

# Apostila Completa e Detalhada de Estatística Aplicada

## Apostila Completa e Detalhada de Estatística Aplicada

### 1. Introdução à Estatística

O que é Estatística?

Importância da Estatística

### 2. Tipos de Dados e Variáveis

Dados Quantitativos

Dados Qualitativos (Categóricos)

### 3. Organização e Apresentação dos Dados

Tabelas de Frequência

Gráficos

### 4. Medidas de Tendência Central

Média Aritmética

Mediana

Moda

### 5. Medidas de Dispersão

Variância e Desvio Padrão

Amplitude

Coefficiente de Variação

### 6. Medidas de Posição

Quartis, Quintis, Decis, Percentis

Boxplot

### 7. Probabilidade

Conceitos Básicos

Regras de Probabilidade

Teorema de Bayes

### 8. Distribuições de Probabilidade

Distribuição Discreta

Distribuição Contínua

### 9. Amostragem e Estimação

Tipos de Amostragem

Estimação Pontual e Intervalar

## 10. Testes de Hipóteses

### Passos do Teste

### Testes Comuns

## 11. Regressão e Correlação

### Correlação

### Regressão Linear Simples

## 12. Análise de Variância (ANOVA)

## 13. Séries Temporais e Previsão

## 14. Estatística Não Paramétrica

## 15. Aplicações Práticas e Software Estatístico

## Exemplos, Exercícios Comentados e Estudo de Caso em Estatística Aplicada

### Exemplos Práticos

#### Exemplo 1 — Cálculo de Média, Mediana e Moda

#### Exemplo 2 — Desvio Padrão

### Exercícios Comentados

#### Exercício 1 — Medidas de Tendência Central

#### Exercício 2 — Probabilidade

#### Exercício 3 — Teste de Hipóteses

## Estudo de Caso: Análise da Satisfação do Cliente em uma Loja

### Contexto

#### Passo 1 — Organizar os dados

#### Passo 2 — Calcular média, mediana e moda

#### Passo 3 — Interpretar resultados

## Guia Prático para Software Estatístico: Excel, R e Python

### 1. Microsoft Excel

#### Por que usar?

#### Passos básicos

### 2. R

#### Por que usar?

#### Passos básicos

### 3. Python (com bibliotecas pandas, numpy e matplotlib)

#### Por que usar?

#### Passos básicos

# 1. Introdução à Estatística

## O que é Estatística?

Estatística é o ramo da matemática que coleta, organiza, analisa, interpreta e apresenta dados para ajudar na tomada de decisões baseadas em informações quantitativas.

## Importância da Estatística

- Auxilia em decisões em áreas como economia, engenharia, saúde, marketing, entre outras.
- Permite inferir características de uma população a partir de uma amostra.
- Ajuda a identificar padrões, tendências e relações entre variáveis.

# 2. Tipos de Dados e Variáveis

## Dados Quantitativos

- **Discretos:** Valores inteiros, contáveis (ex: número de alunos).
- **Contínuos:** Valores em um intervalo contínuo (ex: altura, peso).

## Dados Qualitativos (Categóricos)

- **Nominais:** Categorias sem ordem (ex: cores, tipos de produto).
- **Ordinais:** Categorias com ordem (ex: níveis de satisfação).

# 3. Organização e Apresentação dos Dados

## Tabelas de Frequência

- Listagem dos valores das variáveis com suas frequências absolutas, relativas e acumuladas.

## Gráficos

- **Barras:** Para dados categóricos.
- **Histograma:** Para dados quantitativos agrupados em classes.
- **Setores (pizza):** Para proporções.
- **Boxplot:** Para visualização da dispersão e identificação de outliers.

# 4. Medidas de Tendência Central

## Média Aritmética

- Soma dos valores dividida pelo número de observações.

## **Mediana**

- Valor que divide a amostra em duas partes iguais.

## **Moda**

- Valor que ocorre com maior frequência.

# **5. Medidas de Dispersão**

## **Variância e Desvio Padrão**

- Medem a variabilidade dos dados em torno da média.

## **Amplitude**

- Diferença entre o maior e o menor valor.

## **Coefficiente de Variação**

- Razão entre o desvio padrão e a média, útil para comparar variabilidades.

# **6. Medidas de Posição**

## **Quartis, Quintis, Decis, Percentis**

- Dividem o conjunto de dados em partes iguais.

## **Boxplot**

- Representa graficamente as medidas de posição e dispersão.

# **7. Probabilidade**

## **Conceitos Básicos**

- Experimento, espaço amostral, eventos.

## **Regras de Probabilidade**

- Regra da adição, multiplicação, probabilidade condicional.

## **Teorema de Bayes**

- Atualização da probabilidade após obtenção de nova informação.

# **8. Distribuições de Probabilidade**

## **Distribuição Discreta**

- **Binomial:** Número de sucessos em  $n$  ensaios.
- **Poisson:** Número de eventos em intervalo fixo.

## **Distribuição Contínua**

- **Normal (Gaussiana):** Distribuição simétrica, mais comum.
- **Exponencial:** Tempo entre eventos em um processo de Poisson.
- **t-Student, qui-quadrado, F:** Distribuições usadas em testes estatísticos.

## **9. Amostragem e Estimação**

### **Tipos de Amostragem**

- Aleatória simples, estratificada, sistemática, por conglomerados.

### **Estimação Pontual e Intervalar**

- Estimadores, intervalos de confiança para média, proporção e variância.

## **10. Testes de Hipóteses**

### **Passos do Teste**

- Formulação das hipóteses ( $H_0$  e  $H_1$ ).
- Escolha do nível de significância.
- Cálculo da estatística de teste.
- Decisão (rejeitar ou não  $H_0$ ).

### **Testes Comuns**

- Teste t para média, teste qui-quadrado para independência, teste Z para proporções.

## **11. Regressão e Correlação**

### **Correlação**

- Grau e direção da relação linear entre duas variáveis.

### **Regressão Linear Simples**

- Modelo para explicar uma variável dependente a partir de uma independente.

## **12. Análise de Variância (ANOVA)**

- Testa diferenças entre médias de três ou mais grupos.

## **13. Séries Temporais e Previsão**

- Componentes da série (tendência, sazonalidade, ciclo).
- Métodos de suavização.
- Modelos ARIMA.

## 14. Estatística Não Paramétrica

- Testes que não assumem distribuição específica dos dados (ex: teste de Wilcoxon, teste de Kruskal-Wallis).

## 15. Aplicações Práticas e Software Estatístico

- Utilização de Excel, R, SPSS, Python (pandas, statsmodels).
- Exemplo básico de análise no Excel.
- Interpretação de resultados.

# Exemplos, Exercícios Comentados e Estudo de Caso em Estatística Aplicada

## Exemplos Práticos

### Exemplo 1 — Cálculo de Média, Mediana e Moda

Conjunto de dados: 3, 7, 7, 2, 9, 10, 7, 4

- **Média:**  $(3 + 7 + 7 + 2 + 9 + 10 + 7 + 4) / 8 = 49 / 8 = 6,125$
- **Mediana:** Ordenando os dados: 2, 3, 4, 7, 7, 7, 9, 10  
Como há 8 valores, mediana =  $(4^{\circ} \text{ valor} + 5^{\circ} \text{ valor}) / 2 = (7 + 7) / 2 = 7$
- **Moda:** 7 (aparece 3 vezes, mais que os outros valores)

### Exemplo 2 — Desvio Padrão

Dados: 5, 8, 10, 7, 6

1. Calcular a média:  $(5+8+10+7+6)/5 = 36/5 = 7,2$
2. Calcular as diferenças ao quadrado:  
 $(5-7,2)^2 = 4,84$   
 $(8-7,2)^2 = 0,64$   
 $(10-7,2)^2 = 7,84$   
 $(7-7,2)^2 = 0,04$   
 $(6-7,2)^2 = 1,44$
3. Variância (amostral): soma das diferenças<sup>2</sup> / (n-1) =  
 $(4,84+0,64+7,84+0,04+1,44)/4 = 14,8/4 = 3,7$
4. Desvio padrão =  $\sqrt{3,7} \approx 1,92$

## Exercícios Comentados

### Exercício 1 — Medidas de Tendência Central

Conjunto de dados: 12, 15, 20, 20, 18, 15, 22, 17

**Pergunta:** Calcule a média, mediana e moda.

**Solução Comentada:**

- Média =  $(12+15+20+20+18+15+22+17) / 8 = 139 / 8 = 17,375$
- Ordenando: 12, 15, 15, 17, 18, 20, 20, 22
- Mediana =  $(4^{\circ} + 5^{\circ})/2 = (17+18)/2 = 17,5$
- Moda = 15 e 20 (bimodal)

## **Exercício 2 — Probabilidade**

Uma caixa tem 5 bolas vermelhas, 3 azuis e 2 verdes. Retira-se uma bola ao acaso.

**Pergunta:** Qual a probabilidade de sair uma bola azul?

**Solução Comentada:**

- Total de bolas =  $5 + 3 + 2 = 10$
- Probabilidade = número de bolas azuis / total =  $3/10 = 0,3$  ou 30%

## **Exercício 3 — Teste de Hipóteses**

Uma fábrica afirma que a média da vida útil de suas lâmpadas é 1000 horas. Uma amostra de 10 lâmpadas apresentou média de 980 horas com desvio padrão de 40 horas. A um nível de significância de 5%, aceite ou rejeite a hipótese da fábrica.

**Solução Comentada:**

- $H_0: \mu = 1000$
- $H_1: \mu \neq 1000$
- Estatística t =  $(980 - 1000) / (40 / \sqrt{10}) = -20 / (40/3,16) = -20 / 12,65 \approx -1,58$
- Valor crítico t para 9 graus de liberdade e  $\alpha=0,05$  (bicaudal)  $\approx \pm 2,262$
- Como  $-1,58 > -2,262$ , não rejeitamos  $H_0$ .
- Conclusão: Não há evidência suficiente para dizer que a média difere de 1000 horas.

## **Estudo de Caso: Análise da Satisfação do Cliente em uma Loja**

### **Contexto**

Uma loja de eletrônicos quer avaliar a satisfação dos clientes com base numa escala de 1 a 5 (1 = muito insatisfeito, 5 = muito satisfeito). Foram coletadas as seguintes notas de 15 clientes:

3, 4, 5, 3, 2, 4, 4, 5, 3, 4, 2, 5, 5, 4, 3

### **Passo 1 — Organizar os dados**

**Nota Frequência**

2	2
3	4
4	5
5	4

### **Passo 2 — Calcular média, mediana e moda**

- Média =  $(2 \times 2 + 3 \times 4 + 4 \times 5 + 5 \times 4) / 15 = (4 + 12 + 20 + 20) / 15 = 56 / 15 \approx 3,73$
- Mediana: Com 15 dados, o 8º valor ordenado. Ordenando:  
2,2,3,3,3,3,4,4,4,4,4,5,5,5,5  $\rightarrow$  8º valor = 4
- Moda = 4 (mais frequente)

### **Passo 3 — Interpretar resultados**

- A média indica satisfação geral moderada a alta (3,73 de 5).
- A mediana e moda indicam que a maior parte dos clientes dá notas 4, evidenciando satisfação consistente.

# **Guia Prático para Software Estatístico: Excel, R e Python**

## **1. Microsoft Excel**

### **Por que usar?**

- Fácil acesso e interface intuitiva.
- Funções básicas e ferramentas de análise estatística.
- Gráficos simples para visualização.

### **Passos básicos**

#### **a) Inserir dados**

- Digite seus dados em uma coluna (ex: células A1:A10).

#### **b) Calcular média, mediana, moda**

- Média: =MÉDIA (A1:A10)
- Mediana: =MED (A1:A10)
- Moda: =MODA.UNIMODA (A1:A10)

#### **c) Desvio padrão**

- =DESVPAD (A1:A10) para amostra
- =DESVPADP (A1:A10) para população



#### **d) Criar um histograma**

- Selecione os dados.
- Vá em Inserir > Gráficos > Histograma.

#### **e) Testes básicos**

- Instale o suplemento “Análise de Dados” (Opções > Suplementos > Gerir > Suplementos do Excel > Análise de Dados).
- Use para executar testes t, regressão, ANOVA, etc.

## **2. R**

### **Por que usar?**

- Gratuito, poderoso e amplamente usado na academia e indústria.
- Enorme biblioteca de pacotes para estatística e gráficos.

### **Passos básicos**

#### **a) Instalar R e RStudio (ambiente gráfico recomendado)**

- R: <https://cran.r-project.org/>
- RStudio: <https://rstudio.com/products/rstudio/download/>

#### **b) Importar dados**

```
dados <- c(3,7,7,2,9,10,7,4)
```

Ou importar CSV:

```
dados <- read.csv("caminho/do/arquivo.csv")
```

#### **c) Calcular estatísticas básicas**

```
mean(dados)      # Média
median(dados)     # Mediana
sd(dados)         # Desvio padrão
var(dados)        # Variância
```

#### **d) Criar um histograma**

```
hist(dados, main="Histograma dos Dados", xlab="Valores",
      ylab="Frequência")
```

#### e) Teste t

```
t.test(dados, mu=5) # Testa se a média é igual a 5
```

### 3. Python (com bibliotecas pandas, numpy e matplotlib)

#### Por que usar?

- Flexível e popular para análise de dados e machine learning.
- Grande comunidade e muitas bibliotecas.

#### Passos básicos

##### a) Instalar bibliotecas

No terminal ou Jupyter:

```
pip install numpy pandas matplotlib scipy
```

##### b) Importar bibliotecas

```
import numpy as np  
  
import pandas as pd  
  
import matplotlib.pyplot as plt  
  
from scipy import stats
```

##### c) Criar array de dados

```
dados = np.array([3,7,7,2,9,10,7,4])
```

##### d) Estatísticas básicas

```
print("Média:", np.mean(dados))  
  
print("Mediana:", np.median(dados))  
  
print("Desvio padrão:", np.std(dados, ddof=1)) # ddof=1 para amostra
```

##### e) Histograma

```
plt.hist(dados, bins=5, edgecolor='black')  
plt.title('Histograma dos Dados')  
plt.xlabel('Valores')  
plt.ylabel('Frequência')  
plt.show()
```

#### **f) Teste t**

```
resultado = stats.ttest_1samp(dados, 5)  
print("Estatística t:", resultado.statistic)  
print("p-valor:", resultado.pvalue)
```