

Estatística: Conceitos e Organização de Dados

Introdução
Conceitos
Método Estatístico
Dados Estatísticos
Tabulação de Dados
Gráficos

Introdução

- O que é Estatística?
 - É a parte da matemática aplicada que se ocupa em obter conclusões a partir de dados observados

Introdução

- O que é Estatística?
 - É uma metodologia ou conjunto de técnicas que utiliza:
 - Coleta de dados
 - Classificação de dados
 - Apresentação ou representação dos dados
 - Análise e interpretação



Introdução

- Por que estudar?
 - Entender grandes volumes de informação (pesquisas de mercado, índices populacionais, acessos a sites)
 - Qualidade de processos (CEP)
 - Previsões confiáveis (projeções financeiras e populacionais, vida útil de equipamentos)
 - Planejamento (coleta de dados, definição de amostras, planos de contingência)

Introdução

- Ramos da Estatística:
 - Descritiva ou Dedutiva
 - Descreve e analisa grupos de observações (amostras)
 - Uso de medidas e formas de representação (tabelas, gráficos, curvas, ...)
 - Inferencial ou Indutiva
 - Processo de generalização a partir da análise e interpretação de dados amostrais

Introdução

- Ramos da Estatística
 - Probabilidade
 - Parte da Estatística que utiliza métodos e técnicas apropriadas para o estudo de processos com margem de incerteza

Conceitos

- População
 - Maior conjunto tomado como referência na observação de um fenômeno
 - Pode ser finita ou infinita, concreta ou abstrata
- Amostra
 - Subconjunto não-vazio de uma população
 - Excetuando-se a própria população
 - Seu uso gera economia e rapidez dos resultados
 - Deve ser representativa da população

Conceitos

- Amostragem
 - Técnica para escolher amostras que garanta o acaso na escolha
 - Ou seja, qualquer elemento da população tem a mesma probabilidade de ser escolhido
 - Uma boa técnica de amostragem garante representatividade da amostra

Conceitos

- Tipos de amostragem:
 - Amostragem casual ou aleatória
 - Técnica de amostragem onde cada elemento da população tem a mesma probabilidade de ser incluído na amostra (Ex: sorteios)
 - Amostragem proporcional estratificada
 - Quando a população é dividida em estratos e a amostragem é proporcional a eles
 - Ex: Uma amostra de 10% de alunos em uma classe de 40 homens e 60 mulheres, consistirá na escolha aleatória de 4 homens e 6 mulheres

Conceitos

- Amostragem sistemática
 - Quando os elementos da população já estão ordenados, faz-se uma amostragem sistemática
 - Ex: Para se obter uma amostra de 50 elementos em 15 lotes de 100 produtos, temos os passos a seguir:

Conceitos

- 1) $15 \times 100 = 1500$ produtos
- 2) $1500/50 = 30$, logo será selecionado um produto a cada 30 unidades
- 3) Um número de 1 a 30 é escolhido aleatoriamente (por exemplo, o número 7)
- 4) Serão escolhidos os produtos 7, 37, 67, 97, 127, ...

Conceitos

- Variáveis
 - Conjunto de resultados possíveis de um fenômeno
 - Podem ser:
 - Qualitativas: valores expressos por atributos não numéricos (Ex: cor, forma, profissão, ...)
 - Quantitativas:valores expresso por números (Ex: resistência, peso, idade, ...)

Conceitos

- As variáveis quantitativas se dividem:
 - **Contínuas** – variáveis podem assumir infinitos valores num intervalo (Ex: peso de uma pessoa, tamanho do pé de uma pessoa)
 - **Discretas** – variáveis só podem assumir finitos valores num intervalo (Ex: número de alunos numa sala, tamanho do calçado de uma pessoa)

Conceitos

Cuidado com as afirmações:

- Dados ou variáveis contínuas são sempre infinitos (Verdadeiro)
- Dados ou variáveis discretas são sempre finitos (Nem sempre)

Método Estatístico

- Técnica que visa estruturar e organizar as fases ou etapas que devem ser estabelecidas na abordagem de uma observação estatística:
 - Definição do problema
 - Planejamento
 - Coleta de Dados
 - Apuração dos Dados
 - Apresentação dos Dados
 - Análise e interpretação



Planejamento

- Determinação do objetivo da pesquisa
- Definição de métodos a serem utilizados
- Caracterização da Amostra
- Definição do método de aquisição e processamento de dados

Coleta de Dados

- Coleta direta
 - Contínua
 - Periódica
 - Ocasional
- Coleta indireta (inferência a partir da coleta direta)

**Exemplo: Coleta de média de reprovação
do ensino médio a partir da coleta dos
registros diários de freqüência**

Dados Estatísticos

- Organização
 - Brutos – dados não organizados
 - Rol – dados organizados em ordem crescente ou decrescente

Organização de Dados

- Dados discretos

- Freqüência simples
 - Freqüência Acumulada
-
- ```
graph LR; A[Dados discretos] --> B[Freqüência simples]; A --> C[Freqüência Acumulada]; B --> D[Absoluta]; B --> E[Relativa (%)]; C --> F[Absoluta]; C --> G[Relativa (%)]
```

# Organização de Dados

- Dados contínuos
  - Usando intervalos de classes

## Custo de produção da Indústria X

| Custo de produção (em milhões) | Freqüência simples absoluta | Freqüência simples relativa | Freqüência acumulada absoluta | Freqüência acumulada relativa |
|--------------------------------|-----------------------------|-----------------------------|-------------------------------|-------------------------------|
| 10 – 19                        | 10                          | 9,1%                        | 10                            | 9,1%                          |
| 20 – 29                        | 20                          | 18,2%                       | 30                            | 27,3%                         |
| 30 – 39                        | 50                          | 45,4%                       | 80                            | 72,7%                         |
| 40 – 49                        | 20                          | 18,2%                       | 100                           | 90,9%                         |
| 50 – 59                        | 10                          | 9,1%                        | 110                           | 100,0%                        |

# Exemplo

- Complete a tabela
  - Número de alunos nas faixas de peso

| <b>Faixa de Peso</b> | <b>Freqüência simples absoluta</b> | <b>Freqüência simples relativa</b> | <b>Freqüência acumulada absoluta</b> | <b>Freqüência acumulada relativa</b> |
|----------------------|------------------------------------|------------------------------------|--------------------------------------|--------------------------------------|
| 40 – 59              | 40                                 |                                    |                                      |                                      |
| 60 – 79              | 95                                 |                                    |                                      |                                      |
| 80 – 99              | 60                                 |                                    |                                      |                                      |
| 100 – 119            | 10                                 |                                    |                                      |                                      |
| 120 – 139            | 8                                  |                                    |                                      |                                      |

# Organização de Dados

- Dados contínuos
  - Usando ponto médio de classes

## Custo de produção Indústria X

| Custo de produção (em milhões) | Freqüência simples absoluta | Freqüência simples relativa | Freqüência acumulada absoluta | Freqüência acumulada relativa |
|--------------------------------|-----------------------------|-----------------------------|-------------------------------|-------------------------------|
| 14,5                           | 10                          | 9,1%                        | 10                            | 9,1%                          |
| 24,5                           | 20                          | 18,2%                       | 30                            | 27,3%                         |
| 34,5                           | 50                          | 45,4%                       | 80                            | 72,7%                         |
| 44,5                           | 20                          | 18,2%                       | 100                           | 90,9%                         |
| 54,5                           | 10                          | 9,1%                        | 110                           | 100,0%                        |

# Exemplo

- Complete a tabela utilizando o exemplo anterior de sobre número de alunos em faixas de pesos

| Faixa de peso<br>(ponto médio) | Freqüência<br>simples<br>absoluta | Freqüência<br>simples<br>relativa | Freqüência<br>acumulada<br>absoluta | Freqüência<br>acumulada<br>relativa |
|--------------------------------|-----------------------------------|-----------------------------------|-------------------------------------|-------------------------------------|
|                                | 50                                |                                   |                                     |                                     |
|                                | 35                                |                                   |                                     |                                     |
|                                | 6                                 |                                   |                                     |                                     |
|                                | 1                                 |                                   |                                     |                                     |
|                                | 1                                 |                                   |                                     |                                     |

# Organização de Dados

- Índices: razões entre duas grandezas diferentes

$$renda \ per \ capta = \frac{renda}{populacao}$$

- Coeficientes: razão entre o número de ocorrências e o total

$$coeficiente \ de \ inadimpl\u00eancia = \frac{inadimplentes}{clientes}$$

- Taxas: Coeficientes multiplicados por uma potência de dez para facilitar leitura

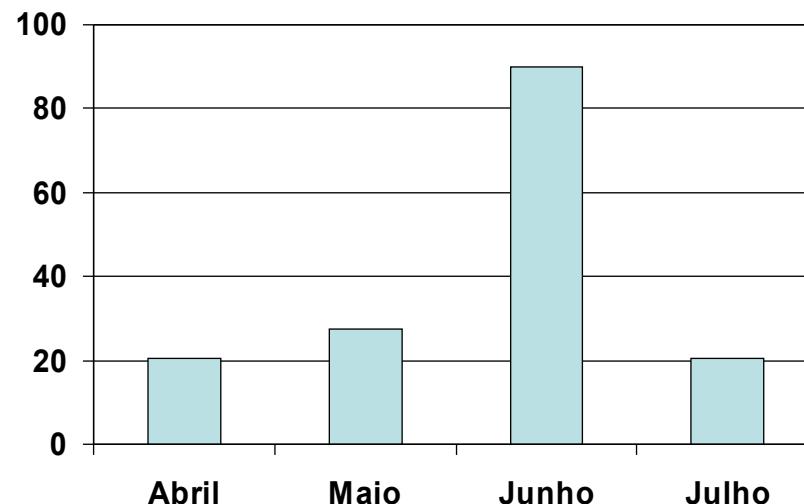
$$taxa \ de \ inadimpl\u00eancia = coeficiente \ de \ inadimpl\u00eancia * 100$$

# Representação Gráfica

- Gráficos estatísticos são formas de apresentação de dados estatísticos com o objetivo de facilitar o entendimento dos fenômenos em estudo
- A seguir são apresentados tipos de gráficos estatísticos

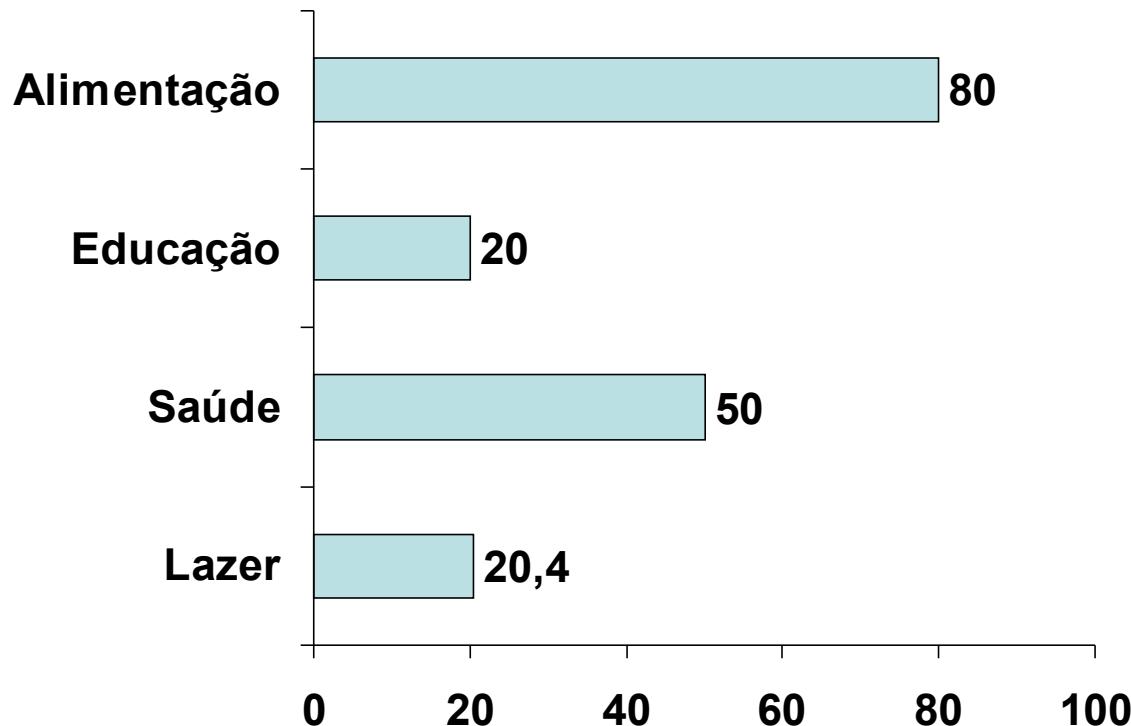
# Representação Gráfica

- Gráfico de Colunas – ideal para comparar dados agrupados em classes ou dados nominais



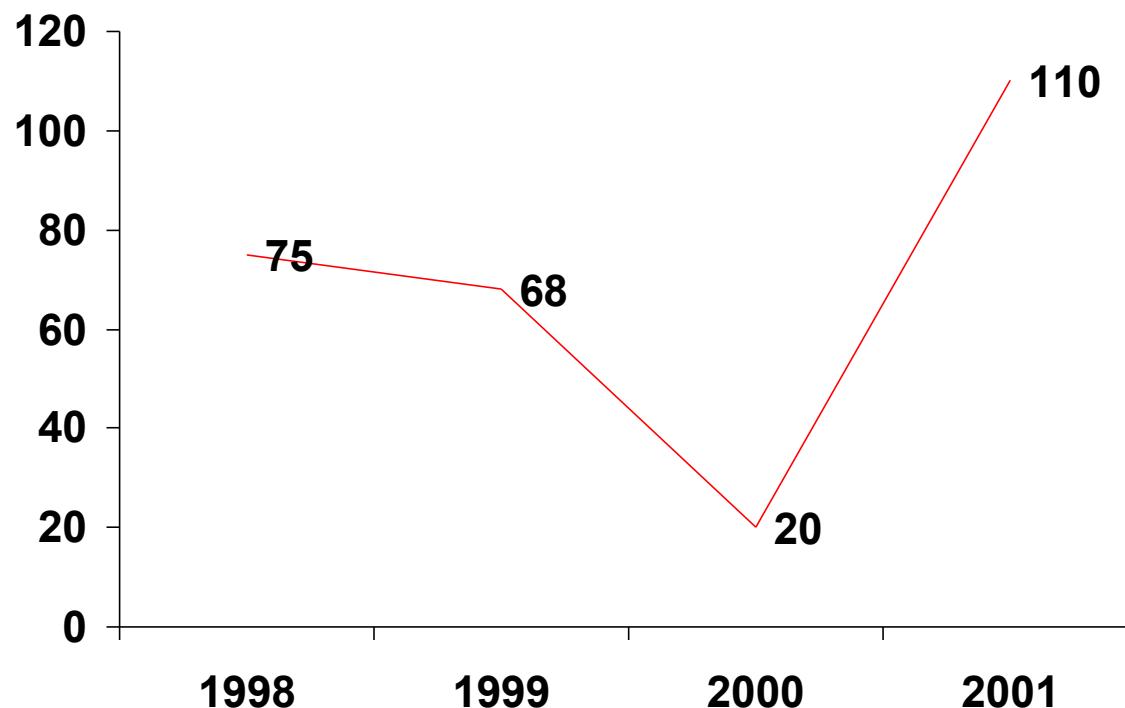
# Representação Gráfica

- Gráfico de Barras – ideal para comparar dados com rótulos longos agrupados em classes ou dados nominais



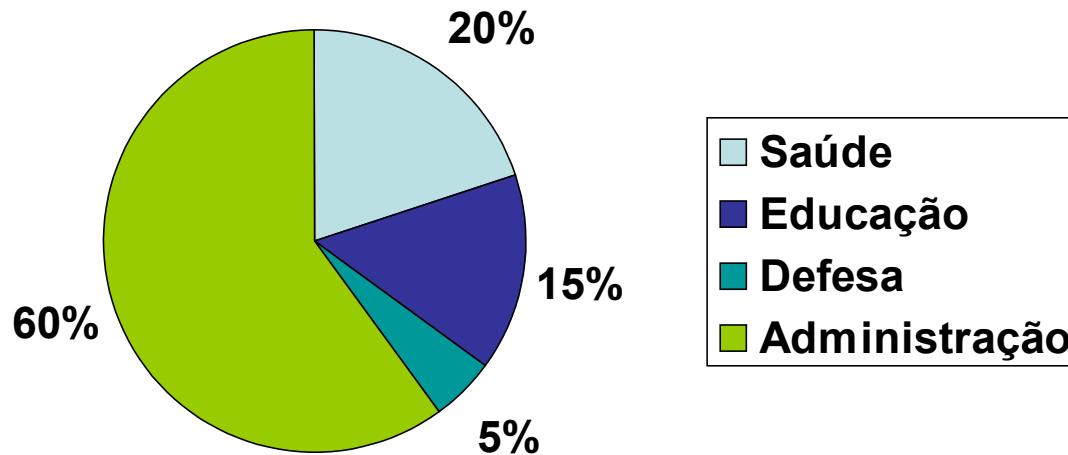
# Representação Gráfica

- Gráfico de Linhas – ideal para indicar variações e tendências de valores de um determinado fenômeno



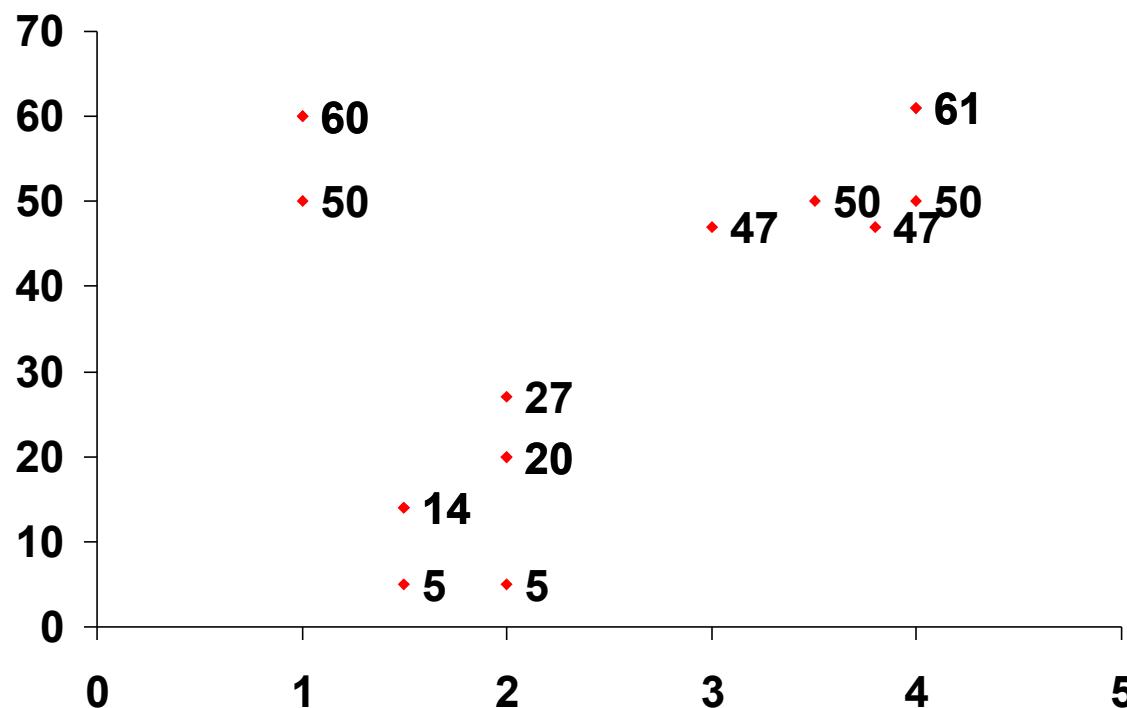
# Representação Gráfica

- Gráfico em Setores ou Pizza – ideal para dados que representam quantidades ou percentuais



# Representação Gráfica

- Gráfico de Dispersão – ideal comparar pares de valores e distribuição de dados



# Exercício

- 1) Elabore os gráficos solicitados utilizando a tabela de custos de produção apresentada anteriormente
  - a) Gráfico de colunas contendo custo X freqüência simples absoluta
  - b) Gráfico de barras contendo custo X freqüência simples relativa
  - c) Gráfico de linhas contendo custo X freqüência simples absoluta
  - d) Gráfico de Pizza contendo custo X freqüência simples relativa

# Exercício

2) Crie um exemplo de cada gráfico a partir de dados obtidos em diversas fontes (revistas, livros, jornais, sites, ...).