considere as notas de um teste de 3 grupos de alunos:

Exemplo 7

Considere as notas de um teste de 3 grupos de alunos:

- Grupo 1: 3, 4, 5, 6, 7

Exemplo 7

Considere as notas de um teste de 3 grupos de alunos:

- Grupo 1: 3, 4, 5, 6, 7
- Grupo 2: 1, 3, 5, 7, 9

Exemplo 7

Considere as notas de um teste de 3 grupos de alunos:

- Grupo 1: 3, 4, 5, 6, 7
- Grupo 2: 1, 3, 5, 7, 9
- Grupo 3: 5, 5, 5, 5, 5

Exemplo 7

Considere as notas de um teste de 3 grupos de alunos:

- Grupo 1: 3, 4, 5, 6, 7
- Grupo 2: 1, 3, 5, 7, 9
- Grupo 3: 5, 5, 5, 5, 5

Note que

Exemplo 7

Considere as notas de um teste de 3 grupos de alunos:

- Grupo 1: 3, 4, 5, 6, 7
- Grupo 2: 1, 3, 5, 7, 9
- Grupo 3: 5, 5, 5, 5, 5

Note que

- $\bar{x}_1 = \bar{x}_2 = \bar{x}_3 = 5$

Exemplo 7

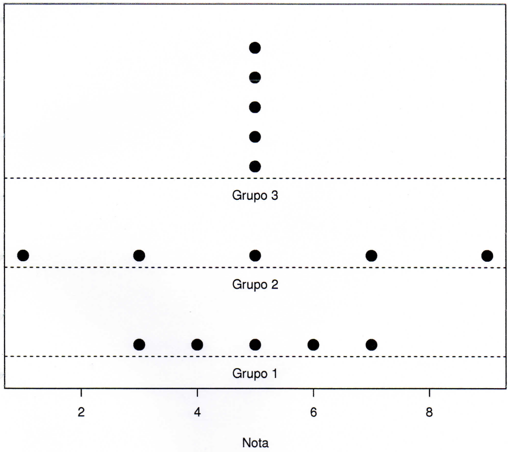
Considere as notas de um teste de 3 grupos de alunos:

- Grupo 1: 3, 4, 5, 6, 7
- Grupo 2: 1, 3, 5, 7, 9
- Grupo 3: 5, 5, 5, 5, 5

Note que

- $\bar{x}_1 = \bar{x}_2 = \bar{x}_3 = 5$
- $\mathrm{Md}_1 = \mathrm{Md}_2 = \mathrm{Md}_3 = 5$

Distribuição das Notas



Medidas de Dispersão

Medidas de Dispersão

Finalidade: encontrar um valor que resuma a variabilidade de um conjunto de dados.

Amplitude

Amplitude

A amplitude é definida por

$$A = \max - \min.$$

Amplitude

A amplitude é definida por

$$A = \text{max} - \text{min}.$$

Para os dados do Exemplo 7 obtemos

Amplitude

A amplitude é definida por

$$A = \max - \min.$$

Para os dados do Exemplo 7 obtemos

- Grupo 1, A=4

Amplitude

A amplitude é definida por

$$A = \text{max} - \text{min}.$$

Para os dados do Exemplo 7 obtemos

- Grupo 1, A=4
- Grupo 2, A=8

Amplitude

A amplitude é definida por

$$A = \max - \min.$$

Para os dados do Exemplo 7 obtemos

- Grupo 1, A=4
- Grupo 2, A=8
- Grupo 3, A=0

Intervalo Interquartil

Intervalo Interquartil

O intervalo interquartil é definido por

$$d = Q3 - Q1.$$

Intervalo Interquartil

O intervalo interquartil é definido por

$$d = Q3 - Q1.$$

Dados Exemplo 5: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7.

Intervalo Interquartil

O intervalo interquartil é definido por

$$d = Q3 - Q1.$$

Dados Exemplo 5: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7.

- Q1=2,05

Intervalo Interquartil

O intervalo interquartil é definido por

$$d = Q3 - Q1.$$

Dados Exemplo 5: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7.

- Q1=2,05
- Q3=4,9

Intervalo Interquartil

O intervalo interquartil é definido por

$$d = Q3 - Q1.$$

Dados Exemplo 5: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7.

- Q1=2,05
- Q3=4,9
- d=4,9 - 2,05 = 2,85

Desvio Médio

Desvio Médio

Observando-se os valores $x_1, \dots, x_n$ para uma variável aleatória X quantitativa, o desvio médio é definido por

$$\mathrm{DM} = \frac{|x_1 - \bar{x}| + \cdots + |x_n - \bar{x}|}{n} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}.$$

Cálculo do desvio médio para os grupos do Exemplo 7

Cálculo do desvio médio para os grupos do Exemplo 7

Grupo 1

$$\text{DM}=\frac{|3-5|+|4-5|+|5-5|+|6-5|+|7-5|}{5}=\frac{6}{5}=1,2.$$

Cálculo do desvio médio para os grupos do Exemplo 7

Grupo 1

$$DM = \frac{|3-5|+|4-5|+|5-5|+|6-5|+|7-5|}{5} = \frac{6}{5} = 1,2.$$

Grupo 2

$$DM = \frac{|1-5|+|3-5|+|5-5|+|7-5|+|9-5|}{5} = \frac{12}{5} = 2,4.$$

Cálculo do desvio médio para os grupos do Exemplo 7

Grupo 1

$$DM = \frac{|3-5|+|4-5|+|5-5|+|6-5|+|7-5|}{5} = \frac{6}{5} = 1,2.$$

Grupo 2

$$DM = \frac{|1-5|+|3-5|+|5-5|+|7-5|+|9-5|}{5} = \frac{12}{5} = 2,4.$$

Grupo 3

$$DM = \frac{|5-5|+|5-5|+|5-5|+|5-5|+|5-5|}{5} = 0.$$

Variância

Variância

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a variância amostral é definida por

$$s^2 = \frac{(x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}.$$

Variância

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a variância amostral é definida por

$$s^2 = \frac{(x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}.$$

A divisão da variância amostral por (n-1) leva a uma estimativa não tendenciosa da variância populacional.

Variância

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a variância amostral é definida por

$$s^2 = \frac{(x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}.$$

A divisão da variância amostral por (n-1) leva a uma estimativa não tendenciosa da variância populacional.

Desvio Padrão

Variância

Observando-se os valores $x_1, \ldots, x_n$ para uma variável aleatória X quantitativa, a variância amostral é definida por

$$s^2 = \frac{(x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}.$$

A divisão da variância amostral por (n-1) leva a uma estimativa não tendenciosa da variância populacional.

Desvio Padrão

O desvio padrão amostral é definido por

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

Cálculo da variância amostral para os grupos do Exemplo 7

Cálculo da variância amostral para os grupos do Exemplo 7

Grupo 1

$$s^2 = \frac{(3-5)^2+(4-5)^2+(5-5)^2+(6-5)^2+(7-5)^2}{4} = \frac{10}{4} = 2,5$$

Cálculo da variância amostral para os grupos do Exemplo 7

Grupo 1

$$s^2 = \frac{(3-5)^2+(4-5)^2+(5-5)^2+(6-5)^2+(7-5)^2}{4} = \frac{10}{4} = 2,5$$

$$\Rightarrow s = \sqrt{2,5} \cong 1,58.$$

Cálculo da variância amostral para os grupos do Exemplo 7

Grupo 1

$$s^2 = \frac{(3-5)^2+(4-5)^2+(5-5)^2+(6-5)^2+(7-5)^2}{4} = \frac{10}{4} = 2,5$$

$$\Rightarrow s = \sqrt{2,5} \cong 1,58.$$

Grupo 2

$$s^2 = \frac{(1-5)^2+(3-5)^2+(5-5)^2+(7-5)^2+(9-5)^2}{4} = \frac{40}{4} = 10$$

Cálculo da variância amostral para os grupos do Exemplo 7

Grupo 1

$$s^2 = \frac{(3-5)^2+(4-5)^2+(5-5)^2+(6-5)^2+(7-5)^2}{4} = \frac{10}{4} = 2,5$$

$\Rightarrow s = \sqrt{2,5} \cong 1,58$.

Grupo 2

$$s^2 = \frac{(1-5)^2+(3-5)^2+(5-5)^2+(7-5)^2+(9-5)^2}{4} = \frac{40}{4} = 10$$

$\Rightarrow s = \sqrt{10} \cong 3,16$.

Cálculo da variância amostral para os grupos do Exemplo 7

Grupo 1

$$s^2 = \frac{(3-5)^2+(4-5)^2+(5-5)^2+(6-5)^2+(7-5)^2}{4} = \frac{10}{4} = 2,5$$

$\Rightarrow s = \sqrt{2,5} \cong 1,58$.

Grupo 2

$$s^2 = \frac{(1-5)^2+(3-5)^2+(5-5)^2+(7-5)^2+(9-5)^2}{4} = \frac{40}{4} = 10$$

$\Rightarrow s = \sqrt{10} \cong 3,16$.

Grupo 3

$$s^2 = \frac{(5-5)^2+(5-5)^2+(5-5)^2+(5-5)^2+(5-5)^2}{4} = 0$$

Cálculo da variância amostral para os grupos do Exemplo 7

Grupo 1

$$s^2 = \frac{(3-5)^2+(4-5)^2+(5-5)^2+(6-5)^2+(7-5)^2}{4} = \frac{10}{4} = 2,5$$

$\Rightarrow s = \sqrt{2,5} \cong 1,58$.

Grupo 2

$$s^2 = \frac{(1-5)^2+(3-5)^2+(5-5)^2+(7-5)^2+(9-5)^2}{4} = \frac{40}{4} = 10$$

$\Rightarrow s = \sqrt{10} \cong 3,16$.

Grupo 3

$$s^2 = \frac{(5-5)^2+(5-5)^2+(5-5)^2+(5-5)^2+(5-5)^2}{4} = 0$$

$\Rightarrow s = \sqrt{0} = 0$.

Variância

Variância

Fórmula alternativa para cálculo da variância

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

Variância

Fórmula alternativa para cálculo da variância

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

Por exemplo, para o Grupo 1 do Exemplo 7 temos que

Variância

Fórmula alternativa para cálculo da variância

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

Por exemplo, para o Grupo 1 do Exemplo 7 temos que
$\sum x_i^2 = 3^2 + 4^2 + 5^2 + 6^2 + 7^2 = 9 + 16 + 25 + 36 + 49 = 135.$

Variância

Fórmula alternativa para cálculo da variância

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

Por exemplo, para o Grupo 1 do Exemplo 7 temos que $\sum x_i^2 = 3^2 + 4^2 + 5^2 + 6^2 + 7^2 = 9 + 16 + 25 + 36 + 49 = 135$. Além disso $\bar{x} = 5$.

Variância

Fórmula alternativa para cálculo da variância

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

Por exemplo, para o Grupo 1 do Exemplo 7 temos que $\sum x_i^2 = 3^2 + 4^2 + 5^2 + 6^2 + 7^2 = 9 + 16 + 25 + 36 + 49 = 135$. Além disso $\bar{x} = 5$. Portanto

$$s^2 = \frac{135 - 5 \times (5)^2}{4} = \frac{10}{4} = 2,5.$$

Previsão PIB de 2016

Exercício

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2015.

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2015.
Dados: 0,0; 0,5; 0,2; 1,0; 0,4; 0,8; 0,3; 0,5; 0,7.

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2015.
Dados: 0,0; 0,5; 0,2; 1,0; 0,4; 0,8; 0,3; 0,5; 0,7.

Com base nessa amostra:

Exercício

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2015.
Dados: 0,0; 0,5; 0,2; 1,0; 0,4; 0,8; 0,3; 0,5; 0,7.

Com base nessa amostra:

- qual o desvio padrão?

Exercício

Previsão PIB de 2016

Supor que $n = 9$ Economistas brasileiros, escolhidos aleatoriamente, foram consultados a respeito do crescimento (em %) previsto para o PIB brasileiro em 2015.
Dados: 0,0; 0,5; 0,2; 1,0; 0,4; 0,8; 0,3; 0,5; 0,7.

Com base nessa amostra:

- qual o desvio padrão?
- qual a diferença interquartil?

Coeficiente de Variação

Coeficiente de Variação

O coeficiente de variação é uma medida de dispersão relativa que elimina o efeito da magnitude dos dados e exprime a variabilidade em relação à média.

Coeficiente de Variação

O coeficiente de variação é uma medida de dispersão relativa que elimina o efeito da magnitude dos dados e exprime a variabilidade em relação à média. O coeficiente de variação amostral é definido por

$$\mathrm{CV} = \frac{s}{\bar{x}} \times 100\%.$$

Exemplo 8

Altura (em cm) de uma amostra de recém-nascidos e de uma amostra de adolescentes

Exemplo 8

Altura (em cm) de uma amostra de recém-nascidos e de uma amostra de adolescentes

Grupo	Média	D. Padrão	CV
Recém-Nascidos	50	6	12%
Adolescentes	160	16	10%

Exemplo 8

Altura (em cm) de uma amostra de recém-nascidos e de uma amostra de adolescentes

Grupo	Média	D. Padrão	CV
Recém-Nascidos	50	6	12%
Adolescentes	160	16	10%

Conclusão

Exemplo 8

Altura (em cm) de uma amostra de recém-nascidos e de uma amostra de adolescentes

Grupo	Média	D. Padrão	CV
Recém-Nascidos	50	6	12%
Adolescentes	160	16	10%

Conclusão

Em relação às médias, as alturas dos adolescentes e dos recém-nascidos apresentam variabilidades muito parecidas.

Exemplo 9

Altura (em m) e Peso (em kg) de uma amostra de alunos.

Exemplo 9

Altura (em m) e Peso (em kg) de uma amostra de alunos.

Variável	Média	D. Padrão	CV
Altura	1,50	0,05	3,3%
Peso	50	3,50	7%

Exemplo 9

Altura (em m) e Peso (em kg) de uma amostra de alunos.

Variável	Média	D. Padrão	CV
Altura	1,50	0,05	3,3%
Peso	50	3,50	7%

Conclusão

Exemplo 9

Altura (em m) e Peso (em kg) de uma amostra de alunos.

Variável	Média	D. Padrão	CV
Altura	1,50	0,05	3,3%
Peso	50	3,50	7%

Conclusão

Os alunos são, aproximadamente, duas vezes mais dispersos quanto ao peso do que quanto à altura.