

Estatística

**Prof. Antonio Estanislau Sanches
2018**

ESTATÍSTICA: ramo da matemática aplicada que fornece **métodos** para coleta, organização, descrição, análise e interpretação de dados e para sua utilização na tomada de decisões.

MÉTODO: é um meio mais eficaz para atingir determinada meta.

MÉTODOS CIENTÍFICOS: método experimental e o método estatístico.

MÉTODO EXPERIMENTAL: consiste em manter constante todas as causas, menos uma, que sofre variação para se observar seus efeitos, caso existam. Ex: Estudos da Química, Física, etc.

MÉTODO ESTATÍSTICO: diante da impossibilidade de manter as causas constantes (*nas ciências sociais*), admitem todas essas causas presentes variando-as, registrando essas variações e procurando determinar, no resultado final, que influências cabem a cada uma delas. Ex: Quais as causas que definem o preço de uma mercadoria quando a sua oferta diminui?

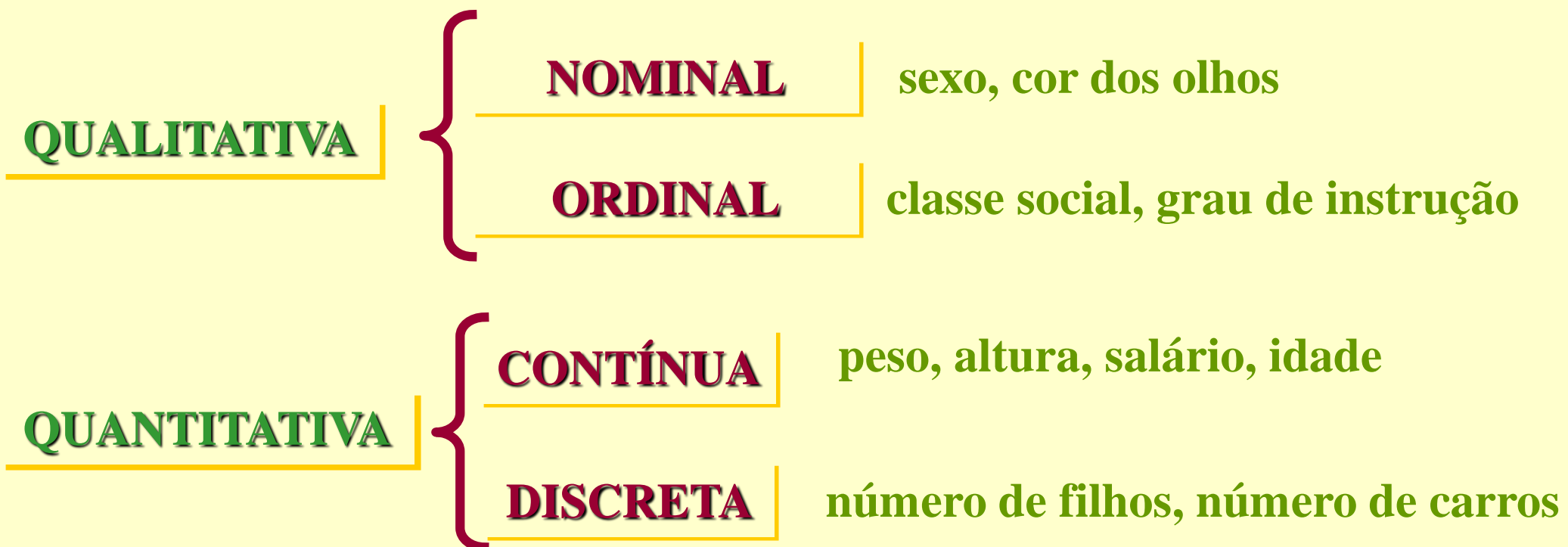
ESTATÍSTICA DESCRITIVA compreende a coleta, organização, descrição dos dados, cálculo e a interpretação de coeficientes.

ESTATÍSTICA INDUTIVA ou INFERENCIAL disponibiliza a análise e a interpretação dos dados, associados a uma margem de incerteza, também chamada como a medida da incerteza, baseada em métodos fundamentados na teoria da probabilidade.

Variável:

Qualquer característica associada a uma população.

Classificação das variáveis:



Classificação das variáveis

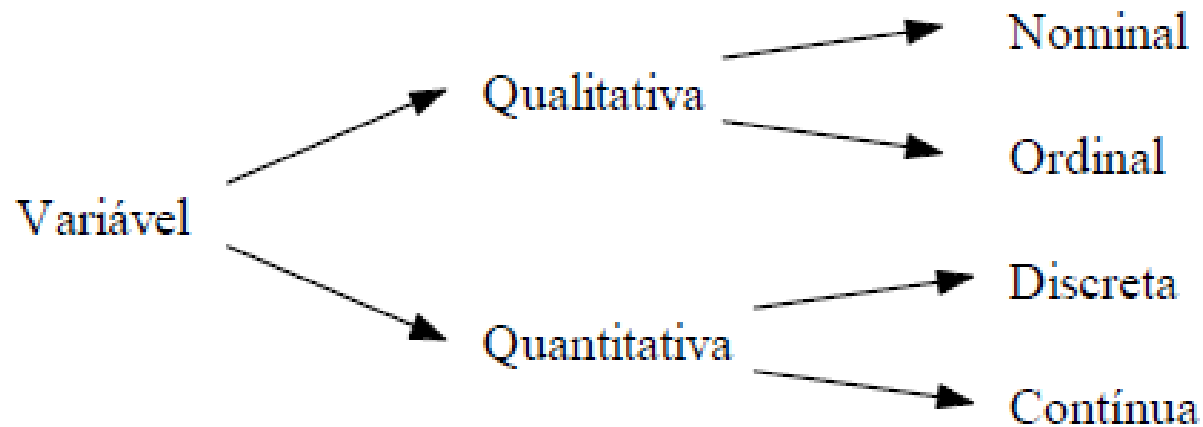
- Qualitativa ou atributos/ Cor, beleza
- Quantitativa discreta/ Número de alunos
- Quantitativa contínua/ Altura, peso

Medições \Rightarrow Var. Contínuas

Contagens \Rightarrow Var. Discretas

Tipos de variáveis

Os dados coletados em uma primeira fase podem ser definidos como variáveis qualitativas ou quantitativas, de acordo com a seguinte figura:





Classificação das variáveis por nível de mensuração



Classificação das variáveis por nível de manipulação

Ex.: Para uma população de peças produzidas em um determinado processo, poderíamos ter:

Variável

Tipo

Estado: Perfeita ou defeituosa

Qualidade: 1^a, 2^a ou 3^a categoria

N^o de peças defeituosas

Diâmetro das peças

Ex.: Para uma população de peças produzidas em um determinado processo, poderíamos ter:

| <i>Variável</i> | <i>Tipo</i> |
|--|-----------------------|
| <i>Estado</i> : Perfeita ou defeituosa | Qualitativa Nominal |
| <i>Qualidade</i> : 1 ^a , 2 ^a ou 3 ^a categoria | Qualitativa Ordinal |
| <i>Nº de peças defeituosas</i> | Quantitativa Discreta |
| <i>Diâmetro das peças</i> | Quantitativa Contínua |

EXERCÍCIO - Classifique as variáveis em qualitativas ou quantitativas (contínuas ou discretas):

- . Cor dos olhos das alunas...
- . Índice liquidez nas indústrias PIM...
- . Produção de café no Brasil...
- . Número defeitos em aparelhos de TV...
- . Comprimento dos pregos produzidos por uma empresa...
- . Valor de cada jogada de um dado...

EXERCÍCIO - Classifique as variáveis em qualitativas ou quantitativas (contínuas ou discretas):

- . Cor dos olhos das alunas... **Resp: qualitativa nominal**
- . Índice liquidez nas indústrias PIM...
- . Produção de café no Brasil...
- . Número defeitos em aparelhos de TV...
- . Comprimento dos pregos produzidos por uma empresa...
- . Valor de cada jogada de um dado...

EXERCÍCIO - Classifique as variáveis em qualitativas ou quantitativas (contínuas ou discretas):

- . Cor dos olhos das alunas... **Resp:qualitativa nominal**
- . Índice liquidez nas industrias PIM... **Resp:quantitativa contínua**
- . Produção de café no Brasil...
- . Número defeitos em aparelhos de TV...
- . Comprimento dos pregos produzidos por uma empresa...
- . Valor de cada jogada de um dado...

EXERCÍCIO - Classifique as variáveis em qualitativas ou quantitativas (contínuas ou discretas):

- . Cor dos olhos das alunas... **Resp:qualitativa nominal**
- . Índice liquidez nas industrias PIM... **Resp:quantitativa contínua**
- . Produção de café no Brasil... **Resp:quantitativa contínua**
- . Número defeitos em aparelhos de TV...
- . Comprimento dos pregos produzidos por uma empresa...
- . Valor de cada jogada de um dado...

EXERCÍCIO - Classifique as variáveis em qualitativas ou quantitativas (contínuas ou discretas):

- . Cor dos olhos das alunas... **Resp:qualitativa nominal**
- . Índice liquidez nas industrias PIM... **Resp:quantitativa contínua**
- . Produção de café no Brasil... **Resp:quantitativa contínua**
- . Número defeitos em aparelhos de TV... **Resp:quantitativa discreta**
- . Comprimento dos pregos produzidos por uma empresa...
- . Valor de cada jogada de um dado...

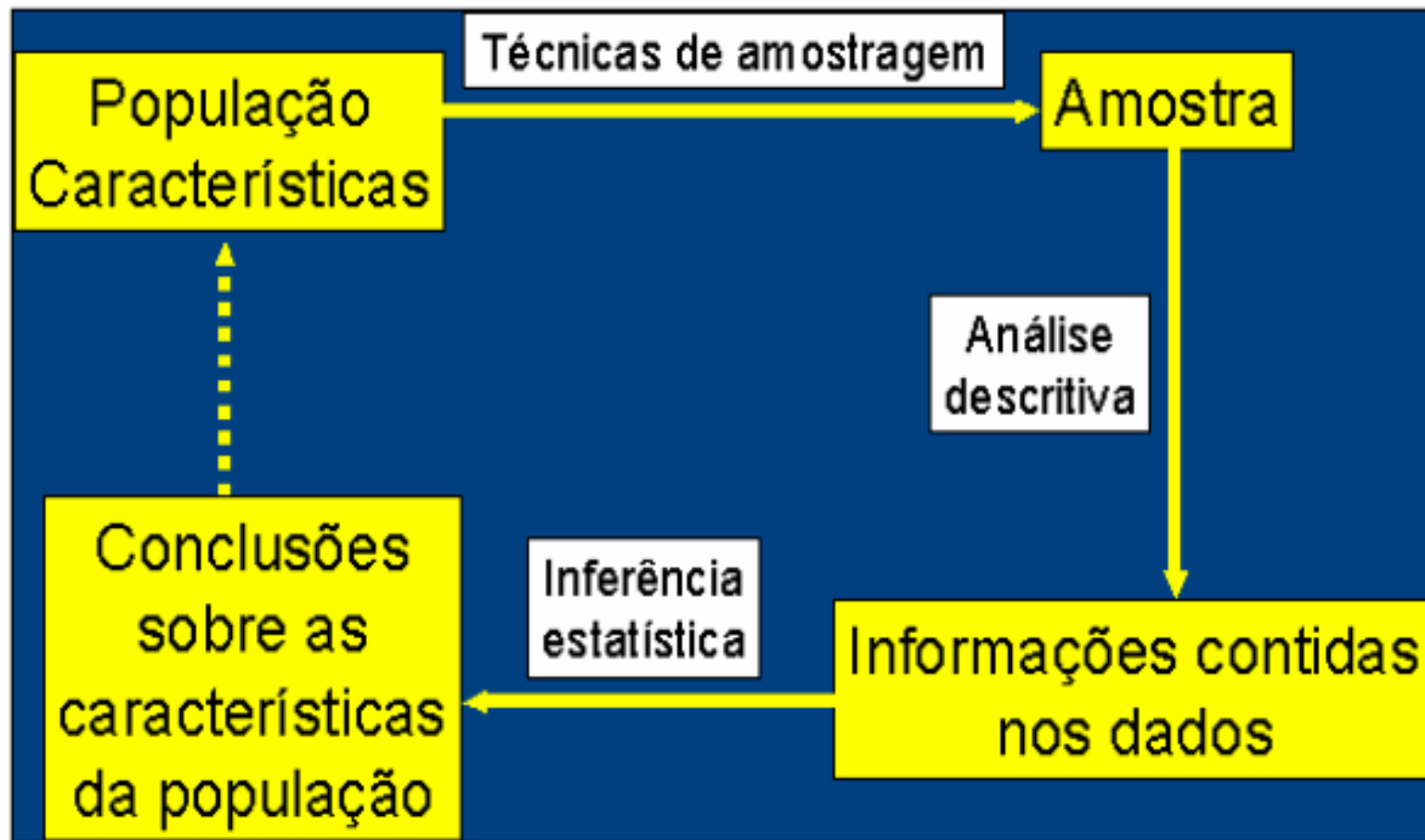
EXERCÍCIO - Classifique as variáveis em qualitativas ou quantitativas (contínuas ou discretas):

- . Cor dos olhos das alunas... **Resp:qualitativa nominal**
- . Índice liquidez nas indústrias PIM... **Resp:quantitativa contínua**
- . Produção de café no Brasil... **Resp:quantitativa contínua**
- . Número defeitos em aparelhos de TV... **Resp:quantitativa discreta**
- . Comprimento dos pregos produzidos por uma empresa...
Resp:quantitativa contínua
- . Valor de cada jogada de um dado...

EXERCÍCIO - Classifique as variáveis em qualitativas ou quantitativas (contínuas ou discretas):

- . Cor dos olhos das alunas... **Resp:qualitativa nominal**
- . Índice liquidez nas indústrias PIM... **Resp:quantitativa contínua**
- . Produção de café no Brasil... **Resp:quantitativa contínua**
- . Número defeitos em aparelhos de TV... **Resp:quantitativa discreta**
- . Comprimento dos pregos produzidos por uma empresa...
Resp:quantitativa contínua
- . Valor de cada jogada de um dado... **Resp:quantitativa discreta**

Etapas da Análise Estatística

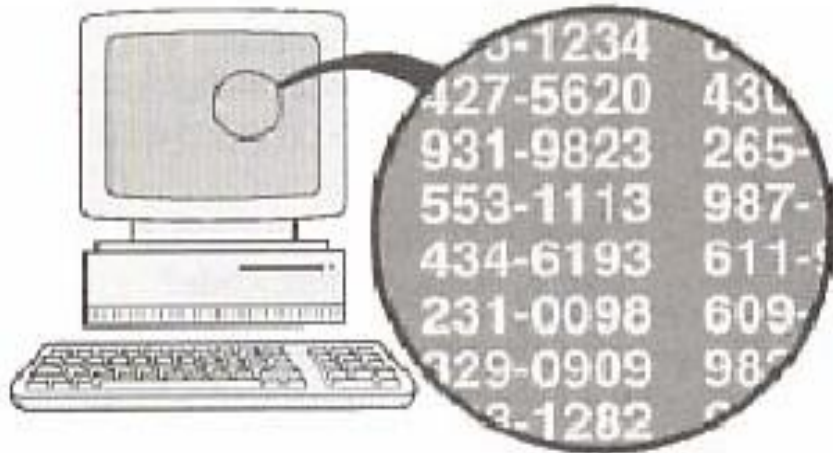


AMOSTRAGEM

Uma área importante em muitas aplicações Estatísticas é a da **Tecnologia de Amostragem**.

Exemplos de Aplicação:

- Pesquisa de mercado,
- Pesquisa de opinião,
- Avaliação do processo de produção,
- Praticamente em todo experimento.



Amostragem Aleatória

Cada elemento da população tem a mesma chance de ser escolhido.



Amostragem Estratificada

Classificar a população em, ao menos dois estratos e extrair uma amostra de cada um.



Amostragem Sistemática

Escolher cada elemento de ordem k .

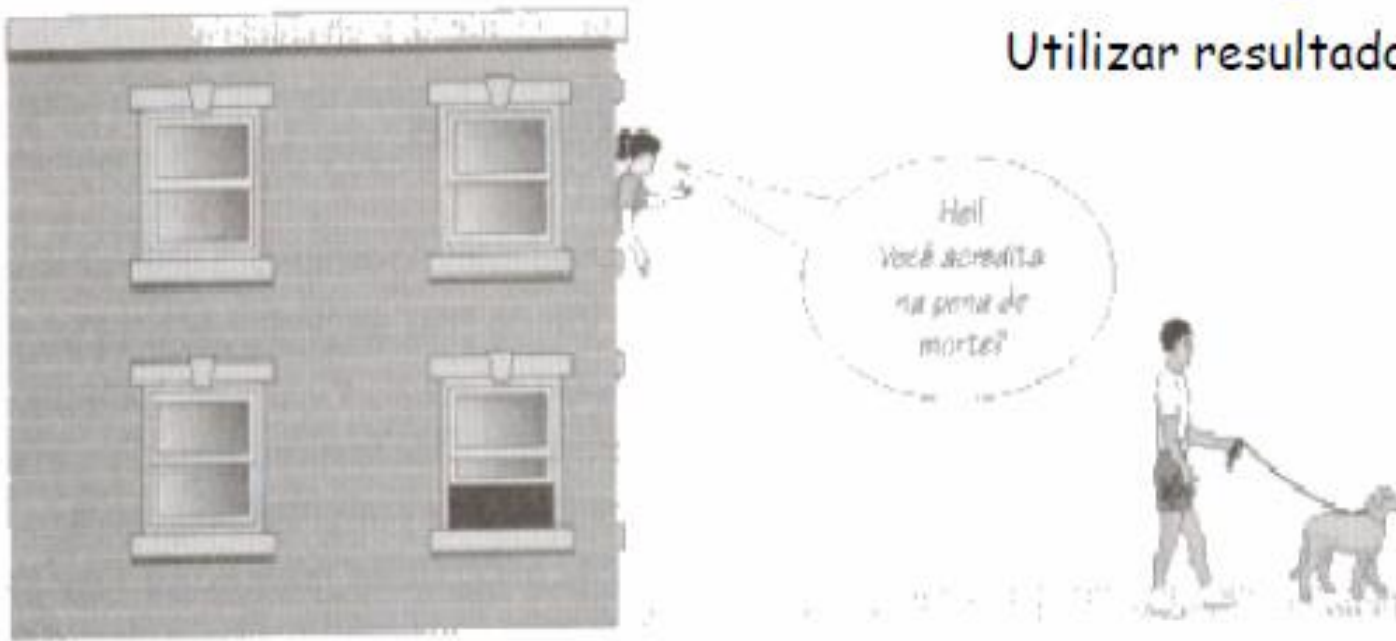
Amostragem por Conglomerados

Dividir em seções a área populacional, selecionar aleatoriamente algumas dessas seções e tomar todos os elementos das mesmas.



Amostragem de Conveniência

Utilizar resultados de fácil acesso.



Conceitos Básicos

População x Amostra

Censo x Amostragem

Dado x Variável

Parâmetros x Estatísticas

| | | | | |
|------------|------|----------------------|------|-----------|
| μ | ———— | Média | ———— | \bar{X} |
| σ^2 | ———— | Variância | ———— | S^2 |
| σ | — | Desvio-padrão | — | S |

Estatística Descritiva

Variáveis Quantitativas

MEDIDAS DE POSIÇÃO:

Mínimo, Máximo, Moda, Média, Mediana, Percentis

MEDIDAS DE DISPERSÃO:

Amplitude, Intervalo-Interquartil, Variância, Desvio Padrão, Coeficiente de Variação.

Medidas de Posição

- **Máximo (max)**: a maior observação
- **Mínimo (min)**: a menor observação
- **Moda (mo)**: é o valor (ou atributo) que ocorre com maior frequência.

Dados: 4, 5, 4, 6, 5, 8, 4

max = ?

min = ?

mo = ?

Medidas de Posição

- **Máximo (max)**: a maior observação
- **Mínimo (min)**: a menor observação
- **Moda (mo)**: é o valor (ou atributo) que ocorre com maior frequência.

Dados: 4, 5, 4, 6, 5, 8, 4

$$\text{max} = 8$$

$$\text{min} = 4$$

$$\text{mo} = 4$$

•Média:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

Dados: 2, 5, 3, 7, 8

$$\bar{x} = ?$$

• Média:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

Dados: 2, 5, 3, 7, 8

$$\bar{x} = \frac{2 + 5 + 3 + 7 + 8}{5} = 5$$

- **Mediana:**

A mediana é o valor da variável que ocupa a **posição central** de um conjunto de n **dados ordenados**.

Posição da mediana: $\frac{n+1}{2}$

Exemplos:

Dados: 2, 6, 3, 7, 8 $\Rightarrow n = 5$ (ímpar)

Dados ordenados: 2 3 6 7 8

Dados: 4, 8, 2, 1, 9, 6 $\Rightarrow n = 6$ (par)

Dados ordenados: 1 2 4 6 8 9

Posição da mediana: $\frac{n+1}{2}$

Exemplos:

Dados: 2, 6, 3, 7, 8 $\Rightarrow n = 5$ (ímpar)

Dados ordenados: 2 3 6 7 8 $\Rightarrow \frac{5+1}{2} = 3 \Rightarrow Md = 6$

Posição da Mediana \uparrow
3º lugar

Dados: 4, 8, 2, 1, 9, 6 $\Rightarrow n = 6$ (par)

Dados ordenados: 1 2 4 6 8 9

Exemplos:

Dados: 2, 6, 3, 7, 8 $\Rightarrow n = 5$ (ímpar)

Dados ordenados: 2 3 6 7 8 $\Rightarrow \frac{5+1}{2} = 3 \Rightarrow Md = 6$

Posição da Mediana \uparrow
3º lugar

Dados: 4, 8, 2, 1, 9, 6 $\Rightarrow n = 6$ (par)

Dados ordenados: 1 2 4 6 8 9 $\Rightarrow \frac{6+1}{2} = 3,5$

Posição da Mediana $\uparrow Md \uparrow$
3º e 4º lugares

$$Md = (4 + 6) / 2 \Rightarrow Md = 5$$

MODA (DADOS ISOLADOS)

- **Moda:**

(MODA DE UM CONJUNTO DE DADOS (=VALORES)) = (VALOR DO CONJUNTO QUE APARECE MAIS VEZES, ISTO É, O VALOR AO QUAL ESTEJA ASSOCIADA A FREQUÊNCIA ABSOLUTA MAIS ALTA.)

Calcular a moda de:

8, 2, 18, 8, 10, 8, 12, 10, 6, 8, 12

CHAMANDO A MODA DE M_o , A VARIÁVEL DE X E AS FREQUÊNCIAS DE n_i , VEM:

| X_i | f_i |
|---------------------|----------------------|
| 2 | 1 |
| 6 | 1 |
| $M_o \rightarrow 8$ | 4 ← FREQUÊNCIA MAIOR |
| 10 | 2 |
| 12 | 2 |
| 18 | 1 |

Na prática acontece o mesmo... é a moda! Se estão todos a ouvir os clássicos, dizemos que a “a musica clássica está na moda”. Se os chapéus entram na moda, então....



Percentis:

O percentil de ordem $p \times 100$ ($0 < p < 1$), em um conjunto de dados de tamanho n , é o valor da variável que ocupa a posição $p \times (n + 1)$ do conjunto de dados ordenados.

Casos particulares:

percentil 50 = mediana ou segundo quartil (Md)

percentil 25 = primeiro quartil (Q_1)

percentil 75 = terceiro quartil (Q_3)

percentil 10 = primeiro decil

Fórmula p/
Calc. Quartil

$$Q_j = X_k + \left(\frac{j(n+1)}{4} - k \right) (X_{k+1} - X_k).$$

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7 $\Rightarrow n=10$

$$Q_j = X_k + \left(\frac{j(n+1)}{4} - k \right) (X_{k+1} - X_k).$$

Cálculo do Q-1

$$j = 1$$

$$n = 10$$

$$j*(n+1)/4 =$$

$$K =$$

$$Q1 =$$

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7 $\Rightarrow n=10$

$$Q_j = X_k + \left(\frac{j(n+1)}{4} - k \right) (X_{k+1} - X_k).$$

Cálculo do Q-1

$$j = 1$$

$$n = 10$$

$$j*(n+1)/4 = 2,75$$

$$K = 2$$

$$Q_1 = 2,075$$

$$k = \frac{j(n+1)}{4} \rightarrow \frac{1(10+1)}{4} \rightarrow k = 2,75 \text{ e } k_{\text{inteiro}} = 2$$

$$Q_1 = 2,0 + \left[\frac{1(10+1)}{4} - 2 \right] * (2,1 - 2,0) \rightarrow Q_1 = 2,075$$

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7 $\Rightarrow n=10$

$$Q_j = X_k + \left(\frac{j(n+1)}{4} - k \right) (X_{k+1} - X_k).$$

Cálculo do Q-1

$$j = 1$$

$$n = 10$$

$$j*(n+1)/4 = 2,75$$

$$K = 2$$

$$Q1 = 2,075$$

$$k = \frac{j(n+1)}{4} \rightarrow \frac{1(10+1)}{4} \rightarrow k = 2,75 \text{ e } k_{\text{inteiro}} = 2$$

$$Q_1 = 2,0 + \left[\frac{1(10+1)}{4} - 2 \right] * (2,1 - 2,0) \rightarrow Q_1 = 2,075$$

Cálculo do Q-2=Mediana

$$j = 2$$

$$n = 10$$

$$j*(n+1)/4 =$$

$$K =$$

$$Q3 =$$

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7 $\Rightarrow n=10$

$$Q_j = X_k + \left(\frac{j(n+1)}{4} - k \right) (X_{k+1} - X_k).$$

Cálculo do Q-1

$$j = 1$$

$$n = 10$$

$$j^*(n+1)/4 = 2,75$$

$$K = 2$$

$$Q_1 = 2,075$$

$$k = \frac{j(n+1)}{4} \rightarrow \frac{1(10+1)}{4} \rightarrow k = 2,75 \text{ e } k_{\text{inteiro}} = 2$$

$$Q_1 = 2,0 + \left[\frac{1(10+1)}{4} - 2 \right] * (2,1 - 2,0) \rightarrow Q_1 = 2,075$$

Cálculo do Q-2=Mediana

$$j = 2$$

$$n = 10$$

$$j^*(n+1)/4 = 5,50$$

$$K = 5$$

$$Q_2 = 3,050$$

$$k = \frac{j(n+1)}{4} \rightarrow \frac{2(10+1)}{4} \rightarrow k = 5,50 \text{ e } k_{\text{inteiro}} = 5$$

$$Q_2 = 3,0 + \left[\frac{2(10+1)}{4} - 5 \right] * (3,1 - 3,0) \rightarrow Q_2 = 3,05$$

Cálculo do Q-3

$$j = 3$$

$$n = 10$$

$$j^*(n+1)/4 =$$

$$K =$$

$$Q_3 =$$

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7 $\Rightarrow n=10$

$$Q_j = X_k + \left(\frac{j(n+1)}{4} - k \right) (X_{k+1} - X_k).$$

Cálculo do Q-1

$$j = 1$$

$$n = 10$$

$$j^*(n+1)/4 = 2,75$$

$$K = 2$$

$$Q_1 = 2,075$$

$$k = \frac{j(n+1)}{4} \rightarrow \frac{1(10+1)}{4} \rightarrow k = 2,75 \text{ e } k_{\text{inteiro}} = 2$$

$$Q_1 = 2,0 + \left[\frac{1(10+1)}{4} - 2 \right] * (2,1 - 2,0) \rightarrow Q_1 = 2,075$$

Cálculo do Q-2=Mediana

$$j = 2$$

$$n = 10$$

$$j^*(n+1)/4 = 5,500$$

$$K = 5$$

$$Q_2 = 3,050$$

$$k = \frac{j(n+1)}{4} \rightarrow \frac{2(10+1)}{4} \rightarrow k = 5,50 \text{ e } k_{\text{inteiro}} = 5$$

$$Q_2 = 3,0 + \left[\frac{2(10+1)}{4} - 5 \right] * (3,1 - 3,0) \rightarrow Q_2 = 3,05$$

Cálculo do Q-3

$$j = 3$$

$$n = 10$$

$$j^*(n+1)/4 = 8,25$$

$$K = 8$$

$$Q_3 = 4,30$$

$$k = \frac{j(n+1)}{4} \rightarrow \frac{3(10+1)}{4} \rightarrow k = 8,25 \text{ e } k_{\text{inteiro}} = 8$$

$$Q_3 = 3,7 + \left[\frac{3(10+1)}{4} - 8 \right] * (6,1 - 3,7) \rightarrow Q_3 = 4,30$$

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7 $\Rightarrow n=10$

$$Md = 3,05$$

$$Q_1 = 2,075$$

$$Q_3 = 4,3$$

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6 $n=11$

Cálculo do Q-1

Cálculo do Q-2=Mediana

Cálculo do Q-3

$$j = 1$$

$$j = 2$$

$$j = 3$$

$$n = 11$$

$$n = 11$$

$$n = 11$$

$$j*(n+1)/4 =$$

$$j*(n+1)/4 =$$

$$j*(n+1)/4 =$$

$$K =$$

$$K =$$

$$K =$$

$$Q1 =$$

$$Q2 =$$

$$Q3 =$$

$$Q_j = X_k + \left(\frac{j(n+1)}{4} - k \right) (X_{k+1} - X_k).$$

$$Md = ?$$

$$Q1 = ?$$

$$Q3 = ?$$

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7 $\Rightarrow n=10$

$$Md = 3,05$$

$$Q_1 = 2,075$$

$$Q_3 = 4,3$$

Dados: 0,9 1,0 1,7 2,9 3,1 5,3 5,5 12,2 12,9 14,0 33,6 $n=11$

Cálculo do Q-1

Cálculo do Q-2=Mediana

Cálculo do Q-3

$$j = 1$$

$$j = 2$$

$$j = 3$$

$$n = 11$$

$$n = 11$$

$$n = 11$$

$$j^*(n+1)/4 = 3,00$$

$$j^*(n+1)/4 = 6,00$$

$$j^*(n+1)/4 = 9,00$$

$$K = 3$$

$$K = 6$$

$$K = 9$$

$$Q1 = 1,7$$

$$Q2 = 5,3$$

$$Q3 = 12,9$$

$$Q_j = X_k + \left(\frac{j(n+1)}{4} - k \right) (X_{k+1} - X_k).$$

$$Md = 5,3$$

$$Q1 = 1,7$$

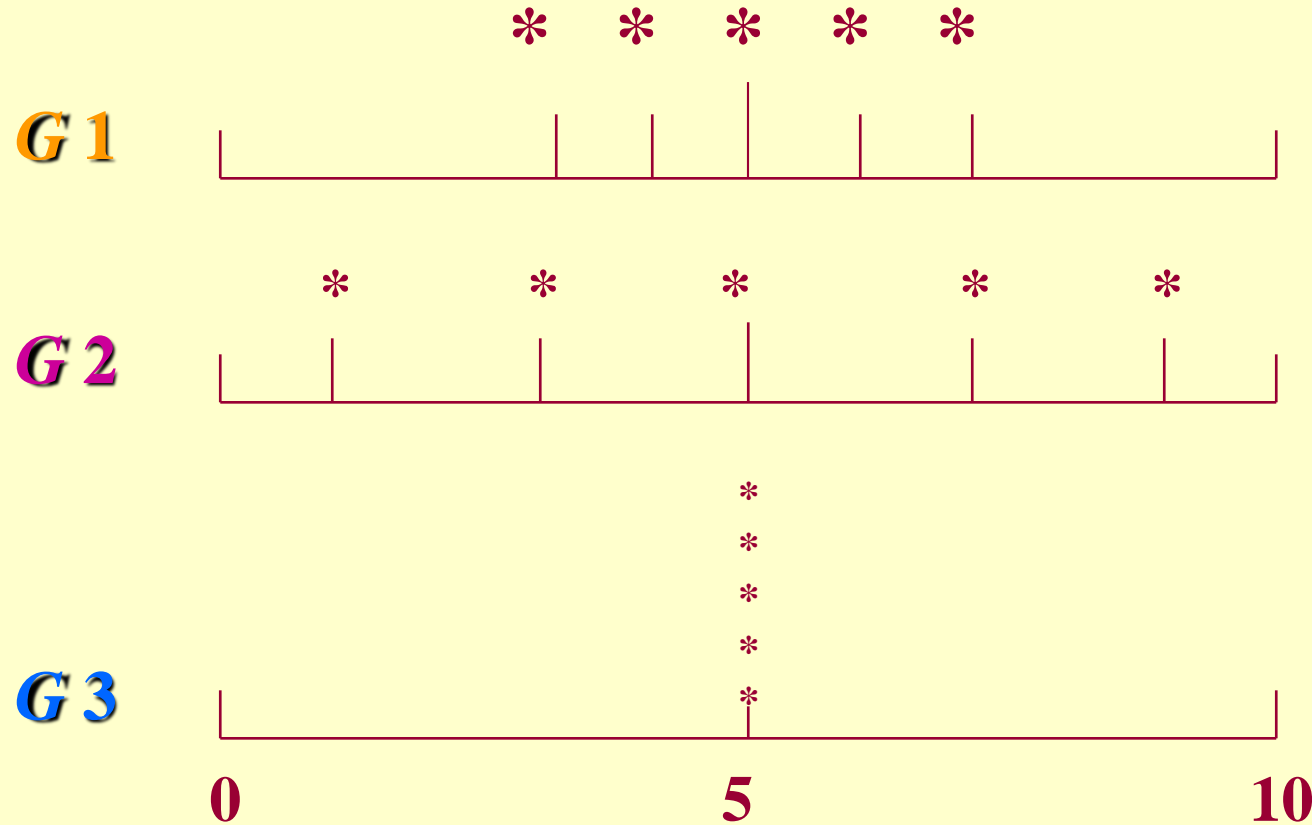
$$Q3 = 12,9$$

Exemplo 2: Considere as notas de um teste de 3 grupos de alunos

Grupo 1: 3,4,5,6,7

Grupo 2: 1, 3, 5, 7, 9

Grupo 3: 5,5,5,5,5



Temos: $\bar{x}_1 = \bar{x}_2 = \bar{x}_3 = 5$ e $md_1 = md_2 = md_3 = 5$

Medidas de Dispersão

Finalidade: encontrar um valor que resuma a variabilidade de um conjunto de dados

• **Amplitude (A):** $A = \text{máx} - \text{min}$

Grupo 1: 3,4,5,6,7

Grupo 2: 1, 3, 5, 7, 9

Grupo 3: 5,5,5,5,5

Para os grupos anteriores, temos:

Grupo 1, $A = 4$

Grupo 2, $A = 8$

Grupo 3, $A = 0$

•Intervalo-Interquartil:

É a diferença entre o terceiro quartil e o primeiro quartil, ou seja, $Q3 - Q1$.

Dados: 1,9 2,0 2,1 2,5 3,0 3,1 3,3 3,7 6,1 7,7

$$Q1 = 2,075 \quad \text{e} \quad Q3 = 4,3$$

$$Q3 - Q1 = 4,3 - 2,075 = 2,225$$

•Variância:

$$\text{Variância} = s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1} = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}$$

•Desvio padrão:

$$\text{Desvio Padrão} = s = \sqrt{\text{Variância}}$$

Cálculo para os grupos:

$$\mathbf{G1: } s^2 = \frac{(3-5)^2 + (4-5)^2 + (5-5)^2 + (6-5)^2 + (7-5)^2}{4}$$

$$\Rightarrow s^2 = 10/4 = 2,5 \quad \Rightarrow s = 1,58$$

$$\mathbf{G2: } s^2 = 10 \Rightarrow s = 3,16$$

$$\mathbf{G3: } s^2 = 0 \Rightarrow s = 0$$

Fórmula alternativa:

$$S^2 = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{(n-1)}$$

Em **G1**: $\sum X_i^2 = 9 + 16 + 25 + 36 + 49 = 135$

$$\Rightarrow S^2 = \frac{135 - 5 \times (5)^2}{4} = 2,5 \quad \Rightarrow s = 1,58$$

• Coeficiente de Variação (*CV*)

- é uma medida de dispersão relativa
- elimina o efeito da magnitude dos dados
- exprime a variabilidade em relação à média

$$CV = \frac{s}{\bar{x}} \times 100\%$$

Exemplo 3:

Altura e peso de alunos

| | Média | Desvio Padrão | Coef. de Variação |
|---------------|---------------|----------------------|--------------------------|
| Altura | 1,143m | 0,063m | 5,5% |
| Peso | 50 kg | 6kg | 12% |

Conclusão: Os alunos são, aproximadamente, duas vezes mais dispersos quanto ao peso do que quanto à altura.

Exemplo 4:

Altura (em *cm*) de uma amostra de recém-nascidos e de uma amostra de adolescentes

| | Média | Desvio padrão | Coef. de variação |
|-----------------------|--------------|----------------------|--------------------------|
| Recém-nascidos | 50 | 6 | 12% |
| Adolescentes | 160 | 16 | 10% |

Conclusão: Em relação às médias, as alturas dos adolescentes e dos recém-nascidos apresentam variabilidade quase iguais.

Tabela de Distribuição de Frequências

Para se descrever os dados coletados, o primeiro passo consiste dispô-los em ordem crescente formando o ROL. Em seguida, determinar sua frequência (*repetição*) dos valores daquela variável, gerando assim, a TABELA de DISTRIBUIÇÃO de FREQUÊNCIA.

Dados Brutos: 45, 41, 42, 41, 42 43, 44, 41 ,50, 46, 50, 46, 60, 54, 52, 58, 57, 58, 60, 51

Rol: 41, 41, 41, 42, 42 43, 44, 45 ,46, 46, 50, 50, 51, 52, 54, 57, 58, 58, 60, 60

| Dados | 41 | 42 | 43 | 44 | 45 | 46 | 50 | 51 | 52 | 54 | 57 | 58 | 60 | TOTAL |
|-------------|----|----|----|----|----|----|----|----|----|----|----|----|----|-------|
| Frequências | 3 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 20 |

Tabela de Distribuição de Frequências: É a simples condensação dos dados conforme as repetições de seu valores.

Para um **ROL** de tamanho razoável a distribuição de frequência poderá se apresentar de forma inconveniente, exigindo muito espaço.

Tabela de Distribuição de Frequências

Tabela de Distribuição de Frequência com intervalos de classe: Quando o tamanho da amostra é elevado, é mais racional efetuar o agrupamento dos valores em intervalos de classe.

Rol: 41, 41, 41, 42, 42 43, 44, 45 ,46, 46, 50, 50, 51, 52, 54, 57, 58, 58, 60, 60

| Classes | | | Freq. |
|--------------|--------|----|-----------|
| 41 | !----- | 45 | 7 |
| 45 | !----- | 49 | 3 |
| 49 | !----- | 53 | 4 |
| 53 | !----- | 57 | 1 |
| 57 | !----- | 61 | 5 |
| TOTAL | | | 20 |

Tamanho da Amostra = N ; no caso, N = 20

Tabela de Distribuição de Frequências

ELEMENTOS de uma DISTRIBUIÇÃO de FREQUENCIA (com intervalos de classe):

| Classes | | | Freq. |
|--------------|--------|----|-----------|
| 41 | !----- | 45 | 7 |
| 45 | !----- | 49 | 3 |
| 49 | !----- | 53 | 4 |
| 53 | !----- | 57 | 1 |
| 57 | !----- | 61 | 5 |
| TOTAL | | | 20 |

CLASSE: são os intervalos de classificação das variáveis, simbolizadas por i . Já o número total de classes é simbolizado por k . Ex: na tabela, $k=5$ e **49 |----- 53** é a **3ª classe**, onde $i=3$;

LIMITES de CLASSE: são os extremos de cada classe. O menor número é o limite inferior de classe (L_i) e o maior número, limite superior de classe (L_s). Ex: em **49 |----- 53** $L_i = 49$ e $L_s = 53$. O símbolo |----- representa um intervalo fechado à esquerda e aberto à direita. O dado **53** do **ROL** não pertence à 3ª classe e sim à 4ª classe representada por **53|---- 57**

AMPLITUDE do INTERVALO de CLASSE: é a diferença entre o limite superior e inferior da classe e é simbolizada por $h_i = L_s - L_i$. Ex: na tabela anterior $h_3 = 53 - 49 = 4$. *Obs: na distribuição de frequência c/ classe o h_i será igual em todas as classes.*

AMPLITUDE TOTAL da DISTRIBUIÇÃO: é a diferença entre o limite superior da última classe e o limite inferior da primeira classe. $A_T = L_s(max) - L_i(min)$. Ex: na tabela anterior $A_T = 61 - 41 = 20$.

AMPLITUDE TOTAL da AMOSTRA (ROL): é a diferença entre o valor máximo e o valor mínimo da amostra (ROL). Onde $A = X_{max} - X_{min}$. no exemplo $A = 60 - 41 = 19$. *Obs: A_T sempre será maior que A .*

PONTO MÉDIO de CLASSE: é o ponto que divide o intervalo de classe em duas partes iguais. Ex: em **49 |----- 53** o ponto médio $X_{med3} = (53+49)/2 = 51$, ou seja $X_{med3} = (L_{s3} + L_{i3})/2$.

Tabela de Distribuição de Frequências

Método prático para construção de uma tabela de distribuição de frequência

- 1º - Organize os dados brutos em um **ROL**;
- 2º - Calcule a amplitude amostral **A** (*Range*);
- 3º - Calcule o número de classes através da "Regra de Sturges" => **$K \approx 1 + 3,3 * \text{Log } N$** ;
- 4º - Decidido o nº de classes, calcule então a amplitude do intervalo de classe **$h \geq A_T/K$** ;
- 5º - Determinar os limites inferiores e superiores das classes para em seguida montar a tabela, com o cuidado de não permitir classes com frequência = 0 (zero).

Estatística Descritiva

Os valores abaixo, representados em porcentagem (%) apresentam a quantidade de agregados “passantes” na peneira Nº 16, cuja trama tem uma espessura de 1,19 mm.

Determine os seguintes elementos:

a) o ROL da amostra; b) média aritmética; c) mediana; d) moda; e) amplitude da amostra; f) variância e g) desvio padrão.

Construir a tabela de distribuição de frequências, utilizando as informações apresentadas em sala de aula.

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 63,10 | 59,50 | 62,80 | 61,40 | 60,10 | 59,31 | 45,14 | 69,07 | 51,76 | 71,16 |
| 58,50 | 53,10 | 55,00 | 57,00 | 55,30 | 56,28 | 42,37 | 65,55 | 49,59 | 68,51 |

Solução Estatística Descritiva

a) o ROL da amostra; b) média aritmética; c) mediana; d) moda; e) amplitude da amostra; f) variância e g) desvio padrão.

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 63,10 | 59,50 | 62,80 | 61,40 | 60,10 | 59,31 | 45,14 | 69,07 | 51,76 | 71,16 |
| 58,50 | 53,10 | 55,00 | 57,00 | 55,30 | 56,28 | 42,37 | 65,55 | 49,59 | 68,51 |

a) ROL

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 42,37 | 45,14 | 49,59 | 51,76 | 53,10 | 55,00 | 55,30 | 56,28 | 57,00 | 58,50 |
| 59,31 | 59,50 | 60,10 | 61,40 | 62,80 | 63,1 | 65,55 | 68,51 | 69,07 | 71,16 |

b) média $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ $\bar{Y} = 58,227$

c) Mediana $(n=10) \leftrightarrow \text{par}$ $M_d = \frac{(X_{10} + X_{11})}{2}$ $Md = 58,905$

d) Moda *não existe* \leftrightarrow *não há repetição de valores*

Solução Estatística Descritiva

e) amplitude da amostra; f) variância e g) desvio padrão.

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 42,37 | 45,14 | 49,59 | 51,76 | 53,10 | 55,00 | 55,30 | 56,28 | 57,00 | 58,50 |
| 59,31 | 59,50 | 60,10 | 61,40 | 62,80 | 63,1 | 65,55 | 68,51 | 69,07 | 71,16 |

Valor Máximo \leftrightarrow ***Máx*** = 71,16

Valor Mínimo \leftrightarrow ***Mín*** = 42,37

e) amplitude **$A = Máx - Mín \Leftrightarrow A = 28,79$**

f) Variância $S^2 = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{(n-1)}$ $\Sigma X_i^2 = 68.902,670$ $n\bar{X}^2 = 67.807,671$ **$S^2 = 57,63$**

d) Desvio Padrão **$s = \sqrt{\text{Variância}}$** **$S = 7,59$**

Tabela de Distribuição de Frequências

Os valores abaixo, representados em porcentagem (%) apresentam a quantidade de agregados “passantes” na peneira Nº 16, cuja trama tem uma espessura de 1,19 mm. Elabore uma tabela de distribuição de frequências.

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 42,37 | 45,14 | 49,59 | 51,76 | 53,10 | 55,00 | 55,30 | 56,28 | 57,00 | 58,50 |
| 59,31 | 59,50 | 60,10 | 61,40 | 62,80 | 63,1 | 65,55 | 68,51 | 69,07 | 71,16 |

Valor Máximo \leftrightarrow ***Máx*** = ***71,16***

Valor Mínimo \leftrightarrow ***Mín*** = ***42,37***

Amplitude Amostral = ***A*** = ***28,79***

Nº de Classes \Leftrightarrow *por Sturges* \Leftrightarrow ***K*** \approx ***1 + 3,3 * Log N*** \Leftrightarrow ***K*** \approx ***5***

*Conhecendo-se o Nº Classes (***K***) e a Amplitude Amostral (***A***), determina-se a Amplitude dos Intervalos de Classe (***hi***) e a consequente Amplitude da Distribuição (***A_T***)*

hi = ***A/K*** \Leftrightarrow ***hi*** \approx ***6***

Mín_T = ***42***

Máx_T = ***72***

A_T = ***Máx_T - Mín_T*** \Leftrightarrow ***A_T*** = ***30***

Tabela de Distribuição de Frequências

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 42,37 | 45,14 | 49,59 | 51,76 | 53,10 | 55,00 | 55,30 | 56,28 | 57,00 | 58,50 |
| 59,31 | 59,50 | 60,10 | 61,40 | 62,80 | 63,1 | 65,55 | 68,51 | 69,07 | 71,16 |

Nº de Classes $\Leftrightarrow K \approx 5$

Intervalos de Classe $\Leftrightarrow h_i \approx 6$

$Min_T = 42$

$Max_T = 72$

Amplitude da Distribuição $\Leftrightarrow A_T = 30$

| | | BIN | |
|-----------------|------|-------|-----------|
| L_i | | L_s | Freq |
| 42 | !--- | 48 | 2 |
| 48 | !--- | 54 | 3 |
| 54 | !--- | 60 | 7 |
| 60 | !--- | 66 | 5 |
| 66 | !--- | 72 | 3 |
| TOTAL... | | | 20 |

Tabela de Distribuição de Frequências

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 42,37 | 45,14 | 49,59 | 51,76 | 53,10 | 55,00 | 55,30 | 56,28 | 57,00 | 58,50 |
| 59,31 | 59,50 | 60,10 | 61,40 | 62,80 | 63,1 | 65,55 | 68,51 | 69,07 | 71,16 |

| | | BIN | |
|-------------------------|------|-------------------------|-------------|
| L_i | | L_s | Freq |
| 42 | !--- | 48 | 2 |
| 48 | !--- | 54 | 3 |
| 54 | !--- | 60 | 7 |
| 60 | !--- | 66 | 5 |
| 66 | !--- | 72 | 3 |
| TOTAL... | | | 20 |

Classe Modal \Leftrightarrow 3ª classe $\Leftrightarrow L_i = 54$ e $L_s = 60 \Leftrightarrow Mo = 57$

FIM