

Bioestatística

Aula 1

Dados, Populações, Amostras, Tabelas e Gráficos

Prof. Tiago A. E. Ferreira

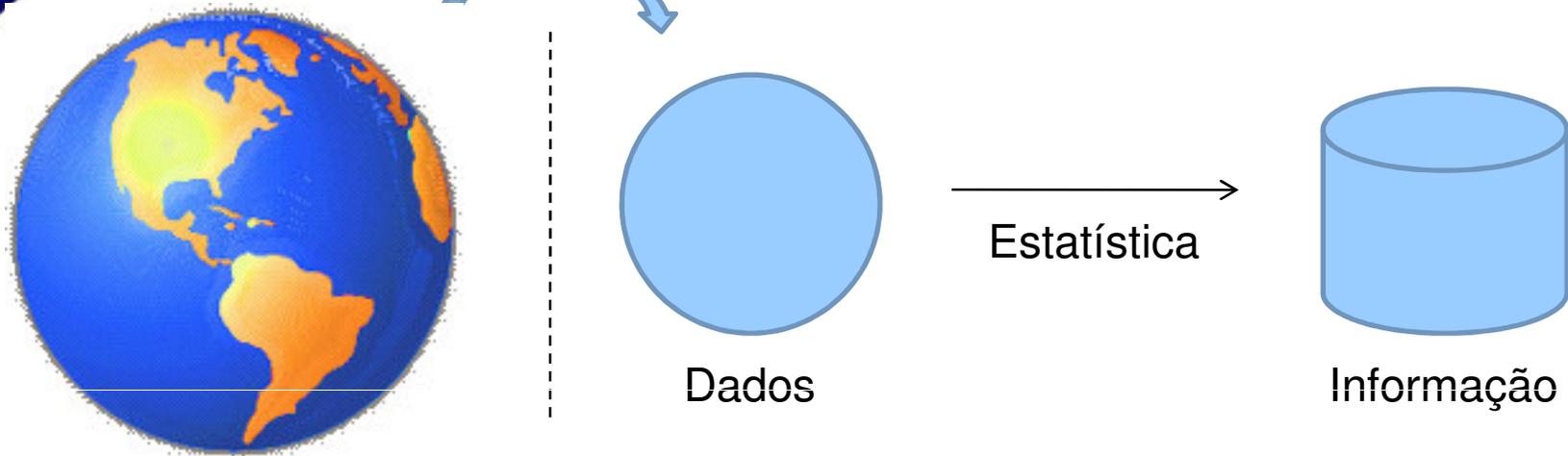


O que é Estatística?



- Segundo a Escola Nacional de Estatística
 - O que modernamente se conhece como Ciências Estatísticas, ou simplesmente Estatística, é um conjunto de técnicas e métodos de pesquisa que entre outros tópicos envolve o planejamento do experimento a ser realizado, a coleta qualificada dos dados, a inferência, o processamento, a análise e a disseminação das informações.

Dados e Informação



Mundo ou Universo

- Dada a existência dos dados, a estatística vem para extrair a informação presente nos dados, de forma clara e mais precisa possível.

Estatística - Ramos



- É possível dividir, grosseiramente, a estatística em partes:
 - **Descritiva:** encarrega-se do levantamento, organização, classificação e descrição dos dados em tabelas, gráficos ou outros recursos visuais, além do cálculo de parâmetros representativos dos dados
 - **Analítica:** trabalha com os dados de forma a estabelecer hipóteses em função desses dados, procede a sua comprovação e, posteriormente, elabora conclusões
 - **Bayesiana:** planejada para reduzir e explicar coerentemente as incertezas que estão presentes na análise de dados através de informações adicionais relevantes

Bioestatística



- A **bioestatística** é a estatística aplicada às ciências que estudam aspectos da vida, ou referentes à vida, como:
 - Medicina
 - Biologia
 - Genética
 - Farmácia
 - Veterinária
 - Agronomia
 - Etc.

População e Amostra



- Na estatística, o termo *população* tem um sentido amplo:
 - **População:** totalidade dos elementos ou de um atributo dos elementos referentes a um certo conjunto em estudo.
 - **Exemplos:**
 - População dos alunos regularmente matriculados no programa de pós-graduação em genética.
 - População dos computadores da UFPE.
 - População de canetas do aluno A da UFPE.

População e Amostra

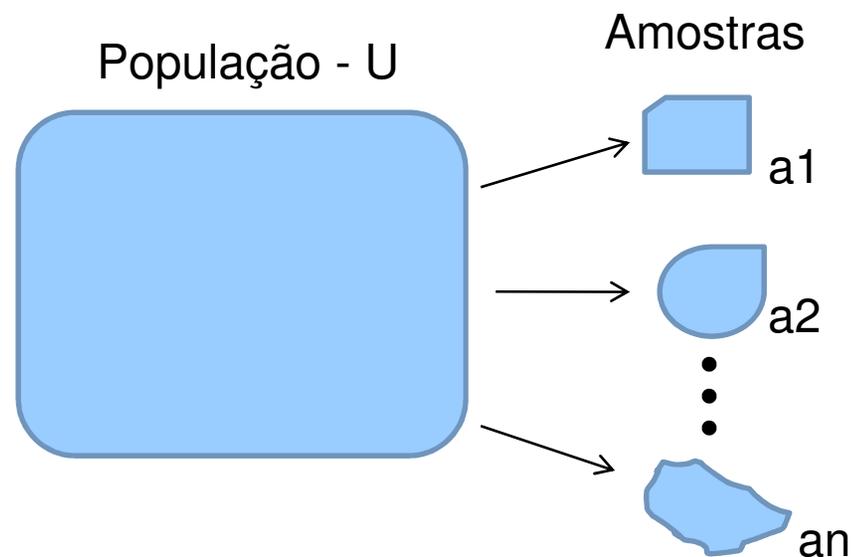


- Uma População pode ser:
 - Enumerável e finita.
 - Inumerável (ou difícil de enumerar) e/ou infinita.
- Dada uma população Inumerável ou Infinita, ou ainda difícil de enumerar, como se pode trabalhar com tal população?
 - Através de partes representativas da população – **AMOSTRAS**.

População e Amostra



- Uma amostra é uma parte da população, ou um subconjunto da população selecionado segundo algum critério



População e Amostra

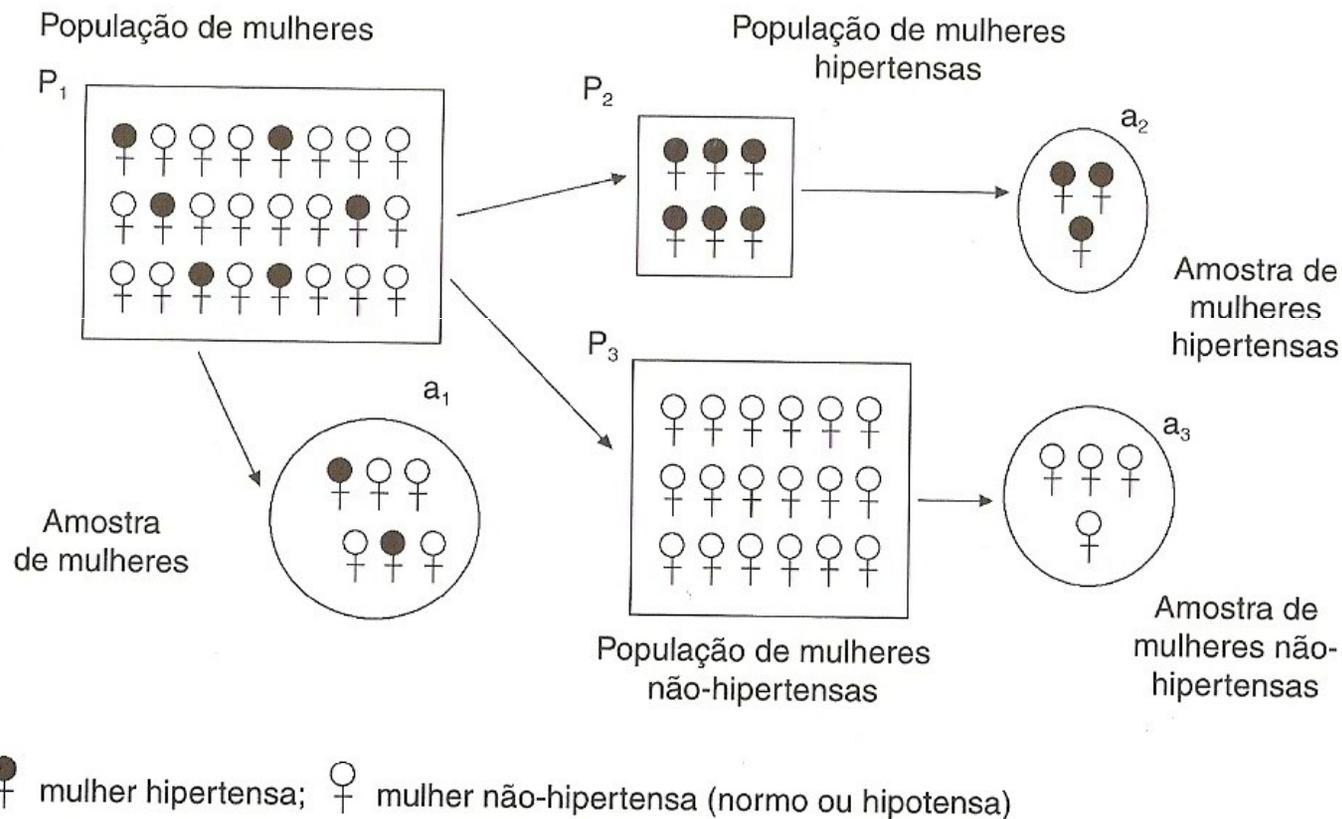


Fig. 1.3 Definição de população e de amostra.

Dados



- ***Dados primários*** são reunidos a partir:
 - Medidas
 - Contagens
 - Experimentos
- Os autores destas medidas, contagens e/ou experimentos são os primeiros a terem acessos a estes dados
 - Espera-se que algum procedimento adequado tenha sido aplicado para garantir a confiabilidade destes dados primários e que as limitações dos mesmo sejam conhecidas

Dados



- Dados secundários são reunidos a partir de diversas fontes, como:
 - Artigos Científicos em Periódicos
 - Eventos Científicos
 - Institutos de Pesquisa e/ou Estatística
 - Etc
- Para este caso, quase sempre, não é possível determinar perfeitamente o seu nível de confiança
 - Espera-se que a fonte garanta um nível mínimo de confiança!
 - Por exemplo, para publicações a sua confiabilidade está geralmente ligada ao seu prestígio.

Censo



- Censo é entendido como o levantamento ou registro de uma certa população, de acordo com um conjunto de critérios.
- Exemplo:
 - Censo demográfico:
 - Inclui informações relacionadas à população de habitantes e habitações, investigando os critérios de tamanho e composição populacional, estrutura familiar, movimentos migratórios, escolaridade, potencial e qualificação de mão de obra, padrões de renda familiar e individual, fecundidade e situação habitacional

Estatísticas no Brasil



- Os recenseamentos são realizados pelo IBGE – *Instituto Brasileiro de Geografia e Estatística* – www.ibge.gov.br
- Os registros de saúde no Brasil podem ser consultados no *site* do Serviço Único de Saúde (SUS) do Ministério da Saúde – www.datasus.gov.br
 - Assistência à Saúde
 - Rede Assistencial
 - Morbidade e Informações Epidemiológicas
 - Estatísticas Vitais e Mortalidade e Nascidos Vivos
 - Recursos Financeiros
 - Informações Demográficas e Socioeconômicas



Conceitos e Informações Preliminares

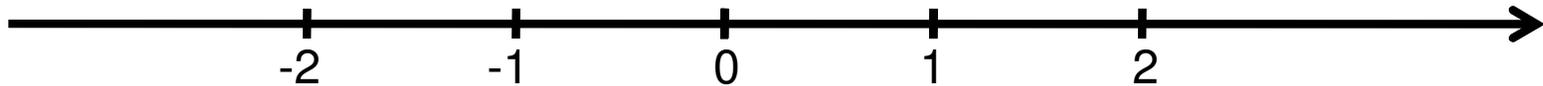


- Para a boa e correta elaboração e análise de dados (de forma estatística) são necessários alguns cuidados quanto:
 - Arredondamento de Dados Numéricos
 - Classificação e Tratamento de Variáveis

Arredondamentos



- Dada a reta dos números reais:



- Tem-se uma quantidade infinita de números entre 1 e 2, por exemplo
 - Isto é equivalente a afirmar que é necessário ter um na quantidade infinita de casas decimais para poder se representar qualquer número real.

Arredondamentos



- Na prática é impossível se trabalhar com uma quantidade infinita de casas decimais, logo defini-se uma quantidade finita de casas decimais para se realizar medidas numéricas!
 - Um efeito desta definição é que apenas um grupo finito de números poderão ser representados exatamente!

Arredondamentos



- Por exemplo,
 - Defini-se utilizar uma casa decimal! Como é possível representar o número 17,342?
 - Este número não poderá ser representado exatamente, mas poderá ser arredondado!
- Questão:
 - Como arredondar este número?

Arredondamentos



- Existem várias formas de arredondamento:
 - Arredondamento Piso ou Inferior:
 - Utiliza-se a representação imediatamente inferior ao número a ser arredondado
 - $\lfloor 17,342 \rfloor = 17,3$
 - Arredondamento Teto ou Superior
 - Utiliza-se a representação imediatamente superior ao número a ser arredondado
 - $\lceil 17,342 \rceil = 17,4$

Arredondamentos



– Arredondamento Simétrico

- Utiliza-se a representação mais próxima ao número a ser arredondado
 - $[17,342] = 17,3$
- Para este tipo de arredondamento, se o número estiver exatamente eqüidistante de duas representações possíveis, como procede o arredondamento?
- Há duas possibilidades:
 - Arredonda-se o número para o número par mais próximo.
 - » 17,25 fica 17,2 e 17,75 fica 17,8
 - Ou, defini-se a regra que se o algarismo logo após o algarismo menos significativo for maior ou igual a 5 o arredondamento é feito para a representação teto, caso contrário para a representação piso.

Classificação de Variável



- Uma variável é uma dimensão ou característica de um dado conjunto de elementos, desde que esta apresenta mais de uma classificação possível.
 - $X = \{x_1, x_2, \dots, x_n\}$
- Caso a dimensão ou característica possua apenas uma classificação é chamada de ***constante***.

Classificação de Variáveis



- É possível dividir as variáveis como:
 - Variáveis quantitativas ou numéricas
 - São avaliadas de forma numérica
 - Exemplo: temperatura = 25° C.
 - Variáveis qualitativas ou literais
 - São expressas como “palavras” ou atributos, ou geram ordenamentos (1º, 2º, ...)
 - Exemplo: Cor do Cabelo = Castanho

Classificação de Variáveis



- As variáveis numéricas ainda podem ser divididas como:
 - Discretas: admitem apenas números inteiros
 - Contínuas: admitem números fracionados, ou seja, números reais.
- Há ainda as variáveis binárias, que tanto podem ser variáveis quantitativas (0 ou 1) como qualitativas (masculino ou feminino)

Estudos Estatísticos



- Dado um estudo científico é possível ter:
 - Estudos observacionais
 - Não há intervenção do pesquisador, ocorrendo apenas uma descrição a respeito de uma dado problema.
 - Método estatístico empregado: **Intervalos de Confiança**
 - Estudos experimentais
 - Há intervenção do pesquisado mediante a adoção de algum tratamento ou alteração da situação em estudo e comparação dos resultados obtidos
 - Método estatístico empregado: **Testes de Significância**

Escolha das Variáveis



- Dado um estudo, como definir as variáveis ou fatores relevantes?
 - Fatores de Interesse direto (relacionadas *a priori*)
 - Fatores que podem modificar a ação dos fatores principais ou que podem elucidar o funcionamento dos fatores principais
 - Fatores relacionados com a técnica experimental utilizada
 - Fatores de classificação, sugeridos por agrupamentos naturais das unidades experimentais
 - Variações proporcionais introduzidas nas unidades experimentais

Levantamento de Dados



- O sucesso das conclusões tiradas de uma dada população depende de como esta população foi selecionada!
 - Dados mal coletados irão carregar distorções para qualquer análise!

Representatividade



- Critério da Proporcionalidade
 - Ao se realizar uma amostragem em uma dada população, procura-se reproduzir as características observáveis da população.
- Dado que o critério da proporcionalidade é obedecido, então a amostra é dita ser ***representativa.***

Representatividade



- Um grande problema prático é a falta de informações sobre a população para a garantia do critério de proporcionalidade.
 - No caso de falta de informação, é utilizado o critério de **aleatoriedade**.
 - Contudo, caso os atributos de uma população seja conhecido, porém ignora ou manipulado, a amostra é considera **tendenciosa**, gerando conclusões **sem consistência**.

Fidedignidade



- A **fidedignidade** dos dados os das informações está relacionada com a **precisão dos dados**, ou a sua **qualidade**.
- Há várias fontes para a queda da precisão dos dados:
 - Instrumentos de medidas.
 - Erro Humano.
 - Sonegação de informação (questionários).

Efeitos Indesejados



- É possível citar alguns efeitos indesejados no levantamento de dados:
 - Efeito Placebo
 - Efeito Rosenthal
 - Efeito Hawthorne
 - Ilusão de Müller-Lyer
 - Efeito Complacencia
 - Efeito Meméria
 - Efeito Desmascaramento

Efeito Placebo



- Ocorre quando um indivíduo participa de um experimento, mas não tratado, acredita está recebendo o tratamento e passa a relatar melhoras em seus sintomas.
 - **Placebo** é uma substância neutra, que não altera o estado do indivíduo ou população.

Efeito Rosenthal



- O efeito Rosenthal ou efeito do experimentador é qualquer mudança ou alteração do padrão de resposta do indivíduo pesquisado provocada, involuntariamente, pelo pesquisador ou experimentador
 - Este efeito pode ocorrer na aplicação de questionário, por exemplo.

Efeito Hawthorne



- O efeito Hawthorne ocorre quando um indivíduo participante de um experimento e tratado passa a responder de forma diferente devido ao fato de estar participando da experiência.

Ilusão de Müller-Lyer



- A Ilusão de Müller-Lyer é verificada quando a simples alteração da ordem das questões de uma pesquisa altera as respostas dos pesquisados.
- Exemplo:
 - Considera a ingestão excessiva de açúcar prejudicial à saúde?
 - Consome muito açúcar?

Efeito Complacência



- O efeito complacência está relacionado com a tendência que os entrevistados apresentam de responder de forma positiva as questões.
- Exemplo:
 - Concorda que o dirigente agiu de forma honesta?
 - Concorda que o dirigente agiu de forma desonesta?

Efeito Memória



- O efeito memória ocorre quando a ordem dos itens das questões pode alterar os resultados de um levantamento de dados.
 - Quanto mais longa a questão maior a exposição a este efeito!
- Exemplo:
 - Considera que o **tabagismo** contribui mais do que o **alcoolismo** para a piora das condições de saúde?
 - Considera que o **alcoolismo** contribui mais do que o **tabagismo** para a piora das condições de saúde?

Efeito Desmascaramento



- O efeito desmascaramento ocorre pela desconfiança do sigilo das suas informações e da sua identidade.

Obtendo Amostras



- Dado que o levantamento de dados foi estruturado com sucesso, existem basicamente dois problemas para a obtenção de amostras:
 - A forma de realizar a amostragem.
 - A determinação do tamanho da amostra.

Formas de Amostragem



- Amostragem aleatória simples
 - Ocorre quando é garantido a mesma chance de seleção para todos os elementos da população – também chamada **amostragem casual**
 - Tentativa de eliminação de viés estatístico
 - Este tipo de amostragem não garante que a amostra obtida seja semelhante a população (critério de proporcionalidade), mas garante a maior probabilidade para que isso ocorra

Formas de Amostragem



- Amostragem Sistemática (Seqüencial)
 - São selecionados elementos de ordem k da população.
 - Exemplo: Dada um classe com 90 alunos ($N=90$, população)
 - Deseja-se uma amostra de 30 alunos ($n=30$). Assim poderiam ser escolhidos 30 alunos seqüencialmente com a ordem $k=3$.
 - Assim, seriam escolhidos os alunos 3,6,9,...,90 ($1k, 2k, 3k, \dots, 30k$)

Formas de Amostragem



- Amostragem Estratificada

- É utilizada quando a população é previamente classificada em estratos, ou classes pré estabelecidas

- Desta forma, uma população com N indivíduos terá coeficientes de proporcionalidade k_j entre cada estrato j , de modo que cada estrato j participe com $n_j = k_j N$

Formas de Amostragem



- Amostragem por Conglomerados ou Área
 - Consiste em efetuar subdivisões em área da população, e compor a amostra com a totalidade dos indivíduos de alguns desses conglomerados
 - Há um grande risco de tendenciamento, porém tem baixo custo na prática.

Formas de Amostragem



- Amostragem por Conveniência
 - Ocorre quando a amostra é formada por dados levantados de fácil obtenção, sem a preocupação da observação dos critérios já discutidos.
 - Exemplos: Deseja-se a amostra da população de crianças de um dado bairro
 - Vai-se a uma creche do bairro e se realiza o levantamento das crianças ali presentes!

Tamanho de uma Amostra



- O tamanho de uma amostra, ou o dimensionamento amostral, é relacionado a uma série de fatores. Os principais são:
 - Erro desejado
 - Risco da conclusão
 - Previsibilidade da variáveis

Representação dos Dados



- De forma básica, os dados podem ser representados de duas formas:
 - Tabelas
 - Gráficos

Dados em Tabelas



- As normas para a apresentação de dados em tabelas podem ser encontradas em:
 - Normas de Apresentação Tabular do Conselho Nacional de Estatística. Revista Brasileira de Estatística, vol. 24, p. 42-60, 1963
 - IBGE – Normas de apresentação tabular. 3ª edição, Rio de Janeiro, 1993.

Dados em Tabelas



- Algumas Regras Gerais:
 - A tabela deve ser simples
 - A tabela deve ser auto-explicativa
 - Não deve haver campo em branco
 - Uniformidade de casas decimais

Exemplo de Tabela



Título

Tabela 1. Número de médicos na população

Cabeçalho

Corpo

Rodapé

| País | Habitantes por Médico |
|-------------|------------------------------|
| Chile | 1.230 |
| Brasil | 1.080 |
| França | 320 |
| EUA | 470 |
| Argentina | 370 |

Fonte: Arango, Hectro G.

Nota: Dados do relatório sobre o desenvolvimento mundial 1990, Banco Mundial, FGV



Organizando os dados



- Ao termino do levantamento dos dados tem-se uma massa de dados
 - Esta massa é chamada de **Dados Brutos e Rol**
- Inicialmente, os dados brutos podem ser ordenados (crescente ou decrescente) para uma melhor visualização (Rol).
 - A partir de então, pode-se definir faixas de variação
 - classes
 - Cada intervalo de classe terá um limite inferior e superior.

Organizando os dados



- Dados Brutos ou Rol
 - Exemplo:
 - Medida Original – Dados Brutos

| Indivíduo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|
| Diâmetro Abdominal (cm) | 88 | 76 | 105 | 94 | 82 | 65 | 72 | 80 | 86 | 122 |

- Ordenamento - Rol

| Indivíduo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|
| Diâmetro Abdominal (cm) | 65 | 72 | 76 | 80 | 82 | 86 | 88 | 94 | 105 | 122 |

Classes



- Uma vez que temos o Rol fica mais clara a visualização da distribuição dos dados
 - Cria-se faixas de variação para os dados, chamadas de **classes**.
 - Tem a finalidade de saber qual o número (ou quantidade) de dados que se situa em cada faixa!
 - Cada Classe têm
 - **Intervalo de classe = *limite superior de classe* – *limite inferior de classe***
 - Normalmente as classes possuem o mesmo intervalo de Classe – **Classes Homogêneas**

Quantidade de Classes



- Não há uma regra totalmente aceita para a determinação da quantidade de classes para uma massa de dados brutos, porém é possível utilizar a fórmula de Struges,

$$C = 1 + \log_2 N$$

- Ou ainda,

$$C = \sqrt{N}$$

Número de Classes

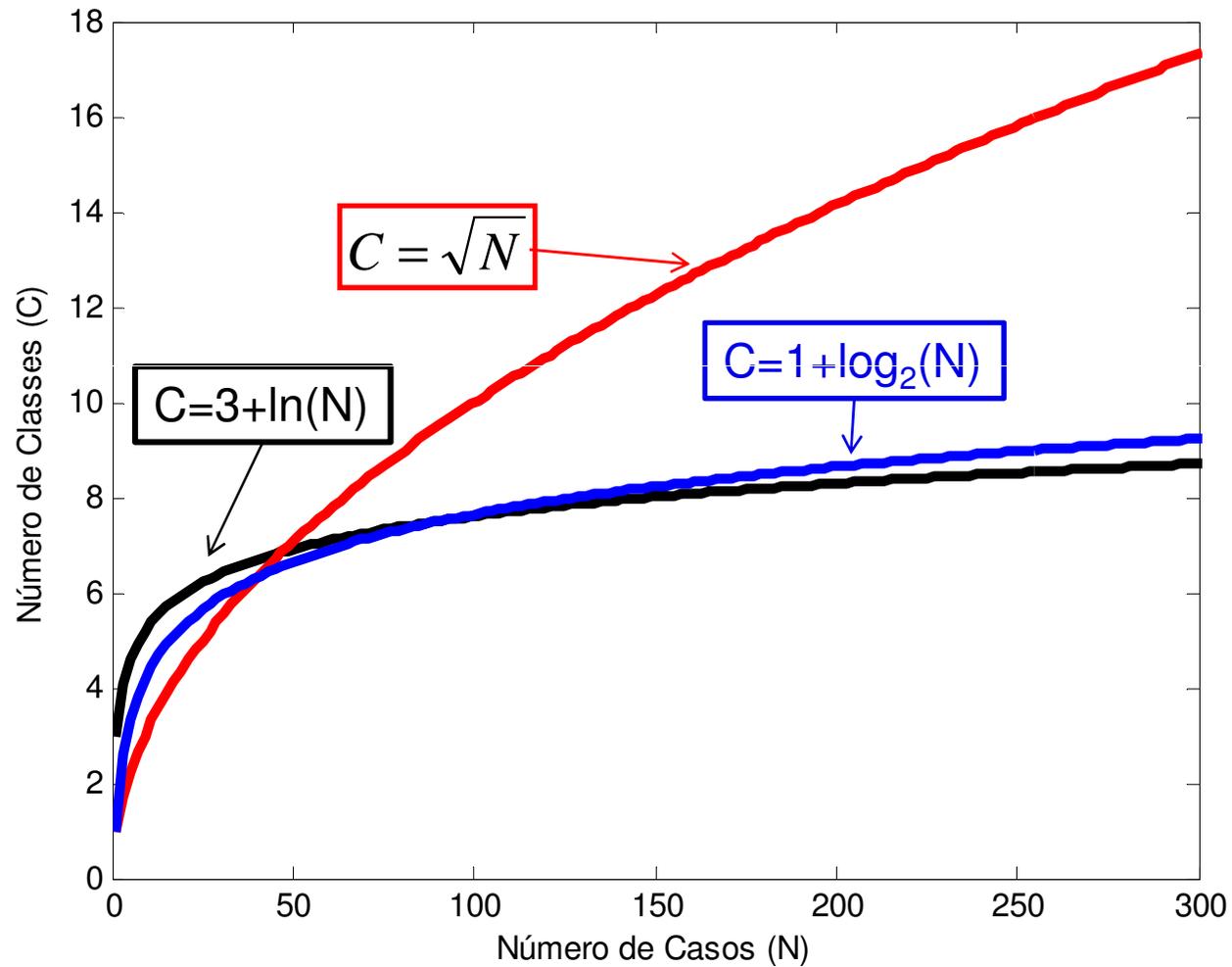


Tabela 2. Número de classes baseado em Sturges

| Número de Casos (N) | Sturges ($C = 1 + \log_2 N$) |
|----------------------------|--|
| 1 | 1 |
| 2 | 2 |
| 3-5 | 3 |
| 6-11 | 4 |
| 12-22 | 5 |
| 23-45 | 6 |
| 46-90 | 7 |
| 91-181 | 8 |
| 182-362 | 9 |
| 365-725 | 10 |

Fonte: Arango, Hectro G.

Número de Classes



Determinação do Número de Classes



- Método p/ Variáveis Discretas (Valores Inteiros)

- Procura-se C em Função de N

- Tenta-se satisfazer:

- $C.I \geq AT + 1$

- $R = \text{Mínimo}$

- $I > 1$

- Onde:

- » $AT = x_n - x_1$ (**A**mplitude **T**otal dos dados no Rol)

- » I é a Amplitude do Intervalo de Classe

- » $R = C.I - (AT+1)$

| N | C |
|-------------|--------|
| 1 até 100 | 4 a 8 |
| 101 até 250 | 5 a 10 |
| + de 250 | 7 ou + |

Determinação do Número de Classes



Exemplo:

– Dado o Rol:

- 10,10,10,11,11,12,12,12,12,13,13,14,14,14,15,15,16,18,19,22

- Temos **$N=20$** .

- **$AT = 22-10 = 12$**

- **Estimar o Número de Classes**

– **Sturges: $C=1+\log_2 20 \approx 5.34 \approx 5$**

– **Pelo Método proposto: C deve ser de 4 a 8!**

» **Como $C \cdot I \geq AT+1$, ou $C \cdot I \geq 13$**

» **E como R deve ser mínimo**

» **Resp.: $C=5$**

| C | I | R |
|---|------------------|---|
| 6 | $13/6 \approx 4$ | 5 |
| 5 | $13/5 \approx 3$ | 2 |
| 4 | $13/4 \approx 4$ | 3 |
| 3 | $13/3 \approx 5$ | 2 |
| 2 | $13/2 \approx 7$ | 1 |

Determinação do Número de Classes



- Método p/ Variáveis Contínuas (Valores Fracionados)
 - Tenta-se satisfazer:
 - $C.I \geq AT + u$
 - $R = \text{Mínimo}$
 - $I > u$
 - Onde u é a menor unidade em que são expressos os valores de AT

Determinação do Número de Classes



- Rotina para a determinação do Número de Classes C (Valores Fracionados)
 - Leia os valores x (Dados Brutos)
 - Ordene os valores de x (Rol)
 - Determine AT
 - Escolha C entre 4 e 8, ou inicie com 4
 - Calcule I , tal que $C \cdot I > AT$
 - Calcule o resto R
 - Se R é par, centralize (o mesmo para cima e para baixo)
 - Se R é ímpar, deixe a maior sobra por último
 - Repita todo o procedimento até encontrar o menor R possível.
 - Em caso de empate, use o valor de C mais próximo do calculado pela fórmula de Sturges

Determinação do Número de Classes



- Exemplo:

| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| X | 0,245 | 0,232 | 0,287 | 0,230 | 0,261 | 0,277 | 0,253 | 0,280 | 0,245 | 0,260 | 0,251 | 0,244 | 0,271 | 0,268 | 0,261 |

- Note que $x_n = 0,287$ e $x_1 = 0,230$
 - Logo $AT = 0,287 - 0,230 = 0,057$
 - $u = 0,001$ (algarismo menos significativo)
 - Para $C=4$, temos $I=0,015$, visto que $C.I \geq AT + u$
 - $C.I = 4 \cdot 0,015 = 0,060$ e $AT+u = 0,057+0,001=0,058$
 - Neste caso: $R=C.I-(AT+u) = 0,060-0,058=0,002$
 - Observe que R é par, logo $n_R = R/u = 2$
 - » Assim, irão “sobrar” os números $0,229$ e $0,288 (\pm u)$

Determinação do Número de Classes



- Continuação do exemplo:
 - Para $C=5$, temos $I=0,012$, visto que $C.I \geq AT + u$
 - $C.I = 5 \cdot 0,012 = 0.060$ e $AT+u = 0,057+0,001=0,058$
 - Neste caso: $R=C.I-(AT+u) = 0.060-0.058=0.002$
 - Observe que R é par, logo $n_R = R/u = 2$
 - Assim, irão “sobrar” os números $0,229$ e $0,288 (\pm u)$
 - Para $C=6$, temos $I=0,010$, visto que $C.I \geq AT + u$
 - $C.I = 5 \cdot 0,012 = 0.060$ e $AT+u = 0,057+0,001=0,058$
 - Neste caso: $R=C.I-(AT+u) = 0.060-0.058=0.002$
 - Observe que R é par, logo $n_R = R/u = 2$
 - Assim, irão “sobrar” os números $0,229$ e $0,288 (\pm u)$

Determinação do Número de Classes



- Continuação do Exemplo:
 - Para $C=7$, temos $I=0,009$, visto que $C.I \geq AT + u$
 - $C.I = 7 \cdot 0,009 = 0.063$ e $AT+u = 0,057+0,001=0,058$
 - Neste caso: $R=C.I-(AT+u) = 0.063-0.058=0.005$
 - Observe que R é par, logo $n_R = R/u = 5$
 - Assim, irão “sobrar” os números 0,228 e 0,290 ($-2u$ e $+3u$)
 - Para $C=8$, temos $I=0,008$, visto que $C.I \geq AT + u$
 - $C.I = 8 \cdot 0,008 = 0.064$ e $AT+u = 0,057+0,001=0,058$
 - Neste caso: $R=C.I-(AT+u) = 0.064-0.058=0.006$
 - Observe que R é par, logo $n_R = R/u = 6$
 - Assim, irão “sobrar” os números 0,227 e 0,290 ($-3u$ e $+3u$)

Determinação do Número de Classes



- Continuação do exemplo
 - Assim temos os menores erros para $C=4, 5$ e 6 .
 - A fórmula de Sturges dá: $C = 1 + \log_2 15 = 4.9069 \cong 5$
 - Assim, $C=5$.

| Classes | fa |
|---------------|----|
| 0,229 – 0,240 | 2 |
| 0,241 – 0,252 | 4 |
| 0,253 – 0,264 | 4 |
| 0,265 – 0,276 | 2 |
| 0,277 – 0,288 | 3 |
| Total | 15 |

Alguns Cuidados



- A representação das Classes deve ser clara!

| Classes | fa |
|-----------------|----|
| 0,229 – 0,240 | 2 |
| 0,241 – 0,252 | 4 |
| 0,253 – 0,264 | 4 |
| 0,265 – 0,276 | 2 |
| 0,277 – 0,288 | 3 |
| Total | 15 |

| Classes | fa |
|----------------|----|
| 0,229 – 0,241 | 2 |
| 0,241 – 0,253 | 4 |
| 0,253 – 0,265 | 4 |
| 0,265 – 0,277 | 2 |
| 0,277 – 0,288 | 3 |
| Total | 15 |

| Classes | fa |
|----------------|----|
| 0,229 – 0,240 | 2 |
| 0,240 – 0,252 | 4 |
| 0,252 – 0,264 | 4 |
| 0,264 – 0,276 | 2 |
| 0,276 – 0,288 | 3 |
| Total | 15 |

Frequências



- A **frequência absoluta** (fa) de uma classe é o número de observações de valores compreendidos em seu intervalo
- Também é possível calcular a **frequência relativa** (fr):

$$fr_j = \frac{fa_j}{\sum_{i=1}^c fa_i}$$

- Pode-se calcular a **frequência acumulada** (FA)

$$FA_j = FA_{j-1} + fa_j, \text{ onde } FA_0 = 0$$

Gráficos

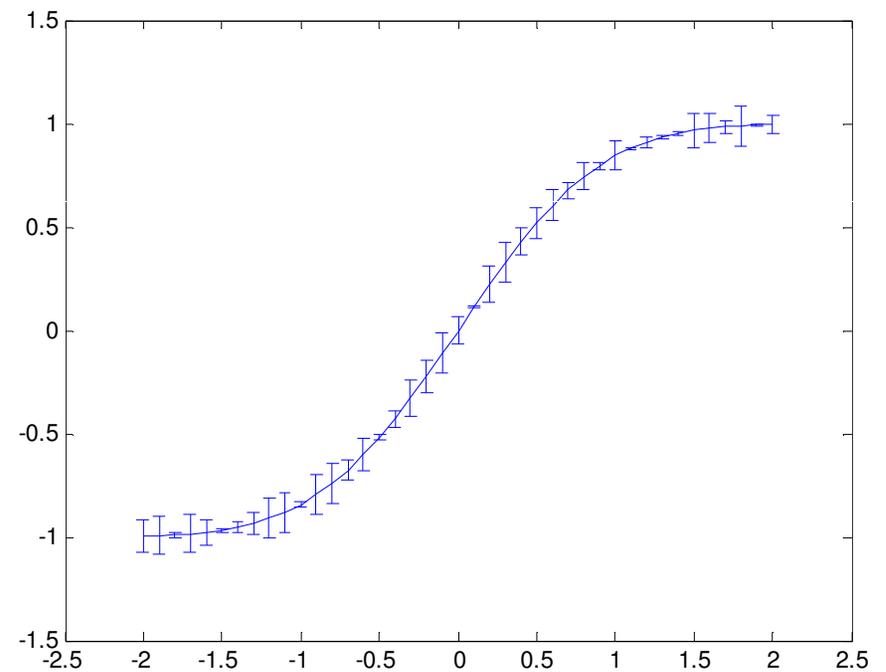
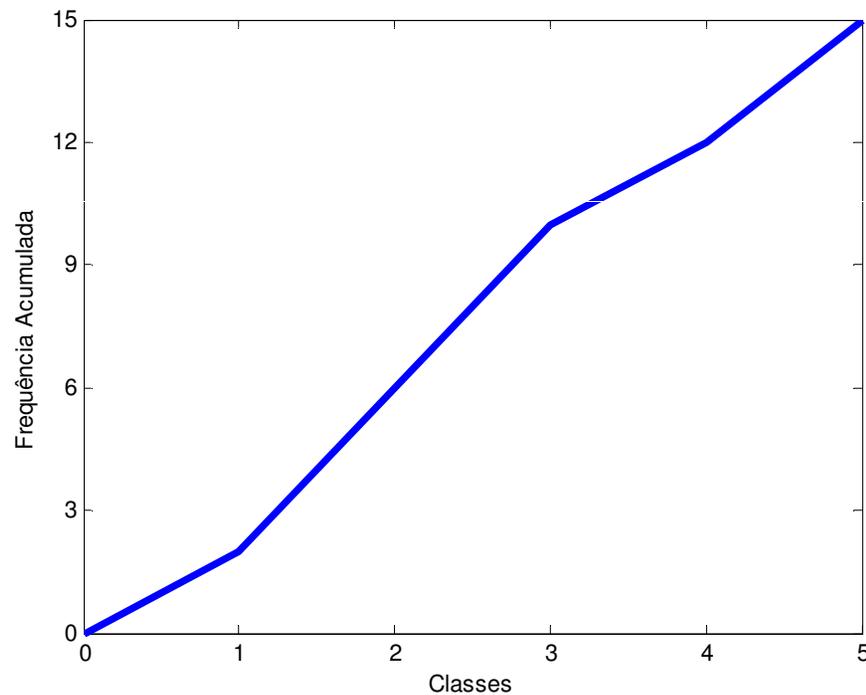


- Outra forma bastante útil e muito popular para a apresentação de informações é através dos gráficos
 - Transmite uma ideia visual da informação!
 - Existe uma grande variedade de gráficos, inclusive para a representação simbólica e cognitiva de informação.
 - Os mais comuns são:
 - Cartesiano
 - Circular
 - Caixa ou *Box Plot*
 - Ramos e folhas ou *stem-and-leaf*

Gráficos - Exemplos



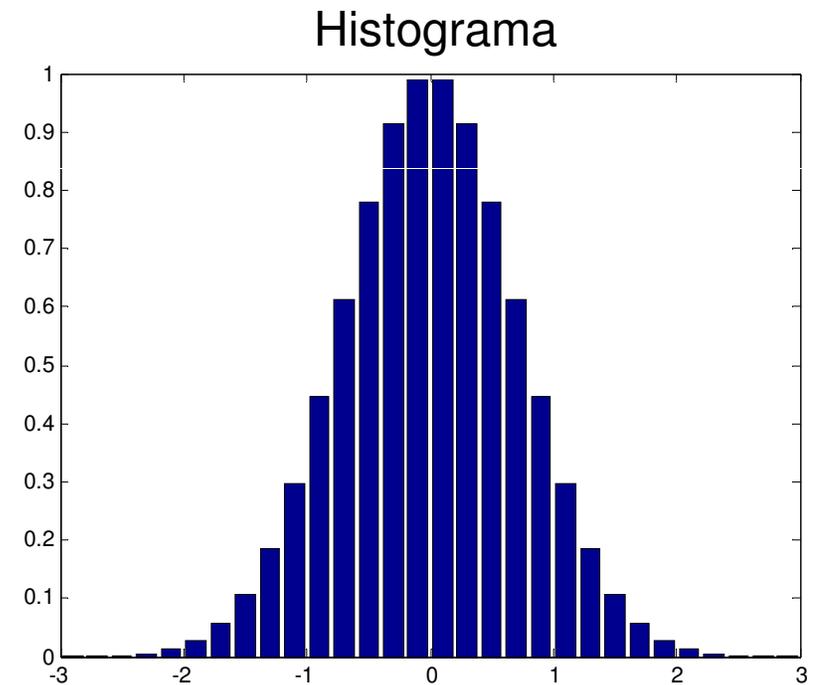
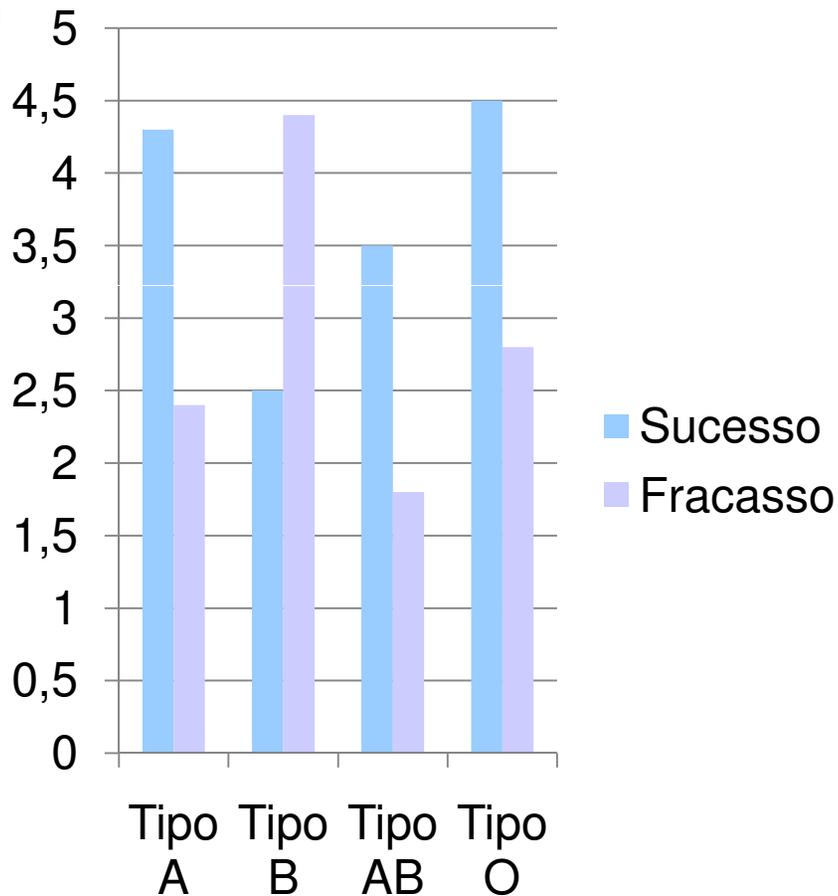
- Cartesiano



Gráficos - Exemplos



Gráficos de Barras

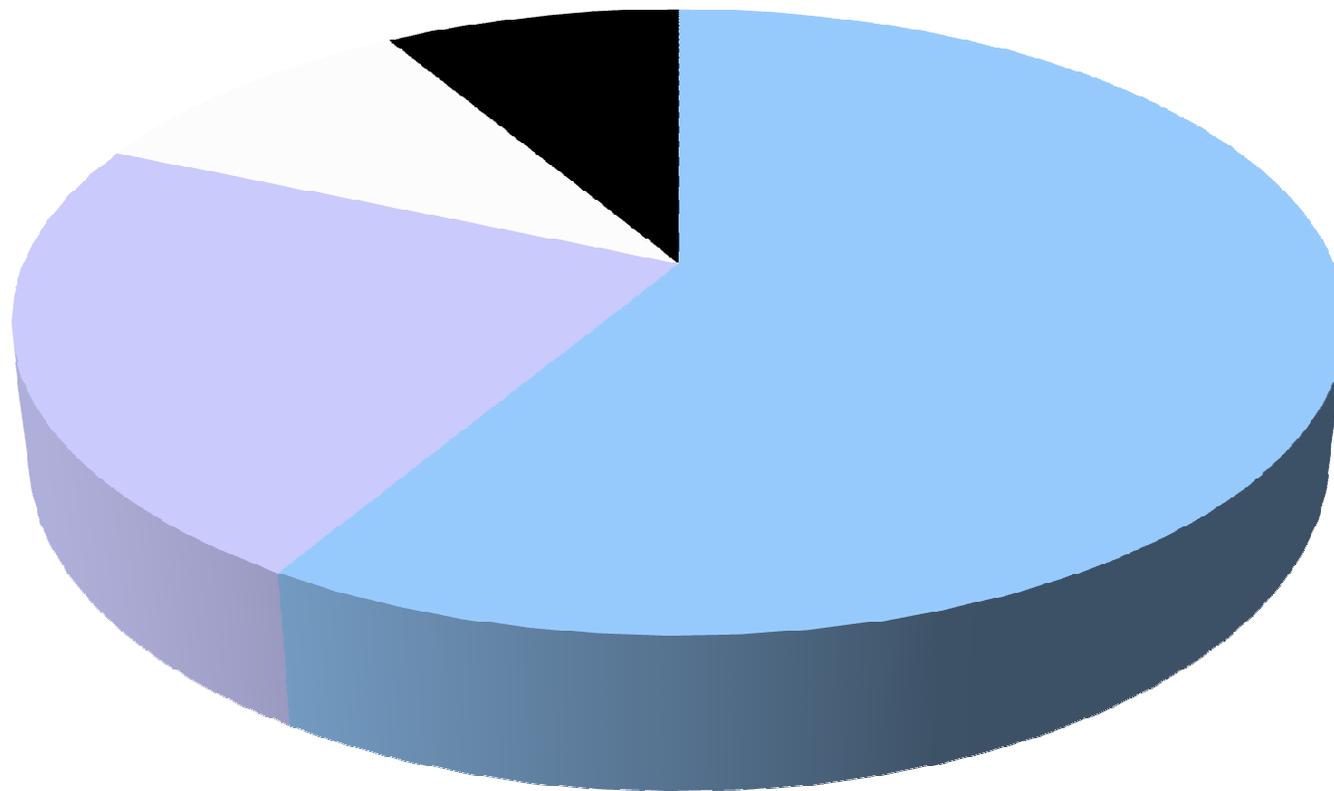


Gráficos - Exemplos



Gráfico Circular

Amostras



- 1º Tri
- 2º Tri
- 3º Tri
- 4º Tri

Gráficos - Exemplos

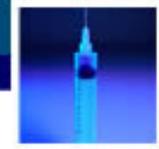
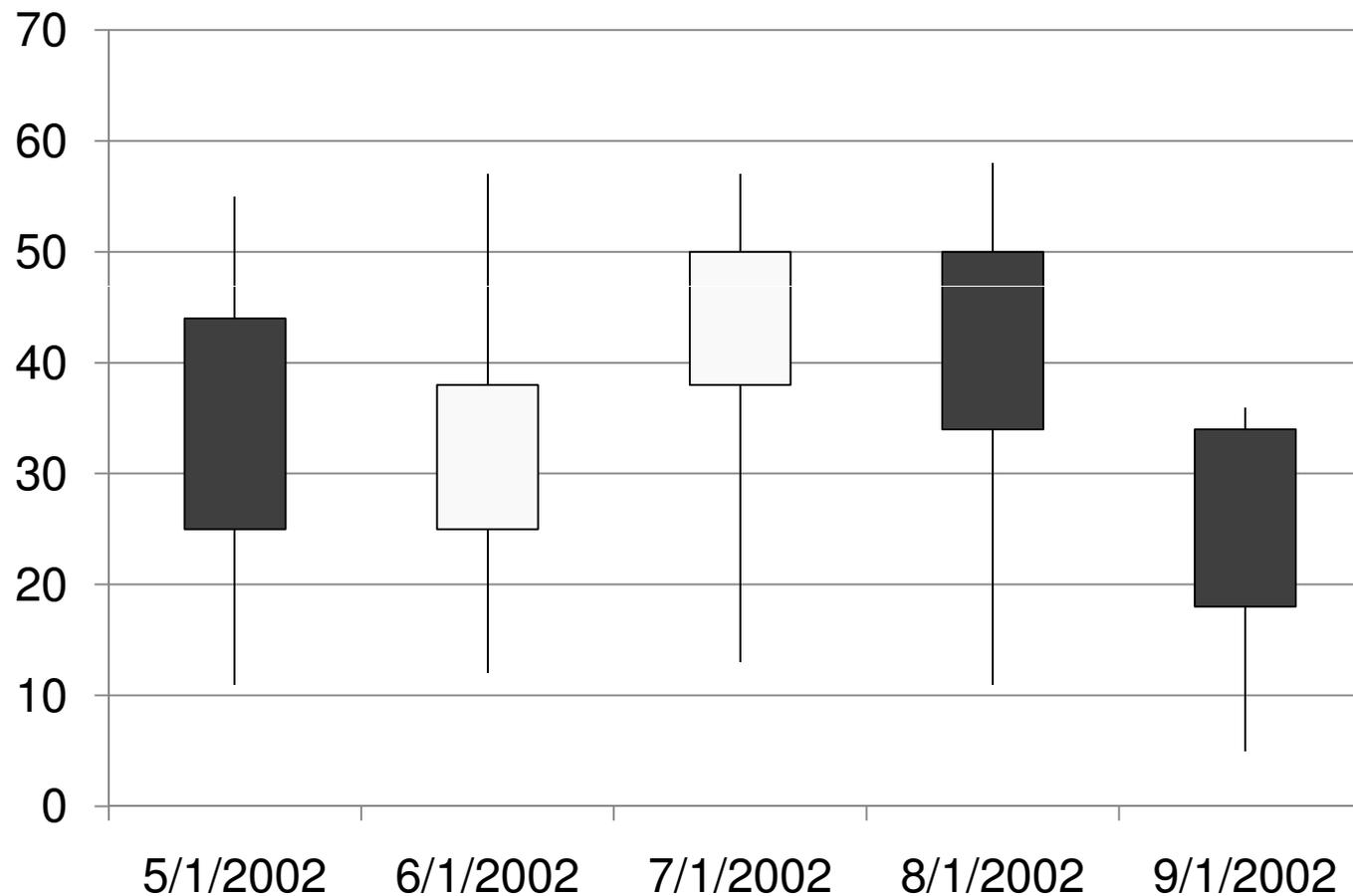


Gráfico de Caixas ou *Box Plot*



Max. Obs.
Int. Sup.
Int. Inf.
Fechar

Gráficos - Exemplos



Ramos e Folhas

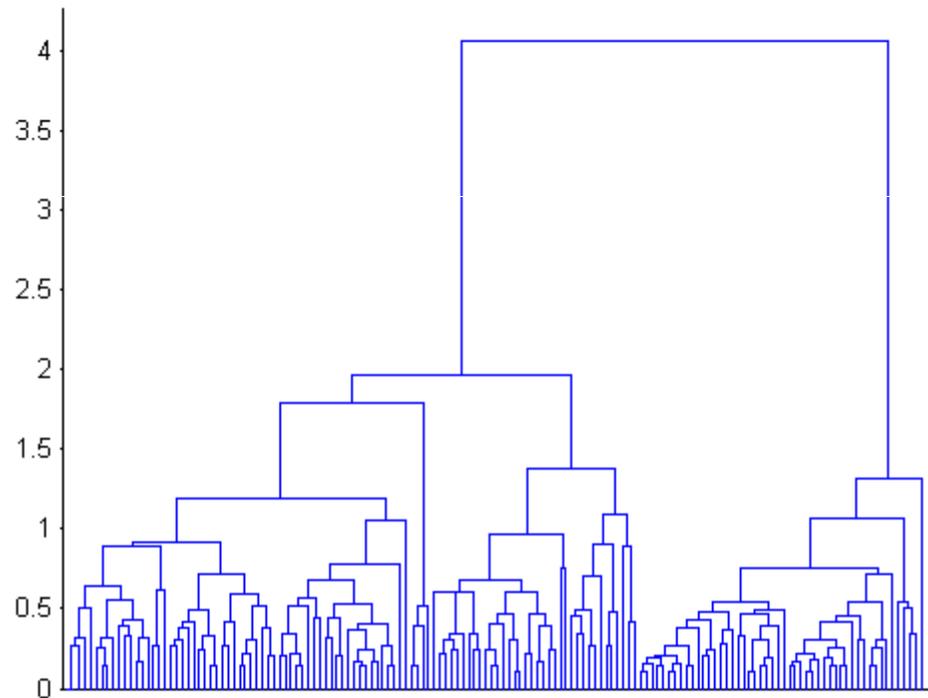
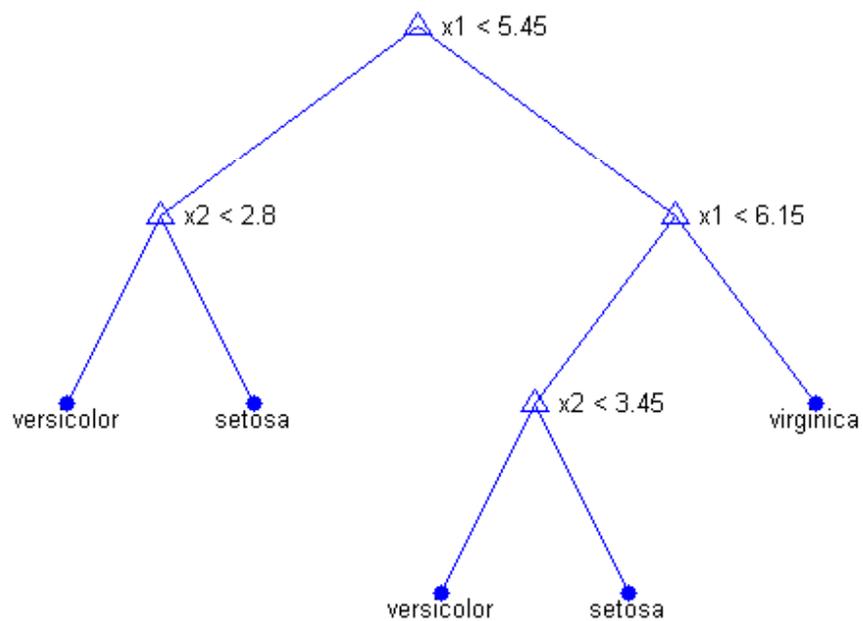
Construção de um gráfico ramos-folhas para a renda diária de empregados de uma empresa. À esquerda coloca-se o ramo e a direita a folha. Assim, rendas de R\$ 4,00 e R\$ 4,56 o ramo é 4 3 as folhas são 00 e 56.

| | | | | | | | | | |
|----|--|----|----|----|----|----|----|--|--|
| 4 | | 00 | 56 | | | | | | |
| 5 | | 25 | 73 | | | | | | |
| 6 | | 26 | 66 | 86 | | | | | |
| 7 | | 39 | 44 | 59 | | | | | |
| 8 | | 12 | 46 | 74 | 95 | 96 | 99 | | |
| 9 | | 13 | 35 | 77 | 80 | | | | |
| 10 | | 53 | 76 | | | | | | |
| 11 | | 06 | 59 | | | | | | |

Gráficos - Exemplos



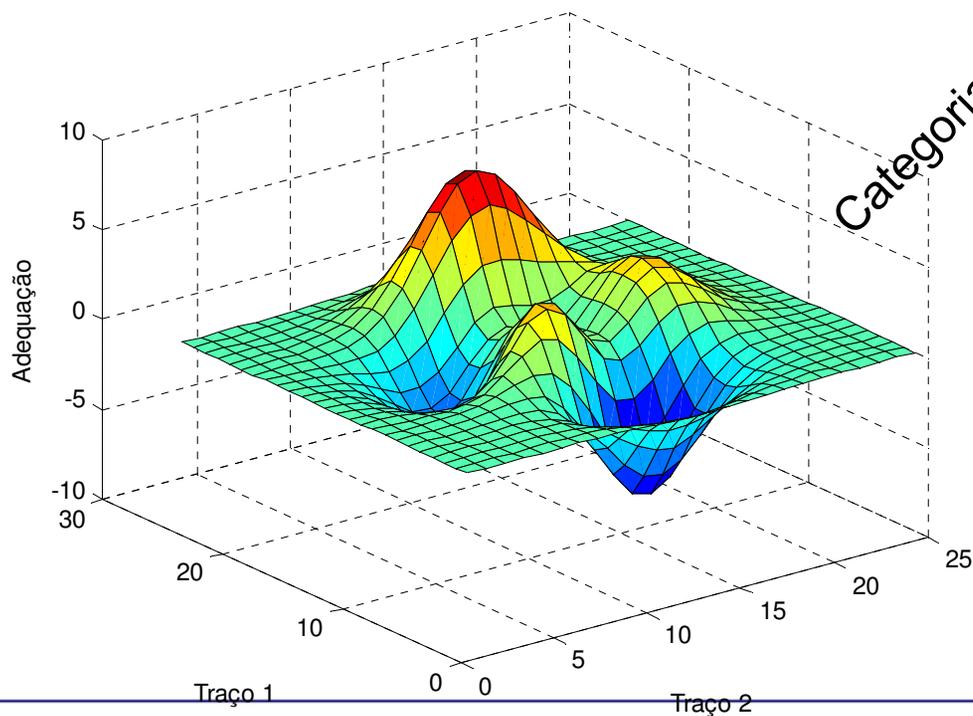
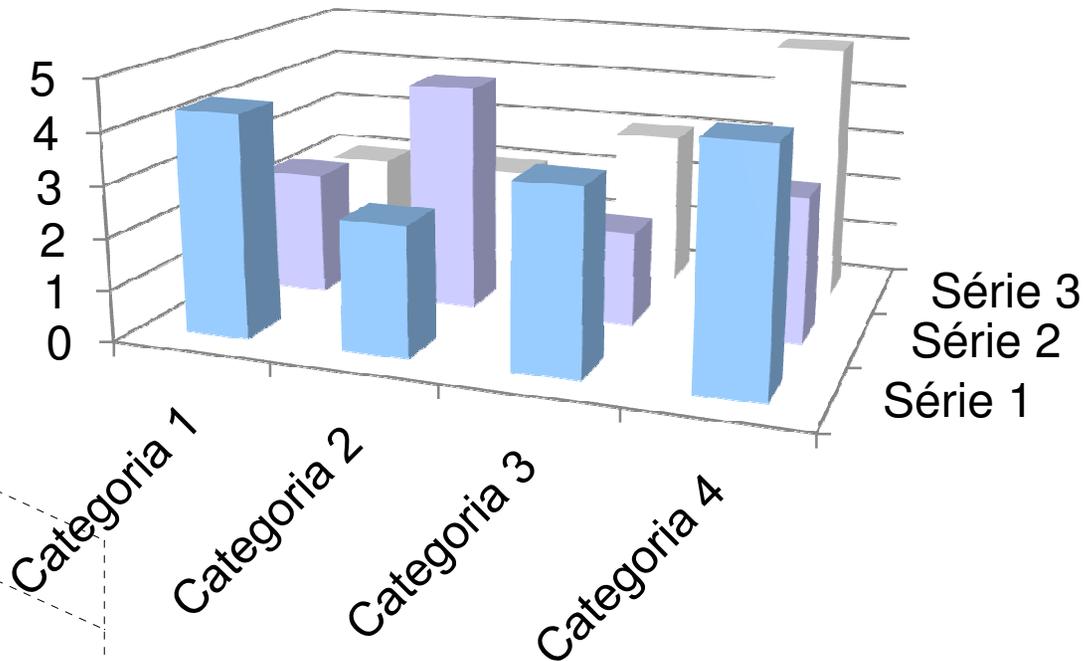
Árvores de Decisão e Agrupamentos



Gráficos - Exemplos



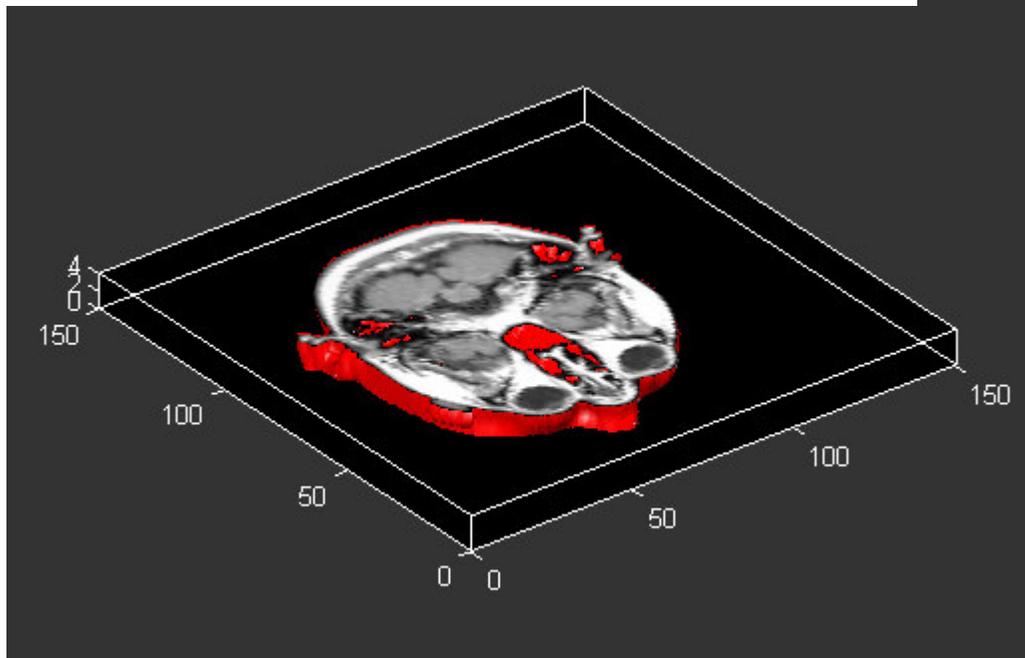
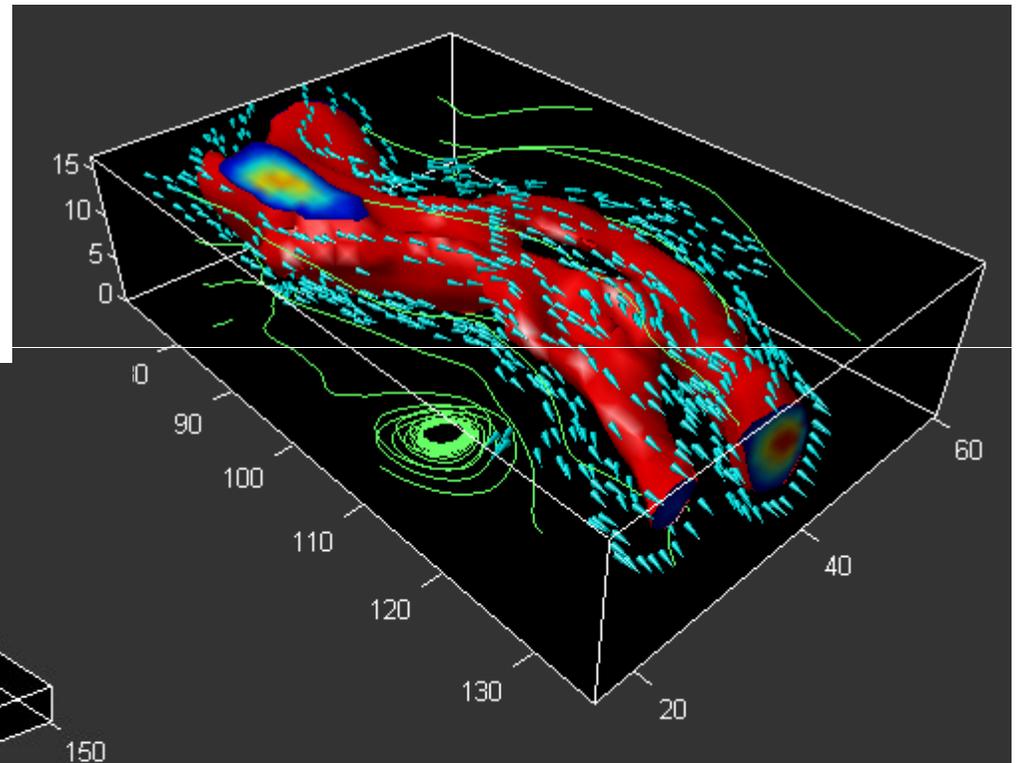
- Gráficos 3D



Gráficos - Exemplos



Reconstrução de imagens a partir de medidas de dados.



Referências da Aula



- Livro Texto: Bioestatística. Teoria e Computacional (Héctor G. Arango). Guanabara
 - Capítulos 1, 2 e 3
- Leituras Complementares:
 - Livro: Bioestatística. Princípios e Aplicações (Siada M. Callegari-Jacques). Artmed
 - Capítulo 1
 - Livro: Estatística Básica (Wilton de O. Bussab e Pedro A. Morettin). Saraiva
 - Capítulos: 2 e 10
 - Livro: Probabilidade. Aplicações à Estatística (Paul L. Meyer). LTC
 - Seções: 2.1, 13.1 e 13.2