



ESTATÍSTICA APLICADA

MARIA ALICE V. F. DE SOUZA



Pensar estatisticamente será um dia, para a eficiente prática da cidadania, tão necessário como a habilidade de ler e escrever.

Herbert George Wells



Estatística – origens

Censo – latim *censere* – significa taxar

Estatística – latim *status* – significa estado.

A Estatística nas mãos de governantes constituiu-se em ferramenta administrativa.



Essência da Estatística: a observação

Objetivo básico: inferência (deduzir, concluir)

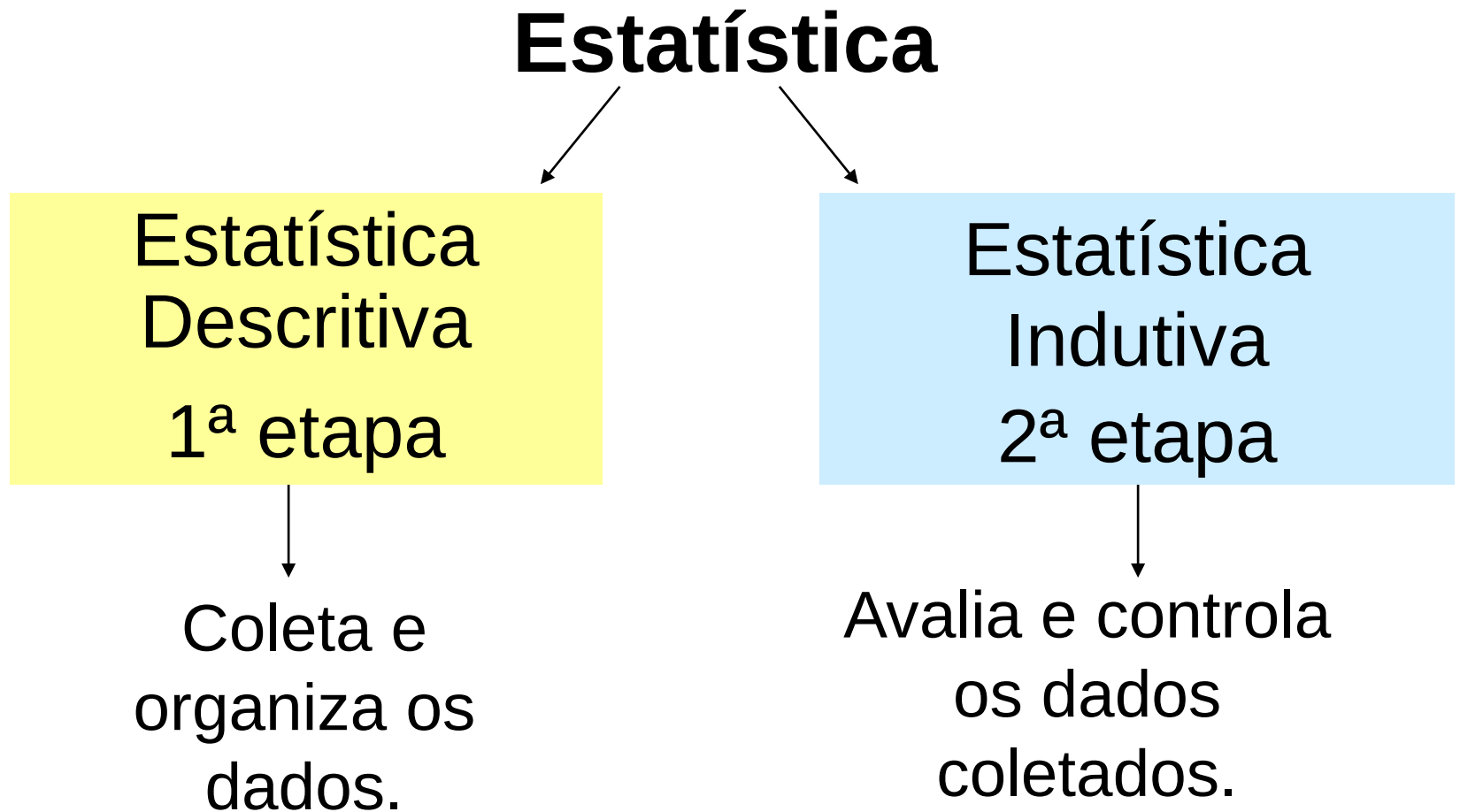
Importância:

- Avaliar o grau de dependência entre duas variáveis como o tempo médio de alguém digitando e sintomas de dores nos dedos;
- Avaliar o tempo médio de duração de uma aula e o nível de atenção;

- 
- O fator previdenciário da Previdência Social;
 - As previsões das pesquisas eleitorais;
 - Conhecer o tempo médio de duração de um sinal de trânsito vermelho e a sua paciência.
 - Controlar a qualidade na produção de componentes eletrônicos;
 - Usos estatísticos pelas seguradoras;

Estatística

É um ramo da Matemática que trata dos métodos de coleta, organização, resumo, apresentação e análise de dados.



Variável

Qualitativa

Nominal

**Não pode
ordenar:**

- Estado civil
- Religião
- Sexo
- região

Ordinal

**Pode
ordenar:**

- Nível de educação
- Classe social

Quantitativa

Discreta

Contáveis:

- Idade
- Capacidade de passageiros
- N° de filhos

Contínua

**Não
Contáveis:**

- Peso
- Altura
- Largura
- Medida de capacidade

População e Amostra

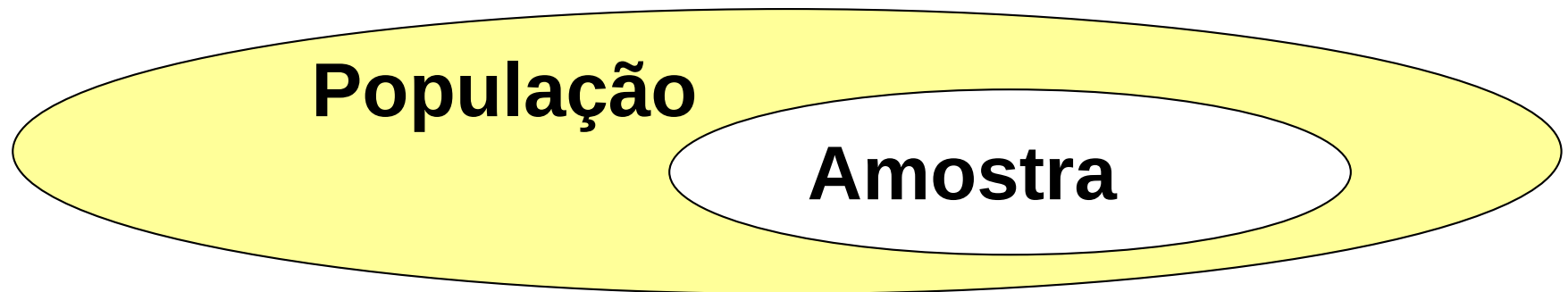
População (ou Universo): é qualquer conjunto de informações que tenham, entre si, uma característica comum.

Ex.: conj. de todas as estaturas – população de estaturas

conj. de todas as cores de olhos – população de cores de olhos

Amostra : redução representativa da população, ou seja, sem perda das características essenciais.

Escolha dos números – números aleatórios (tabelas, sorteios etc.)



TIPOS DE AMOSTRAGEM

É o conjunto de técnicas utilizadas para a seleção de uma amostra. Esse conjunto de técnicas pode ser subdividido em dois grupos básicos:

AMOSTRAGEM ALEATÓRIA

- amostragem aleatória simples ou ao acaso;
- amostragem sistemática;
- amostragem estratificada
- amostragem por conglomerados.

AMOSTRAGEM NÃO ALEATÓRIA

- amostragem intencional;
- amostragem voluntária.



Não permitem o controle de variabilidade amostral, o que inviabiliza o controle de qualidade da estimação.

AMOSTRAGEM ALEATÓRIA OU PROBABILÍSTICA

AMOSTRAGEM ALEATÓRIA SIMPLES OU AO ACASO

É aquela em que se atribui aos grupos de mesma quantidade de elementos, a mesma probabilidade de participar da amostra. Cada elemento da população tem a mesma probabilidade de participar da amostra.

Para se obter uma amostra aleatória simples, caso a população seja finita, podemos atribuir a cada elemento um número.

Ex: Fichas numeradas podem ser misturadas em uma urna. O sorteio das fichas identificam os elementos que deverão participar da amostra, garantindo a mesma chance para cada um deles.

AMOSTRAGEM SISTEMÁTICA

Quando se conhece uma listagem dos elementos da população pode-se obter uma amostra aleatória de n elementos dividindo-se o número de elementos da população pelo tamanho da amostra.

Ex: Auditoria em notas fiscais de uma empresa.

AMOSTRAGEM ESTRATIFICADA

Pode ocorrer que a população seja formada por subgrupos diferentes, mas cada um deles **homogêneo**. Neste caso, vamos selecionar aleatoriamente uma quantidade de cada grupo para formar a amostra, proporcional ao tamanho desse grupo.

Exemplo: Dividir a população em 3 estratos: renda alta, renda média, renda baixa, e proceder à amostragem estratificada.

Ex.: analisar indivíduos de todas as idades presentes na população.

AMOSTRAGEM POR CONGLOMERADOS

Em algumas situações, podemos identificar um grupo de elementos que tenha aproximadamente a mesma composição de população. Neste caso, pode ser interessante realizar amostragem usando somente os elementos desse grupo.

Ex.: Algumas empresas, quando pretendem avaliar a aceitação de um produto no eixo Rio-São Paulo, lançam o produto em Curitiba, cuja população se comporta como uma miniatura desse mercado.

Ex.: quarteirões em um bairro.

AMOSTRAGEM NÃO ALEATÓRIA OU NÃO PROBABILÍSTICA

AMOSTRAGEM INTENCIONAL

Ocorre quando o pesquisador seleciona intencionalmente os componentes da amostra;

Ex.: Em um estudo sobre automóveis, o pesquisador procura apenas oficinas; Em uma pesquisa sobre preferência por determinado cosmético, o pesquisador entrevista os frequentadores de um grande salão de beleza.

AMOSTRAGEM NÃO ALEATÓRIA OU NÃO PROBABILÍSTICA

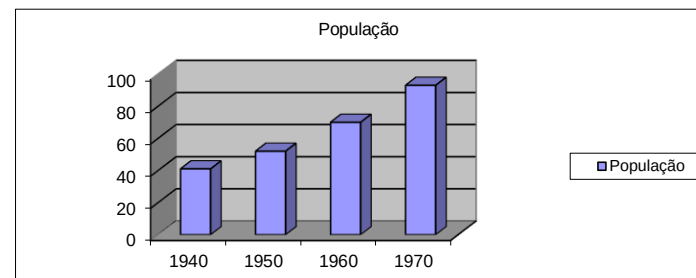
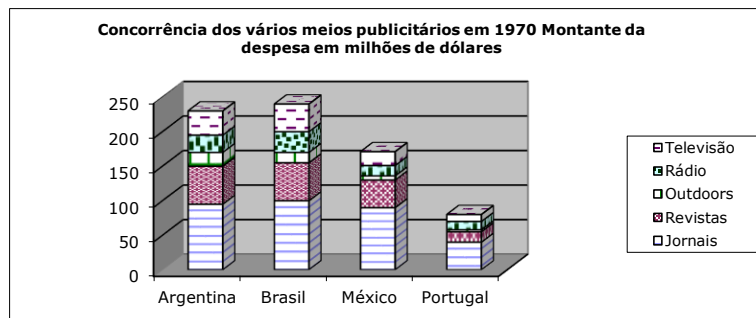
AMOSTRAGEM VOLUNTÁRIA OU ACIDENTAL

Ocorre quando o componente da população se oferece voluntariamente para participar da amostra independentemente do julgamento do pesquisador.

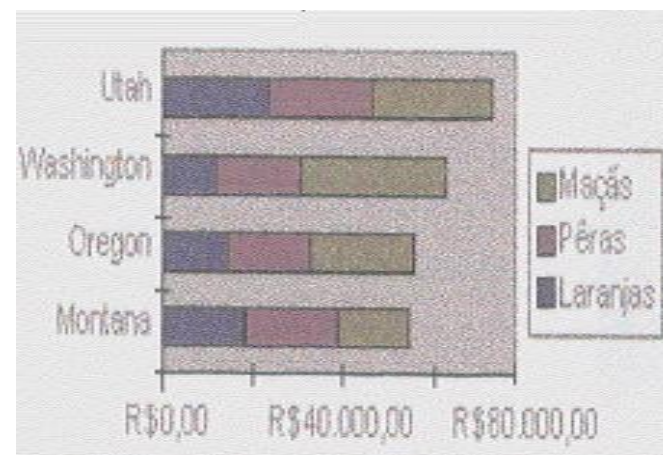
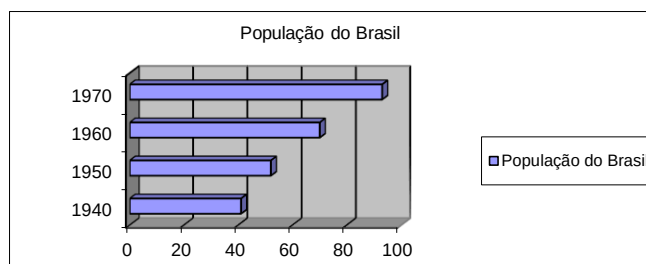
Ex.: Coleta por amostragem voluntária, de sangue para identificação de doenças sexuais curáveis; Pesquisas de opinião em praças públicas, ruas movimentadas de grandes cidades, etc.

Gráficos

Colunas

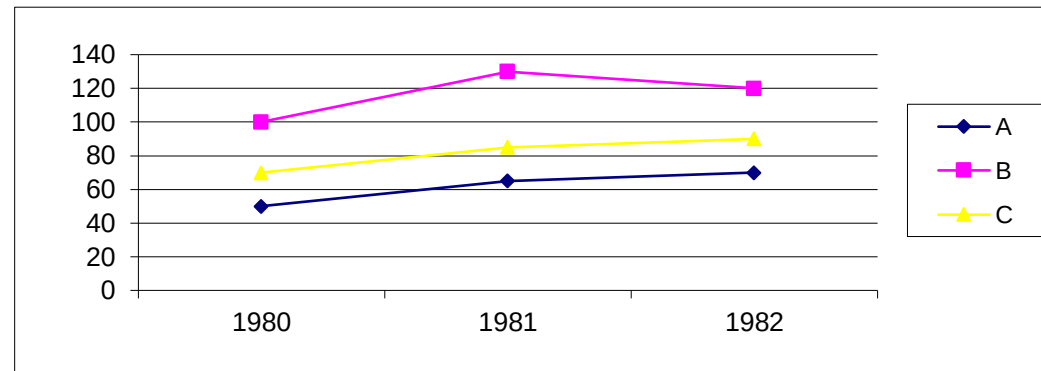
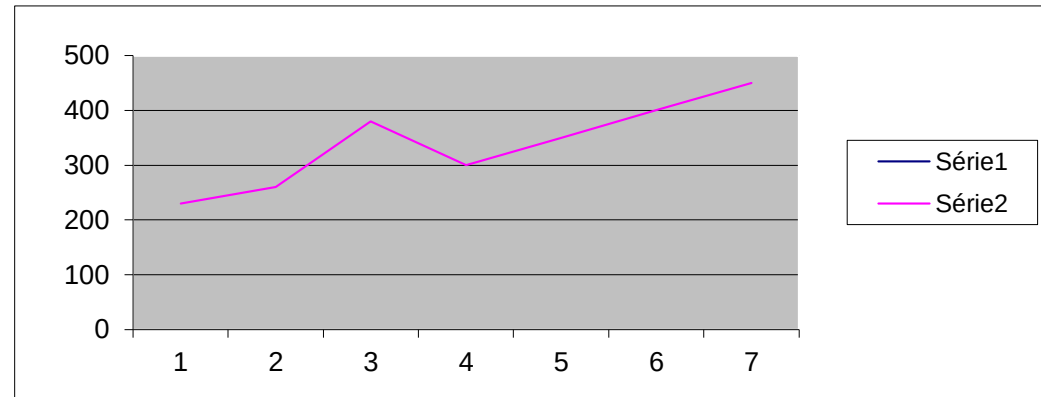


Barras



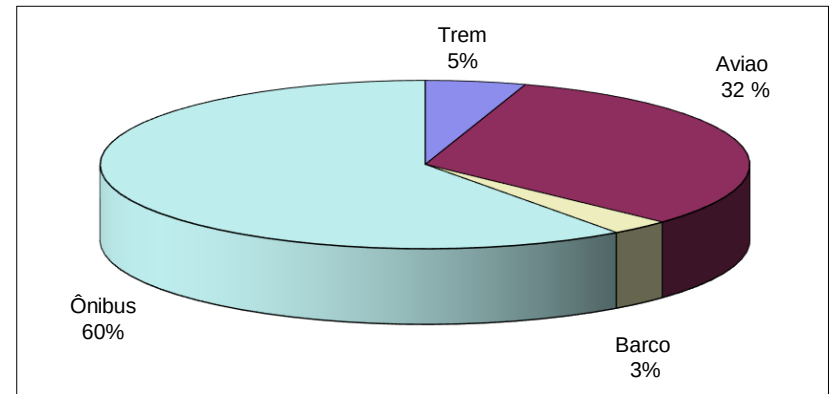
Gráficos

Linhas

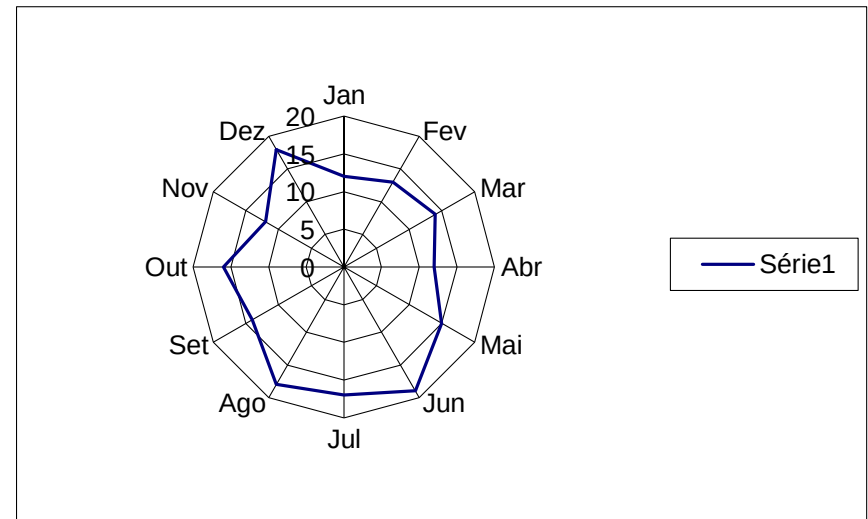


Gráficos

Setores



Polar



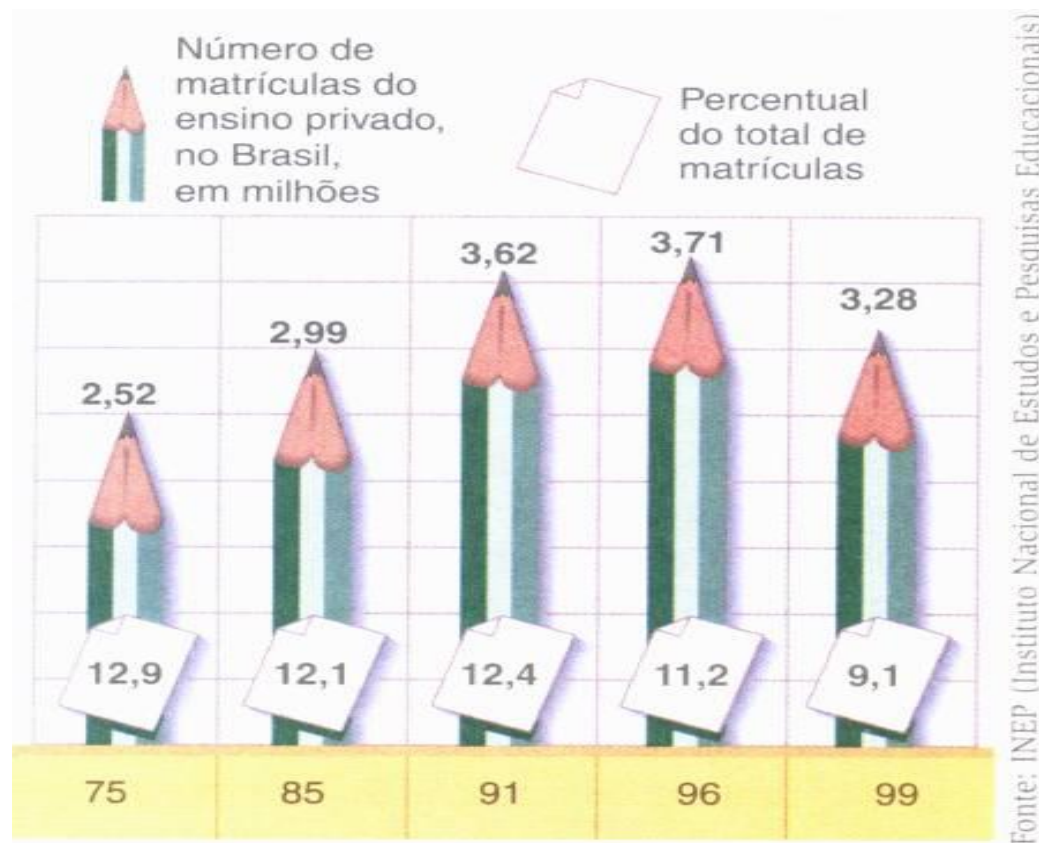
Gráficos

Pictórico



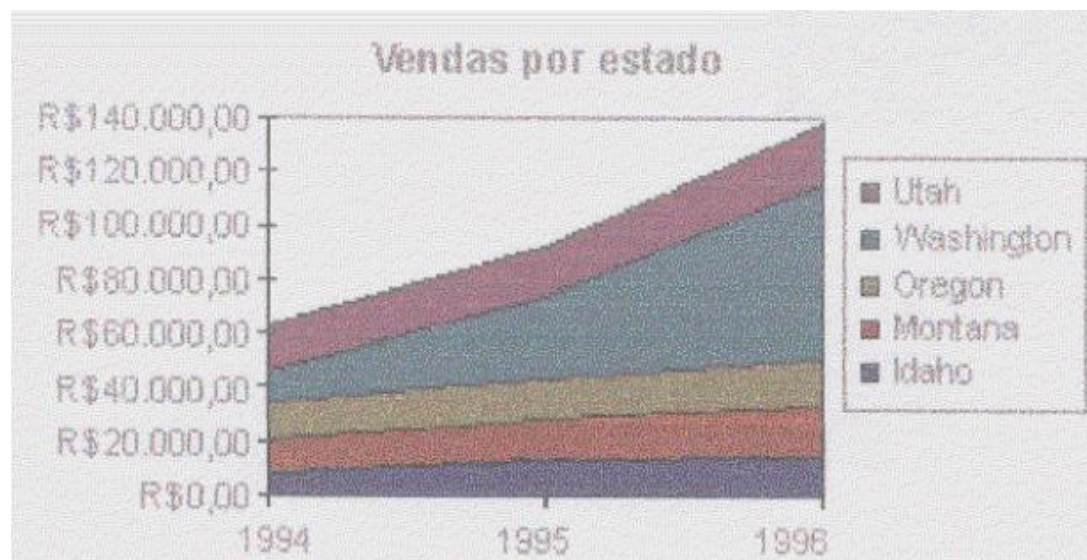
Gráficos

Pictórico



Outros Gráficos

Área



Gráficos

Superfície

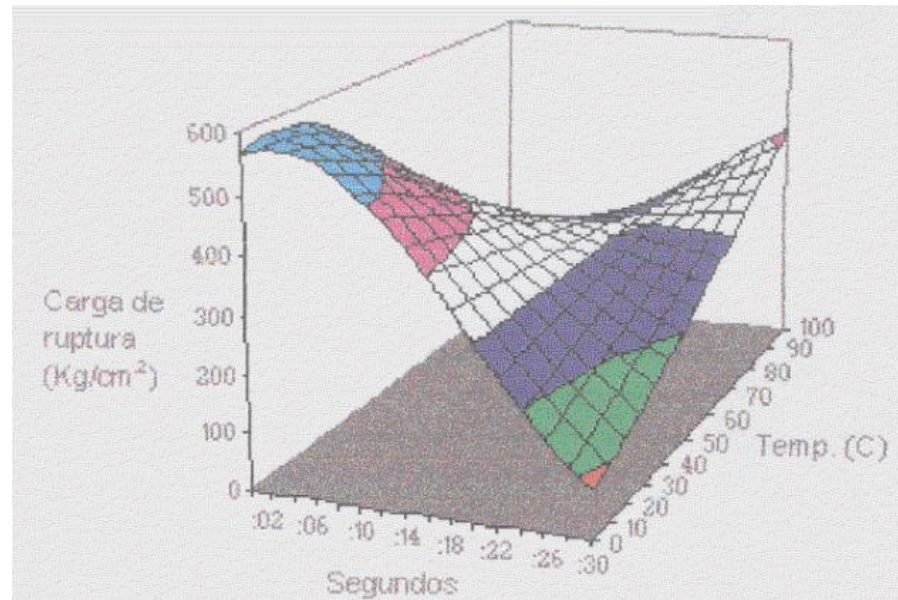
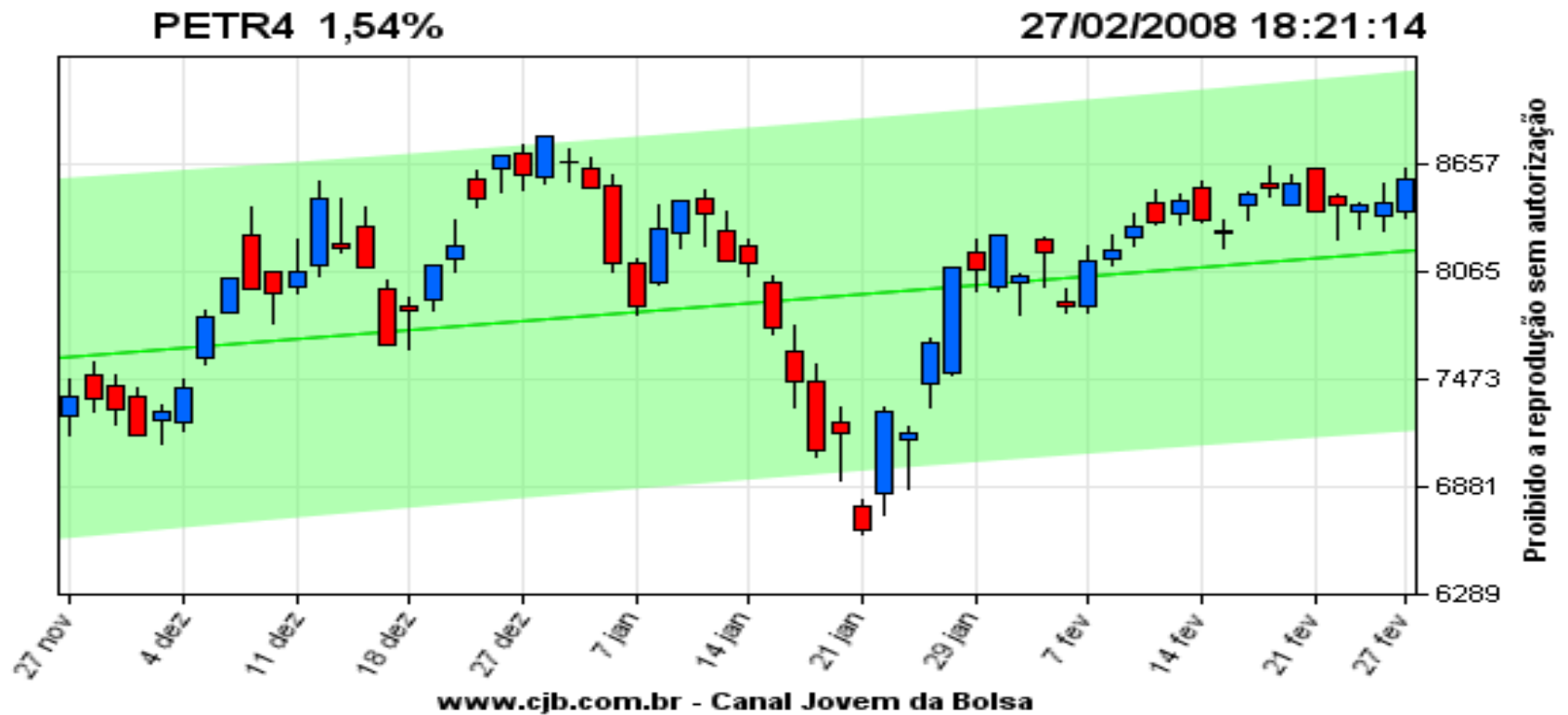


gráfico de candlesticks/ box plot



Fonte: site da bolsa de ações japonesa.



ESTATÍSTICA DESCRITIVA

Distribuições de Frequência

PROF^a MARIA ALICE V. F. DE SOUZA

Distribuição de Frequência

Título	
x_i	f_i
21	3
22	2
23	2
24	1
25	4
Total	12

Fonte:

Título	
Classe	f_i
45 - 55	15
55 - 65	30
65 - 75	35
75 - 85	15
85 - 95	5
Total	100

Fonte:

Amplitude Total (R): diferença entre o maior e o menor valor observado nos dados.

Título	
x_i	f_i
21	3
22	2
23	2
24	1
25	4
Total	12

Fonte:

$$R = 25 - 21 = 4$$

Título	
Classe	f_i
45 - 55	15
55 - 65	30
65 - 75	35
75 - 85	15
85 - 95	5
Total	100

Fonte:

$$R = 95 - 45 = 50$$

Distribuição de Frequência

Exemplo: Notas de Matemática de 80 estudantes de Engenharia da Universidade X no ano de 1990.

68	84	75	82	68	90	62	88	76	93
73	79	88	73	60	93	71	59	85	75
61	65	75	87	74	62	95	78	63	72
66	78	82	75	94	77	69	74	68	60
96	78	89	61	75	95	60	79	83	71
79	62	67	97	78	85	76	65	71	75
65	80	73	57	88	78	62	76	53	74
86	67	73	81	72	63	76	75	85	77

Distribuição de Frequência

Dados Brutos: dados desorganizados

Rol: são os dados brutos organizados em ordem crescente ou decrescente.

53	57	59	60	60	60	61	61	62	62
62	62	63	63	65	65	65	66	67	67
68	68	68	69	71	71	71	72	72	73
73	73	73	74	74	74	75	75	75	75
75	75	75	76	76	76	76	77	77	78
78	78	78	78	79	79	79	80	81	82
82	83	84	85	85	85	86	87	88	88
88	89	90	93	93	94	95	95	96	97

**Notas de
Matemática -
Universidade X-
1990 - Engenharia**

Notas	f_i
50 - 55	1
55 - 60	2
60 - 65	11
65 - 70	10
70 - 75	12
75 - 80	21
80 - 85	6
85 - 90	9
90 - 95	4
95 - 100	4
Σ	80

Fonte: hipotética

Distribuição de Frequência

f_i – frequência absoluta – valores absolutos;

f_r – frequência relativa – é a freq. absoluta da classe em termos percentuais;

f_a – frequência acumulada – freq. total de todos os valores inferiores ao limite superior de um dado intervalo de classe;

f_{ar} – frequência acumulada relativa – é a freq. acumulada em termos relativos ou percentuais.

Distribuição de Frequência

**Notas de Matemática - Universidade X -
1990 - Engenharia**

Notas	f_i	f_r (%)	f_a	f_{ar} (%)
50 - 55	1	1,25	1	1,25
55 - 60	2	2,50	3	3,75
60 - 65	11	13,75	14	17,50
65 - 70	10	12,50	24	30,00
70 - 75	12	15,00	36	45,00
75 - 80	21	26,25	57	71,25
80 - 85	6	7,50	63	78,75
85 - 90	9	11,25	72	90,00
90 - 95	4	5,00	76	95,00
95 - 100	4	5,00	80	100,00
Σ	80	100,00		

Fonte: hipotética

Distribuição de Frequência

Ponto Médio de uma Classe: É o ponto intermediário do intervalo da classe e é obtido pela média aritmética entre o limite inferior e o superior de cada classe.

**Notas de Matemática -
Universidade X - 1990 -
Engenharia**

Notas	f_i	ponto médio
50 - 55	1	52,5
55 - 60	2	57,5
60 - 65	11	62,5
65 - 70	10	67,5
70 - 75	12	72,5
75 - 80	21	77,5
80 - 85	6	82,5
85 - 90	9	87,5
90 - 95	4	92,5
95 - 100	4	97,5
Σ	80	

Fonte: hipotética



ESTATÍSTICA DESCRITIVA

Medidas de Posição

PROF^a MARIA ALICE V. F. DE SOUZA

Medidas de Posição

Estas medidas nos orientam quanto à posição da distribuição no eixo x , possibilitando comparações de séries de dados entre si pelo confronto desses números. São também conhecidas como medidas de tendência central, pois representam os fenômenos pelos seus valores médios, em torno dos quais tendem a concentrar-se os dados.

Média Aritmética ($\bar{X}$)

Para dados **não agrupados**:

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

$n = n^{\circ}$ de elementos do conjunto

Ex.: $\bar{X}$ de 3, 7, 8, 10, 11

$$\bar{X} = \frac{\sum x}{n} = \frac{3 + 7 + 8 + 10 + 11}{5} = 7,8$$

Para dados **agrupados**:

$$\bar{X} = \frac{\sum_{i=1}^n x_i \cdot f_i}{n}$$

$n = n^{\circ}$ de elementos do conjunto

$f_i =$ frequência simples

Exemplo 1:

x_i	f_i
1	1
2	3
3	5
4	1
Total	10

$$\sum f_i = n$$

$$\bar{X} = \frac{1.1 + 2.3 + 3.5 + 4.1}{10} = \frac{26}{10} = 2,6$$

Exemplo 2:

Renda familiar (em milhares \$)	f_i	ponto médio
2 - 4	5	3
4 - 6	10	5
6 - 8	14	7
8 - 10	8	9
10 - 12	3	11
Total	40	

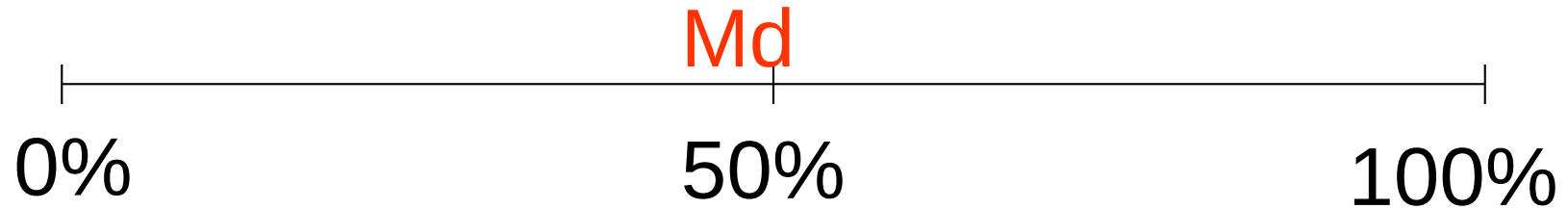
$$\bar{X} = \frac{5.3 + 10.5 + 14.7 + 8.9 + 3.11}{40} = \frac{268}{40} = 6,7$$

Propriedades da Média Aritmética

A média aritmética é influenciada pela soma, subtração, multiplicação e divisão.

Mediana (Md)

Colocados os dados em ordem crescente, mediana é o valor que divide a amostra, ou população, em duas partes iguais.



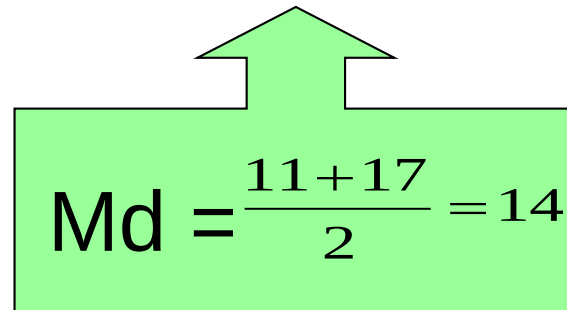
Para dados **isolados**:

Se n for ímpar: a Md será o elemento central;

Se n for par: a Md será a média aritmética entre os elementos centrais.

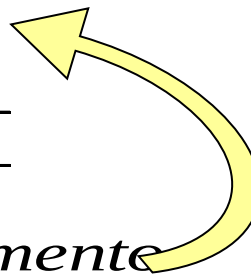
2 4 7 9 11 17 23 24 24

2 4 7 9 11 17 23 24 24 25


$$Md = \frac{11 + 17}{2} = 14$$

Para dados **isolados agrupados**:

x_i	f_i	f_a
1	1	1
2	3	4
3	5	9
4	2	11
Total	11	

$$Md = \frac{n + 1}{2} = 6^{\circ} \text{elemento}$$


O 6° elemento se encontra na 3ª classe; $Md = 3$.

x_i	f_i	f_a
82	5	5
85	10	15
87	15	30
89	8	38
90	4	42
Total	42	

$$Md = \frac{21^{\circ} + 22^{\circ}}{2} = \frac{87 + 87}{2} = 87$$

Para dados **agrupados em classes**:

soma das freq. anteriores
à classe da Md.

$$Md = l_{i_{Md}} + \frac{(\frac{n}{2} - \sum f) \cdot h}{f_{Md}}$$

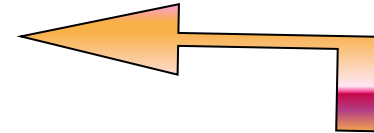
amplitude classe Md

limite inferior da classe Md

freqüência da classe Md

Exemplo:

Classes	f_i	f_a
35 - 45	5	5
45 - 55	12	17
55 - 65	18	35
65 - 75	14	49
75 - 85	6	55
85 - 95	3	58
Total	58	



Classe Md

(58/2=29)

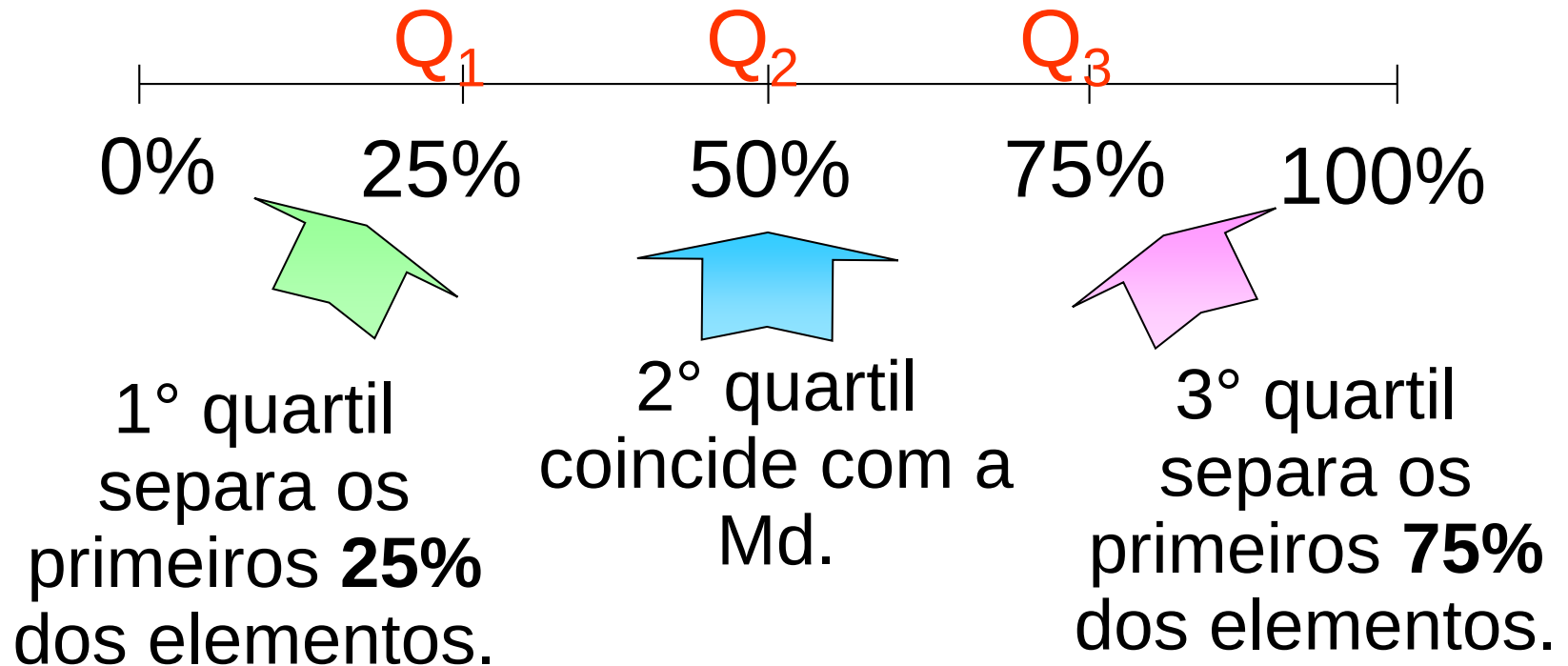
$$Md = l_{i_{Md}} + \frac{(\frac{n}{2} - \sum f).h}{f_{Md}}$$

$$Md = 55 + \frac{(\frac{58}{2} - 17).10}{18} = 55 + \frac{120}{18} \approx 61,67$$

Separatrizes { Quartis
Decis
Percentis

Quartis

Os quartis dividem um conjunto de dados em 4 partes iguais.



Cálculo do 1º Quartil Q_1 :

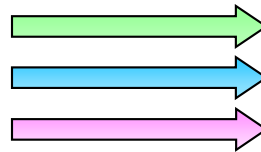
$$Q_1 = l_{Q_1} + \frac{(\frac{n}{4} - \sum f).h}{f_{Q_1}}$$

Cálculo do 3º Quartil Q_3 :

$$Q_3 = l_{Q_3} + \frac{(\frac{3n}{4} - \sum f) \cdot h}{f_{Q_3}}$$

Exemplo:

Classes	f_i	f_a
7 - 17	6	6
17 - 27	15	21
27 - 37	20	41
37 - 47	10	51
47 - 57	5	56
Total	56	



classe do Q_1
classe da Md
classe do Q_3

classe do Q_1

classe da Md

classe do Q_3

$$\Rightarrow \frac{n}{4} = 14^{\circ} \text{elemento}$$

$$\Rightarrow \frac{n}{2} = 28^{\circ} \text{elemento}$$

$$\Rightarrow \frac{3n}{4} = 42^{\circ} \text{elemento}$$

$$Q_1 = 17 + \frac{(\frac{56}{4} - 6).10}{15} = 22,33$$

$$Md = 27 + \frac{(\frac{56}{2} - 21).10}{20} = 30,5$$

$$Q_3 = 37 + \frac{(\frac{3.56}{4} - 41).10}{10} = 38$$

Moda (Mo)

É o valor mais freqüente da distribuição. Para distribuições simples (sem agrupamentos em classes), verifica-se a moda pela simples observação do elemento que apresenta maior freqüência.

Ex.: 2 - 4 - 4 - 5 - 7 - 8 - 10

$$M_o = 4$$

Ex.:

x_i	f_i
243	7
245	17
248	23
251	20
307	8
Total	75

$$M_o = 248$$

Para dados agrupados:

- identificar a classe modal
- fórmula de Czuber

limite inferior da
classe M_o

amplitude classe
 M_o

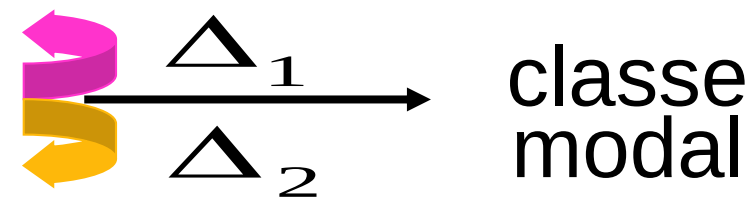
$$M_o = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot h$$

diferença entre a
frequência da
classe modal e a
imediatamente
anterior.

diferença entre a frequência
da classe modal e a
imediatamente **posterior**.

Ex.:

classes	f_i
0 - 1	3
1 - 2	10
2 - 3	17
3 - 4	8
4 - 5	5
Total	43



classe modal

$$M_o = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot h$$

$$M_o = 2 + \frac{(17 - 10)}{(17 - 10) + (17 - 8)} \cdot 1$$

$$M_o = 2 + \frac{7}{7 + 9} \cdot 1 = 2 + \frac{7}{16} = 2,44$$

Se a classe modal cair na 1ª ou na última classe, fazer a diferença com zero.



ESTATÍSTICA DESCRITIVA

Outros Gráficos Estatísticos

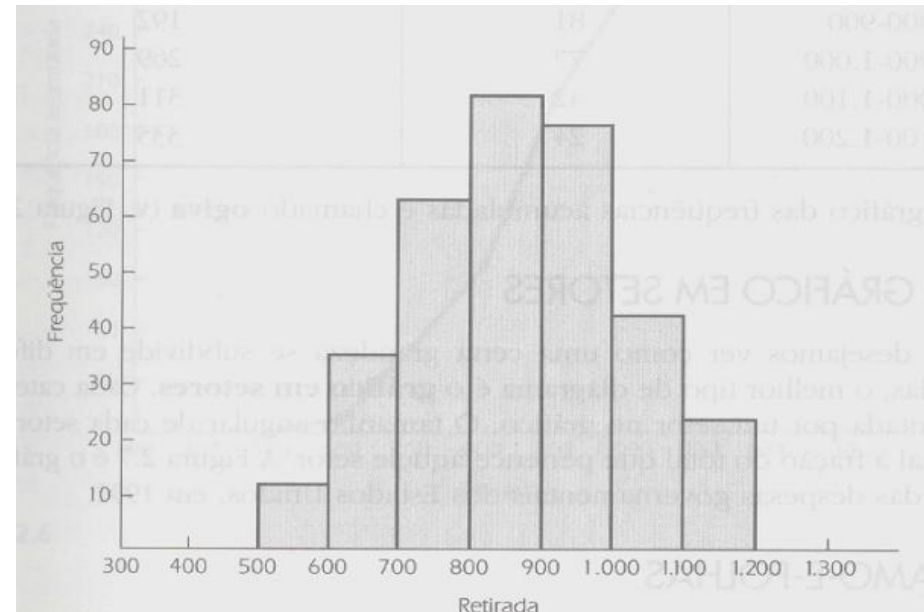
PROF^a MARIA ALICE V. F. DE SOUZA

Outros Gráficos: Histograma

Retiradas diárias de um banco

Retiradas (R\$)	Frequência	pt. médio
500 - 600	12	550
600 - 700	36	650
700 - 800	63	750
800 - 900	81	850
900 - 1000	77	950
1000 - 1100	42	1050
1100 - 1200	24	1150
Σ	335	

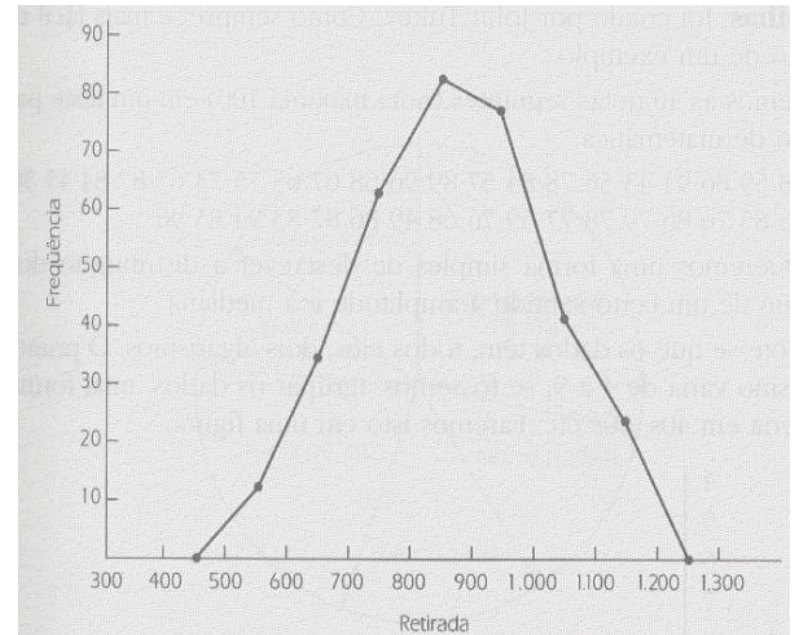
Fonte: fictícia



Outros Gráficos: Polígono de frequência

Retiradas diárias de um banco		
Retiradas (R\$)	Frequência	pt. médio
500 - 600	12	550
600 - 700	36	650
700 - 800	63	750
800 - 900	81	850
900 - 1000	77	950
1000 - 1100	42	1050
1100 - 1200	24	1150
Σ	335	

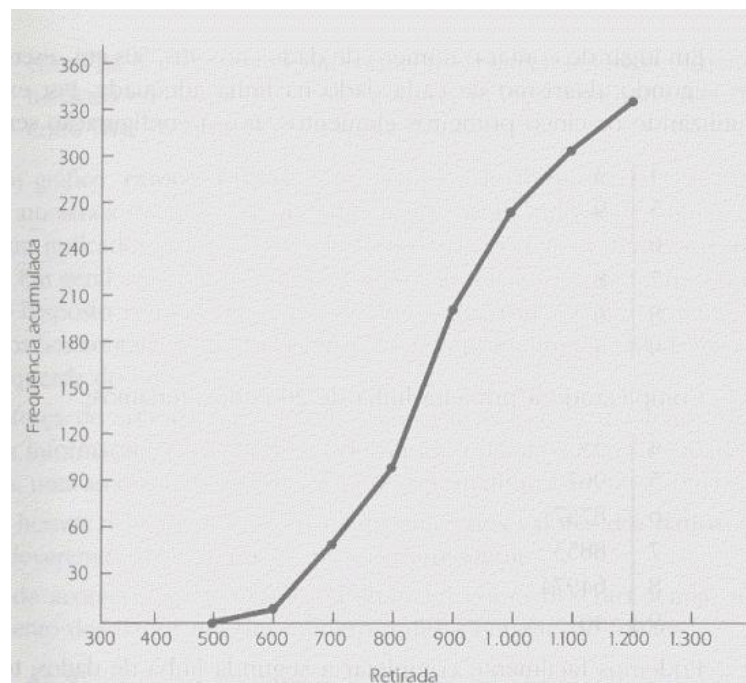
Fonte: fictícia



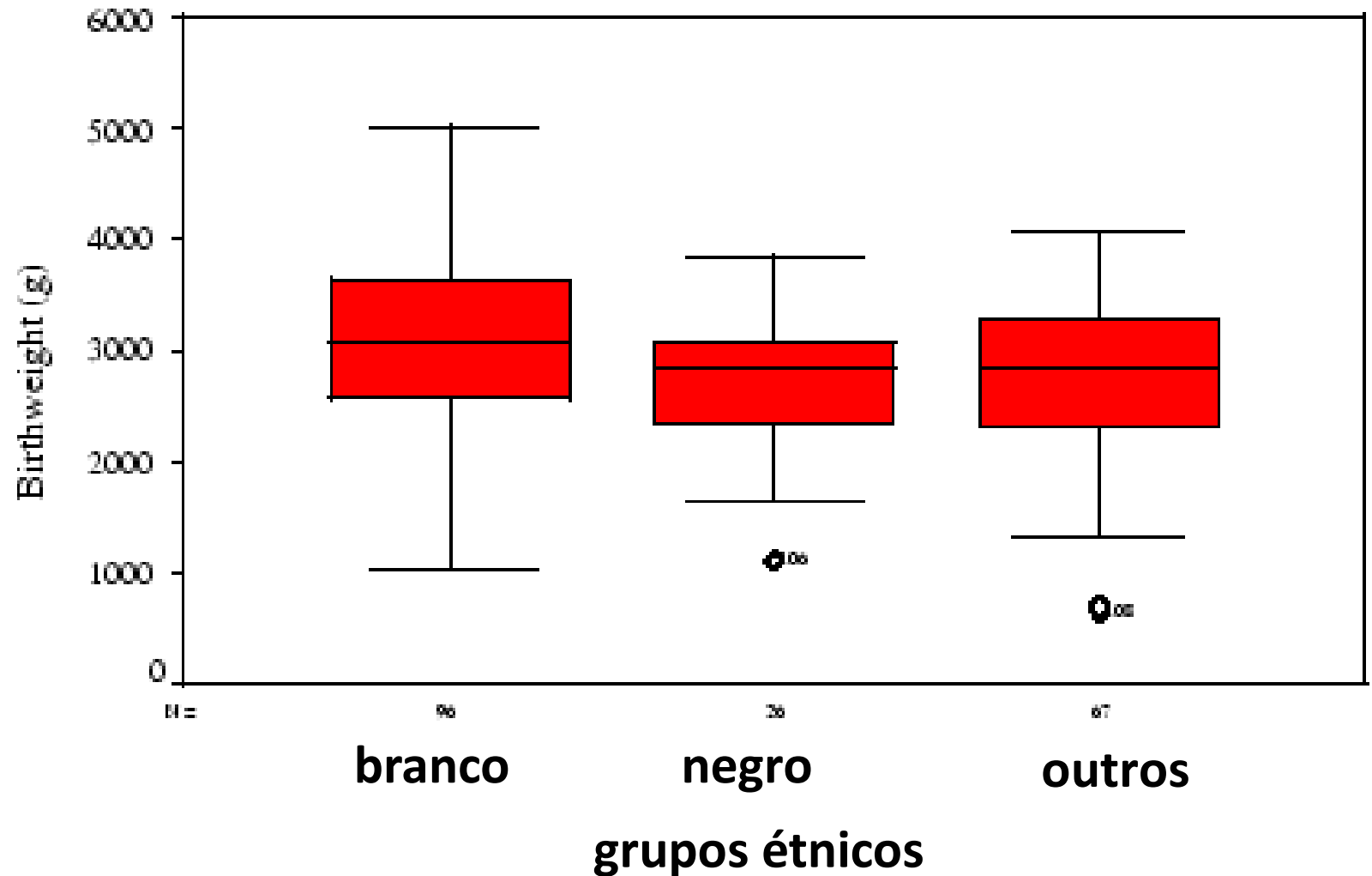
Outros Gráficos: Ogiva de Galton

Retiradas diárias de um banco		
Retiradas (R\$)	Frequência	pt. médio
500 - 600	12	550
600 - 700	36	650
700 - 800	63	750
800 - 900	81	850
900 - 1000	77	950
1000 - 1100	42	1050
1100 - 1200	24	1150
Σ	335	

Fonte: fictícia



Outros Gráficos: Box Plot

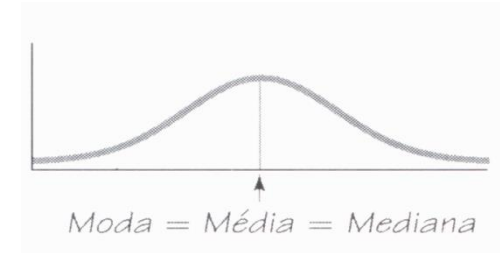
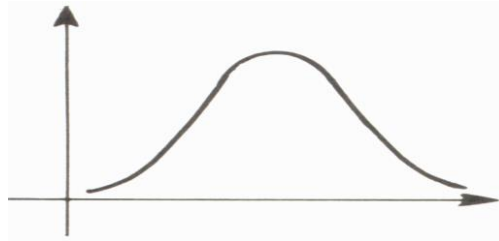


Outros Gráficos: Ramo e Folhas

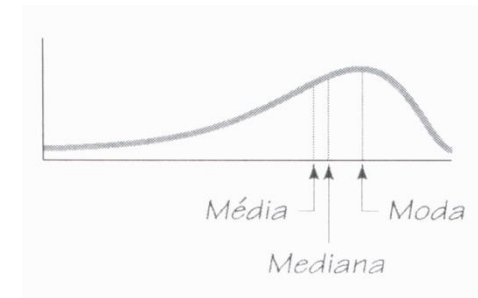
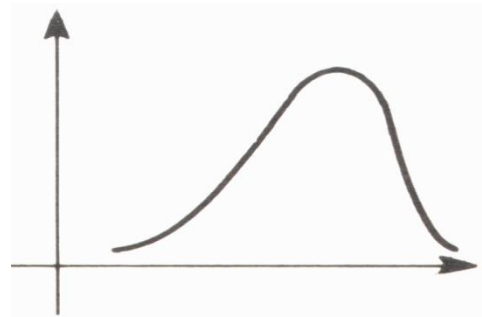
0	9
1	0 1 3 3 5 7 8 8 8 9
2	0 0 1 2 2 2 3 3 3 5 7 7 8 9
3	2 2 3 6 6 8 9
4	0 1
5	
6	2

Tipos de Curvas

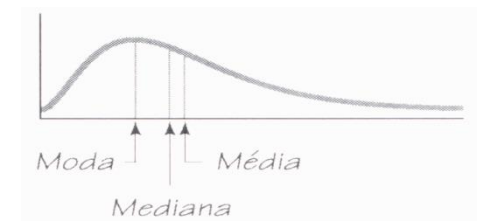
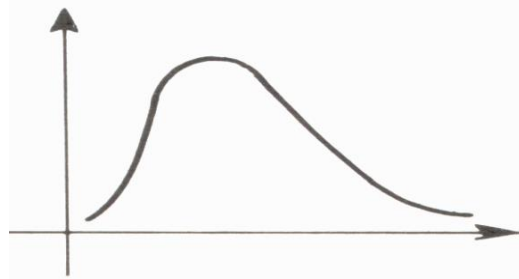
CURVA SIMÉTRICA



CURVA ASSIMÉTRICA NEGATIVA



CURVA ASSIMÉTRICA POSITIVA





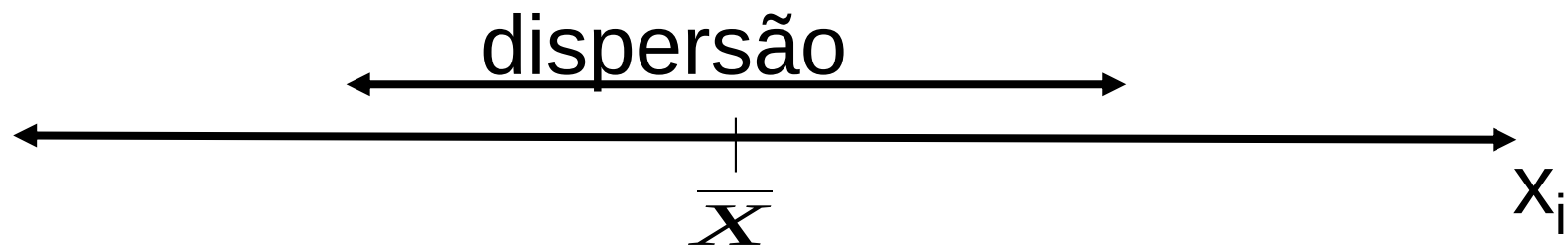
ESTATÍSTICA DESCRITIVA

Medidas de Dispersão

PROF^a MARIA ALICE V. F. DE SOUZA

Medidas de Dispersão

São medidas estatísticas utilizadas para avaliar o grau de variabilidade, ou dispersão, dos valores em torno da média. O termo dispersão indica o grau de afastamento de um conjunto de números em relação à sua média.



Exemplos:

a) 20, 20, 20 $\bar{X}_a = 20$;

b) 15, 10, 20, 25, 30 $\bar{X}_b = 20$

Amplitude Total (Range - R)

Diferença entre o maior e o menor dos valores da série.

$$R = x_{\text{máx.}} - x_{\text{mín.}}$$

Ex.: para a série

10, 12, 20, 22, 25, 33, 38

$$R = 38 - 10 = 28$$

A amplitude total é uma medida de dispersão limitada pois não leva em consideração a dispersão dos valores internos.

Variância (σ^2)

Mede as variações ocorridas. É calculada a partir das diferenças entre cada elemento e a média do conjunto.

$$\sigma^2 = \frac{\sum [(x_i - \bar{x})^2 \cdot f_i]}{\sum f_i}$$

Desvio Padrão (σ)

É também uma medida capaz de medir variação.

$$\sigma = \sqrt{\sigma^2}$$

$$\sigma = +\sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

$$\sigma = +\sqrt{\frac{\sum (x - \bar{x})^2 \cdot f_i}{\sum f_i}}$$

Propriedades do Desvio-padrão e da Variância

O desvio-padrão e a variância **não** são influenciados pela soma e subtração, mas **sim** pela multiplicação e divisão.

Caso os dados sejam de uma amostra, as fórmulas passam a ser:

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

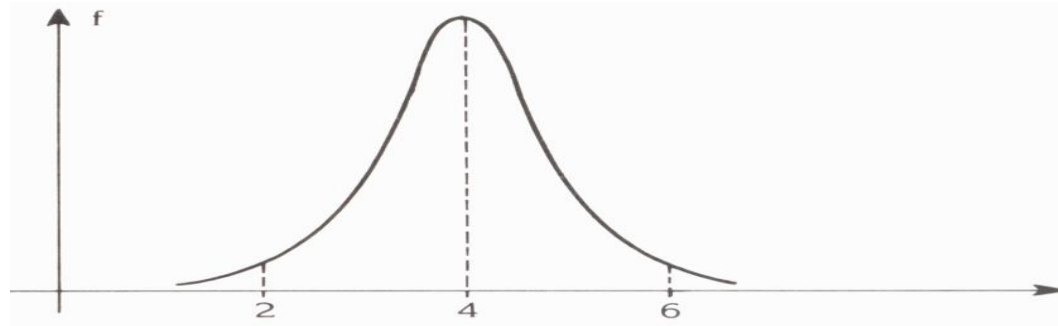
tamanho da amostra menos 1.

ou

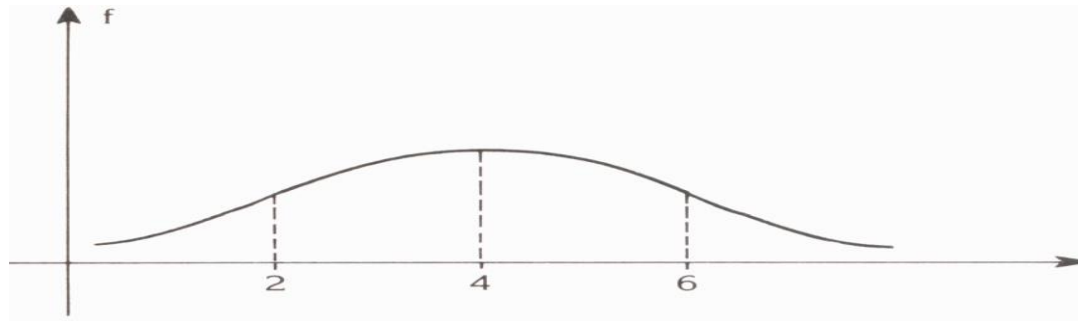
$$s^2 = \frac{1}{(\sum f_i) - 1} \left[\sum x_i^2 f_i - \frac{(\sum x_i f_i)^2}{\sum f_i} \right]$$

tamanho da amostra menos 1.

Grau de Curtose

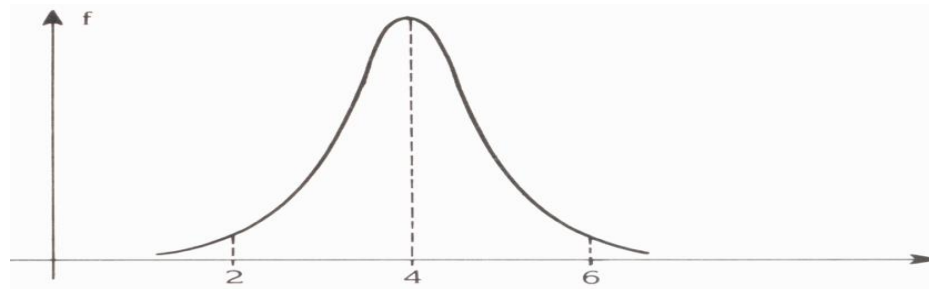


MENOR DISPERSÃO; MAIOR
CONCENTRAÇÃO DE DADOS

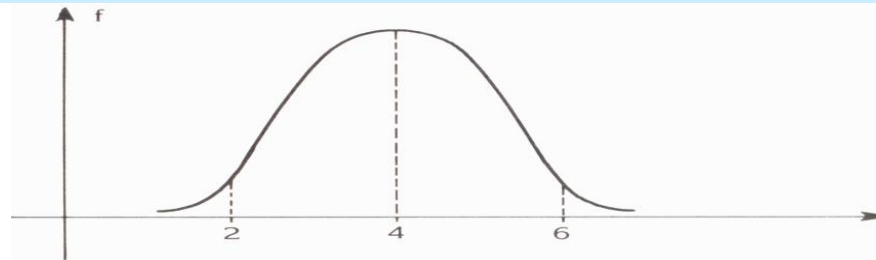


MAIOR DISPERSÃO; MENOR
CONCENTRAÇÃO DE DADOS

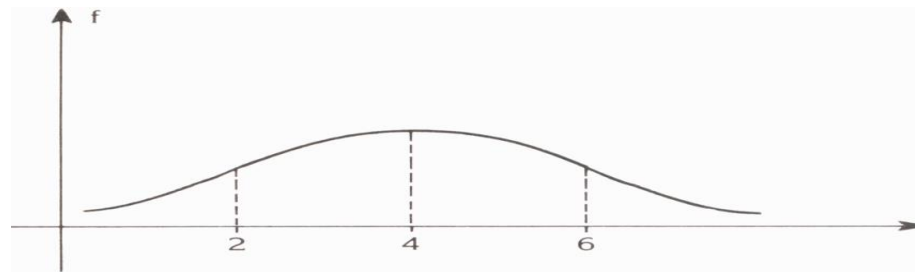
Distribuição delgada; LEPTOCÚRTICA



Distribuição nem chata, nem delgada; MESOCÚRTICA



Distribuição achatada; PLATICÚRTICA



Cálculo do Grau de Curtose

$$K = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$$

$Q_3 \rightarrow 3^\circ \text{quartil}$

$Q_1 \rightarrow 1^\circ \text{quartil}$

$P_{90} \rightarrow 90^\circ \text{percentil}$

$P_{10} \rightarrow 10^\circ \text{percentil}$

Se $k < 0,263$ LEPTOCÚRTICA

Se $K = 0,263$ MESOCÚRTICA

Se $K > 0,263$ PLATICÚRTICA

Cálculo do Coeficiente de Assimetria

$$AS = \frac{\bar{X} - Mo}{\sigma}$$

$$AS = \frac{Q_1 + Q_3 - 2Md}{Q_3 - Q_1}$$

$AS = 0 \rightarrow$ Distribuição é **Simétrica**

$AS > 0 \rightarrow$ Distribuição é **Assimétrica Positiva**

$AS < 0 \rightarrow$ Distribuição é **Assimétrica Negativa**

Coeficiente de Variação (CV)

Trata-se de uma medida relativa de dispersão útil para a comparação em **termos relativos** do grau de concentração em torno da média de séries distintas.

$$CV = \frac{\sigma}{\bar{X}} . 100$$

Em geral, a distribuição possui:

baixa dispersão: $CV \leq 15\%$

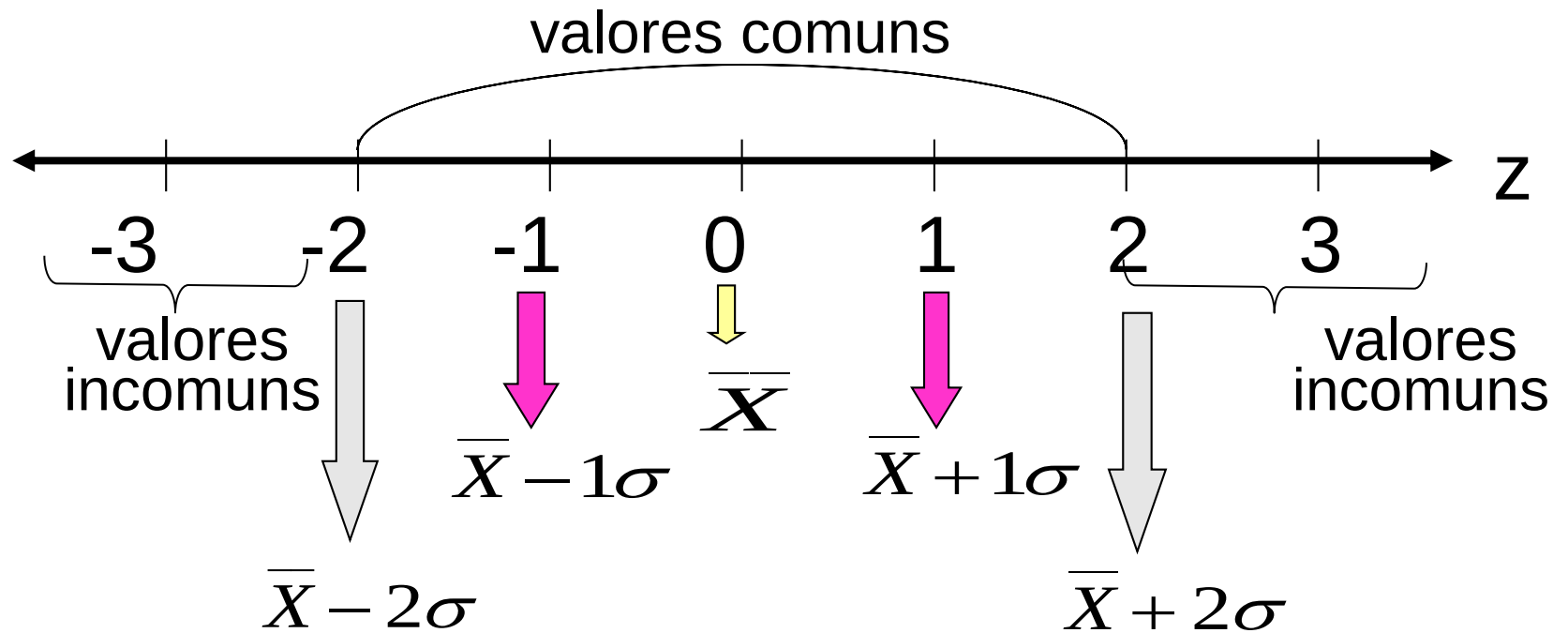
média dispersão: $15\% < CV < 30\%$

alta dispersão: $CV \geq 30\%$

Escore z

É o n° de desvios-padrão pelo qual um valor x dista da média (para mais ou para menos).

$$Z = \frac{x - \bar{x}}{\sigma}$$



É uma escala que permite que você meça quantos desvios-padrão o valor em questão está acima/abaixo da $\bar{X}$.

Exemplo 1:

As alturas da população de homens adultos têm $\bar{X}=69,0$ in, desvio-padrão $\sigma=2,8$ in (in=pol). O jogador Michael Jordan ganhou reputação de gigante, mas com 78 in, ele pode ser considerado excepcionalmente alto, comparado com a população geral de homens adultos americanos?

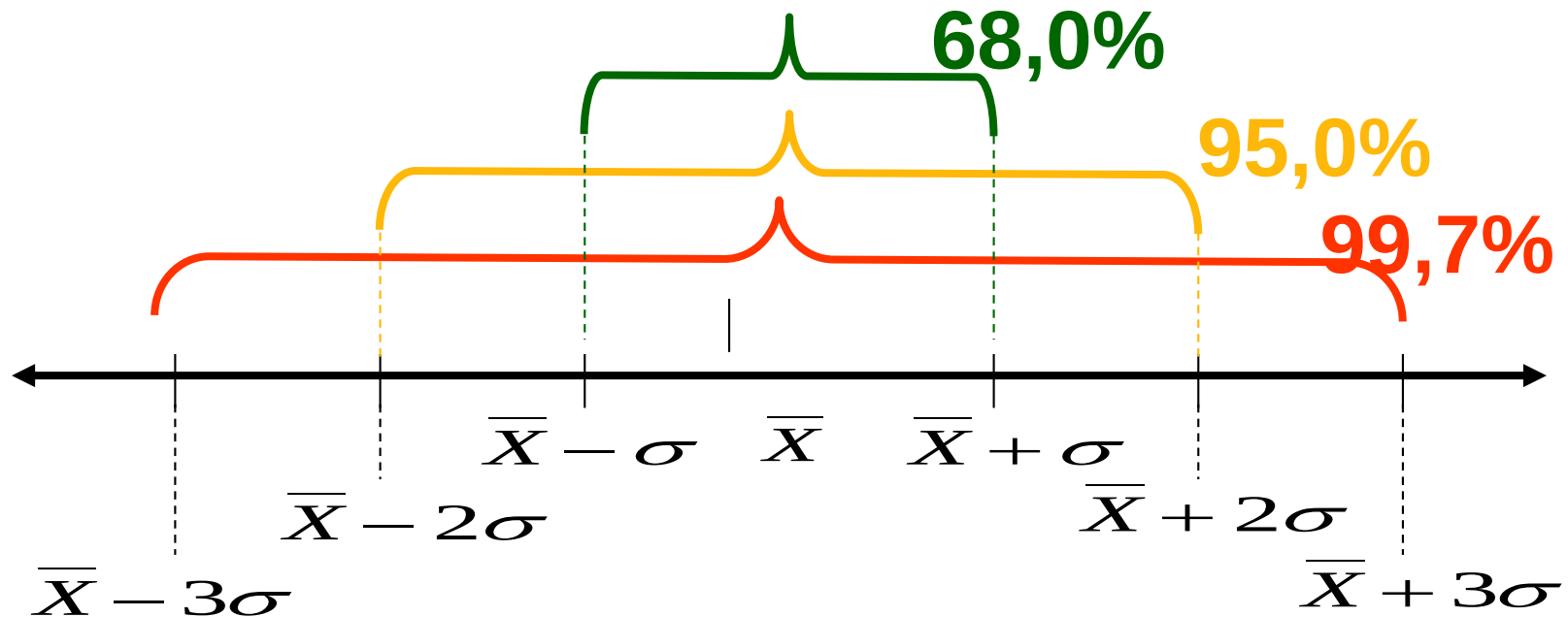
$$z = \frac{x - \bar{x}}{\sigma} = \frac{78 - 69}{2,8} = 3,21$$

A altura de Michael Jordan está 3,21 desvios-padrão acima da média. Sim, ele é alto para os padrões americanos.

Vamos investigar como é essa curva.

$$CV = \frac{\sigma}{\bar{X}} \cdot 100 = \frac{2,8}{69} \cong 0,0406 \cong 4,06\%$$

$CV \leq 15\%$ — baixa dispersão



A altura de Michael Jordan está contido nos 0,15% da população.



PROBABILIDADE

PROF^a MARIA ALICE V. F. DE SOUZA

A probabilidade surgiu com os jogos de azar como jogos de cartas, de roletas, de números etc, no século XVI. Mais tarde, sua utilização foi ampliada para estudos de Genética por Mendel que, até hoje, continuam sendo válidos.

Os resultados previsíveis são chamados de **determinísticos**. Por exemplo, o tempo gasto para um objeto atingir o solo sendo lançado de uma certa altura. Caso não saibamos o resultado final de um experimento com antecedência, chamamos de **aleatório**. Por exemplo, lançar uma moeda; jogar na loteria de números, jogar um dado etc.

Em Probabilidade estudamos **experimentos aleatórios** que tenham a mesma chance de ocorrência, isto é, a chance de acontecer os diferentes resultados é a mesma.

Espaço Amostral (S)

Para cada experimento aleatório E defini-se **Espaço Amostral – S** - o conjunto de todos os possíveis resultados desse experimento.

Ex.: Experimento E_1 = jogar um dado e observar o número da face de cima.

$$S = \{1, 2, 3, 4, 5, 6\}$$

Ex.: Experimento E_2 = jogar 2 moedas e observar o resultado.

c – cara

$$S = \{(c,c), (c,k), (k,c), (k,k)\}$$

k - coroa

Evento

O evento se caracteriza por ser qualquer subconjunto do conjunto S .

Ex.: E_1 = lançar um dado e observar o número de cima.

$$S = \{1, 2, 3, 4, 5, 6\}$$

Evento A = ocorrer múltiplo de 2 = $\{2, 4, 6\}$

Observações:

- 1- Quando o evento for igual ao espaço amostral, o evento é dito **CERTO**;
- 2- Quando o evento for ~~$\emptyset$~~ , o evento é considerado **IMPOSSÍVEL**.

Probabilidade

$$p(\text{Evento}) = \frac{n^{\circ} \text{ elementos} - \text{evento}}{n^{\circ} \text{ elementos} - \text{espaço} - \text{amostral}}$$

$$p(Ev) = \frac{n^{\circ}(Ev)}{n^{\circ}(S)}$$

Exemplos:

1- Seja o experimento aleatório: nascimento de 3 filhotes de um casal de gatos.

Determine:

a) O **S** deste experimento:

$$S = \{(MFF), (MFM), (MMF), (MMM), (FMM), (FMF), (FFM), (FFF)\}$$

b) A probabilidade do evento: exatamente 2 fêmeas.

$$p(Ev) = \frac{n(Ev)}{n(S)} = \frac{(MFF)(FMF)(FFM)}{8} = \frac{3}{8}$$

c) A probabilidade do evento: nascimento de exatamente 3 machos.

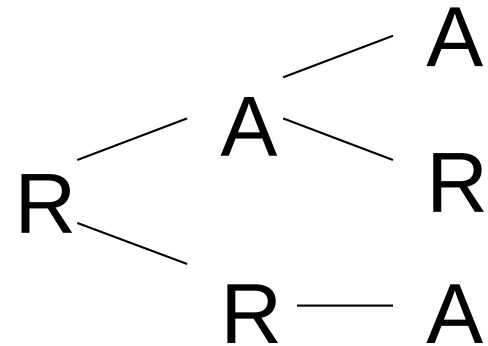
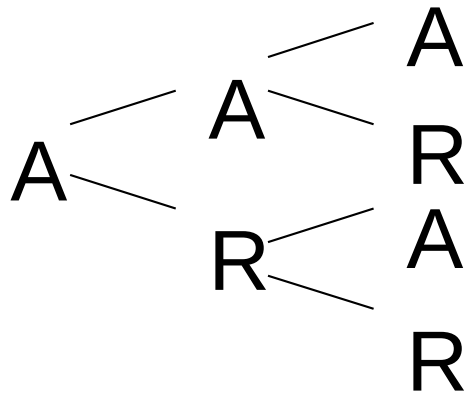
$$p(Ev) = \frac{n(Ev)}{n(S)} = \frac{1}{8}$$

d) A probabilidade do evento: nascimento de, pelo menos, 2 fêmeas.

$$p(Ev) = \frac{4}{8} = \frac{1}{2} = 50\%$$

2- Numa urna existem 4 bolas amarelas e 2 bolas rosas. Qual é o S retirando-se 3 bolas sucessivamente?

$S = \{(AAA), (AAR), (ARA), (ARR), (RAA), (RAR), (RRA)\} \Rightarrow n(S)=7$

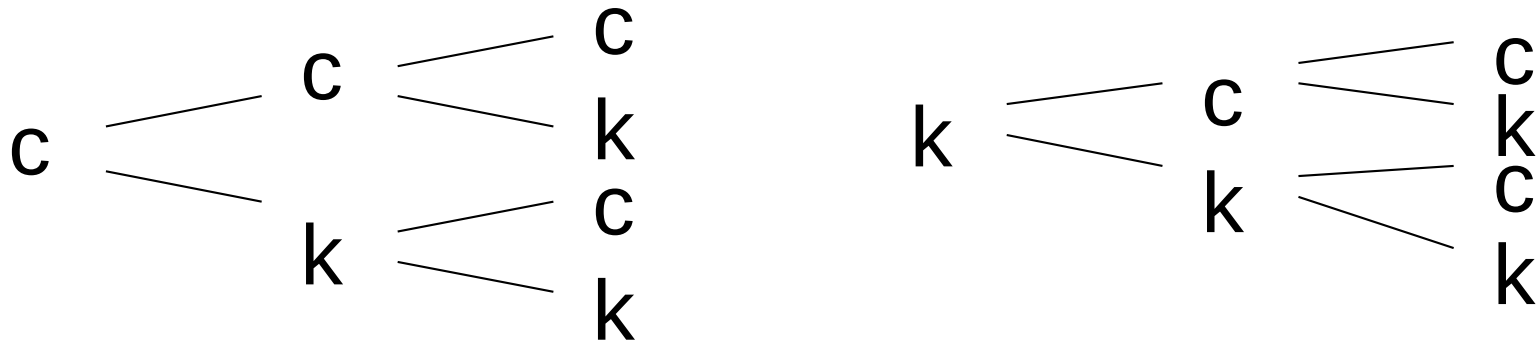


3- Qual o espaço amostral no lançamento de 2 moedas simultaneamente?

	c	k
c	cc	ck
k	kc	kk

$$S = \{(cc), (ck), (kc), (kk)\}$$

4- Qual o S do experimento aleatório lançamento simultâneo de 3 moedas diferentes?



$$S = \{(ccc), (cck), \dots (kkk)\}$$

$$n(S) = 8$$

4.1- Calcule a probabilidade do evento: no mínimo 1 cara:

$$p(Ev) = \frac{7}{8}$$

4.2- Calcule a probabilidade do evento: pelo menos 2 coroas:

$$p(Ev) = \frac{4}{8} = \frac{1}{2}$$

4.3- Calcule a probabilidade do evento: no máximo 1 cara:

$$p(Ev) = \frac{4}{8} = \frac{1}{2}$$

4.4- Calcule a probabilidade do evento: exatamente 2 coroas:

$$p(Ev) = \frac{3}{8}$$

Alguns Tipos de Eventos

1- Eventos Mutuamente Exclusivos ou Excludentes

Dois eventos são chamados mutuamente exclusivos, se eles não puderem ocorrer simultaneamente, isto é, evento A e evento B são disjuntos, ou seja,

$$A \cap B = \emptyset$$

Ex.: lançar um dado e observar se vai dar par ou ímpar.

$$S = \{1, 2, 3, 4, 5, 6\}$$

$$Ev(A) = \text{ocorrer n}^\circ \text{ par} = \{2, 4, 6\}$$

$$Ev(B) = \text{ocorrer n}^\circ \text{ ímpar} = \{1, 3, 5\}$$

$$A \cap B = \emptyset$$

2- Eventos Complementares $\overline{A}$

O complemento de um evento A , denotado por $\overline{A}$, consiste em todos os resultados em que o evento A não ocorre.

Ex.: Se $p(\text{chuva})=0,4$, determine $p(\text{ñ-chuva})$.

$$p(\text{ñ-chuva}) = 1 - p(\text{chuva}) =$$

$$1 - 0,4 = 0,6$$

3- Eventos Independentes

Dois eventos A e B são independentes se a ocorrência de um deles não afeta a probabilidade de ocorrência do outro. (Analogamente, vários eventos são independentes se a ocorrência de qualquer um deles não afeta as probabilidades de ocorrência dos outros).

Ex.: a jogada de uma moeda e a jogada de um dado são eventos independentes, porque o resultado da moeda não afeta a probabilidade do resultado do dado.

Então:

$$p(A \cap B) = p(A).p(B)$$

4- Eventos Condicionais ou Probabilidade Condicional

Se A e B são eventos de um espaço amostral S , com $p(B)$ diferente de zero, então a probabilidade condicional do evento A , tendo ocorrido o evento B , é indicada por **$p(A/B)$** .

$$p(A / B) = \frac{p(A \cap B)}{p(B)}$$

Ex.: Dois dados são lançados. Considere os eventos:

$$A = \{(x_1, x_2) \mid x_1 + x_2 = 10\}$$

$$B = \{(x_1, x_2) \mid x_1 > x_2\}$$

onde x_1 é o resultado do dado 1 e x_2 é o resultado do dado 2.

Calcule $p(A/B)$ e $p(B/A)$

Solução:

$$A = \{(4,6), (5,5), (6,4)\}$$

$$B = \{(2,1), (3,1), (3,2), (4,1), (4,2), (4,3), (5,1), (5,2), (5,3), (5,4), (6,1), (6,2), (6,3), (6,4), (6,5)\}$$

$$p(A/B) = \frac{p(A \cap B)}{p(B)} = \frac{\{(6,4)\}}{\{(2,1), (3,1), \dots, (6,5)\}} =$$

$$\frac{\frac{1}{36}}{\frac{15}{36}} = \frac{1}{15}$$

Ainda podemos calcular a probabilidade condicional da seguinte maneira:

Considere o primeiro evento como sendo um espaço amostral reduzido:

$S_{\text{reduzido}} = \{(2,1), (3,1), \dots, (6,5)\} = 15$
elementos

E o segundo evento passa a ser um evento desse S_{reduzido} .

$$\text{Ev: } x_1 + x_2 = 10 : \{(6,4)\} \implies \frac{1}{15}$$

$$p(B / A) = \frac{p(B \cap A)}{p(A)} = \frac{\{(6,4)\}}{\{(4,6), (5,5), (6,4)\}} =$$

$$\frac{\frac{1}{15}}{\frac{3}{15}} = \frac{1}{3}$$

Ou, pelo S_{reduzido} :

$$S_{\text{reduzido}}: \{(4,6), (5,5), (6,4)\}$$

$$\text{Evento: } x_1 > x_2 = \{(6,4)\}$$

$$p(\text{Ev}) = \frac{1}{3}$$

Notem que, em geral,
 $p(A/B)$ é diferente de $p(B/A)$.

Resumo:

$$p(A/B) = \frac{p(A \cap B)}{p(B)}$$

interseção do nº elementos
de **A** e **B** dividido pelo S
total.

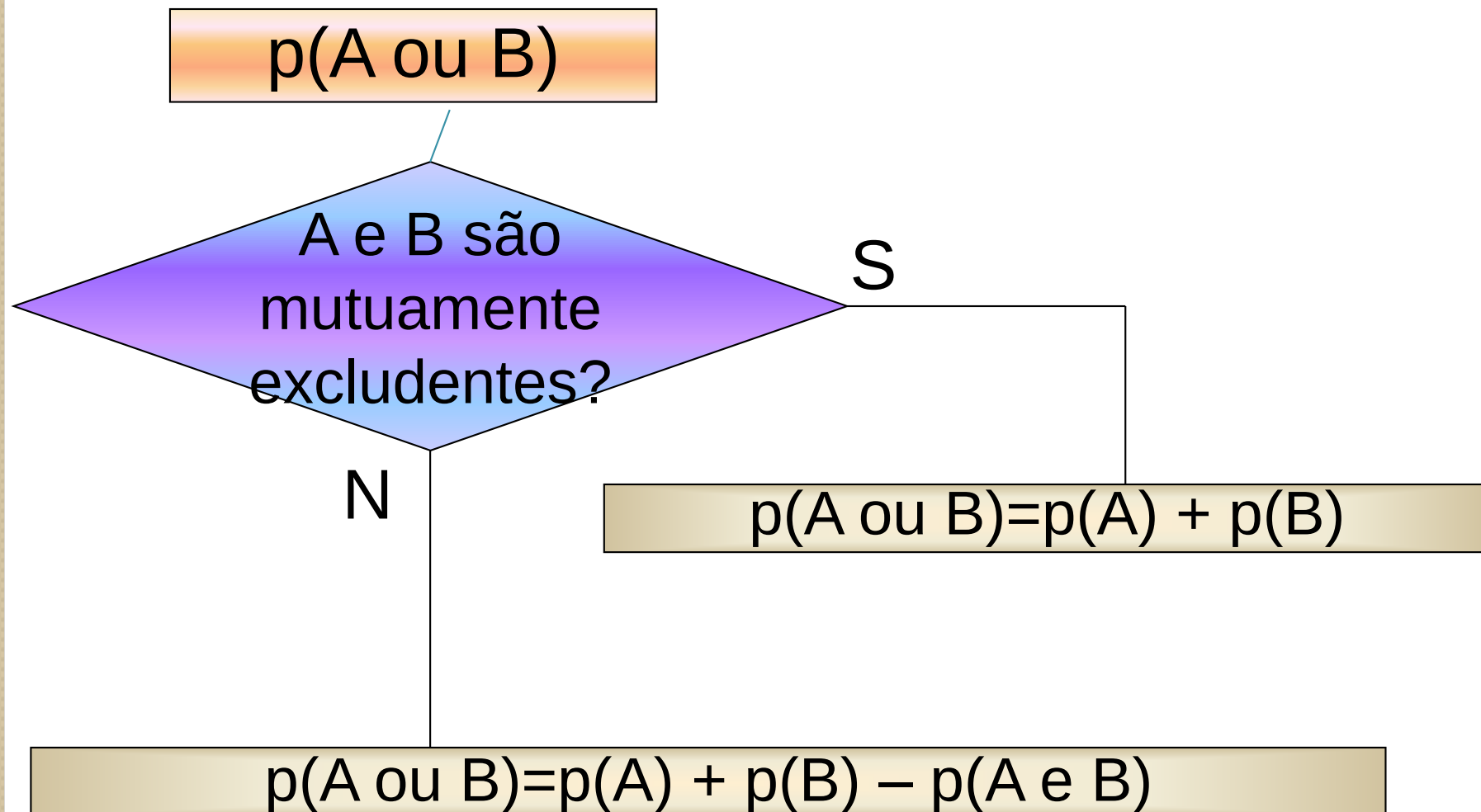
probabilidade de **B**
dividido pelo total.

OU

$p(B)$ se torna o S_{reduzido} e verificamos
quantos elementos do conjunto **A** estão em
B.

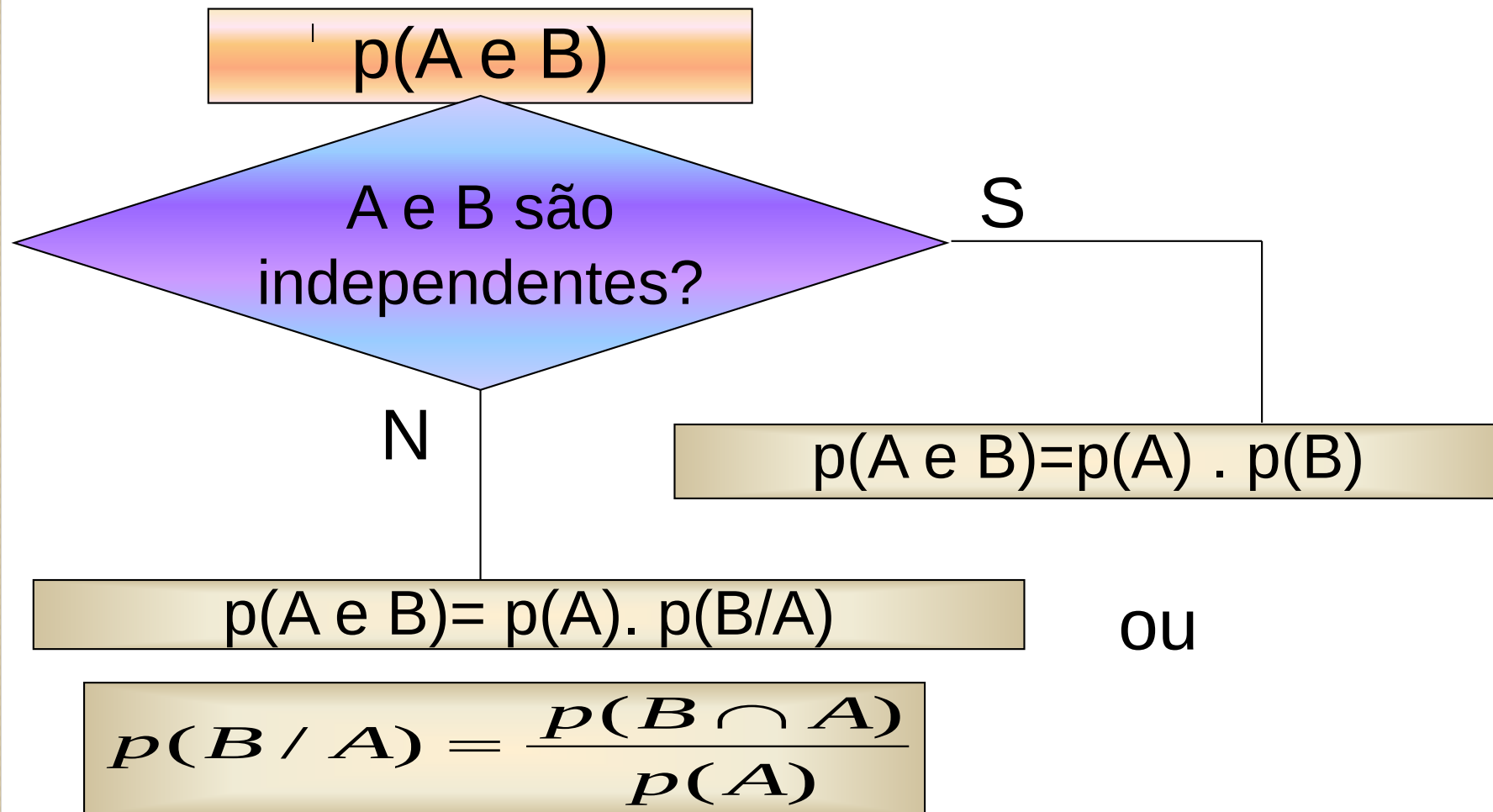
RESUMO

Regra da Adição



RESUMO

Regra da Multiplicação ou do Produto



Teorema de Bayes

Sejam os eventos $A_1, A_2, \dots, A_k$, mutuamente excludentes e seja B um evento qualquer. Se quisermos saber a probabilidade $p(A_1/B)$, conhecidas as probabilidades $p(B/A_1), p(B/A_2), \dots, p(B/A_k)$ e $p(A_1), p(A_2), \dots, p(A_k)$, então:

$$p(A_1 / B) = \frac{p(B / A_1) p(A_1)}{p(B / A_1) p(A_1) + \dots + p(B / A_k) p(A_k)}$$

para $p(B) > 0$

Ex.: A probabilidade de que um teste médico identifique corretamente alguém com uma doença, dando positivo é 0,99; e a probabilidade de que o teste identifique corretamente alguém sem a doença, dando negativo, é 0,95. A incidência da doença na população em geral é 0,0001. Você fez o teste e o resultado foi positivo. Qual a probabilidade de que você tenha a doença?
Seja D = ter doença

$\bar{D}$ = não ter doença $p(D) = 0,0001 \Rightarrow p(\bar{D}) = 0,9999$

S = teste positivo

$\bar{S}$ = teste negativo

$$p(S / D) = 0,99$$

$$p(\bar{S} / \bar{D}) = 0,95 \Rightarrow p(S / \bar{D}) = 0,05$$

$$p(D / S) = ???$$

$$p(D / S) = \frac{p(S / D) \cdot p(D)}{p(S / D) \cdot p(D) + p(S / \bar{D}) \cdot p(\bar{D})}$$

$$p(D / S) = \frac{0,99 \cdot 0,0001}{0,99 \cdot 0,0001 + 0,05 \cdot 0,9999}$$

$$p(D / S) = \frac{0,000099}{0,000099 + 0,049995}$$

$$p(D / S) = \frac{0,000099}{0,050094} = 0,001976... \cong 0,002$$

$$p(D / S) = 0,2\%$$