# BIOESTATÍSTICA BÁSICA

## AULA 3

# ARREDONDAMENTO E AJUSTE DE DADOS EM ESTATÍTISCA

**Quando for necessário, em cálculos e tabelas estatísticas, o arredondamento de dados, devemos proceder de acordo com a resolução número 886/66, da Fundação IBGE.**

I) Quando o primeiro algarismo a ser abandonado for 0, 1,2,3 ou 4, fica inalterado o último algarismo a permanecer.

exemplos:

a) 53,24 arredondado ao décimo mais próximo será 53,2.
b) 12,473 arredondado ao centésimo mais próximo será 12,47.

II) Quando o primeiro algarismo a ser abandonado for 6,7,8 ou 9, aumenta-se de uma unidade o último algarismo a permanecer.
exemplos:

a) 42,87 arredondado ao décimo mais próximo será 42,9.
b) 13,7 arredondado ao inteiro mais próximo será 14.

III) Quando o primeiro algarismo a ser abandonado for o 5, há duas situações a considerar:

A) Se ao 5 seguir, em qualquer ordem, um algarismo diferente de zero, aumenta se de uma unidade o algarismo a permanecer.

exemplos:

a) 2,352 arredondado a décimos será 2,4

b) 14,325004 arredondado a centésimos será 14,33.

B) Se o 5 for o último algarismo , ou ao 5 só se seguirem zeros, o último algarismo a ser conservado só será aumentado de uma unidade se for ímpar.

exemplos:

a) 24,75 arredondado a décimos será 24,8.

b) 14,7850 arredondado a centésimos será 14,78.

# Compensação ou Ajuste de Dados:

Quando houver arredondamento em somatórios e precisarmos "ajustar" alguns elementos, de modo a acertar a soma, usamos **"descarregar"** a diferença na maior ou nas maiores parcelas.

Exemplo:
25,32 + 17,85 + 10,44 + 31,17 = 84,78.

Se arredondássemos todas as parcelas e a soma, para décimos, teríamos:
25,3 + 17,8 + 10,4 + 31,2 = 84,7.

No entanto esta soma agora é **84,7** que não é o arredondamento correto de 84,78. A compensação faríamos na maior parcela (31,2), tornando-a 31,3, logo,

teríamos:
25,3 + 17,8 + 10,4 + 31,3 = 84,8.

# MEDIDAS DE TENDÊNCIA CENTRAL E

# MEDIDAS DE SEPARATRIZES

## OBJETIVOS

- CALCULAR AS MEDIDAS DE TENDÊNCIA CENTRAL E SEPARATRIZES;
- INTERPRETAR OS VALORES DESSAS MEDIDAS EM DIFERENTES TIPOS DE DISTRIBUIÇÃO.

## MÉDIA ( $\bar{X}$ )

**A MÉDIA ARITMÉTICA PODE SER CLASSIFICADA SEM INTERVALO DE CLASSE E COM INTERVALO DE CLASSE.**

## SEM INTERVALO DE CLASSE

**A MÉDIA É A MEDIDA DE TENDÊNCIA CENTRAL MAIS SIMPLES QUE TEMOS EM ESTATÍSTICA.**

$$\bar{X} = \frac{\sum X_i}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

**ONDE:**

- **x É O VALOR MÉDIO DA VARIÁVEL;**
- **x OS VALORES POSSÍVEIS DA VARIÁVEL;**
- **n É A QUANTIDADE DE VALORES.**

**OBS.**

**A MÉDIA SÓ TEM SENTIDO QUANDO OS DADOS OBEDECEM À DISTRIBUIÇÃO NORMAL OU DE GAUSS.**

# EXEMPLO

**OS DADOS A SEGUIR REPRESENTAM A IDADE DE UMA POPULAÇÃO VOLUNTÁRIA NA PESQUISA DE INGESTÃO DE BEBIDA ALCOÓLICA E SEU EFEITO AO DIRIGIR VEÍCULOS AUTOMOTORES: 20, 25, 18, 32, 21, 27, 19, 18, 23, 21. DETERMINE A MÉDIA ARITMÉTICA DA IDADE DOS VOLUNTÁRIOS.**

# SOLUÇÃO

$$\bar{X} = \frac{\sum X_i}{n} =$$

$$\bar{X} = \frac{20 + 25 + 18 + 32 + 21 + 27 + 19 + 18 + 23 + 21}{10}$$

$$\bar{X} = 22,4 \text{ ANOS}$$

## EXEMPLO 2

**EM 10 CIDADES BRASILEIRAS OS ÍNDICES PLUVIOMÉTRICOS(EM mm) REGISTRADOS NO MÊS DE JULHO DE 2010, CORRESPONDEM A: 110, 158, 117, 220, 190, 25, 89, 90, 240, 348. DETERMINAR A MÉDIA PLUVIOMÉTRICA APURADA.**

# SOLUÇÃO

$$\bar{X} = \frac{\sum X_i}{n} =$$

$$\bar{X} = \frac{110 + 158 + 117 + 220 + 190 + 25 + 89 + 90 + 240 + 348}{10}$$

$$\bar{X} = 158,70 \text{ mm}$$

# DESVIO MÉDIO

**O DESVIO EM RELAÇÃO À MÉDIA É À DIFERENÇA ENTRE CADA ELEMENTO E A MÉDIA.**

$$d_i = x - \bar{X}$$

**O SOMATÓRIO DOS DESVIOS MÉDIOS SERÁ IGUAL A UM, QUANDO OS DADOS FOREM DISTRIBUÍDOS NORMALMENTE, OU SEJA, QUANDO DA MÉDIA FOR REPRESENTATIVA.**

# MÉDIA PONDERADA ($\bar{X}_P$)

$$\bar{X} = \frac{\sum X_i \; P_i}{\sum P_i} =$$

$$\bar{X} = \frac{X_1 P_1 + \ldots\ldots + X_n P_n}{\sum P_i}$$

# EXEMPLO

**DE ACORDO COM AS NOTAS DE UMA AVALIAÇÃO E SEUS PESOS APRESENTADOS NA PLANILHA A SEGUIR, ENCONTRE A MÉDIA.**

| NOTAS | PESOS |
|---|---|
| 5,0 | 1 |
| 8,0 | 2 |
| 7,0 | 3 |
| 6,0 | 4 |

# SOLUÇÃO

$$\bar{X} = \frac{5\,.1 + 8.2 + 7\,.\,3 + 6.4}{1 + 2 + 3 + 4}$$

$$\bar{X} = 6{,}6$$

# EXEMPLO 2

**NUM CONCURSO, FORAM APLICADAS TRÊS PROVAS: MATEMÁTICA, COM PESO 5; PORTUGUÊS, COM PESO 3; E INFORMÁTICA, COM PESO 2. CADA PROVA VALE 100 PONTOS E A MÉDIA PARA APROVAÇÃO É 70. UM CANDIDATO TIROU 84 EM MATEMÁTICA E 72 EM PORTUGUÊS. QUAL DEVE SER SUA NOTA MÍNIMA EM INFORMÁTICA PARA O CANDIDATO SER APROVADO**

# MÉDIA COM INTERVALO DE CLASSE

**DEFINIMOS UM VALOR MÉDIO (MÉDIA ARITMÉTICA) PARA O INTERVALO DA CLASSE E, ENTÃO, DETERMINAMOS A MÉDIA ARITMÉTICA ATRAVÉS DA EQUAÇÃO:**

$$\bar{X} = \frac{\sum X_c f_i}{\sum f_i}$$

# MÉDIA COM INTERVALO DE CLASSE

**ONDE ( $\overline{X}$ ) É O VALOR DA MÉDIA PONDERADA DA VARIÁVEL, $x_C$ E A MÉDIA ARITMÉTICA DE CADA CLASSE, $f_i$ A FREQÜÊNCIA RELATIVA ÀQUELA CLASSE E $\Sigma f_i$ O SOMATÓRIO DOS DAS FREQÜÊNCIAS.**

$$\overline{X} = \frac{\sum X_c \, f_i}{\sum f_i}$$

# EXEMPLO

| ESTATURA (cm) | $f_i$ | $x_c$ | $x_c \cdot f_i$ |
|---|---|---|---|
| 150 ⊢ 160 | 4 | 155 | 620 |
| 160 ⊢ 170 | 10 | 165 | 1650 |
| 170 ⊢ 180 | 6 | 175 | 1050 |
| | $\Sigma = 20$ | | $\Sigma = 3320$ |

# SOLUÇÃO

$$\bar{X} = \frac{\sum X_c f_i}{\sum f_i}$$

$$\bar{X} = \frac{3320}{20}$$

$$\bar{X} = 166\,\text{cm}$$

# EXEMPLO

**DETERMINAR A MÉDIA PARA OS DADOS AGRUPADOS DA DISTRIBUIÇÃO DAS ESTATURAS, EM CENTÍMETROS, DE ALUNOS DO CURSO COLEGIAL:**

| CLASSE | VALOR CENTRAL | FREQUÊNCIA |
|---|---|---|
| 135 ⊢ 145 | 140 | 15 |
| 145 ⊢ 155 | 150 | 150 |
| 155 ⊢ 165 | 160 | 250 |
| 165 ⊢ 175 | 170 | 70 |
| 175 ⊢ 185 | 180 | 10 |
| 185 ⊢ 195 | 190 | 5 |

# SOLUÇÃO

| CLASSES | $f_i$ | $x_c$ | $x_c \cdot f_i$ |
|---|---|---|---|
| 135 ⊢ 145 | 15 | 140 | 2.100 |
| 145 ⊢ 155 | 150 | 150 | 22.500 |
| 155 ⊢ 165 | 250 | 160 | 40.000 |
| 165 ⊢ 175 | 70 | 170 | 11.900 |
| 175 ⊢ 185 | 10 | 180 | 1.800 |
| 185 ⊢ 195 | 5 | 190 | 950 |
| | $\Sigma = 500$ | | $\Sigma = 79.250$ |

# MODA

**É O VALOR DE MAIOR FREQÜÊNCIA EM UMA DISTRIBUIÇÃO DE VALORES, OU SEJA, O VALOR QUE MAIS SE REPETE. ESSE VALOR PODE SER ENCONTRADO DE DUAS FORMAS DIFERENTES, SEM INTERVALO DE CLASSE E COM INTERVALO DE CLASSE.**

## SEM INTERVALO DE CLASSE

**DISTRIBUIÇÃO DE DADOS:**

| | | | | | |
|---|---|---|---|---|---|
| 12 | 18 | 18 | 19 | 18 | 16 |
| 17 | 16 | 18 | 18 | 19 | 18 |
| 18 | 18 | 16 | 19 | 17 | 11 |
| 18 | 12 | 18 | 19 | 17 | 12 |
| 21 | | | | | |

**Mo = 18**

## COM INTERVALO DE CLASSE

**A TABELA ABAIXO APRESENTA A DISTRIBUIÇÃO DAS ESTATURAS DE 20 PLANTAS DE UMA PESQUISA. CALCULAR A MODA DAS ALTURAS DESSAS PLANTAS.**

| i | ESTATURAS (cm) | $f_i$ |
| --- | --- | --- |
| 1 | 150 ⊢ 160 | 4 |
| 2 | 160 ⊢ 170 | 10 |
| 3 | 170 ⊢ 180 | 6 |
| | | Σ = 20 |

$$M_o = \frac{160 + 170}{2} = 165$$

**CLASSE MODAL 2**

# MEDIANA (Md)

**É O VALOR DO ELEMENTO QUE OCUPA A POSIÇÃO CENTRAL DA DISTRIBUIÇÃO, OU SEJA, ELA DIVIDE UM CONJUNTO ORDENADO DE DADOS EM DOIS GRUPOS IGUAIS. A METADE DOS VALORES MENORES QUE A MEDIANA E A OUTRA METADE TERÁ VALORES SUPERIORES À MEDIANA.**

## SEM INTERVALO DE CLASSE

**NA MEDIANA SEM INTERVALO DE CLASSE, TÊM-SE DOIS MÉTODOS: QUANDO A QUANTIDADE É IMPAR E QUANDO A QUANTIDADE É PAR.**

## NÚMERO ÍMPAR DE VALORES

**A MEDIANA É O VALOR CENTRAL. DE ACORDO COM A DISTRIBUIÇÃO DE DADOS ABAIXO, VAMOS ENCONTRAR A MEDIANA.**

**1, 12, 15, 24, 2, 3, 45, 25, 32, 65, 11, 14, 63, 24, 50**

**1, 2, 3, 11, 12, 14, 15, 24, 24, 25, 32, 45, 50, 63, 65**

$$Md = \frac{n+1}{2} \Leftrightarrow Md = \frac{15+1}{2} \Leftrightarrow Md = 8$$

LOGO, A $Md = 24$

**NÚMERO PAR DE VALORES – É MÉDIA ARITMÉTICA ENTRE OS DOIS VALORES CENTRAIS**

**2, 6, 7, 10, 12, 13, 18, 21**

$$\mathbf{Md} = \frac{12 + 10}{2} = 11$$

# COM INTERVALO DE CLASSE

**CONSISTE NA DETERMINAÇÃO DA CLASSE EM QUE SE ENCONTRA A MEDIANA. PARA DETERMINAR A CLASSE MEDIANA, UTILIZA-SE A FREQÜÊNCIA ACUMULADA IMEDIATAMENTE MAIOR QUE A ENCONTRADA ATRAVÉS DA RELAÇÃO:**

$$\sum \frac{f_i}{2}$$

**A TABELA ABAIXO APRESENTA A DISTRIBUIÇÃO DAS ESTATURAS DE 48 ALUNOS DE UMA UNIVERSIDADE ESTADUAL. PODEMOS, COM ESSAS INFORMAÇÕES, CALCULAR A MEDIANA DAS ESTATURAS DOS ESTUDANTES DESSA UNIVERSIDADE.**

| i | ESTATURA (cm) | $f_i$ | $f_{ac}$ |
|---|---|---|---|
| 1 | 150 ⊢ 160 | 4 | 4 |
| 2 | 160 ⊢ 170 | 10 | 14 |
| 3 | 170 ⊢ 180 | 20 | 34 |
| 4 | 180 ⊢ 190 | 6 | 40 |
| 5 | 190 ⊢ 200 | 6 | 46 |
| 6 | 200 ⊢ 210 | 2 | 48 |
| | | Σ = 48 | |

$$\sum \frac{f_i}{2} = \frac{48}{2} = 24$$

**APÓS OBTER O VALOR 24 ACIMA, TEM-SE QUE OBSERVAR NA TABELA QUAL O PRIMEIRO VALOR MAIOR QUE 24, NESTE CASO, FOI O NÚMERO 34, CLASSE MEDIANA FOI A TERCEIRA CLASSE. O PRÓXIMO PASSO É UTILIZAR A FÓRMULA ABAIXO COM OS DADOS DA TABELA.**

$$Md = L_i + \frac{\left( \frac{\sum f_i}{2} - f_k \right) . h}{f_i}$$

$$Md = 170 + \frac{(24 - 14) . 10}{20}$$

$$Md = 175 \text{ cm}$$

# DADA A DISTRIBUIÇÃO AMOSTRAL, DETERMINAR A MEDIANA:

| i | INTERVALOS DAS CLASSES | $f_i$ | $f_k$ |
|---|---|---|---|
| 1 | 35 ⊢ 45 | 5 | 5 |
| 2 | 45 ⊢ 55 | 12 | 17 |
| 3 | 55 ⊢ 65 | 18 | 35 |
| 4 | 65 ⊢ 75 | 14 | 49 |
| 5 | 75 ⊢ 85 | 6 | 55 |
| 6 | 85 ⊢ 95 | 3 | 58 |
| | | $\Sigma = 58$ | |

# QUARTIL (Q)

**SÃO VALORES QUE DIVIDEM A SÉRIE EM QUATRO PARTES IGUAIS. TEMOS TRÊS QUARTIS:**

**PRIMEIRO QUARTIL ($Q_1$) – SÃO OS PRIMEIROS 25% DOS DADOS QUE SÃO MENORES QUE OS OUTROS 75% RESTANTES.**

# QUARTIL (Q)

**SEGUNDO QUARTIL ($Q_2$) – É O VALOR QUE COINCIDE COM A MEDIANA.**

**TERCEIRO QUARTIL ($Q_3$) – É O VALOR QUE CORRESPONDE AOS 75% DOS DADOS QUE SÃO MENORES QUE OS 25% RESTANTES.**

**PARA OS RESPECTIVOS CÁLCULOS, UTILIZAMOS PROCEDIMENTO SIMILAR AO UITILIZADO PARA O CÁLCULO DA MEDIANA.**

# EXEMPLO

**DETERMINAR O PRIMEIRO E TERCEIRO QUARTIL, DA DISTRIBUIÇÃO ABAIXO**

| i | ESTATURA (cm) | $f_i$ | $f_k$ |
|---|---|---|---|
| 1 | 150 ⊢ 160 | 4 | 4 |
| 2 | 160 ⊢ 170 | 10 | 14 |
| 3 | 170 ⊢ 180 | 20 | 34 |
| 4 | 180 ⊢ 190 | 6 | 40 |
| 5 | 190 ⊢ 200 | 6 | 46 |
| 6 | 200 ⊢ 210 | 2 | 48 |
| | | $\Sigma = 48$ | |

# PRIMEIRO QUARTIL

$$\sum \frac{K \,.\, f_i}{4} = \frac{1 \,.\, 48}{4} = 12$$

$$Q_1 = L_i + \frac{\left(\sum \frac{f_i}{4} - f_k\right) . \, h}{f_i}$$

$$Q_1 = 160 + \frac{(12 - 4) \,.\, 10}{10}$$

$$Q_1 = 168 \text{ cm}$$

# TERCEIRO QUARTIL

$$\sum \frac{K \,.\, f_i}{4} = \frac{3 \,.\, 48}{4} = 36$$

$$Q_1 = L_i + \frac{\left(\sum \frac{f_i}{4} - f_k\right) . \, h}{f_i}$$

$$Q_1 = 180 + \frac{(36 - 34) \,.\, 10}{6}$$

$$Q_1 = 183{,}34 \text{ cm}$$

# DECIL

**PARA DETERMINARMOS O DECIL, USAREMOS O MESMO RACIOCÍNIO DO CÁLCULO DO QUARTIL, APENAS SUBSTITUINDO O DENOMINADOR 4 POR 10, NA EQUAÇÃO ABAIXO.**

$$\sum \frac{K \cdot f_i}{10}$$

# PERCENTIL

**OS PERCENTIS SÃO OS NOVENTA E NOVE VALORES QUE SEPARAM UMA SÉRIE EM 100 PARTES IGUAIS. DE MANEIRA ANÁLOGA, AO CÁLCULO DA MEDIANA USAREMOS O SEGUINTE PROCEDIMENTO:**

$$\sum \frac{K \, . \, f_i}{100}$$

# OBSERVAÇÕES

$$P_{50} = Md$$

$$P_{25} = Q_1$$

$$P_{75} = Q_3$$

DADA A TABELA A SEGUIR, DETERMINAR O 15º PERCENTIL.

| i | ESTATURA (cm) | $f_i$ | $f_k$ |
|---|---|---|---|
| 1 | 150 ⊢ 160 | 4 | 4 |
| 2 | 160 ⊢ 170 | 10 | 14 |
| 3 | 170 ⊢ 180 | 20 | 34 |
| 4 | 180 ⊢ 190 | 6 | 40 |
| 5 | 190 ⊢ 200 | 6 | 46 |
| 6 | 200 ⊢ 210 | 2 | 48 |
| | | $\Sigma = 48$ | |

# BIOESTATÍSTICA BÁSICA

# AULA 5

# DISPERSÃO OU

# VARIABILIDADE

# OBJETIVOS

- ENCONTRAR O VALOR DO DESVIO PADRÃO PARA UMA VARIÁVEL DISCRETA;
- ENCONTRAR O VALOR DO DESVIO PADRÃO PARA UMA VARIÁVEL CONTÍNUA.

# VARIÂNCIA E DESVIO PADRÃO

A VARIÂNCIA E O DESVIO PADRÃO CONSIDERAM TOTALMENTE OS VALORES DA VARIÁVEL EM ESTUDO. A BASE DA VARIÂNCIA GIRA EM TORNO DA MÉDIA ARITMÉTICA, DETERMINANDO ASSIM A MÉDIA ARITMÉTICA DOS QUADRADOS DOS DESVIOS.

# VARIÂNCIA (S)

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

ONDE $x_i$ É O VALOR DE CADA DADO, $x$ É A MÉDIA ARITMÉTICA DOS VALORES E $n$ É A QUANTIDADE DE VALORES.

A VARIÂNCIA É UMA MÉDIA ARITMÉTICA CALCULADA A PARTIR DOS QUADRADOS DOS DESVIOS OBTIDOS ENTRE OS ELEMENTOS DA SÉRIE E A SUA MÉDIA.

O DESVIO PADRÃO É A RAIZ QUADRADA POSITIVA DA VARIÂNCIA, COMO É MOSTRADO NA EQUAÇÃO A SEGUIR:

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

A SÉRIE REPRESENTA UMA POPULAÇÃO, ENTÃO O DESVIO PADRÃO SERÁ ENCONTRADO ATRAVÉS DA EQUAÇÃO:

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

**DETERMINAR O DESVIO PADRÃO DA SEQÜÊNCIA 4, 5, 8, 5.**

**1° PASSO: ENCONTRAR A MÉDIA ARITMÉTICA**

$$\bar{X} = \frac{\sum X_i}{n} = \frac{4+5+8+5}{4} = 5,5$$

**2º PASSO:**

$$(x_i - \bar{x})^2$$

$$(x_i - \bar{x})^2 = (4 - 5{,}5)^2 = 2{,}25$$

$$(x_i - \bar{x})^2 = (5 - 5{,}5)^2 = 0{,}25$$

$$(x_i - \bar{x})^2 = (8 - 5{,}5)^2 = 6{,}25$$

$$(x_i - \bar{x})^2 = (5 - 5{,}5)^2 = 0{,}25$$

$$\sum = 9$$

## 3º PASSO:

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

$$S = \sqrt{\frac{9}{4}}$$

$$S = \frac{3}{2} = 1{,}5 \text{ UNIDADES}$$

# VARIÁVEIS

## QUALITATIVAS
*(são atributos)*

*Exemplos:*

- *sexo*
- *religião*
- *naturalidade*
- *cor dos olhos*
- *faixa etária*

## QUANTITATIVAS
*(são numéricas)*

### DISCRETAS

*Exemplos:*

- *quantidade de estudantes em uma disciplina*
- *quantidade de cômodos em uma residência*

### CONTÍNUAS

*Exemplos:*

- *tempo de vôo entre duas cidades*
- *duração da bateria de telefone celular*

# VARIÁVEL DISCRETA

AQUI SURGE A REPETIÇÃO DOS TERMOS DA SÉRIE, E O DESVIO PADRÃO PASSA A SER UMA MÉDIA PONDERADA DOS QUADRADOS DOS DESVIOS DOS TERMOS DA SÉRIE.

# PRIMEIRA POSSIBILIDADE

A VARIÁVEL DISCRETA REPRESENTA UMA POPULAÇÃO E O DESVIO PADRÃO É CALCULADO ATRAVÉS DA SEGUINTE EQUAÇÃO:

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2 . f_i}{\sum f_i}}$$

NA TABELA DE DADOS A SEGUIR, COM DISTRIBUIÇÃO DE FREQÜÊNCIA, VAMOS DETERMINAR O SEU DESVIO PADRÃO DE ACORDO COM AS INFORMAÇÕES SUGERIDAS ANTERIORMENTE.

| $x_i$ | $f_i$ | $f_i \cdot x_i$ |
|---|---|---|
| 1 | 2 | 2 |
| 2 | 5 | 10 |
| 3 | 8 | 24 |
| 4 | 6 | 24 |
| 5 | 3 | 15 |
| 6 | 1 | 6 |
| | $\Sigma = 25$ | $\Sigma = 81$ |

## CÁLCULO DA MÉDIA:

$$\bar{X} = \frac{\sum X_i \; f_i}{\sum f_i} = \frac{81}{25}$$

$$\bar{X} = 3,24$$

| $x_i$ | $f_i$ | $f_i \cdot x_i$ | $(x_i - \bar{x})^2 \cdot f_i$ |
|---|---|---|---|
| 1 | 2 | 2 | 10,03 |
| 2 | 5 | 10 | 7,68 |
| 3 | 8 | 24 | 0,46 |
| 4 | 6 | 24 | 3,46 |
| 5 | 3 | 15 | 9,29 |
| 6 | 1 | 6 | 7,61 |
| | Σ = 25 | Σ = 81 | Σ = 38,53 |

A VARIÁVEL DISCRETA REPRESENTA UMA AMOSTRA DA POPULAÇÃO E O DESVIO PADRÃO É DETERMINADO POR INTERMÉDIO DA SEGUINTE EQUAÇÃO:

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2 . f_i}{\sum f_i}} = \sqrt{\frac{38,53}{25}}$$

$$S = 1,24$$

# VARIÁVEL CONTÍNUA

PERCEBER QUE A ÚNICA DIFERENÇA QUE HÁ PARA A SEÇÃO ANTERIOR, COMO TEMOS INTERVALOS DE CLASSE, É QUE TEMOS QUE ENCONTRAR O VALOR DA MÉDIA DE CADA CLASSE $X_C$.

## PRIMEIRA POSSIBILIDADE

PARA UMA VARIÁVEL CONTÍNUA REPRESENTATIVA DE UMA POPULAÇÃO, A EQUAÇÃO UTILIZADA PARA O CÁLCULO SERÁ:

$$S = \sqrt{\frac{\sum (x_c - \bar{x})^2 . f_i}{\sum f_i}}$$

DETERMINE O DESVIO PADRÃO PARA A DISTRIBUIÇÃO DE FREQÜÊNCIA A SEGUIR:

| CLASSES | INT. DE CLSSES | $f_i$ | Xo | Xo. fi |
|---|---|---|---|---|
| 1 | $0 \mapsto 4$ | 1 | 2 | 2 |
| 2 | $4 \mapsto 8$ | 3 | 6 | 18 |
| 3 | $8 \mapsto 12$ | 5 | 10 | 50 |
| 4 | $12 \mapsto 16$ | 1 | 14 | 14 |
| | | $\Sigma = 11$ | | $\Sigma = 84$ |

# CÁLCULO DA MÉDIA

$$\bar{X} = \frac{\sum X_c \, f_i}{\sum f_i} = \frac{84}{10}$$

$$\bar{X} = 8,4$$

DETERMINE O DESVIO PADRÃO PARA A DISTRIBUIÇÃO DE FREQÜÊNCIA A SEGUIR:

| CLASSES | INT. DE CLSSES | $f_i$ | Xc | Xc . f i | $\left(x_c - \overline{X}\right)^2 . f_i$ |
| --- | --- | --- | --- | --- | --- |
| 1 | $0 \vdash\!\!\rightarrow 4$ | 1 | 2 | 2 | 40,96 |
| 2 | $4 \vdash\!\!\rightarrow 8$ | 3 | 6 | 18 | 17,28 |
| 3 | $8 \vdash\!\!\rightarrow 12$ | 5 | 10 | 50 | 12,80 |
| 4 | $12 \vdash\!\!\rightarrow 16$ | 1 | 14 | 14 | 31,36 |
| | | $\Sigma = 10$ | | $\Sigma = 84$ | $\Sigma = 102,4$ |

# CÁLCULO DO DESVIO PADRÃO

$$S = \sqrt{\frac{\sum (x_c - \bar{x})^2 . f_i}{\sum f_i}}$$

$$S = \sqrt{\frac{102}{10}}$$

$$S = 3,2$$

COM OS DADOS A SEGUIR, DETERMINAR O DESVIO PADRÃO DA DISTRIBUIÇÃO DE FREQUÊNCIA DO CONSUMO DE ENERGIA ELÉTRICA

| CONSUMO | $f_i$ | $x_o$ |
|---|---|---|
| 5 ⊢ 25 | 4 | 15 |
| 25 ⊢ 45 | 6 | 35 |
| 45 ⊢ 65 | 14 | 55 |
| 65 ⊢ 85 | 26 | 75 |
| 85 ⊢ 105 | 14 | 95 |
| 105 ⊢ 125 | 8 | 115 |
| 125 ⊢ 145 | 6 | 135 |
| 145 ⊢ 165 | 2 | 155 |
| | | |

| CONSUMO | $f_i$ | Xc | Xc . f i | |
|---|---|---|---|---|
| 5 ⊢→ 25 | 4 | 15 | 60 | |
| 25 ⊢→ 45 | 6 | 35 | | |
| 45 ⊢→ 65 | 14 | 55 | | |
| 65 ⊢→ 85 | 26 | 75 | | |
| 85 ⊢→ 105 | 14 | 95 | | |
| 105 ⊢→ 125 | 8 | 115 | | |
| 125 ⊢→ 145 | 6 | 135 | | |
| 145 ⊢→ 165 | 2 | 155 | | |
| | | | | |

# ASSIMETRIA E CURTOSE

# OBJETIVOS

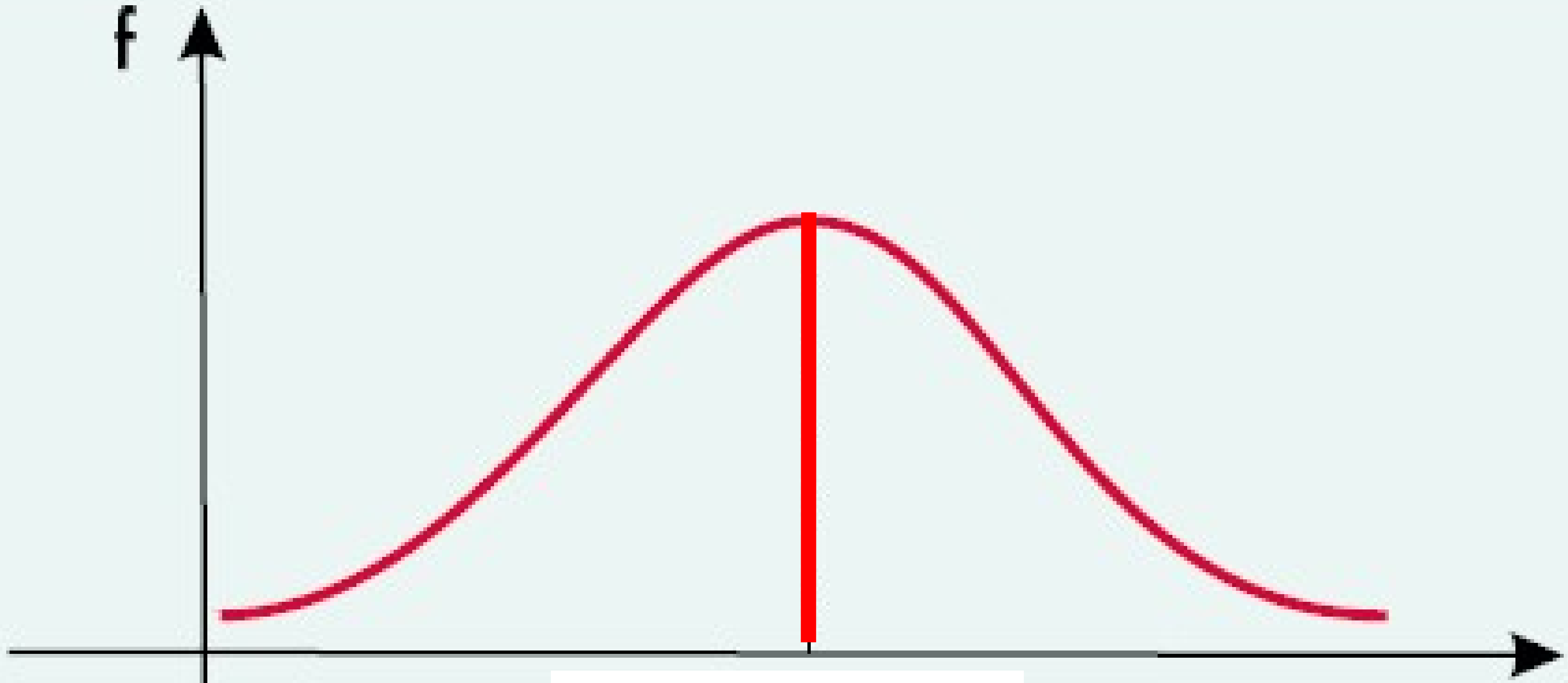
- DETERMINAR A ASSIMETRIA DE UMA DISTRIBUIÇÃO DE FREQÜÊNCIA;

- ENCONTRAR OS VALORES DAS MEDIDAS DE CURTOSE EM UMA DISTRIBUIÇÃO.

# ASSIMETRIA

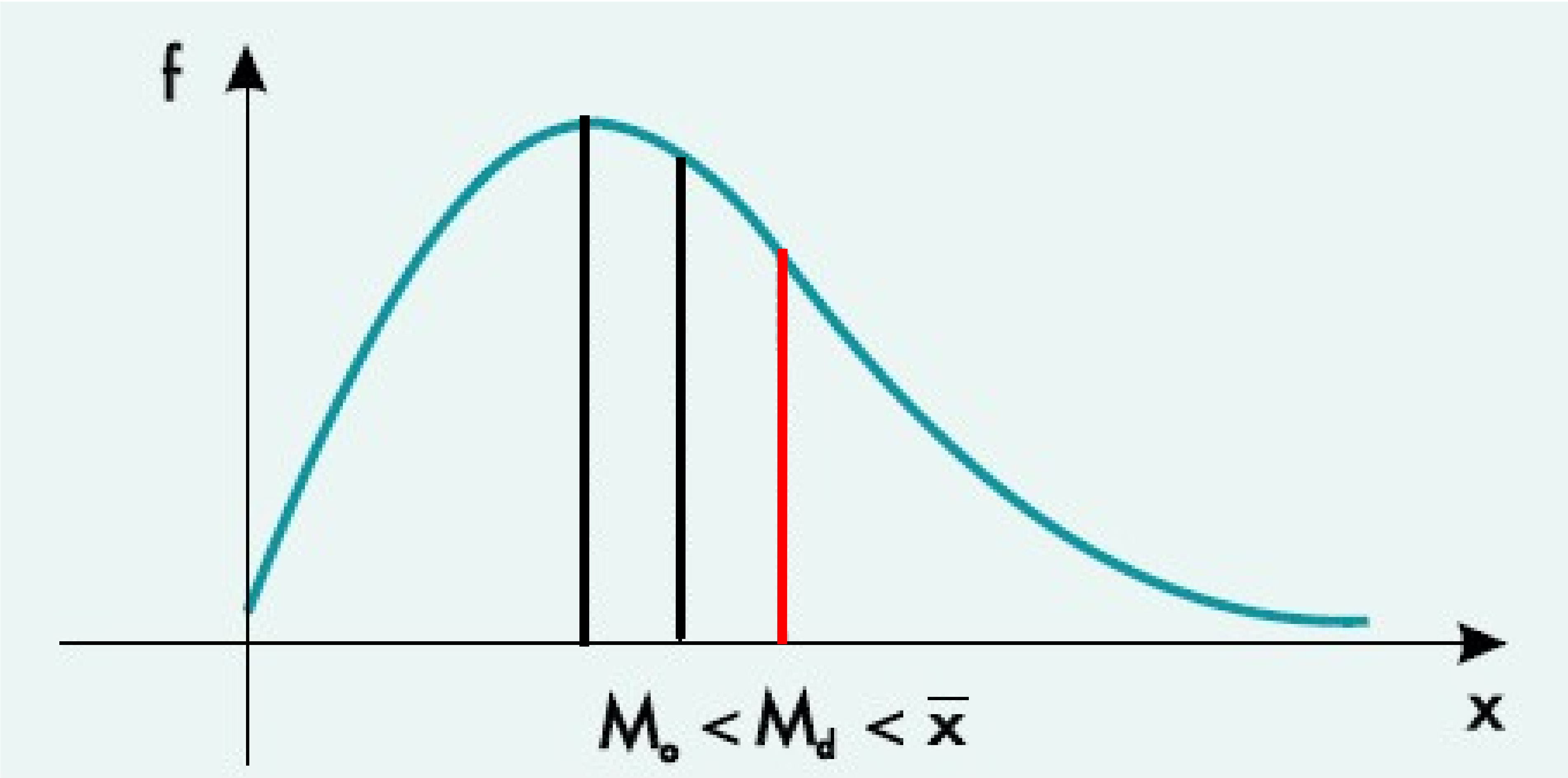
**QUANDO A DISTRIBUIÇÃO FOR SIMÉTRICA, A MÉDIA E A MODA COINCIDEM; QUANDO ELA FOR ASSIMÉTRICA À ESQUERDA OU NEGATIVA, A MÉDIA É MENOR QUE A MODA E, QUANDO FOR ASSIMÉTRICA À DIREITA OU POSITIVA, A MÉDIA É MAIOR QUE A MODA.**

# DISTRIBUIÇÃO SIMÉTRICA

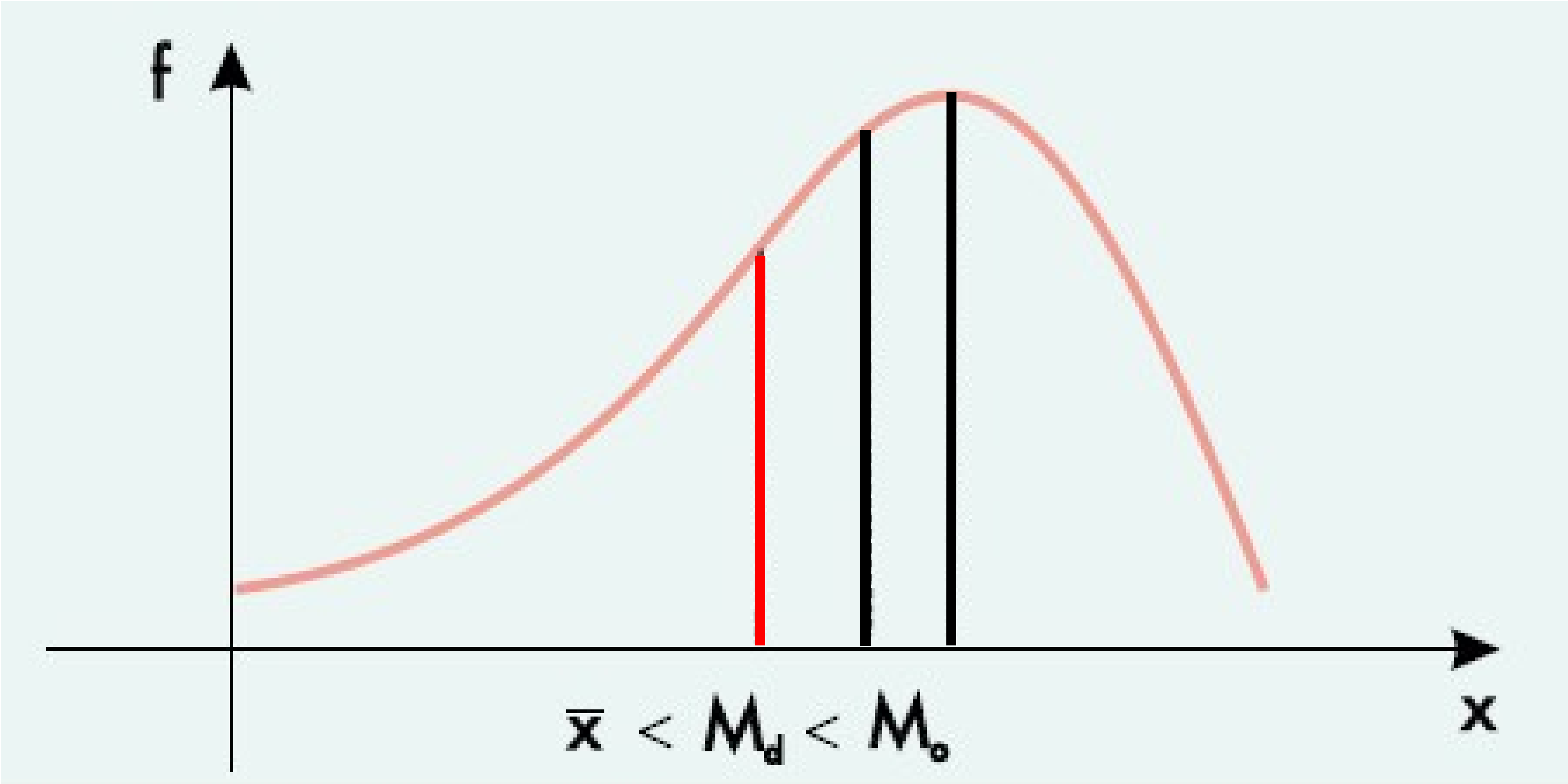



$\bar{X} = Md = Mo$

# ASSIMÉTRICA À DIREITA OU POSITIVA



# ASSIMÉTRICA À ESQUERDA OU NEGATIVA



**O PARÂMETRO QUANTITATIVO A RESPEITO DA ASSIMETRIA, CONSISTE NA DIFERENÇA ENTRE A MÉDIA E A MODA DA DISTRIBUIÇÃO.**

$$\bar{X} - Mo$$

# CLASSIFICAÇÃO

$\overline{X} - Mo = 0 \Rightarrow$ **DISTRIBUIÇÃO SIMÉTRICA.**

$\overline{X} - Mo < 0 \Rightarrow$ **DISTRIBUIÇÃO ASSIMÉTRICA NEGATIVA.**

$\overline{X} - Mo > 0 \Rightarrow$ **DISTRIBUIÇÃO ASSIMÉTRICA POSITIVA.**

# EXEMPLO 1

| i | ESTATURAS (cm) | $f_i$ |
|---|---|---|
| 1 | 150 ⊢ 160 | 4 |
| 2 | 160 ⊢ 170 | 10 |
| 3 | 170 ⊢ 180 | 20 |
| 4 | 180 ⊢ 190 | 10 |
| 5 | 190 ⊢ 200 | 4 |
| | | $\Sigma = 48$ |

# EXEMPLO 1

$\overline{x} = 175$

$Md = 175$

$Mo = 175$

$S = 10{,}4$

$$\overline{x} - Mo = 175 - 175 = 0$$

# DISTRIBUIÇÃO SIMÉTRICA

# EXEMPLO 2

| i | ESTATURAS (cm) | $f_i$ |
|---|---|---|
| 1 | 150 ⊢ 160 | 4 |
| 2 | 160 ⊢ 170 | 10 |
| 3 | 170 ⊢ 180 | 20 |
| 4 | 180 ⊢ 190 | 40 |
| 5 | 190 ⊢ 200 | 4 |
| | | $\Sigma = 78$ |

$\bar{x}$ = 178,5

Md = 181,25

Mo = 185

S = 9,5

$\bar{x}$ – Mo = 178,5 – 185 = - 6,50

**DISTRIBUIÇÃO ASSIMÉTRICA NEGATIVA**

# EXEMPLO 3

| i | ESTATURAS (cm) | $f_i$ |
|---|---|---|
| 1 | 150 ⊢ 160 | 4 |
| 2 | 160 ⊢ 170 | 40 |
| 3 | 170 ⊢ 180 | 20 |
| 4 | 180 ⊢ 190 | 10 |
| 5 | 190 ⊢ 200 | 4 |
| | | $\Sigma = 78$ |

$\bar{x} = 171{,}15$

$Md = 168{,}75$

$Mo = 165$

$S = 9{,}5$

$$\bar{x} - Mo = 171{,}15 - 165 = +\,6{,}50$$

**DISTRIBUIÇÃO ASSIMÉTRICA POSITIVA**

## COEFICIENTE DE ASSIMETRIA (OU COEFICIENTE DE PEARSON)

$$A_v = \frac{3.(\bar{x} - Md)}{S}$$

EXEMPLO 4

| i | ESTATURAS (cm) | $f_i$ |
|---|---|---|
| 1 | 150 ⊢ 160 | 4 |
| 2 | 160 ⊢ 170 | 10 |
| 3 | 170 ⊢ 180 | 20 |
| 4 | 180 ⊢ 190 | 10 |
| 5 | 190 ⊢ 200 | 4 |
| | | $\Sigma = 48$ |

$\bar{x} = 175$

$Md = 175$

$Mo = 175$

$S = 10,4$

$$A_v = \frac{3.(\bar{x} - Md)}{S}$$

$$A_v = \frac{3.(175 - 175)}{10,4} = 0$$

**DISTRIBUIÇÃO SIMÉTRICA**

EXEMPLO 5

| i | ESTATURAS (cm) | $f_i$ |
|---|---|---|
| 1 | 150 ⊢ 160 | 4 |
| 2 | 160 ⊢ 170 | 10 |
| 3 | 170 ⊢ 180 | 20 |
| 4 | 180 ⊢ 190 | 40 |
| 5 | 190 ⊢ 200 | 4 |
| | | $\Sigma = 78$ |

$\bar{x} = 178{,}5$

$Md = 181{,}25$

$Mo = 185$

$S = 9{,}5$

$$A_v = \frac{3.(\bar{x} - Md)}{S}$$

$$A_v = \frac{3.(178{,}5 - 181{,}25)}{9{,}5} = -0{,}75$$

DISTRIBUIÇÃO ASSIMÉTRICA NEGATIVA

EXEMPLO 6

| i | ESTATURAS (cm) | $f_i$ |
|---|---|---|
| 1 | 150 ⊢ 160 | 4 |
| 2 | 160 ⊢ 170 | 40 |
| 3 | 170 ⊢ 180 | 20 |
| 4 | 180 ⊢ 190 | 10 |
| 5 | 190 ⊢ 200 | 4 |
| | | $\Sigma = 78$ |

$\bar{x} = 171{,}15$

$Md = 168{,}75$

$Mo = 165$

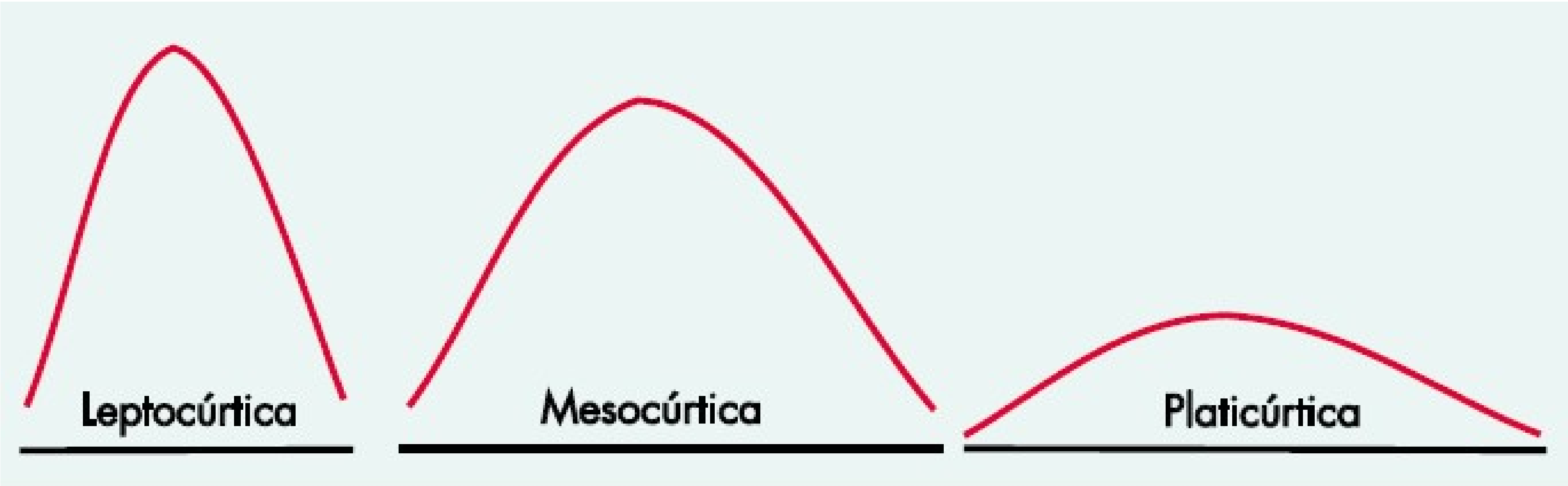
$S = 9{,}5$

$$A_v = \frac{3.(\bar{x} - Md)}{S}$$

$$A_v = \frac{3.(171{,}15 - 168{,}75)}{9{,}5} = +0{,}75$$

**DISTRIBUIÇÃO ASSIMÉTRICA POSITIVA**

# CURTOSE



# COEFICIENTE DE CURTOSE

$$C = \frac{Q_3 - Q_1}{2 \,.\, (P_{90} - P_{10})}$$

COEFICIENTE PERCENTÍLICO PARA CURVA NORMAL É DADO PELO VALOR 0,263

# INTERPRETAÇÃO

**$C = 0{,}263 \Rightarrow$ CURVA MESOCÚRTICA;**

**$C < 0{,}263 \Rightarrow$ CURVA LEPTOCÚRTICA;**

**$C > 0{,}263 \Rightarrow$ CURVA PLATICÚRTICA.**

# EXEMPLOS

| DISTRIBUIÇÕES | $Q_1$ | $Q_3$ | $P_{10}$ | $P_{90}$ |
|---|---|---|---|---|
| 1 | 28,8 | 45,6 | 20,5 | 49,8 |
| 2 | 814 | 935 | 772 | 1012 |
| 3 | 63,7 | 80,3 | 55 | 86,6 |

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

## DISTRIBUIÇÃO 1

$$C = \frac{Q_3 - Q_1}{2 \, . \, (P_{90} - P_{10})}$$

$$C = \frac{45{,}6 - 28{,}8}{2 \, . \, (49{,}8 - 20{,}5)}$$

$$C = 0{,}287 \text{ / PLATICÚRTICA}$$

# DISTRIBUIÇÃO 2

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

$$C = \frac{93,5 - 81,4}{2 . (101,2 - 77,2)}$$

$$C = 0,252 / \text{ LEPTOCÚRTICA}$$

## DISTRIBUIÇÃO 3

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

$$C = \frac{80,3 - 63,7}{2 \cdot (86,6 - 55)}$$

$$C = 0,263 \text{ / MESOCÚRTICA}$$

# OBSERVAÇÃO

**A ASSIMETRIA MOSTRA O QUANTO A POPULAÇÃO SE AFASTA DA SUA SIMETRIA, OU SEJA, É O GRAU DE DEFORMAÇÃO DE UMA CURVA DE FREQÜÊNCIA. JÁ A MEDIDA DE CURTOSE OU EXCESSO, INDICA ATÉ ONDE A CURVA DE FREQUÊNCIA DA DISTRIBUIÇÃO SE APRESENTA MAIS AFILADA OU MAIS ACHATADA EM COMPARAÇÃO COM UMA CURVA PADRÃO.**

# EXEMPLO 2

| PESOS | 50 ⊢ 58 | 58 ⊢ 66 | 66 ⊢ 74 | 74 ⊢ 82 | 82 ⊢ 90 | 90 ⊢ 98 |
|---|---|---|---|---|---|---|
| Nº DE OPERÁRIOS | 10 | 15 | 25 | 24 | 16 | 10 |

**DETERMINAR O GRAU DE ASSIMETRIA E CURTOSE DA DISTRIBUIÇÃO ACIMA.**

VERIFICAR O TIPO DE ASSIMETRIA DAS DISTRIBUIÇÕES A SEGUIR E DETERMIAR O COEFICIENTE DE CURTOSE:

1)

| CLASSES | $f_i$ |
|---|---|
| 10 ⊢ 20 | 5 |
| 20 ⊢ 30 | 10 |
| 30 ⊢ 40 | 15 |
| 40 ⊢ 50 | 20 |
| 50 ⊢ 60 | 5 |
| | $\Sigma = 55$ |

**2)**

| CLASSES | $f_i$ |
|---|---|
| 10 ⊢ 20 | 5 |
| 20 ⊢ 30 | 10 |
| 30 ⊢ 40 | 15 |
| 40 ⊢ 50 | 10 |
| 50 ⊢ 60 | 5 |
| | $\Sigma = 45$ |

3)

| CLASSES | $f_i$ |
|---|---|
| 10 ⊢ 20 | 5 |
| 20 ⊢ 30 | 20 |
| 30 ⊢ 40 | 15 |
| 40 ⊢ 50 | 10 |
| 50 ⊢ 60 | 5 |
| | $\Sigma = 55$ |

# Quando usar então?!!!

| QUANDO UTILIZAR MODA, MÉDIA OU MEDIANA | |
|---|---|
| Moda | Série é unimodal |
| Média | Variável é contínua |
| | Série não contém valores anômalos |
| | Série continha valores extremos incorretos que foram corrigidos |
| | Série continha valores extremos que não puderam ser corrigidos e que foram retirados do banco de dados |
| Mediana | Variável é discreta e $n$ é ímpar |
| | Série continha valores extremos que não foram retirados do banco de dados porque estavam corretos e/ou havia interação entre variáveis |

# CORRELAÇÃO

# E REGRESSÃO

# OBJETIVOS

- AVALIAR O GRAU DE DEPENDÊNCIA ENTRE DUAS VARIÁVEIS;

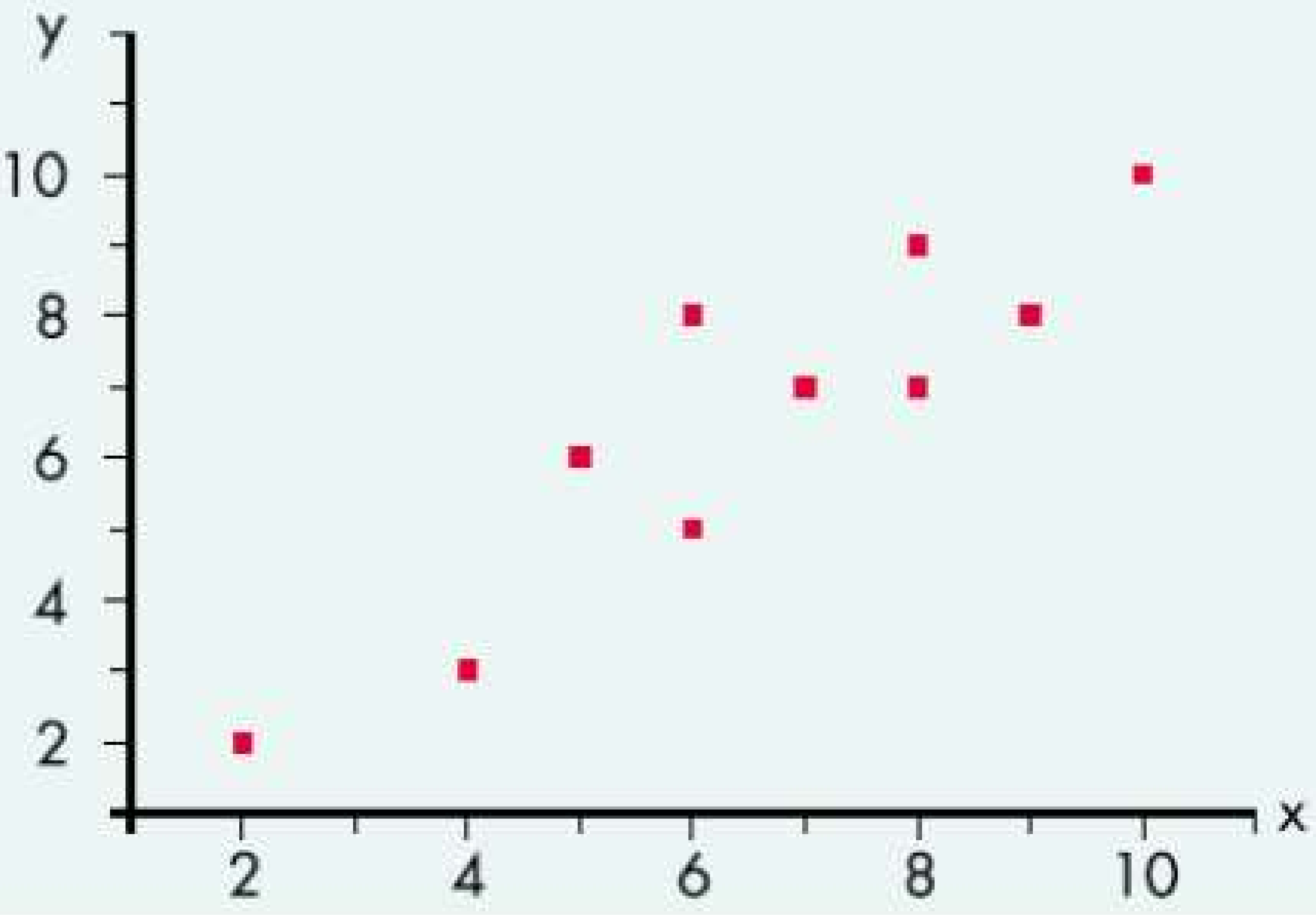
- ENCONTRAR A EQUAÇÃO DE CORRELAÇÃO ENTRE DUAS VARIÁVEIS.

A ANÁLISE DA CONTA DE ÁGUA, EM RESIDÊNCIAS URBANAS, ESTÁ INTIMAMENTE RELACIONADA AO CONSUMO DA MESMA EM M3; QUANTO MAIOR O CONSUMO, MAIOR A CONTA, FICANDO AQUI CARACTERIZADA A RELAÇÃO FUNCIONAL. AO RELACIONAR A INCIDÊNCIA DE UMA DETERMINADA DOENÇA EM UMA POPULAÇÃO, AQUI AS RELAÇÕES EM GERAL NÃO SÃO IMEDIATAS E PRECISAS; PORTANTO ESSA RELAÇÃO É CHAMADA DE RELAÇÃO ESTATÍSTICA; NESSE CASO, AS VARIÁVEIS ESTÃO ENVOLVIDAS EM UM PROCESSO DE CORRELAÇÃO.

# DIAGRAMA DE DISPERSÃO

**CONSIDEREMOS UMA AMOSTRA ALEATÓRIA DE MÉDIAS DE 10 ALUNOS DA EDUCAÇÃO A DISTÂNCIA DA UNITINS DAS DISCIPLINAS DE MATEMÁTICA E ESTATÍSTICA.**

| ALUNOS | MATEMÁTICA ( X ) | ESTATÍSTICA ( Y ) |
|---|---|---|
| 01 | 2,0 | 2,0 |
| 02 | 6,0 | 8,0 |
| 03 | 4,0 | 3,0 |
| 04 | 8,0 | 9,0 |
| 05 | 7,0 | 7,0 |
| 06 | 5,0 | 6,0 |
| 07 | 10,0 | 10,0 |
| 08 | 8,0 | 7,0 |
| 09 | 9,0 | 8,0 |
| 10 | 6,0 | 5,0 |

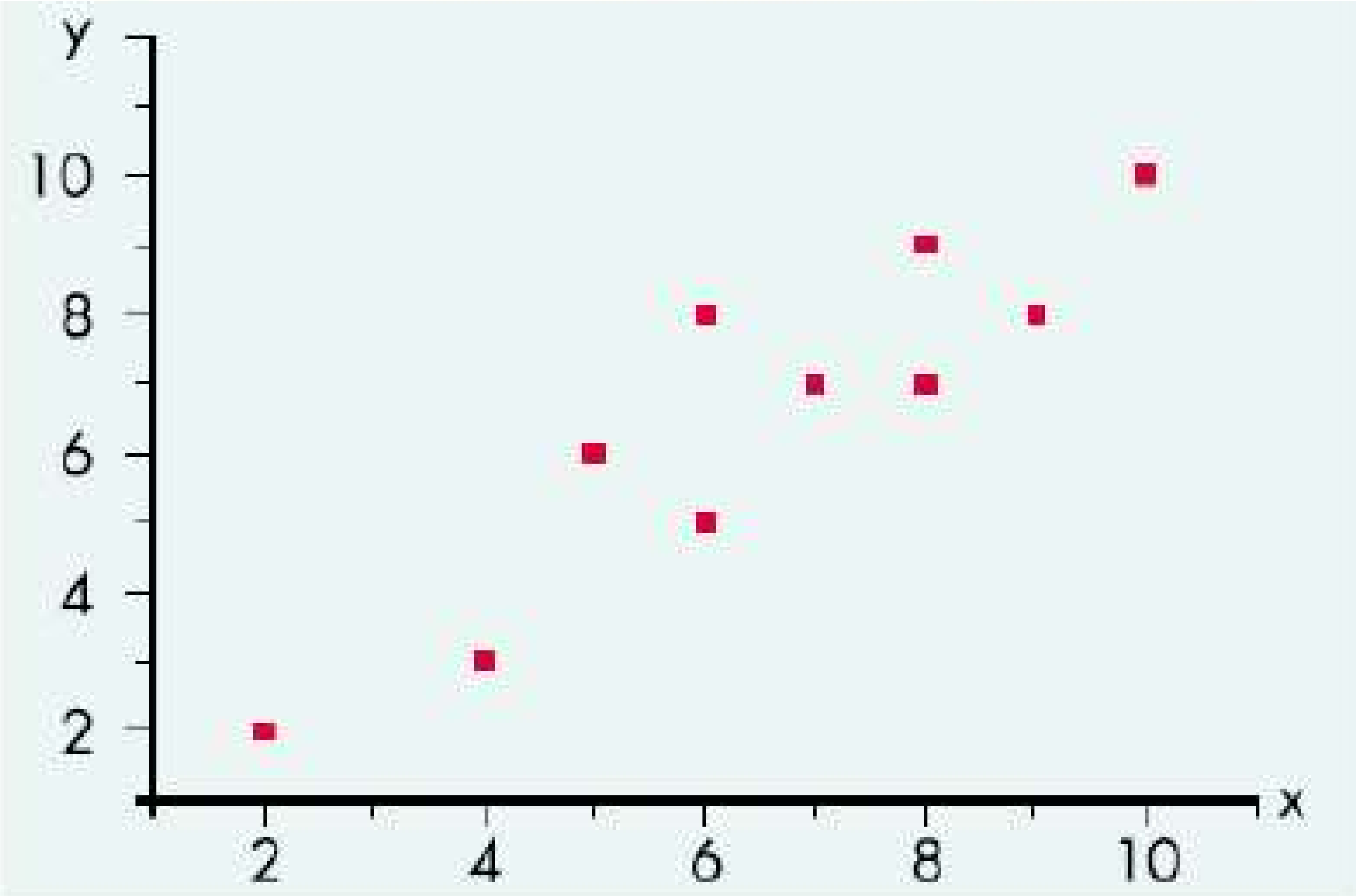

y
10
8
6
4
2
x
2
4
6
8
10

NOTEM QUE O DIAGRAMA MOSTRADO ANTERIORMENTE TRAZ A RELAÇÃO DAS NOTAS DAS DISCIPLINAS DE MATEMÁTICA (EIXO X) E FÍSICA (EIXO Y), ONDE CADA PONTO (PAR ORDENADO) REPRESENTA A COMBINAÇÃO DAS NOTAS DE CADA ALUNO.

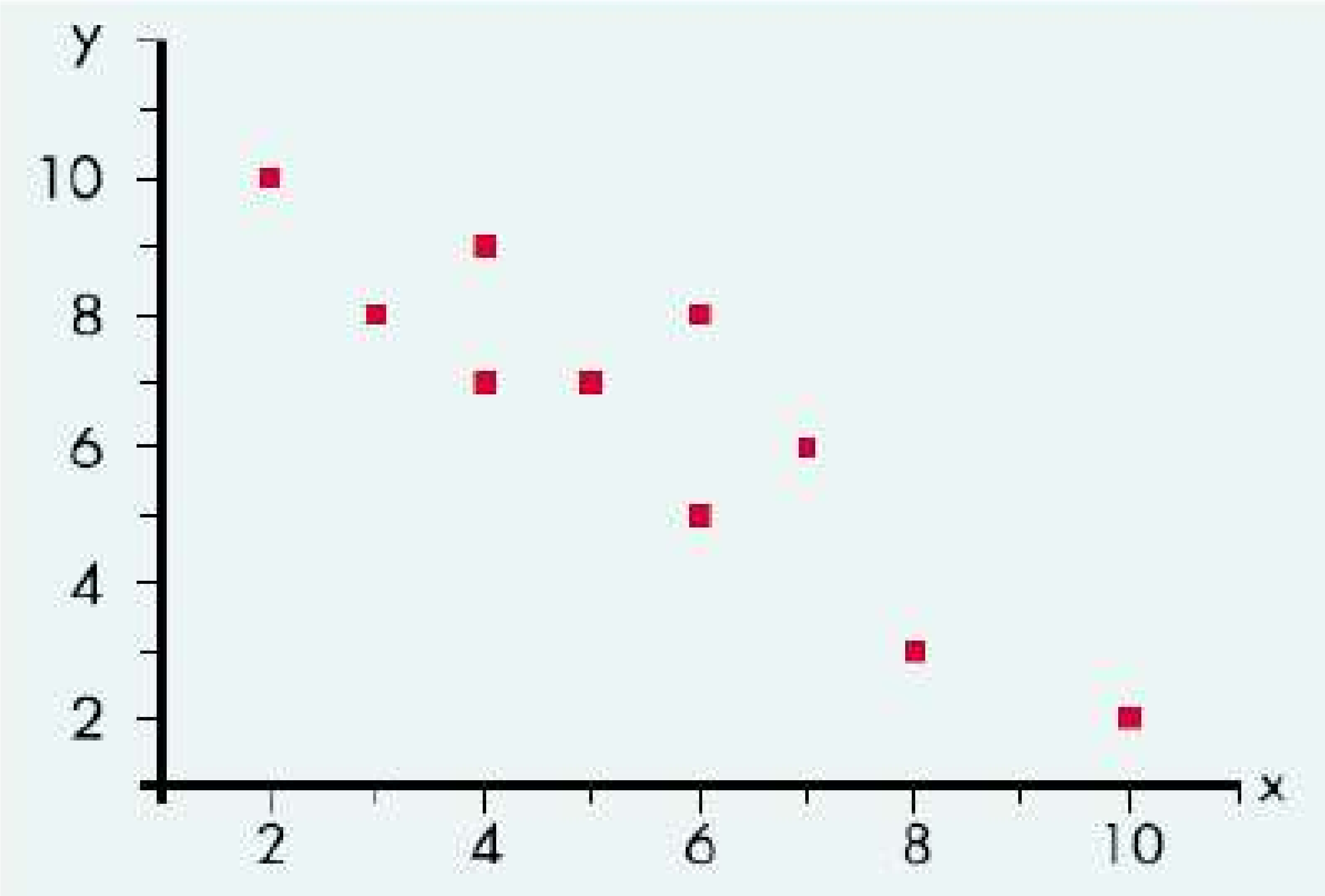
# CORRELAÇÃO LINEAR

AO LOCALIZARMOS TODOS OS PONTOS NO PLANO CARTESIANO, OS PONTOS ESTARÃO DISPERSOS EM UMA FORMA ALEATÓRIA; PODEREMOS CHAMAR ESSA CORRELAÇÃO DE UMA CORRELAÇÃO ALEATÓRIA OU QUALQUER. POR UM MÉTODO DE AJUSTE DE CURVA VAMOS SUPOR QUE ESSA DISTRIBUIÇÃO DOS PONTOS SE APROXIMA DE UMA RETA E ENTÃO A CHAMAREMOS DE CORRELAÇÃO LINEAR.

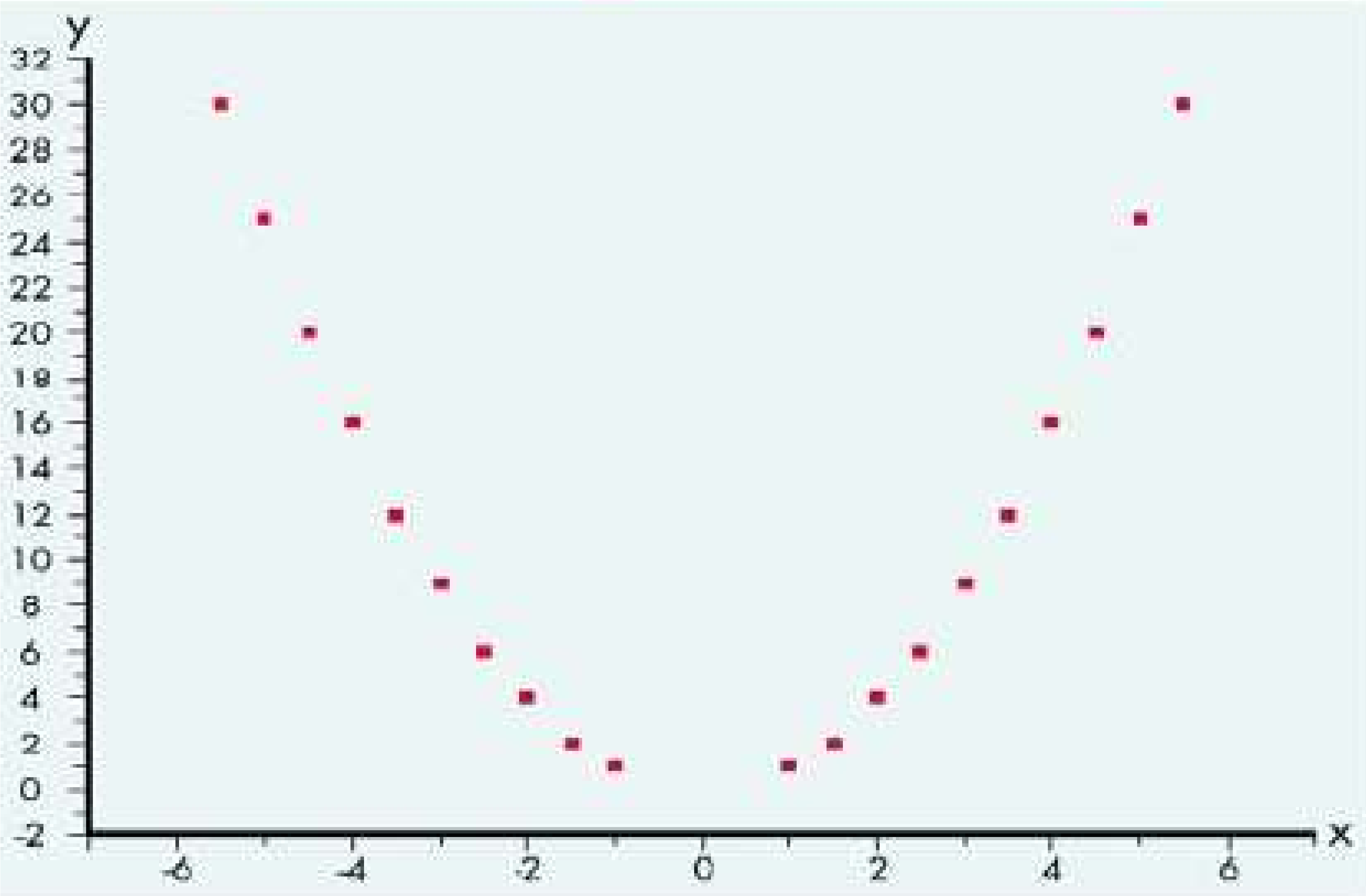
# CORRELAÇÃO LINEAR POSITIVA



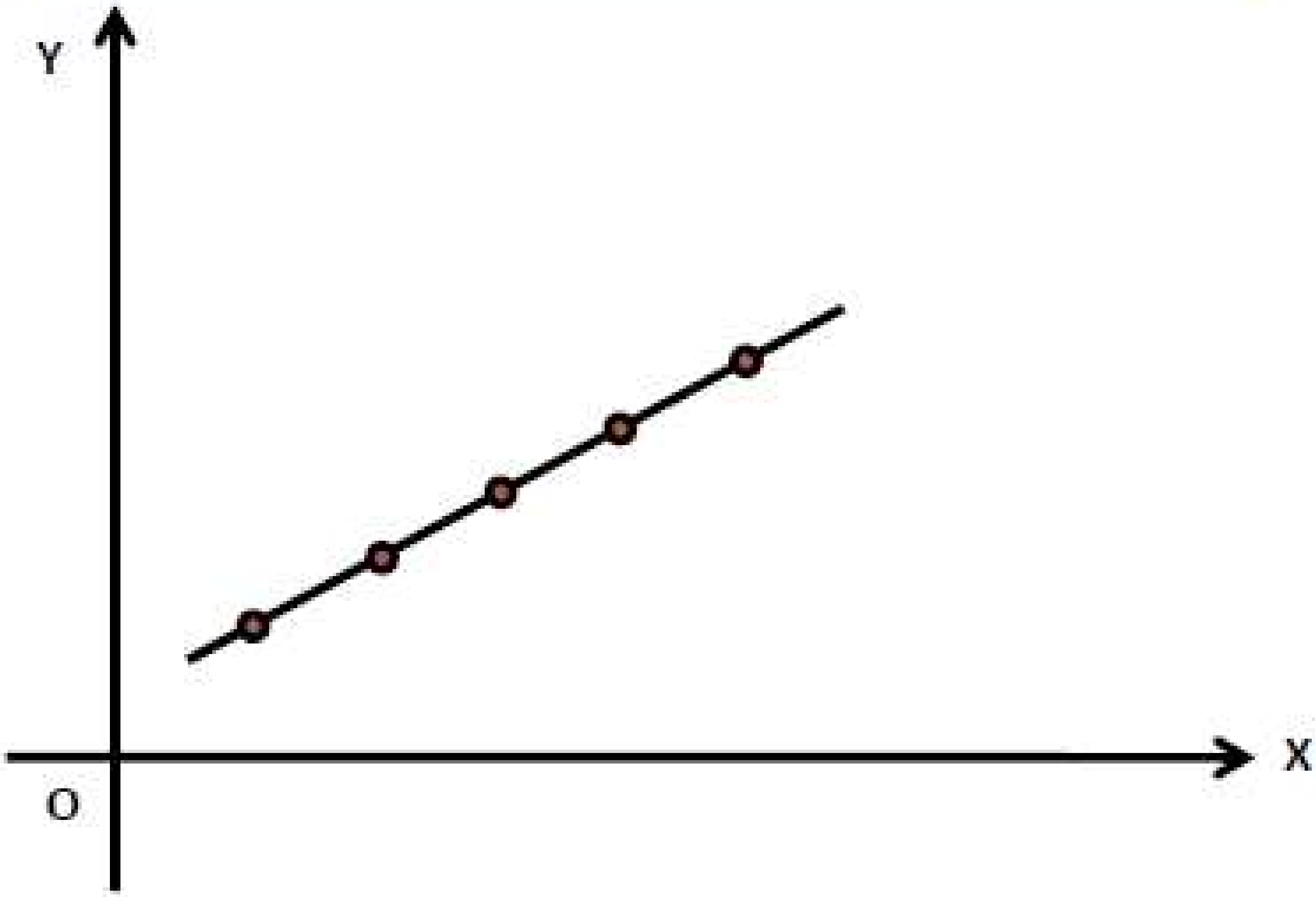
# CORRELAÇÃO LINEAR NEGATIVA



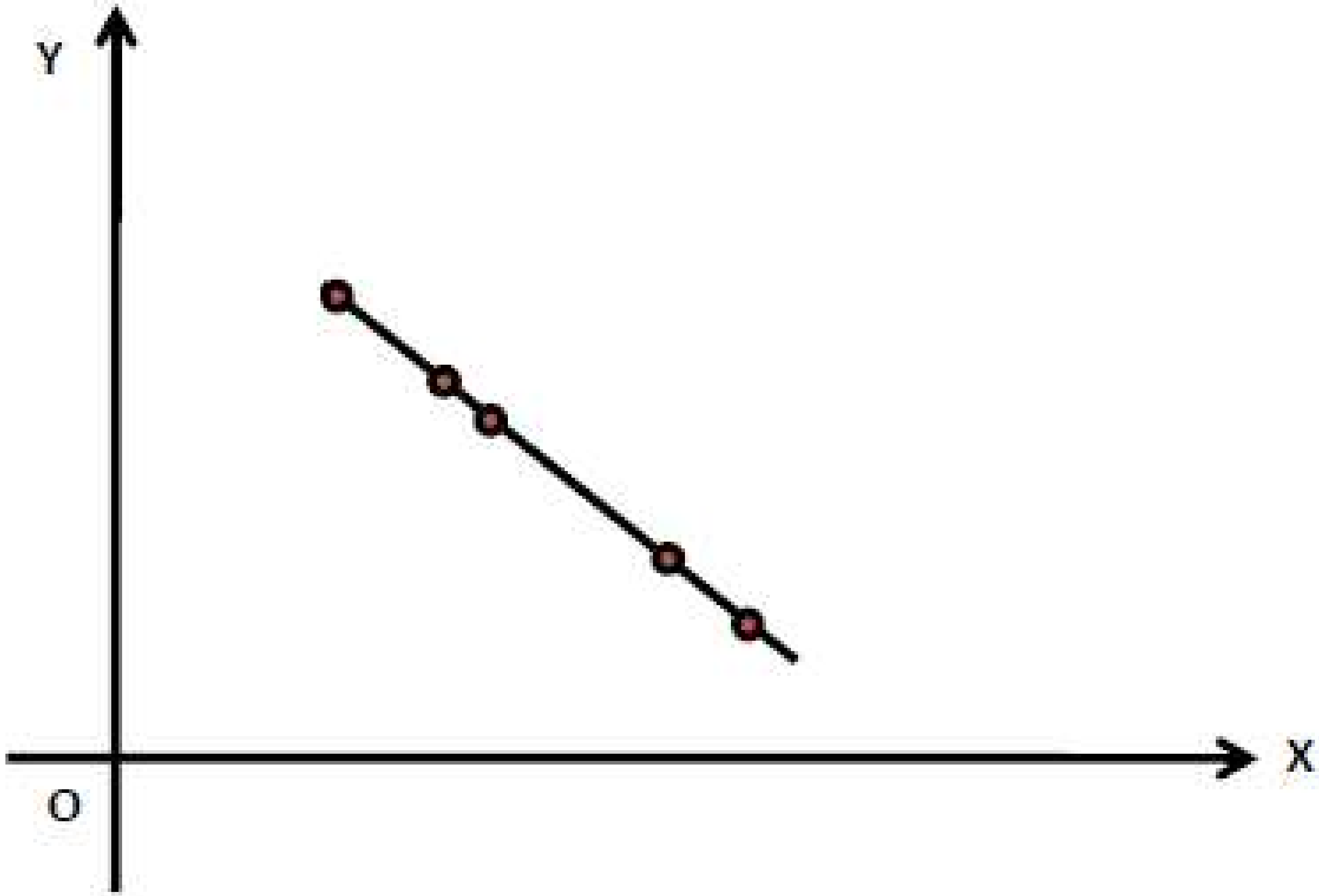
# CORRELAÇÃO NÃO-LINEAR



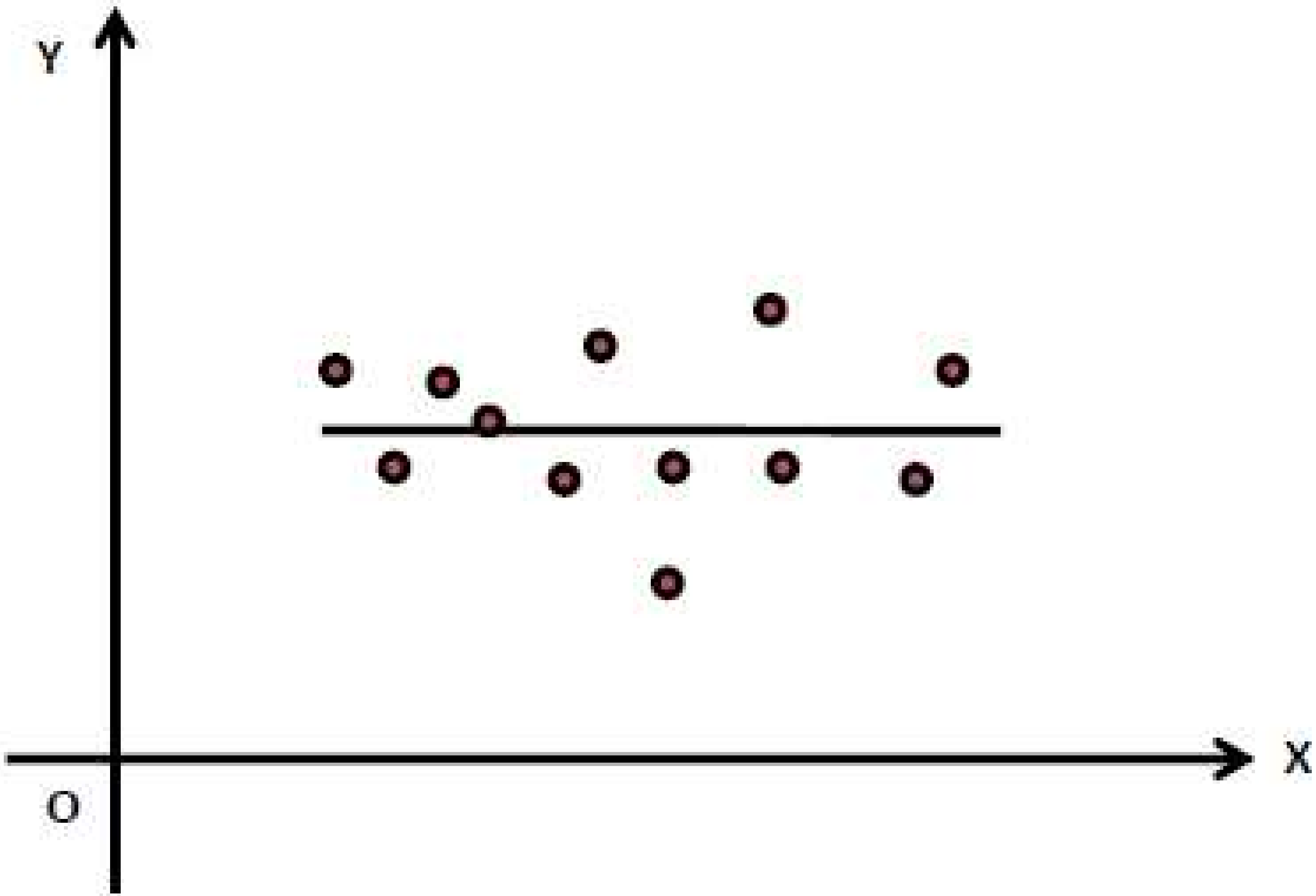
# CORRELAÇÃO PERFEITA POSITIVA



# CORRELAÇÃO PERFEITA NEGATIVA



# CORRELAÇÃO NULA



## COEFICIENTE DE CORRELAÇÃO LINEAR (CORRELAÇÃO DE PEARSON)

ESTE COEFICIENTE INDICA O GRAU DE INTENSIDADE DA CORRELAÇÃO ENTRE AS DUAS VARIÁVEIS E SE É POSITIVA OU NEGATIVA.

$$r = \frac{n \cdot \sum x \cdot y - \left(\sum x\right) \cdot \left(\sum y\right)}{\sqrt{\left[n \cdot \sum x^2 - \left(\sum x\right)^2\right]\left[n \cdot \sum y^2 - \left(\sum y\right)^2\right]}}$$

# CLASSIFICAÇÃO

SE $r = +1$, CORRELAÇÃO PERFEITA E POSITIVA.

SE $r = -1$, CORRELAÇÃO PERFEITA E NEGATIVA.

SE $r = 0$, NÃO HÁ CORRELAÇÃO OU A CORRELAÇÃO NÃO É LINEAR.

$$0,6 \leq | r | \leq 1$$

| ALUNOS | MAT ( X ) | ESTAT ( Y ) | X . Y | $X^2$ | $Y^2$ |
|---|---|---|---|---|---|
| 01 | 2 | 2 | 4 | 4 | 4 |
| 02 | 6 | 8 | 48 | 36 | 64 |
| 03 | 4 | 3 | 12 | 16 | 9 |
| 04 | 8 | 9 | 72 | 64 | 81 |
| 05 | 7 | 7 | 49 | 49 | 49 |
| 06 | 5 | 6 | 30 | 25 | 36 |
| 07 | 10 | 10 | 100 | 100 | 100 |
| 08 | 8 | 7 | 56 | 64 | 49 |
| 09 | 9 | 8 | 72 | 81 | 64 |
| 10 | 6 | 5 | 30 | 36 | 25 |
| | $\sum = 65$ | $\sum = 65$ | $\sum = 473$ | $\sum = 475$ | $\sum = 481$ |

# CÁLCULO DE “r”

$$r = \frac{n \cdot \sum x \cdot y - \left(\sum x\right) \cdot \left(\sum y\right)}{\sqrt{[n \cdot \sum x^2 - (\sum x)^2][n \cdot \sum y^2 - (\sum y)^2]}}$$

$$r = \frac{10 \cdot 473 - 65 \cdot 65}{\sqrt{[10 \cdot 475 - (65)^2] \cdot [10 \cdot 481 - (65)^2]}}$$

$r = 0{,}91$

## CORRELAÇÃO LINEAR POSITIVA

Níveis de correlação

| *Valor* | *Correlação* |
|---|---|
| R = 0 | nula |
| 0 < \|R\| < 0.30 | fraca |
| 0.30 < \|R\| < 0.60 | média |
| 0.60 < \|R\| < 0.90 | forte |
| 0.90 < \|R\| < 1 | fortíssima |
| \|R\|= 1 | perfeita |

CALCULAR O COEFICIENTE DE CORRELAÇÃO LINEAR ENTRE RENDA FAMILIAR E A POUPANÇA DA DISTRIBUIÇÃO.

A TABELA A SEGUIR MOSTRA OS RESULTADOS DE UMA PESQUIS COM 10 FAMÍLIAS DE UMA DETERM INADA REGIÃO

| FAMÍLIAS | RENDA x 100 | POUPANÇA x 1.000 |
|---|---|---|
| A | 10 | 4 |
| B | 15 | 7 |
| C | 12 | 5 |
| D | 70 | 20 |
| E | 80 | 20 |
| F | 100 | 30 |
| G | 20 | 8 |
| H | 830 | 8 |
| I | 10 | 3 |
| J | 60 | 15 |
| | $\sum = 1.207$ | $\sum = 120$ |

# REGRESSÃO

UMA ANÁLISE DE REGRESSÃO TEM COMO OBJETIVO DESCREVER, COM O USO DE UM MODELO MATEMÁTICO, A RELAÇÃO ENTRE DUAS VARIÁVEIS COM N OBSERVAÇÕES DAS MESMAS. ESSE MODELO MATEMÁTICO PODERÁ NOS DIZER, PARA UM DADO VALOR DE UMA VARIÁVEL, QUAL O POSSÍVEL VALOR DA OUTRA VARIÁVEL E VICE-VERSA.

ESSA ANÁLISE APRESENTA DUAS VARIÁVEIS; A VARIÁVEL EM ESTUDO É A VARIÁVEL DEPENDENTE, E A OUTRA É A VARIÁVEL INDEPENDENTE. A RELAÇÃO MATEMÁTICA QUE ENVOLVE AS DUAS VARIÁVEIS É DADA PELA EQUAÇÃO DO PRIMEIRO GRAU A SEGUIR:

$$y = ax + b$$

**ONDE:**

$$a = \frac{n \cdot \sum x \cdot y - \left(\sum x\right) \cdot \left(\sum y\right)}{n \cdot \sum x^2 - \left(\sum x\right)^2}$$

$$b = \bar{y} - a\bar{x}$$

$$\bar{x} = \frac{\sum x}{n} \quad e \quad \bar{y} = \frac{\sum y}{n}$$

| ALUNOS | MAT ( X ) | ESTAT ( Y ) | X . Y | $X^2$ |
|---|---|---|---|---|
| 01 | 2 | 2 | 4 | 4 |
| 02 | 6 | 8 | 48 | 36 |
| 03 | 4 | 3 | 12 | 16 |
| 04 | 8 | 9 | 72 | 64 |
| 05 | 7 | 7 | 49 | 49 |
| 06 | 5 | 6 | 30 | 25 |
| 07 | 10 | 10 | 100 | 100 |
| 08 | 8 | 7 | 56 | 64 |
| 09 | 9 | 8 | 72 | 81 |
| 10 | 6 | 5 | 30 | 36 |
| | $\sum = 65$ | $\sum = 65$ | $\sum = 473$ | $\sum = 475$ |

SOLUÇÃO

$$a = \frac{n.\sum x.y - \left(\sum x\right).\left(\sum y\right)}{n.\sum x^2 - \left(\sum x\right)^2}$$

$$a = \frac{10\ .\ 473 - (65\ .\ 65)}{10\ .475 - 65^2}$$

$$a = 0{,}96$$

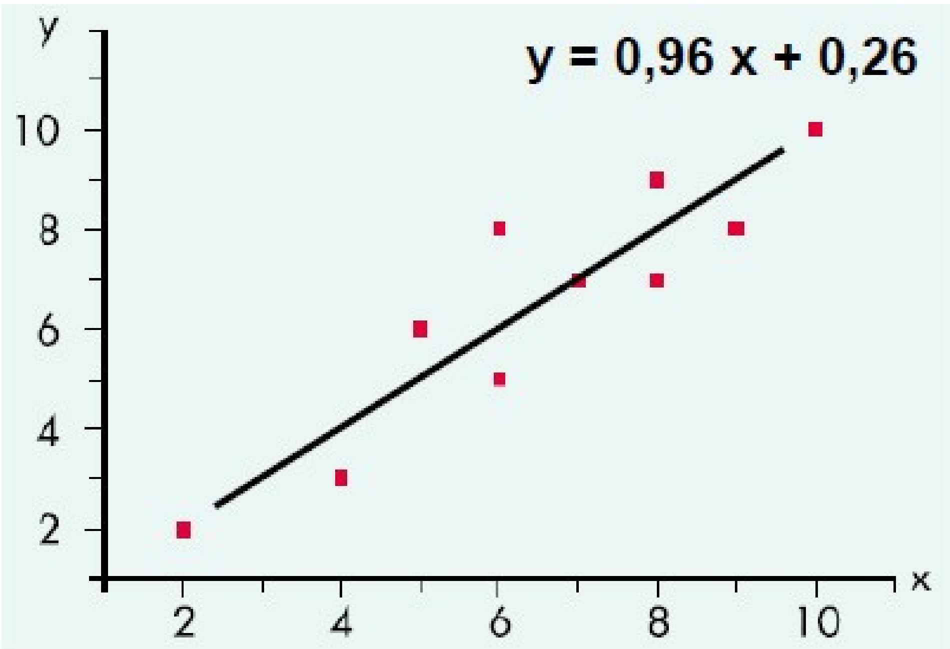
SOLUÇÃO

$$\overline{x} = \frac{\sum X}{n} = \frac{65}{10} = 6{,}5$$

$$\overline{y} = \frac{\sum y}{n} = \frac{65}{10} = 6{,}5$$

$$b = \overline{y} - a\overline{x} \Leftrightarrow b = 6{,}5 - 0{,}96.6{,}5$$

$$b = 0{,}26$$



## 1. ELABORAR O DIAGRAMA DE DISPERSÃO

## 2. DETERMINAR "a" E "b"

$$a = \frac{n . \sum x . y - \left(\sum x\right) . \left(\sum y\right)}{n . \sum x^2 - \left(\sum x\right)^2}$$

$$\overline{x} = \frac{\sum x}{n} \quad e \quad \overline{y} = \frac{\sum y}{n}$$

$$b = \overline{y} - a\overline{x}$$

# EXEMPLOS

CERTA EMPRESA, ESTUDANDO A VARIAÇÃO DA DEMANDA DE SEU PRODUTO EM RELAÇÃO À VARIAÇÃO DE PREÇO DE VENDA, OBTEVE A TABELA A SEGUIR:

| PREÇO | 38 | 42 | 50 | 56 | 59 | 63 | 70 | 80 | 95 | 110 |
|---|---|---|---|---|---|---|---|---|---|---|
| DEMANDA | 350 | 325 | 297 | 270 | 256 | 246 | 238 | 223 | 215 | 208 |

DETERMINE O COEFICIENTE DE CORRELAÇÃO "r" E DETERMINE OS COEFICIENTES "a" E "b" DA RETA.

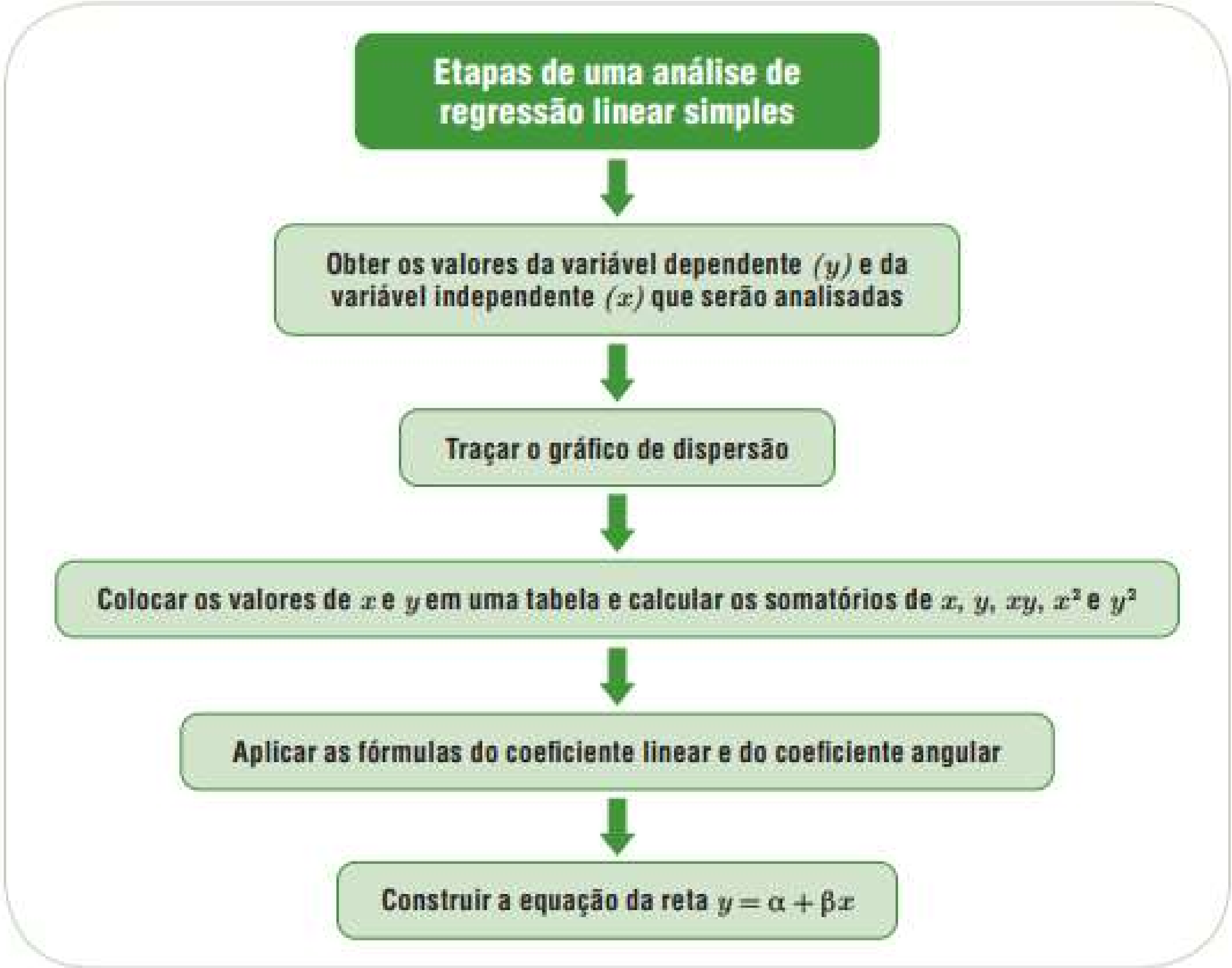



**Figura 2 –** Etapas para a realização de uma análise de regressão linear simples