

## 5

ANÁLISE E  
INTERPRETAÇÃO  
DOS DADOS

## 5.1 MEDIDAS DE POSIÇÃO

As medidas de posição, também chamadas parâmetros de posição ou medidas de tendência central, constituem-se em "um dos procedimentos para a redução dos dados, expressando valores que se encontram situados entre os extremos de uma série ou distribuição (Hofmann, 1974:312). Referem-se a dados não tabulados e a dados tabulados.

## 5.1.1 Dados Não Tabulados

5.1.1.1 MÉDIA ( $\bar{X}$ )

Média é a medida de posição mais usada nos procedimentos estatísticos. A média de uma distribuição equivale à média aritmética. Quando os dados não são tabulados, a média aritmética é calculada pela fórmula:

$$M = \frac{\sum X_1}{N}$$

$M$  = média aritmética

$\Sigma$  (sigma) = soma

$X_1$  = valores

$N$  = número de valores

*Exemplo:* Calcular a média aritmética das seguintes notas: 20, 80, 40, 60, 50. Nesta série há 5 valores; logo,  $N = 5$ .

- 1º – em primeiro lugar a fórmula pede a soma ( $\Sigma$ ) dos valores. Por isso, antes de tudo, deve-se somar  $20 + 80 + 40 + 50 + 60 = 250$ .
- 2º – em seguida, a fórmula pede que a soma obtida seja dividida por  $N$ . Como  $N = 5$ , tem-se:

$$\frac{20 + 80 + 40 + 60 + 50}{5} = \frac{250}{5} = 50$$

Resposta: a média aritmética das notas é 50 ou  $\bar{X} = 50$ .

5.1.1.2 MEDIANA ( $Md$ )

Mediana é o valor central, situado exatamente no centro do rol. Antes da  $Md$  encontram-se 50% da distribuição e depois da  $Md$  os outros 50%. É uma medida de posição, mais do que de grandeza.

Para se encontrar a mediana de dados não tabulados, basta localizar o valor central.

*Exemplos:*

1º) 1, 3, 5, 6, 7, 9, 10	$Md = 6$
2º) 15, 17, 20, 22, 23, 25, 29, 30, 32, 33, 35	$Md = 25$
3º) -5, -4, -3, -1, 0, 2, 3, 4, 5	$Md = 0$

Procedimentos para contar a mediana em valores de seriação (dados agrupados):

1. Ordenar os valores hierarquicamente (do menor para o maior ou vice-versa).
2. Se o número de valores for ímpar, a mediana é o valor que se encontra no meio da ordenação.
3. Se o número de valores for par, a mediana é um valor médio entre os dois valores centrais.

*Exemplos:*

1º) 10, 20, 30, 40,   50, 60, 70, 80	$Md = 45$
2º) 20, 30, 40,   60, 70, 80	$Md = 50$
3º) 5, 6, 8, 10,   12, 15, 17, 20	$Md = 11$
4º) -5, -4, -2,   2, 5, 7	$Md = 0$

A mediana não é afetada pelos extremos da série.

Exemplos:

$$1^{\text{a}}) 7, 6, 5, 4, 3, 2, 1 \quad Md = 4$$

$$2^{\text{a}}) 1, 2, 3, 4, 5, 29, 50 \quad Md = 4$$

Para facilitar a localização da mediana, pode-se usar a fórmula  $\frac{N+1}{2}$ , que dá a posição de valor mediano.

Exemplos:

$$1^{\text{a}}) 1, 3, 5, 6, 6, 7, 8 = 7 \text{ valores}$$

$$\frac{7+1}{2} = \frac{8}{2} = 4$$

$$(\text{ou } 4^{\text{a}} \text{ valor}) = 6$$

$$2^{\text{a}}) 14, 17, 20, 22, 23, 25, 29, 30, 32, 33, 39 = 11 \text{ valores}$$

$$\frac{11+1}{2} = \frac{12}{2} = 6$$

$$(\text{ou } 6^{\text{a}} \text{ valor}) = 25$$

$$3^{\text{a}}) -5, -4, -3, -1, 0, 2, 3, 4, 5 = 9 \text{ valores}$$

$$\frac{9+1}{2} = \frac{10}{2} = 5 \text{ (ou } 5^{\text{a}} \text{ valor)} = 0$$

Considerando-se a curva correspondente a certas distribuições, o valor da mediana será obtido, graficamente, traçando-se uma perpendicular ao eixo X, a partir do ponto médio da distribuição (indivíduo mediano).

Esta perpendicular divide a distribuição em duas partes iguais: 50% de cada lado do valor considerado, ou seja, 50% dos valores maiores e 50% dos valores menores que o valor da mediana.

A mediana, em certos tipos de análise, pode refletir com mais exatidão a posição da tendência central do que a média.

Exemplo:

Ordenados de professores de certa escola:

Professor A	R\$ 1.500,00	
B	R\$ 1.300,00	
C	R\$ 950,00	Média = R\$ 1.008,00
D	R\$ 800,00	Mediana = R\$ 950,00
E	R\$ 490,00	

$$M = \frac{\sum X_i}{N} = \frac{152.000,00}{5} = 30.400,00$$

O ordenado médio deste grupo está mais bem representado pela mediana dos ordenados (26.000,00 do Professor C) do que a média dos ordenados (30.400,00), que é superior aos dos cinco professores.

### 5.1.1.3 MODA ( $M_o$ )

Moda ou norma é o valor mais freqüente em uma distribuição. É apenas uma medida de posição que, como a mediana, não pode entrar, posteriormente, nas relações matemáticas. Moda é, portanto, o valor que se repete em maior número de vezes.

Quando os dados não são tabulados, a moda é encontrada por simples inspeção: basta verificar qual o valor mais repetido.

$$1^{\text{a}}) 1, 2, 3, 5, 5, 5, 7, 8 \quad M_o = 5$$

$$2^{\text{a}}) 20, 35, 35, 40, 45, 50 \quad M_o = 35$$

$$3^{\text{a}}) 5, 9, 10, 10, 14, 15, 16, 16, 17 \quad M_o = 16$$

### 5.1.2 Dados Tabulados

#### 5.1.2.1 MÉDIA ARITMÉTICA

A média aritmética pode ser efetuada de dois modos: processo longo e processo abreviado.

a. *Processo longo*. O cálculo da média aritmética, pelo processo longo, é feito pela fórmula:

$M$  ou  $\bar{X}_1$  = média aritmética (Ler: X-Barra)

$\Sigma$  = soma

$X_1$  = valores (pontos médios das classes)

$n_1$  = freqüências

$N$  = número de valores

$$\bar{X} = \frac{\sum X_1 n_1}{N}$$

QUADRO 3

Classes	$n_1$	$X_1$	$X_1 n_1$
0  — 10	1	5	5
10  — 20	3	15	45
20  — 30	13	25	325
30  — 40	6	35	210
40  — 50	2	45	90
	25		675

Exemplo: Calcular a média aritmética da distribuição do Quadro 3.

- 1º – Organizar uma coluna com pontos médios das classes (3ª coluna do Quadro 3).
- 2º – Destacar, na fórmula, a expressão:  $X_1 n_1$ . Ora,  $X_1$  são os pontos médios e  $n_1$  são as freqüências. O fato de estarem colocados um ao lado do outro, sem qualquer sinal, quer dizer que devem ser multiplicados. Logo,  $X_1 n_1$  significa: multiplicar os pontos médios pelas freqüências (como na 4ª coluna). A 4ª coluna foi obtida multiplicando-se cada freqüência (2ª coluna) pelo ponto médio correspondente (3ª coluna).
- 3ª – Antes da expressão  $X_1 n_1$  há um sinal  $\Sigma$ . Isso significa que os números da 4ª coluna devem ser somados. O resultado da soma é 675.
- 4ª – Finalmente, na fórmula, aparece o  $N$  como denominador. Isto quer dizer que o resultado da soma dos  $X_1 n_1$  deve ser dividido por  $N$ , número de valores obtido pela soma das freqüências.

Neste exemplo,  $N = 25$ . Tem-se, então:

$$M = \frac{\Sigma X_1 n_1}{N} = \frac{675}{25} = 27$$

- b. **Processo abreviado.** O cálculo da média aritmética, pelo processo abreviado, é feito com o auxílio da seguinte fórmula:

$$M = A + \frac{h \Sigma d_1 n_1}{N}$$

$A$  = ponto médio escolhido arbitrariamente (ver 1º passo)  
 $h$  = amplitude de classe (intervalos)  
 $\Sigma$  = soma  
 $d_1$  = desvios (ver 2º passo)  
 $n_1$  = freqüências  
 $N$  = número de valores  
 $M$  = valor da média real

QUADRO 4

Classes	$X_1$	$n_1$	$d_1$	$d_1 n_1$
0  — 10	5	1	-2	-2
10  — 20	15	3	-1	2
20  — 30	25	13	0	0
30  — 40	35	6	1	6
40  — 50	45	2	2	4
		25		5

Exemplo: Calcular, pelo processo abreviado, a média aritmética da distribuição dada no exemplo acima.

- 1º – Escolher, em primeiro lugar, o valor arbitrário. Há inteira liberdade nessa escolha, mas o cálculo será mais simplificado se for escolhido um ponto médio. A escolha recairá então no ponto médio que tenha maior freqüência ou que esteja mais ou menos no meio da distribuição. Nesse exemplo, deve-se escolher o ponto médio da classe 20 |— 30, à qual corresponde a maior freqüência. O ponto médio dessa classe é:

$$\frac{30 + 20}{2} = \frac{50}{2} = 25. \text{ Portanto, de agora em diante, } A = 25.$$

- 2º – Deve-se calcular os desvios ( $d_1$ ).

Nesse cálculo emprega-se a fórmula:  $d_1 = \frac{X_1 - A}{h}$ , isto é, de cada

ponto médio,  $X_1$ , tira-se  $A$  e divide-se o resultado por  $h$  (amplitude de classe). Procedendo-se dessa forma, acham-se os valores da 4ª

coluna do Quadro 4:  $d_1 = \frac{5 - 25}{10} = \frac{-20}{10} = -2$  etc.

Há, todavia, um cálculo menos trabalhoso. Os  $d_1$  podem ser achados diretamente, da seguinte forma: em correspondência à classe cujo ponto médio foi escolhido para  $A$ , tem-se  $d_1 = 0$ . Nas classes acima dela (de baixo para cima) têm-se os seguintes valores sucessivos para  $d_1$ : -1, -2...

Nas classes abaixo dela têm-se, sucessivamente (de cima para baixo): 1, 2... Isso acontece sempre. Portanto, toda vez que se precisar encontrar  $d_1$ , procede-se conforme indicado.

- 3º – Destaca-se agora, na fórmula, a expressão  $d_1 n_1$ , que quer dizer: cada desvio deve ser multiplicado pela respectiva freqüência. Fazendo isso, obtém-se a 5ª coluna.

- 4º – Antes de  $d_1 n_1$  há o símbolo  $\Sigma$ . A 5ª coluna, portanto, deve ser somada. Nessa coluna há números positivos e negativos. A soma deve ser algébrica, isto é, valem os sinais. Logo, soma dos números positivos menos soma dos números negativos. A soma dos números positivos é 10; a dos negativos é 5. O resultado final dessa soma será:  $10 - 5 = 5$ .
- 5º – Substituir os valores conhecidos na fórmula, ficando:

$$\Sigma d_1 n_1 = 5 \text{ (como no 4º passo)}$$

$$h = 10 \text{ (amplitude de classe)}$$

$$N = 25$$

$$A = 25 \text{ (como no 1º passo)}$$

$$M = A + \frac{h \Sigma d_1 n_1}{N} = 25 + \frac{10 \times 5}{25} = 25 + \frac{50}{25} = 25 + 2 = 27$$

Chega-se, assim, ao mesmo resultado obtido pelo processo longo.

### 5.1.2.2 MEDIANA

A fórmula para o cálculo da Mediana é a seguinte:

$$Md = \ell_i + \frac{\frac{N}{2} - Fa}{n_1} \times h$$

$Md$  = mediana  
 $\ell_i$  = limite inferior da classe mediana  
 $N$  = número de valores (obtido pela soma das frequências)  
 $Fa$  = frequência acumulada da soma anterior à da classe mediana  
 $n_1$  = frequência absoluta da classe mediana  
 $h$  = amplitude de classe

QUADRO 5

Classes	$n_i$	$N_i$
0 — 3	2	2
3 — 6	4	6
6 — 9	5	11
9 — 12	6	17
12 — 15	10	27
15 — 18	8	35
18 — 21	6	41
21 — 24	4	45
24 — 27	3	48
27 — 30	2	50
	50	

Exemplo: Calcular a mediana da distribuição do Quadro 5.

- 1º – O primeiro cuidado deve ser: organizar uma coluna com frequências acumuladas, tal como foi feito na 3ª coluna do Quadro 5.
- 2º – Todos os elementos da fórmula serão conhecidos sabendo-se qual é a classe em que está a mediana. Para descobrir essa classe, deve-se destacar, na fórmula, a expressão  $\frac{N}{2}$ , que permitirá localizar a mediana. Nessa distribuição,  $N = 50$ .

$$\text{Logo, } \frac{N}{2} = \frac{50}{2} = 25$$

- 3º – Achado  $\frac{N}{2} = 25$ , procura-se, na coluna de frequência acumulada, uma que seja igual a 25. Não existindo, toma-se a que esteja logo acima. Consultando a 3ª coluna, verifica-se que 27 é a frequência acumulada que mais convém. Em correspondência a ela encontra-se a classe 12 — 15.

- 4º – Identificada a classe onde está a mediana, todos os valores da fórmula ficam identificados.

$$\ell_i = 12 \text{ (limite inferior da classe 12 — 15)}$$

$$Fa = 17 \text{ (frequência acumulada da classe anterior)}$$

$$n_1 = 10 \text{ (frequência absoluta da classe 12 — 15)}$$

$$h = 3 \text{ (amplitude de classe)}$$

Podem-se, portanto, fazer as substituições na fórmula:

$$Md = \ell_i + \frac{\frac{N}{2} - Fa}{n_1} \times h$$

$$Md = 12 + \frac{25 - 17}{10} \times 3$$

$$Md = 12 + \frac{8}{10} \times 3$$

$$Md = 12 + 0,8 \times 3$$

$$Md = 12 + 2,4$$

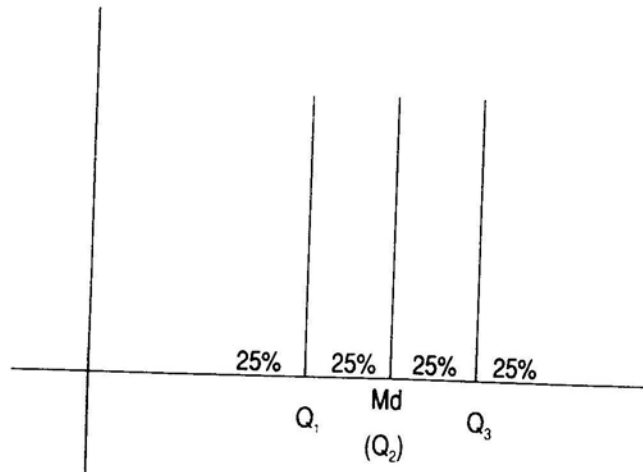
$$Md = 14,4$$

### 5.1.2.3 QUARTIS

Se a mediana divide a distribuição em duas partes iguais, os quartis dividem-na em quatro partes iguais. O primeiro quartil ( $Q_1$ ) tem abaixo de si 25%



da distribuição e acima de si 75%. O terceiro quartil ( $Q_3$ ) tem abaixo de si 75% da distribuição e acima de si 25%. O segundo quartil ( $Q_2$ ) é a própria mediana.



Para o cálculo do primeiro quartil ( $Q_1$ ) tem-se a seguinte fórmula:

$$Q_1 = \ell_i + \frac{\frac{N}{4} - Fa}{n_1} \times h$$

$Q_1$  = primeiro quartil  
 $\ell_i$  = limite inferior da classe onde está  $Q_1$   
 $N$  = número de valores (obtido pela soma das freqüências)  
 $Fa$  = freqüência acumulada da classe anterior à classe onde está  $Q_1$   
 $n_1$  = freqüência absoluta da classe onde está  $Q_1$   
 $h$  = amplitude da classe

Para o cálculo do terceiro quartil ( $Q_3$ ) tem-se a seguinte fórmula:

$$Q_3 = \ell_i + \frac{\frac{3N}{4} - Fa}{n_1} \times h$$

$Q_3$  = terceiro quartil  
 $\ell_i$  = limite inferior da classe onde está  $Q_3$   
 $N$  = número de valores (obtido pela soma das freqüências)  
 $Fa$  = freqüência acumulada da classe anterior à classe onde está  $Q_3$   
 $n_1$  = freqüência absoluta da classe onde está  $Q_3$   
 $h$  = amplitude de classe

A única diferença entre as fórmulas de  $Md$ ,  $Q_1$  e  $Q_3$  está na fração de  $N$ : a fórmula da mediana pede  $\frac{N}{2}$ , a de  $Q_1$  pede  $\frac{N}{4}$  e a de  $Q_3$  pede  $\frac{3N}{4}$ .

Tendo sido achado  $\frac{N}{2}$ , pode-se localizar a classe mediana e, assim, identificar os demais elementos pedidos pela fórmula de  $Md$ .

Tendo sido achado  $\frac{N}{4}$ , pode-se, igualmente, localizar a classe onde está  $Q_1$  e identificar, em função dela, os demais elementos pedidos pela fórmula. Por outro lado, tendo sido achado  $\frac{3N}{4}$ , pode-se localizar a classe onde está  $Q_3$  e identificar, em função dela, os demais elementos da fórmula.

Exemplo: Calcular os 1º e 2º quartis da distribuição do Quadro 6.

QUADRO 6

Classes	$n_i$	$N_i$
0  — 3	2	2
3  — 6	4	6
6  — 9	4	10
9  — 12	7	17
12  — 15	10	27
15  — 18	9	36
18  — 21	6	42
21  — 24	4	46
24  — 27	3	49
27  — 30	1	50
	50	

a) Cálculo de  $Q_1$

- 1º - Achar  $\frac{N}{4}$ . Sabe-se que  $N = 50$ . Portanto,  $\frac{N}{4} = \frac{50}{4} = 12,5$ .
- 2º - Localizar a classe onde está  $Q_1$ . Procura-se na coluna de freqüências acumuladas uma que seja igual a 12,5. Não existindo, toma-se a que fica acima: 17. Em correspondência à freqüência acumulada 17, tem-se a classe 9 |— 12. Portanto,  $Q_1$  está localizado nessa classe.
- 3º - Sabendo-se que a classe onde está  $Q_1$  é a classe 9 |— 12, podem-se identificar os demais elementos da fórmula. Então:  $\ell_i = 9$ ,  $Fa = 10$ ,  $n_1 = 7$ ,  $h = 3$ .

Fazendo-se as substituições na fórmula, tem-se:

$$Q_1 = 9 + \frac{12,5 - 10}{7} \times 3$$

$$Q_1 = 9 + \frac{2,5}{7} \times 3$$

$$Q_1 = 9 + 0,357 \times 3$$

$$Q_1 = 9 + 1,071$$

$$Q_1 = 10,07$$

#### b) Cálculo de Q3

$$\bullet 1^\circ - \text{Achar } \frac{3N}{4} = \frac{3 \times 50}{4} = \frac{150}{4} = 37,5$$

$\bullet 2^\circ$  - Localizar a classe onde está Q<sub>3</sub>. A frequência acumulada que mais se aproxima de 37,5 é 42, que corresponde à classe 10 |— 21 (ver Quadro 6).

$\bullet 3^\circ$  - Sabendo que a classe onde está Q<sub>1</sub> é a classe 18 |— 21, pode-se identificar:  $l_i = 18$ ,  $Fa = 36$ ,  $n_1 = 6$ ,  $h = 3$ .

Substituindo-se na fórmula, vem:

$$Q_3 = 18 + \frac{37,5 - 36}{6} \times 3$$

$$Q_3 = 18 + \frac{1,5}{6} \times 3$$

$$Q_3 = 18 + 0,25 \times 3$$

$$Q_3 = 18 + 0,75$$

$$Q_3 = 18,75$$

#### 5.1.2.4 DECIL

Esta separatriz abrange nove elementos:

1º decil, 2º decil, 3º decil . . . . . até 9º decil.

$$D_1 = \frac{N}{10}; D_2 = \frac{2N}{10}; D_3 = \frac{3N}{10} \text{ etc.}; \text{ até } \frac{9N}{10}$$

As aplicações são feitas da mesma forma que a usada para os quartis.

#### 5.1.2.5 PERCENTIS

Assim como a mediana divide a distribuição em duas partes e o quartis em quatro, os decis a dividem em dez e os percentis em cem partes iguais. A media-

na, os quartis, os decis e os percentis são conhecidos pelo nome genérico de *separatrizes*. As fórmulas para o cálculo de cada um deles diferem entre si apenas na fração de  $N$ .

A fórmula geral para o cálculo dos percentis é:

$P_r$  = percentil de ordem  $R$

$l_i$  = limite inferior da classe onde está o percentil

$R$  = ordem do percentil

$N$  = número de valores (frequência total)

$Fa$  = frequência acumulada anterior à classe onde está o percentil

$h$  = amplitude de classe

$n_1$  = frequência absoluta da classe onde está o percentil

$$P_r = l_i + \frac{\frac{R \times N}{100} - Fa}{n_1} \times h$$

A única novidade nesta fórmula é o  $R$ . Ele indica a ordem do percentil. Se se quiser achar o décimo quinto percentil ( $P_{15}$ ), isto é, aquele que tem abaixo de si 15% dos valores de distribuição, faz-se  $R = 15$ . Substitui-se esse número na fórmula e procede-se, daí para a frente, seguindo os mesmos passos do cálculo da mediana ou dos quartis.

Se se quiser achar  $P_{67}$ , faz-se  $R = 67$ ; se se quiser  $P_{41}$ , faz-se  $R = 41$ , e assim por diante.

É interessante notar que a fórmula para o cálculo dos percentis serve também para o cálculo dos decis, dos quartis e da mediana.

*Exemplo:*

$$P_r = l_i + \frac{\frac{R \times N}{100} - Fa}{n_1} \times h$$

$\bullet 1^\circ$  - Determinar a posição do elemento que está inserido no percentil dado, por meio da regra de três:

$$100 - N \text{ (frequência total)}$$

$$R - X$$

Portanto,  $100^{\circ} - 50$

$$57^{\circ} - X$$

$$X = \frac{57^{\circ} \times 50}{100}$$

$$X = 28,5^{\circ} = \frac{R \times N}{100}$$

- 2º – Procura-se na coluna da frequência acumulada a classe dele (28,5). Assim, tem-se a 6ª classe, cujo limite inferior é 15, e a frequência acumulada anterior é 27, da 5ª classe (Quadro 6).

Logo, pode-se substituir:

$$P_{57} = 15 + \frac{28,5 - 27}{9} \times 3$$

$$P_{57} = 15 + \frac{1,5}{9} \times 3$$

$$P_{57} = 15 + 0,5$$

$$P_{57} = 15,5$$

Se o percentil (R) for 20, tem-se:

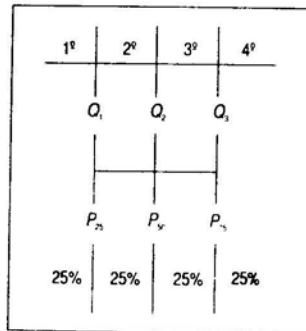
$$100^{\text{a}} - 50$$

$$20^{\text{a}} - X$$

$$X = \frac{20^{\text{a}} \times 50}{100}$$

$$X = 10^{\text{a}} \text{ (lugar)}$$

$$P_{20} = 6 + \frac{10 - 6}{4} \times 3$$



$$P_{20} = 6 + 1 \times 3$$

$$P_{20} = 9$$

Da mesma forma, pode-se demonstrar que:

Cada quarto corresponde a 25% dos casos (Best, 1972:168).

### 5.1.2.6 MODA

A moda é o valor mais freqüente de uma distribuição. Quando os dados não são tabulados, é encontrada por simples inspeção. Entretanto, a moda de uma distribuição de frequência precisa ser calculada.

- a. **Classe Modal.** É a classe em correspondência à qual existe maior frequência.

*Exemplo:* A classe 12 |—— 15 da distribuição do Quadro 6, por apresentar  $n_1 = 10$ .

- b. **Moda Bruta.** É o ponto médio da classe modal.

*Exemplo:* Na distribuição do Quadro 6, a moda bruta é 13,5 (ponto médio da classe 12 |—— 15).

- c. **Antimoda.** Se a moda é caracterizada por uma frequência maior, a antimoda, ao contrário, é caracterizada por uma frequência menor.

Numa curva, a moda é o valor em correspondência ao qual se tem o ponto máximo de uma saliência; a antimoda é o valor em correspondência ao qual se tem o ponto mais baixo de uma depressão.

Para o cálculo da moda empregam-se, freqüentemente, as fórmulas:

$$Mo = 3Md - 2M \quad (Mo = \text{moda}; Md = \text{mediana}; M = \text{média aritmética}).$$

(Fórmula empírica de Pearson).

*Exemplo:* 1, 3, 6, 6, 14.

$$Mo = 3Md - 2M$$

$$Mo = 3 \times 6 - 2 \times 6$$

$$Mo = 18 - 12$$

$$Mo = 6$$

Essa fórmula dá um valor aproximado da moda. Só deve ser usada quando:

- a distribuição for unimodal, isto é, só tem uma moda;
- a distribuição não apresentar assimetria muito acentuada.

- d. **Moda King**

$$MO_k = li + \frac{N_1p}{N_1a + N_1p} \times h, \text{ onde}$$

$MO_k$  = Moda de King

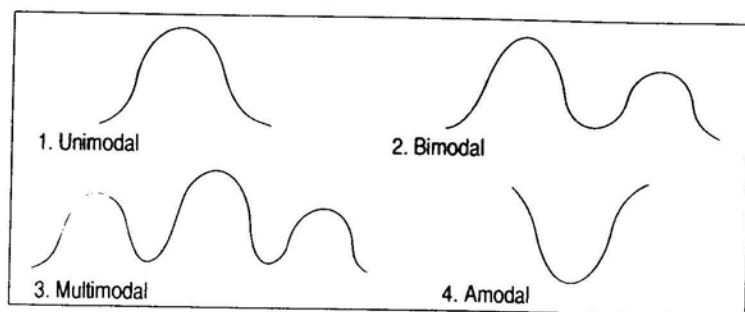
$li$  = limite inferior da classe de maior freqüência modal

$N_{1p}$  = freqüência absoluta da classe *posterior* à da classe modal

$N_{1a}$  = freqüência absoluta da classe *anterior* à da classe modal

A distribuição pode ter mais de uma moda. Quando só tem uma moda, chama-se unimodal; quando tem duas, bimodal, e quando tem mais de duas, multimodal; quando não tem moda, denomina-se amodal.

Tipos de curvas:



### 5.1.2.7 RELAÇÕES ENTRE A MÉDIA ARITMÉTICA, A MEDIANA E A MODA

a. *Influência dos valores extremos.* A média aritmética é a que sofre influência dos valores extremos.

*Exemplo:* Toma-se a seguinte série: 4, 5, 6, 6, 6, 7, 8.

Faz-se o cálculo da  $M$ , da  $Md$  e da  $Mo$ .

$$\text{A média aritmética é: } M = \frac{4+5+6+6+6+7+8}{7} = \frac{42}{7} = 6.$$

$$\text{A mediana é: } \frac{N+1}{2} = \frac{7+1}{2} = \frac{8}{2} = 4 \text{ (4º valor = 6)}$$

Então: 4, 5, 6, 6, 6, 7, 8 ( $Md = 6$ ).

A moda é 6, pois é esse o valor que se repete mais vezes.

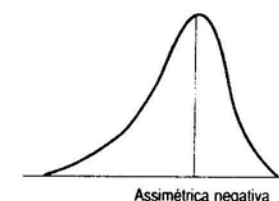
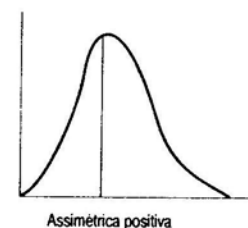
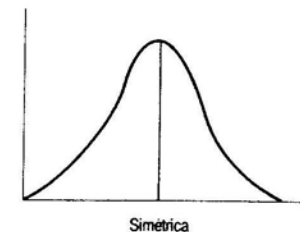
Na série dada, portanto,  $M = 6$ ,  $Md = 6$ ,  $Mo = 6$ .

Substituindo-se, na série dada, o valor 8 pelo valor 18, a série passa a ser: 4, 5, 6, 6, 6, 7, 18.

Calcula-se a  $M$ , a  $Md$  e a  $Mo$  dessa nova série.

$$M = \frac{4+5+6+6+6+7+18}{7} = \frac{52}{7} = 7,4.$$

A mediana e a moda, porém, continuarão sendo 6, como é fácil de verificar. Neste exemplo, a única medida que sofreu influência da introdução de um valor extremo foi a média aritmética: era 6 e passou a ser 7,4. A mediana e a moda não se alteraram.



b. *Distribuições simétricas e assimétricas.*

- *Simétricas.* São aquelas nas quais a freqüência de intervalos correspondentes, nos lados opostos de uma linha média, é igual.
- *Assimétricas.* São aquelas nas quais os valores extremos da distribuição se estendem mais numa direção do que na outra. Se os valores extremos se estendem mais à direita, tem-se uma assimetria positiva; se mais à esquerda, tem-se uma assimetria negativa.

## 5.2 MEDIDAS DE DISPERSÃO (VARIABILIDADE)

As medidas de dispersão (ou de variabilidade) servem para determinar as variações dos valores individuais a partir da média, da mediana e da moda. A oscilação pode ser determinada facilmente se for anotada a diferença entre o item maior e o menor. Se o item menor for 9 e o maior 81, a oscilação é igual a 72.

$$81 - 9 = 72$$

Exemplo: Duas classes de 10 alunos cada uma, com as seguintes notas:

Classe A: 50, 50, 50, 50, 50, 50, 50, 50, 50, 50

$$M = \frac{\sum X_1}{N}$$

$$M = \frac{50 + 50 + 50 + 50 + 50 + 50 + 50 + 50 + 50 + 50}{10} = \frac{500}{10} = 50$$

Classe B: 0, 0, 10, 10, 40, 60, 90, 90, 100, 100

$$M = \frac{\sum X_1}{N}$$

$$M = \frac{0 + 0 + 10 + 10 + 40 + 60 + 90 + 90 + 100 + 100}{10} = \frac{500}{10} = 50$$

Na classe A todas as notas foram iguais à média aritmética; nenhuma se desviou dela. Na classe B, ao contrário, houve grande variabilidade: as notas divergiram bastante da média aritmética. A diferença entre cada valor e a média aritmética ( $X_1 - M$ ) chama-se desvio (afastamento ou discrepância) ao redor de  $M$ . Desvios encontrados nas duas classes:

Classe A: (50 - 50); (50 - 50); (50 - 50) ..... ou seja:

Desvios: 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,

Classe B: (0 - 50); (0 - 50); (10 - 50); (10 - 50); (40 - 50); (60 - 50); (90 - 50); (90 - 50); (100 - 50); (100 - 50), ou seja:

Desvios: -50, -50, -40, -40, -10, 10 (M 10), 40, 40, 50, 50.

Cada desvio indica quanto cada nota se afastou da média aritmética.

Problema: Se são muitos os desvios, um para cada valor, não haverá um número só para indicar a variabilidade da classe em conjunto? Para responder a esta questão há duas soluções:

- 1ª solução. Usar valores absolutos sem consideração pelos sinais.

O índice assim encontrado chama-se *desvio-médio*.

$$DM = \frac{380}{10}$$

$$DM = 38$$

Para calcular o desvio-médio usa-se a fórmula:

$$DM = \frac{\sum (X_1 - M)}{N}$$

$\Sigma$  = soma

$X_1$  = valor (nota)

$M$  = média aritmética

$N$  = número de valores ou desvios

Procede-se da seguinte maneira:

1. Têm-se os desvios ( $X_1 - M$ ).
2. Depois, a soma desses desvios, sem levar em consideração os sinais, isto é, somam-se os módulos dos desvios  $\Sigma (X_1 - M)$ .
3. Divide-se essa soma pelo número de valores dados.

- 2ª solução. Elevar os desvios ao quadrado. Todo número negativo, quando elevado ao quadrado, torna-se positivo. Ter-se-á:

$$(-50)^2, (-50)^2, (-40)^2, (-40)^2, (-10)^2, (10)^2, (40)^2, (40)^2, (50)^2, (50)^2 = 2500, 2500, 1600, 1600, 100, 100, 1600, 1600, 2500, 2500$$

Estão assim desaparecidos os números negativos. Agora, pode-se calcular a média desses números, sem receio de resultado nulo. (Ver estatística inferencial.) Então:

$$2500 + 2500 + 1600 + 1600 + 100 + 100 + 1600 + 1600 + 2500 + 2500 = \frac{16600}{10} = 1660$$

Na primeira fase, elevamos ao quadrado. Para desfazer esta operação, usa-se a operação inversa, que é a da raiz quadrada.

O desvio calculado chama-se *desvio-padrão*. Se não for extraída a raiz quadrada, recebe o nome de variância. (V)

Cálculo para encontrar o desvio-padrão:

1. Em primeiro lugar, encontram-se os desvios: ( $X_1 - M$ )
2. Elevam-se esses desvios ao quadrado: ( $X_1 - M$ )<sup>2</sup>
3. Somam-se esses quadrados:  $\Sigma (X_1 - M)^2$
4. O resultado da soma é dividido por  $N$ :  $\frac{\Sigma (X_1 - M)^2}{N}$
5. Finalmente, extrai-se a raiz quadrada:

$$\sigma = \sqrt{\frac{\Sigma (X_1 - M)^2}{N}}$$

Observação:  $\sigma$  (sigma), letra s minúscula do alfabeto grego, é o símbolo usado para designar o desvio-padrão.

Quando  $N > 30$ , utilizar  $N - 1$

As principais medidas de variabilidade serão vistas a seguir.

### 5.2.1 Amplitude Total

A amplitude total é a diferença entre o maior e o menor valores da série.

$$\begin{aligned} At &= L_s - L_i + 1 & At &= \text{amplitude} \\ \text{ou} & & X_2 \text{ ou } L_s &= \text{limite superior} \\ At &= X_2 - X_1 + 1 & X_1 \text{ ou } L_i &= \text{limite inferior} \end{aligned}$$

Numa série em que todos os valores são iguais, a amplitude total é igual a zero. À medida que os valores se tornam mais diferenciados, a amplitude total aumenta. Por isso, a amplitude total pode ser tomada como medida de variabilidade. Não é medida segura, pelas seguintes razões:

1. O cálculo da amplitude total está unicamente baseado no maior e no menor valores da série. Os demais valores não têm a menor influência. Quer eles se concentrem ao redor da média aritmética, quer se enviem para a direita ou para a esquerda. Nada disso tem influência na amplitude total.

*Exemplos:*

- A. 5, 6, 6, 8, 12, 12, 12, 13, 18, 21, 29
- B. 8, 8, 9, 16, 20, 21, 21, 25, 27, 27, 32

Na série A, a amplitude total é:  $29 - 5 = 24$ . Por sua vez, a média aritmética, obtida pela fórmula  $M = \frac{\sum X_1}{N}$  (ver 5.1.1.1), é 12,91.

Dessa maneira, a maior concentração ocorre à esquerda da média.

Na série B, a amplitude total é também 24 ( $32 - 8 = 24$ ). A média aritmética é 19,45, sendo que a maior concentração se encontra à direita da média.

2. A amplitude total sofre influência do número de casos. Num grupo pequeno de valores há pouca probabilidade de aparecerem valores muito extremos, ocorrendo o contrário, mais freqüentemente, num grupo grande. Sendo as divergências mais acentuadas, a amplitude total, conseqüentemente, aumenta bastante.

### 5.2.2 Amplitude Semiquartil (Q)

A amplitude semiquartil é a diferença entre o 3º e o 1º quartis. (Quadro 6). Nessa distribuição, tem-se:

$$Q_3 = 18,75 \text{ e } Q_1 = 10,07 \text{ (ver 5.1.2.3.a e 5.1.2.3.b)}$$

$$\text{Portanto: } Q = \frac{18,75 - 10,07}{2} = \frac{8,68}{2} = 4,34$$

### 5.2.3 Desvio-padrão ( $\sigma$ )

O desvio-padrão é a medida de variabilidade de mais larga aplicação nos trabalhos estatísticos. Há três fórmulas: (1) para dados não tabulados; (2) para dados tabulados, processo longo; (3) para dados tabulados, processo abreviado.

#### 5.2.3.1 DADOS NÃO TABULADOS

Deduz-se o desvio-padrão de dados não tabulados da mesma fórmula já apresentada no item 5.2:

$$\sigma = \sqrt{\frac{\sum (X_1 - M)^2}{N}}$$

$\sigma$  = desvio-padrão  
 $\Sigma$  = soma  
 $X_1$  = valores  
 $M$  = média aritmética  
 $N$  = número de valores

*Exemplo:*

Calcular o desvio-padrão da série 5, 6, 7, 7, 8, 9.

QUADRO 7

$X_1$	$(X_1 - M)$	$(X_1 - M)^2$
5	-2	4
6	-1	1
7	0	0
7	0	0
8	1	1
9	2	4
	0	10

- 1ª - Calcular a média aritmética:  

$$M = \frac{5 + 6 + 7 + 7 + 8 + 9}{6} = \frac{42}{6} = 7$$
- 2ª - Calcular os desvios  $(X_1 - M)$ . Os desvios estão calculados na 2ª coluna do Quadro 7 (a soma dos desvios é zero).
- 3ª - Elevar os desvios ao quadrado. Isso foi feito na 3ª coluna.  $(-2)^2 = 4$ ;  $(-1)^2 = 1$  etc.
- 4ª - Efetuar a soma dos quadrados dos desvios  $\Sigma (X_1 - M)^2 = 4 + 1 + 0 + 0 + 1 + 4 = 10$  (ver 3ª coluna).

- 5ª – Substituir  $\Sigma (X_1 - M)^2$  e  $N$ , na fórmula:

$$\sigma = \sqrt{\frac{\Sigma (X_1 - M)^2}{N}} = \sqrt{\frac{10}{6}} = \sqrt{1,66} = 1,29$$

### 5.2.3.2 DADOS TABULADOS

#### a. Processo longo

Para esse cálculo tem-se a fórmula:

$$\sigma = \sqrt{\frac{\Sigma (X_1 - M)^2 n_1}{N}}$$

$\sigma$  = desvio-padrão

$\Sigma$  = soma

$X_1$  = pontos médios das classes

$M$  = média aritmética

$n_1$  = freqüência

$N$  = número de valores

Exemplo: Calcular o desvio-padrão da distribuição do Quadro 8.

QUADRO 8

Classes	$n_1$	$X_1$	$n_1 X_1$	$(X_1 - M)$	$(X_1 - M)^2$	$(X_1 - M)^2 n_1$
0  — 10	2	5	10	-26	676	1352
10  — 20	2	15	30	-16	256	512
20  — 30	8	25	200	-6	36	288
30  — 40	7	35	245	4	16	112
40  — 50	4	45	180	14	196	784
50  — 60	2	55	110	24	576	1152
	25		775			4200

O 1º e o 2º passos serão dedicados ao cálculo da média aritmética. Do 3º passo em diante, ter-se-á o cálculo do desvio-padrão propriamente dito.

- 1º – Organizar uma coluna com pontos médios (3ª coluna).
- 2º – Calcular a média aritmética. Para isso, precisa-se multiplicar os pontos médios pelas freqüências ( $X_1 n_1$ ), conforme a 4ª coluna. A soma desses produtos é 775. Então:

$$M = \frac{\Sigma X_1 n_1}{N} = \frac{775}{25} = 31$$

- 3º – Calcular os desvios ( $X_1 n_1$ ). De cada ponto médio, subtrai-se a média aritmética (5ª coluna): (5 - 31) = -26; (15 - 31) = -16; (25 - 31) = -6; e assim por diante.
  - 4º – Elevar os desvios ao quadrado:  $(X_1 - M)^2$ . O quadrado de -26 é 676; de -16 é 256; de -6 é 36; de 4 = 16; e assim por diante (6ª coluna).
  - 5º – Multiplicar os quadrados dos desvios pelas freqüências:  $(X_1 - M)^2 n_1$ : 676 × 2 = 1352; 256 × 2 = 512; 36 × 8 = 288; 16 × 7 = 112; e assim por diante (7ª coluna).
  - 6º – Somar a 7ª coluna:  $\Sigma (X_1 - M)^2 n_1$ . O resultado dessa soma é 4200.
  - 7º – Substituir  $\Sigma (X_1 - M)^2 n_1$  e  $N$  na fórmula:
- $$\sigma = \sqrt{\frac{\Sigma (X_1 - M)^2 n_1}{N}} = \sqrt{\frac{4.200}{25}} = \sqrt{168} = 12,96$$

#### b. Processo abreviado

Para o cálculo do desvio-padrão pelo processo abreviado, tem-se a fórmula:

$\sigma$  = desvio-padrão

$h$  = amplitude de classe

$\Sigma$  = soma

$d_1$  = desvio ao redor de  $A$ , valor arbitrariamente escolhido

$n_1$  = freqüências

$N$  = número de valores

$$\sigma = h \times \sqrt{\frac{\Sigma d_1^2 n_1}{N} - \left( \frac{\Sigma d_1 n_1}{N} \right)^2}$$

Exemplo: Calcular o desvio-padrão da distribuição do Quadro 9, pelo processo abreviado.

QUADRO 9

Classes	$n_1$	$d_1$	$d_1 n_1$	$d_1^2 n_1$
0  — 10	2	-2	-4	8
10  — 20	2	-1	-2	2
20  — 30	8	0	0	0
30  — 40	7	1	7	7
40  — 50	4	2	8	16
50  — 60	2	3	6	18
	25		15	51

- 1ª – Achar os desvios ( $d_1$ ), tal como foi feito no cálculo da média aritmética pelo processo abreviado (ver 5.1.2.1.b). Em correspondência à classe que tem maior frequência, faz-se  $d_1 = 0$ , depois, de baixo para cima, coloca-se  $-1$ ,  $-2$ , e de cima para baixo  $1$ ,  $2$  e  $3$  (3ª coluna).
- 2ª – Multiplica-se o desvio ( $d_1$ ) pela frequência ( $n_1$ ) obtendo:  $2 \times -2 = -4$ ;  $2 \times -1 = -2$ ;  $2 \times 0 = 0$ ;  $7 \times 1 = 7$  etc. (4ª coluna).
- 3ª – Somando esses produtos ( $\Sigma d_1 n_1$ ). Na coluna de  $d_1 n_1$ , (4ª coluna), têm-se dois números negativos ( $-4$  e  $-2$ ) e três números positivos ( $7$ ,  $8$  e  $6$ ). Fazendo-se a soma dos positivos menos a soma dos negativos, tem-se:  $(7 + 8 + 6) - (4 + 2) = 21 - 6 = 15$ .  
Portanto,  $d_1 n_1 = 15$

- 4ª – Dividir essa soma por  $N$ :

$$\frac{\Sigma d_1 n_1}{N} = \frac{15}{25} = 0,6$$

- 5ª – Elevar ao quadrado:

$$\left( \frac{\Sigma d_1 n_1}{N} \right)^2 = (0,6)^2 = 0,36$$

Desta forma, tem-se calculada a expressão que, na fórmula, aparece entre parênteses.

Calcula-se agora:  $\frac{\Sigma d_1^2 n_1}{N}$

- 6ª – Calcular  $d_1 n_1$ . Na 3ª coluna tem-se  $d_1$  e na 4ª coluna,  $d_1 n_1$ . Se os números da 3ª coluna forem multiplicados pelos números correspondentes da 4ª coluna, obtém-se  $d_1^2 n_1$ . Realmente  $d_1 \times d_1 n_1 = d_1^2 n_1$ .

Foi assim que se obteve a 5ª coluna:  $(-2 \times -4) = 8$ ;  $(-1 \times -2) = 2$ ; e assim por diante.

- 7ª – Somar esses produtos:  $\Sigma d_1^2 n_1$ . Realizando a soma da 5ª coluna, obtém-se:  $51$ .
- 8ª – Dividir essa soma por  $N$ :  
 $\frac{\Sigma d_1^2 n_1}{N} = \frac{51}{25} = 2,04$
- 9ª – Fazer as substituições na fórmula. Já se conhece  $h = 10$ .

$$\frac{\Sigma d_1^2 n_1}{N} = 2,04 \left( \frac{\Sigma d_1 n_1}{N} \right)^2 = 0,36$$

Logo, pode-se substituí-los na fórmula:

$$\begin{aligned} \sigma &= h \times \sqrt{\frac{\Sigma d_1^2 n_1}{N} - \left( \frac{\Sigma d_1 n_1}{N} \right)^2} = 10 \times \sqrt{2,04 - 0,36} = \\ &= 10 \times \sqrt{1,68} = 10 \times 1,296 = 12,96 \end{aligned}$$

### 5.3 COMPARAÇÃO DE FREQUÊNCIAS

As cifras absolutas, em Estatística, às vezes, são pouco significativas, surgindo a necessidade de transformar os valores absolutos em relativos. Trabalhar com cifras muito grandes também pode dificultar a sua compreensão e comparação. Por isso, é comum utilizar expressões adequadas em relação a certas grandezas no tempo. São elas: razão, proporção, percentagem e taxa.

#### 5.3.1 Razão

Razão é um método comum e simples para se compararem frequências ou quocientes. “Razão é um meio indicado ou um quociente que relaciona o tamanho de um número a outro” (Beltrão, 1972:440). Sua função é atuar como medida relativa, possibilitando a comparação de números diferentes. A razão, então, seria a relação entre dois quocientes. Pode ser escrito de duas formas:

a. quocientes indicadores:  $25 : 10 = \frac{25}{10} = 5 : 2$

b. quocientes reais:  $25 : 10 = \frac{25}{10} = 2,5$

Exemplos:

- a. Como conhecer a relação da proporção em uma classe de 70 alunos, sendo 50 do sexo masculino e 20 do sexo feminino.

$$R = \frac{50}{20} = 5 : 2 \text{ ou } 2,5$$

A razão indica que para cada 5 rapazes há 2 moças.

- b. Para conhecer a relação de proporção de mortes femininas e masculinas, de várias idades, uma série de razões será instituída:



MORTES POR CENTENA DA POPULAÇÃO DE ACORDO COM O SEXO, SEGUNDO A IDADE

Idade	Homens	Mulheres	Razão de homem para mulher
0 — 5	17,2	13,6	1,26
5 — 10	2,4	1,7	1,41
10 — 15	1,5	1,2	1,25
15 — 20	2,4	1,9	1,26

Os algarismos da 4ª coluna são quocientes. Poderiam ser expressos como 172 a 136; 24 a 17; 15 a 12 e 24 a 19. Porém, o emprego de quocientes reais é mais útil, pois reduz os algarismos da direita a um em cada casa, facilitando a comparação. A 4ª coluna é obtida empregando-se a fórmula:

$$172 : 136 = \frac{17,2}{13,6} = 1,26$$

$$24 : 17 = \frac{2,4}{1,7} = 1,41 \text{ etc.}$$

Não querendo utilizar a razão por quociente, podem-se comparar os valores por meio da proporção. Para se conseguir a proporção, obtém-se uma fração cujo numerador é uma das duas frequências observadas e o denominador a soma das duas:

$$\frac{17,2}{17,2+13,6} = \frac{17,2}{30,8} = 0,558$$

$$\frac{2,4}{2,4+1,7} = \frac{2,4}{4,1} = 0,585 \text{ etc.}$$

Quando as proporções são expressas em múltiplos de 100, representam percentagens.

No exemplo acima, as mortes entre as idades de 0 a 4, de indivíduos do sexo masculino, correspondem a 55,8 por cento; entre 5 e 9 anos, a 58,5 por cento e assim por diante.

“A escolha entre razão, proporção ou porcentagem, para a análise de dados é uma questão de pura preferência e depende da maneira como o pesquisador comunica seus resultados” (Goode e Hatt, 1969:441).

### 5.3.2 Proporção

A proporção é a igualdade de duas razões. Esta medida se constitui em uma fração cujo numerador é uma das duas frequências observadas e o denominador a soma das frequências observadas.

*Exemplo:* Alunos que ingressaram em uma faculdade, no ano de 1981, num total de 105 (70 do sexo masculino e 35 do sexo feminino).

Obtém-se a proporção de um e de outro sexo aplicando a seguinte fórmula:

$$P = \frac{(A)}{N} \text{ ou } P' = \frac{(a)}{N}$$

*P* = proporção de rapazes  
*P'* = proporção de moças  
*A* = número de rapazes  
*a* = número de moças  
*N* = total de alunos

Fazendo-se a substituição na fórmula, tem-se:

$$P = \frac{70}{105} = 0,67 \text{ (proporção de rapazes)}$$

ou

$$P' = \frac{35}{105} = 0,33 \text{ (proporção de moças)}$$

### 5.3.3 Porcentagem

Trata-se de proporções que se multiplicam por 100 ou porção de um valor dado que pode ser determinado, desde que se saiba quanto corresponde a cada 100.

As percentagens, afirmam Goode e Hatt (1969:442):

- servem para dar forma numérica às características qualitativas;
- reduzem duas distribuições por frequência a uma base comum, simplificando muito a comparação.”

*Exemplo:* Classe de 90 alunos, sendo 58 do sexo masculino e 32 do sexo feminino. Calcular a porcentagem de cada sexo.

Aplica-se a fórmula:

$$P = \frac{A \times 100}{N}$$

ou

$$P' = \frac{B \times 100}{N}$$

P = percentagem de rapazes

P' = percentagem de moças

A = número de rapazes

B = número de moças

N = total de alunos

Substituindo-se a fórmula pelos números correspondentes, tem-se:

$$P = \frac{58 \times 100}{90} = \frac{5800}{90} = 64,4\%$$

$$P' = \frac{32 \times 100}{90} = \frac{3200}{90} = 35,6\%$$

Conclui-se que, do total de alunos (90), 64,4% são do sexo masculino e 35,6% do sexo feminino.

Embora a percentagem ajude na comunicação, devido à simplificação, pode conduzir a erros, se os dados significativos não forem evidenciados. Daí a importância da apresentação, nas tabelas de percentagens, dos números brutos que elas representam, indicando, dessa maneira, a base utilizada para o cálculo da percentagem.

Entretanto, há casos em que podem surgir dificuldades:

- a. Quando a tabulação inclui categorias residuais ou mais de uma dimensão.

Exemplos:

1. Residuais. Se for indagado dos entrevistados se são favoráveis à institucionalização do aborto, e forem obtidas as seguintes respostas:

Sim .....	97
Não .....	78
Indecisos .....	44
Não responderam .....	31
TOTAL .....	250

as percentagens podem ser baseadas na "amostra total", quando então os "a favor" totalizarão 38,8%; se forem tomados como base "todos os que responderam" (219), a percentagem dos favoráveis será 44,3%; e, finalmente, fundamentando-se nos que "omitiram uma opinião" (75), os favoráveis alcançarão a percentagem de 55,4.

2. Mais de uma dimensão. Uma indagação apresentada a 250 entrevistados, possibilitando mais de uma opção nas respostas, traz um re-

sultado diferente dos residuais; por exemplo: "quais são os seus programas prediletos na televisão":

Novela .....	101
Noticiário .....	37
Esportes .....	48
Filmes .....	112
Humorismo .....	63
Outros .....	26
Não responderam .....	11
TOTAL .....	398

Possibilidades em relação às percentagens:

- 250 = 100%. Neste caso, os que gostam de novelas totalizam 40,4%. Assim procedendo, o total das percentagens ultrapassa 100%.
- 387 = 100% (387 é o número de respostas obtidas). Agora, os que preferem novelas equivalem a 26,1%.

- b. Quando se usa tabulação de dupla entrada.

Exemplo:

MORTES DEVIDAS AO CÂNCER NOS ESTADOS UNIDOS, POR RAÇA (GOODE E HATT, 1969:444)

Raça	Causa de Morte		
	Câncer	Todas as outras	Total
Branca	139.627	1.055.804	1.195.431
Negra	9.182	169.391	178.573
TOTAL	148.809	1.225.195	1.374.004

As percentagens podem ser representadas de duas maneiras:

1ª maneira

Raça	Câncer	Todas as outras	Total
Branca	93,8	86,2	87,0
Negra	6,2	13,8	13,0
TOTAL	100,0	100,0	100,0

## 2ª maneira

Raça	Câncer	Todas as outras	Total
Branca	11,7	88,3	100,0
Negra	5,1	94,9	100,0
TOTAL	10,7	89,3	100,0

Neste caso, as porcentagens podem ser calculadas em relação aos totais vertical e horizontal.

Este fato demonstra a necessidade de uma análise do que indicam os dados para determinar o sentido em que se devem colocar as porcentagens.

No primeiro caso, os percentuais indicam:

- na primeira coluna – do total de mortos por câncer, quantos são da raça branca e quantos da negra;
- na segunda coluna – o total de mortes por outras causas;
- na terceira coluna – do total de mortes pesquisadas, quantas são da raça branca e quantas da raça negra.

Evidentemente, estas porcentagens não permitem apreender, à primeira vista, qual das duas raças pesquisadas é mais suscetível ao câncer.

No segundo caso, os percentuais apontam:

- na primeira linha, do total de mortos da raça branca investigados, quantos morrem de câncer e quantos de outras causas;
- na segunda linha consta o mesmo para a raça negra;
- na terceira linha verifica-se, do total de casos investigados, quantos faleceram de câncer e quantos de outras causas.

Portanto, no segundo caso, torna-se claro que os elementos da raça branca são mais suscetíveis ao câncer do que os da raça negra.

A diferença pode ser apontada de duas formas:

1. Quanto por cento a porcentagem dos que morreram de câncer, entre os da raça branca, é superior à porcentagem dos que morreram pela mesma causa, da raça negra? Para encontrar essa diferença, emprega-se uma regra de três, em que a porcentagem menor é equivalente a 100:

$$5,1 = 100$$

$$X = \frac{11,7 \times 100}{5,1} = 229,4$$

$$11,7 = X$$

Deste total, retirar 100, ficando a porcentagem procurada, que é 129,4%.

2. Quanto por cento a porcentagem dos que morreram de câncer, entre os de raça negra, é inferior à porcentagem dos de raça branca?

A operação processa-se equivalendo a porcentagem mais elevada a 100, empregando-se também a regra de três:

$$11,7 = 100$$

$$X = \frac{5,1 \times 100}{11,7} = 43,6$$

$$5,1 = X$$

Deve-se subtrair este total de 100, aparecendo a porcentagem real, que é igual a 56,4%.

- c. Quando se usa tabulação com mais de duas entradas.

GRAU DE ORGANIZAÇÃO DA FAMÍLIA MIGRANTE NO QUE CONCERNE A PROGRAMAÇÃO PARA OS GASTOS, SEGUNDO O TEMPO MÉDIO DE PERMANÊNCIA EM CADA LOCALIDADE

Tempo Médio de Permanência	Programação para os Gastos					
	Família Organizada		Família Desorganizada		Total (1)	
	N	%	N	%	N	%
0  — 2	2	10,00	18	90,00	20	100,00
2  — 6	15	35,71	27	64,29	42	100,00
6  — 10	14	42,42	19	57,58	33	100,00
10  — 14	16	61,54	10	38,46	26	100,00
14  —	15	62,50	9	37,50	24	100,00
TOTAL (2)	62	42,76	83	57,24	145	100,00

Tempo Médio de Permanência	Programação para os Gastos					
	Família Organizada		Família Desorganizada		Total (1)	
	N	%	N	%	N	%
0  — 2	2	3,23	18	21,69	20	13,79
2  — 6	15	24,19	27	32,53	42	28,97
6  — 10	14	22,58	19	22,89	33	22,76
10  — 14	16	25,81	10	12,05	26	17,93
14  —	15	24,19	9	10,84	24	16,55
TOTAL (2)	62	100,00	83	100,00	145	100,00

FONTE: DI GIANNI, Victalina Maria Pereira. *A convivência social do idoso francano*. Tese de Mestrado. 1990.

A hipótese que orientou a realização da pesquisa postulava que, quanto menor o tempo de permanência da família migrante por localidade, maior o grau de desorganização da família. Vários indicadores foram utilizados para dividir as famílias pesquisadas em organizadas e desorganizadas.

No exemplo em pauta, o indicador é a programação para os gastos mensais da família.

Colocando-se o total das percentagens no sentido horizontal, como ocorre na primeira Tabela, percebemos no Total (1) que a maioria das famílias desorganizadas tem um baixo grau de tempo médio de permanência: das 20 famílias que se estabeleceram por menos de dois anos em cada localidade, 90,00% são desorganizadas (sob o aspecto em pauta) e 10,00% organizadas. A primeira percentagem (família desorganizada) é 80,00% maior que a segunda (família organizada).

Por outro lado, entre as que permaneceram 14 anos ou mais, 37,50% são desorganizadas e 62,50% organizadas, isto é, 66,67% a mais, no que se refere ao grau de organização.

As percentagens colocadas em sentido vertical indicam apenas no Total (1) o percentual de famílias que permanecem por determinado tempo médio em cada localidade de migração.

Em conseqüência, a comprovação da influência do tempo médio de permanência no grau de organização da família é dada pelo Total (1) das percentagens, no sentido horizontal.

### 5.3.4 Taxas

Taxas, de acordo com Ander-Egg (1978:254), "são razões que têm um caráter dinâmico, por meio das quais se expressa a relação de uma proposição numérica existente entre duas séries de coisas". Taxa de população é, para Goode e Hatt (1969:377), "uma freqüência de ocorrências de um fato por unidade-padrão de uma população-base, durante determinado período de tempo".

São vários os tipos de taxas, sendo os mais empregados os de natalidade, de mortalidade, de nupcialidade, de migração e de crescimento.

#### 5.3.4.1 TAXA DE NATALIDADE (OU DE FECUNDIDADE)

É a freqüência de nascimentos em dada população, durante um ano. Também denominada coeficiente de natalidade, é definida como a relação entre o total de nascidos vivos e o número total de pessoas que a formam.

Para o cálculo da taxa de natalidade emprega-se a fórmula:

$$N = \frac{nv}{P} \times 1.000$$

$N$  = natalidade  
 $nv$  = nascidos vivos  
 $P$  = população

Tanto para investigação de taxas de natalidade quanto para as de mortalidade e de nupcialidade o procedimento é o mesmo.

#### 5.3.4.2 TAXA DE CRESCIMENTO

A fórmula básica usada para medir a taxa de crescimento é a seguinte:

$$P_2 = P_1 + (B - D) + (IM - OM)$$

$P_2$  = população em determinado ano  
 $P_1$  = população do ano anterior  
 $B$  = total de nascidos vivos  
 $D$  = total de mortes  
 $IM$  = total da população imigrante  
 $OM$  = total da população emigrante

As taxas podem ser: bruta, específica e padronizada.

- Taxa Bruta.** É aquela em que o cálculo é realizado sobre o total da população.
- Taxa Específica.** Refere-se especialmente ao aumento ou diminuição da população, tendo como base o número de nascimentos e/ou de mortes. Baseia-se sobre uma população específica (idade ou sexo), para que a medida de fertilidade ou de mortalidade seja mais precisa. Segundo Beltrão (1972:147), as "taxas de fecundidade (específicas) apresentam os nascimentos, não em confronto com toda a população, mas em relação com o total ou uma parte da população feminina em idade de procriação". Esse procedimento evita erros procedentes da diferença na estrutura etária da população. Para calcular a taxa de nascimento, deve-se eliminar os homens, já que só as mulheres concebem filhos, e levar em consideração o fato de que elas não são férteis em todas as idades.

Fórmula empregada:

$$\text{Taxa anual de crescimento} = \frac{\text{Total dos nascidos vivos}}{\text{Total de mulheres na faixa etária de 15 a 45 anos}} \times 1.000$$

- Taxa padronizada.** Quanto a taxa de mortalidade está relacionada com uma distribuição por idade e/ou idade e sexo, em determinada área geográfica, durante determinado ano.

Fórmula empregada:

$$\text{Taxa anual de mortalidade por grupo de idade} = \frac{\text{Mortes em uma faixa etária, em uma área geográfica, durante certo ano}}{\text{População do mesmo grupo de idade, na mesma área, na metade desse ano}} \times 1.000$$

Para comparar a mortalidade em duas cidades dentro de um mesmo Estado, pode-se utilizar, como padrão, a população de todo o Estado (por idade ou por idade e sexo), chegando-se dessa forma a resultados diferentes.

Esse procedimento oferece uma visão mais completa da mortalidade do que os índices simples.

## 5.4 APRESENTAÇÃO DOS DADOS

Para a apresentação dos dados são utilizados cinco procedimentos: série estatística, representação escrita, representação semitabular, tabelas e gráficos.

### 5.4.1 Série Estatística

Série é toda e qualquer coleção de dados referentes a uma mesma ordem de classificação. Os dados de uma série são também denominados itens ou termos da série e sua classificação atende a quatro modalidades principais, que podem caracterizar um fato em observação: tempo, lugar, categoria e intensidade.

Há quatro tipos de série: temporal, geográfica, categórica e ordenada.

#### 5.4.1.1 TEMPORAL, CRONOLÓGICA OU MARCHA

É a série em que os dados são distribuídos de acordo com o tempo em que se produziram, permanecendo fixos os locais e a categoria. Tem a finalidade de "analisar o comportamento de uma variável em sucessivos intervalos de tempo".

*Exemplo:*

**CRESCIMENTO DA POPULAÇÃO BRASILEIRA**  
Série Cronológica

Ano	População
1872	9.930.478
1890	14.333.915
1900	17.438.434
1920	30.635.605
1940	41.236.315
1950	51.944.397
1960	70.191.370
1970	93.139.037
1980	119.002.706
1991	146.825.475
1996	157.079.573

FONTE: IBGE. 1996. Contagem da População de 1996 e Censos Demográficos.

#### 5.4.1.2 GEOGRÁFICA, TERRITORIAL OU REGIONAL

Aqui, os dados são distribuídos por regiões, fixos o tempo e as categorias.

*Exemplo:*

**POPULAÇÕES POR REGIÕES DO BRASIL**  
Série Regional

Regiões	População
Sudeste	67.003.069
Nordeste	44.768.201
Sul	23.516.730
Norte	11.290.093
Centro-Oeste	10.501.480

FONTE: IBGE. 1996. Contagem da População de 1996.

#### 5.4.1.3 SÉRIE CATEGÓRICA OU ESPECIFICADA

Nesta, os dados são distribuídos de acordo com espécies ou categorias, permanecendo fixos o tempo e o local.

*Exemplos:*

### 1. RELAÇÃO ENTRE IDADE MENTAL E NOTAS ESCOLARES Série Especificada

Notas Escolares	Idade Mental			Total
	Retardados	Normais	Adiantados	
Fracas	29	17	1	47
Regulares	25	83	21	129
Boas	0	13	31	44
Ótimas	0	9	39	48
TOTAL	54	122	92	268

### 2. ESTATURA DE ALUNOS EM UMA CLASSE Série Especificada

Estatura	Alunos
Baixos	8
Médios	25
Altos	7
TOTAL	40

#### 5.4.1.4 ORDENADA OU DISTRIBUIÇÃO DE FREQUÊNCIA

É aquela cuja ordem de classificação é formada pelas intensidades ou modalidades de um atributo quantitativo ou qualitativo.

*Exemplo:*

#### ESTATURA DE ALUNOS EM UMA CLASSE

Estatura cm	Frequência
150 — 155	6
155 — 160	11
160 — 165	15
165 — 170	8
TOTAL	40

#### 5.4.2 Representação Escrita

A representação escrita consiste em apresentar os dados coletados em forma de texto. Hoje, é a modalidade mais comum em documentos, livros e informações.

*Exemplo:* Atualmente, a região Sudeste possui 51,8% do total dos estabelecimentos industriais do Brasil; 69,7% do total do pessoal que trabalha na indústria do Brasil; 78,3% do total do valor da produção industrial do Brasil.

**Fonte:** Secretaria da Economia e Planejamento do Estado de São Paulo, 1977.

#### 5.4.3 Semitabela

Este procedimento é empregado quando são incorporadas cifras a um texto, ressaltando-as, de forma a facilitar sua comparação.

*Exemplo:* “Na Suíça, há quatro idiomas nacionais exatamente iguais perante a lei. O Censo de 1941 demonstra (Ander-Egg, 1978:263):

3.097.059 falam alemão	(72,6%)
884.668 falam francês	(20,7%)
220.530 falam italiano	(5,2%)
46.456 falam reto-romano	(1,1%).”

#### 5.4.4 Tabela ou Quadro

Tabela é uma forma de disposição gráfica das séries, de acordo com determinada ordem de classificação. Seu objetivo é sintetizar os dados de observação, tornando-os mais compreensivos. Visa “ajudar o investigador para que distinga semelhanças, diferenças e relações mediante a clareza e o relevo que a distribuição lógica presta à classificação” (Ander-Egg, 1978:150).

Na tabela, os dados numéricos são ordenados em filas ou colunas com as especificações equivalentes à sua natureza.

Assim como as séries, as tabelas são classificadas levando-se em consideração quatro modalidades principais: tempo, lugar, categoria e intensidade.

##### 5.4.4.1 ELEMENTOS DA TABELA

Os elementos da tabela ou quadro são:

- Título:** a palavra *tabela* ou *quadro* é escrita em caracteres maiúsculos, seguida de um número identificador, preferencialmente em algarismos romanos. Em seguida, vem a legenda da tabela (**nome**) em uma, duas ou três linhas, sempre centralizadas em relação à largura da tabela. Exemplo:

**TABELA III  
ESTRUTURA DE ALUNOS EM UMA CLASSE**

**QUADRO V  
SÍNTESE DAS INSTITUIÇÕES, CARACTERÍSTICAS E RAÍZES DOS  
SISTEMAS ECONÔMICOS CONTEMPORÂNEOS**

2. *Corpo*: composto de:

- Cabeçalho: linha horizontal que precede o preenchimento com números da tabela.
- Coluna indicadora: fica à esquerda no quadro ou tabela. Ela também é composta de palavras separadas dos números da tabela por linhas verticais.
- Conteúdo propriamente dito da tabela: disposto por linhas e colunas.

3. *Notas*: qualquer explicação referente à tabela é indicada por asteriscos.

Da mesma forma, as notas colocadas logo após o quadro ou tabela iniciarão com um, dois, três asteriscos, conforme as chamadas que aparecem na tabela ou quadro.

4. *Fonte*: pode ser de outro autor ou oriunda da pesquisa do próprio autor.

**Observações:** 1. uma tabela não é fechada lateralmente por convenção internacional, cujo significado é: existe uma causa anterior que não foi pesquisada e existe uma consequência posterior que não foi examinada. Ou seja: como a tabela é uma representação de um elo de uma cadeia causal de fenômenos, sua representação com laterais abertas indica a exclusão de variável(is) anterior(es) e posterior(es).

2. Diferencia-se tabela de quadro pelos seguintes fatores:

- Geralmente, a tabela é composta de dados da própria autoria e o quadro é transcrição de dados obtidos por outra(s) pessoa(s).
- A tabela é sempre composta por números, ao passo que o quadro pode conter apenas palavras.

3. Os quadros são fechados lateralmente porque:

- São transcrições de dados de outra pessoa, e neste caso cita-se apenas o elo da cadeia causal examinado pelo autor. Se composto de texto, ele não pressupõe antecedentes nem consequentes.

- Se se compõe apenas de palavras, elas podem ser quer da própria autoria quer transcritas de uma fonte.

*Exemplo:*

Título	POPULAÇÃO POR REGIÕES DO BRASIL – 1996			
	Coluna Matriz		Corpo	
Cabeçalho	Regiões	População	% da População Total	Densidade Demográfica
	Sudeste	67.003.069	42,6	72,44
	Nordeste	44.768.201	28,5	28,91
	Sul	23.516.730	15,0	40,70
	Norte	11.290.093	7,2	2,92
	Centro-Oeste	10.501.480	6,7	6,60
	Brasil	157.079.573	100,0	18,45 hab/km <sup>2</sup>

FONTE: IBGE, 1996 – Contagem da População de 1996.

#### 5.4.4.2 NORMAS GERAIS DE TABELAS

As tabelas devem ser designadas com clareza; o título deve apresentar o assunto da tabela, bem como todas as qualificações necessárias. No caso de serem necessárias explicações, esclarecimentos, estes devem ser apresentados no rodapé da tabela. O cabeçalho deve ser composto de expressões curtas e consistentes.

Mesmo que a apresentação esteja clara, na tabela costuma-se apresentar no texto os dados importantes, pois, se o leitor não quiser lê-la, ele poderá obter as devidas informações lendo o texto.

#### 5.4.4.3 TABELAS COMPLEXAS

Referem-se àquelas em que mais de duas dimensões devem ser simultaneamente apresentadas. A análise dos dados, diretamente relacionada com a complexidade das hipóteses ou hipótese, pode criar problemas.

*Exemplo:* Número de alunos, sexo dos alunos, nível econômico.

Uma apresentação complexa exigiria várias tabelas, tornando a apresentação confusa. A solução está na sua simplificação, o que pode ser feito remo-



vendo uma variável. Se qualquer percentagem dicotômica pode ser expressa com apenas um algarismo, é possível transformar qualquer variável tricotômica ou dicotômica simplificando a tabela.

Se 60% dos alunos de uma classe mista são do sexo masculino, já está implícito que os 40% restantes são do sexo feminino; não há, portanto, necessidade de mencionar essa variável, o que leva a uma simplificação.

*Exemplos:*

1. Tabela de uma entrada (dimensão):

TRABALHADORES DESOCUPADOS POR REGIÃO DO BRASIL

Região	Número	%
Norte	275.442	6,1
Nordeste	1.110.122	24,6
Sudeste	2.172.238	48,2
Sul	619.026	13,7
Centro-Oeste	333.005	7,4
Brasil	4.509.833	100,0

FONTE: IBGE, PNAD - 1995.

2. Tabela de duas entradas:

TRABALHADORES TEMPORÁRIOS POR ÁREA DE ATUAÇÃO, SEGUNDO A REGIÃO

Região	Área de Atuação					
	Burocrática		Produção		Total	
	N	%	N	%	N	%
São Paulo	308	72,0	220	62,2	528	67,5
ABC	52	12,1	118	33,3	170	21,7
Rio de Janeiro	68	15,9	16	4,5	84	10,8
TOTAL	428	100,0	354	100,0	782	100,0

3. Tabela de três entradas:

TRABALHADORES TEMPORÁRIOS POR ÁREA DE ATUAÇÃO E SEXO, SEGUNDO A REGIÃO

Região	Burocrática				Produção				Total			
	Masculino		Feminino		Masculino		Feminino		Masculino		Feminino	
	N	%	N	%	N	%	N	%	N	%	N	%
São Paulo	153	73,2	155	70,8	184	60,1	36	75,0	337	65,4	191	71,5
ABC	25	12,0	27	12,3	113	36,9	5	10,4	138	26,8	32	12,0
Rio de Janeiro	31	14,8	37	16,9	9	3,0	7	14,6	40	7,8	44	16,5
TOTAL	209	100,0	219	100,0	306	100,0	48	100,0	515	100,0	267	100,0

4. Tabela de quatro entradas:

TRABALHADORES TEMPORÁRIOS POR ÁREA DE ATUAÇÃO E SEXO, SEGUNDO A REGIÃO E A FAIXA ETÁRIA

Região e Faixa Etária	Burocrática				Produção				Total			
	Masculino		Feminino		Masculino		Feminino		Masculino		Feminino	
	N	%	N	%	N	%	N	%	N	%	N	%
SÃO PAULO												
12   18 anos	21	13,7	12	7,8	21	11,4	4	11,1	42	12,5	16	8,4
18   30 anos	97	63,4	116	74,8	121	65,8	22	61,1	218	64,7	138	72,2
30   45 anos	5	3,3	2	1,3	16	8,7	3	8,3	21	6,2	5	2,6
45												
ABC												
12   18 anos	1	4,0	2	7,4	3	2,6	1	20,0	4	2,9	3	9,4
18   30 anos	15	60,0	19	70,4	75	65,4	3	60,0	90	65,2	22	68,8
30   45 anos	9	36,0	4	14,8	34	30,1	1	20,0	43	31,2	5	15,6
45	-	-	2	7,4	1	0,9	-	-	1	0,7	2	6,2
RIO DE JANEIRO												
12   18 anos	-	-	-	-	1	11,1	-	-	1	2,5	-	-
18   30 anos	20	64,5	18	48,7	3	33,3	4	57,1	23	57,5	22	50,0
30   45 anos	9	29,0	18	48,6	5	55,6	3	42,9	14	35,0	21	47,7
45	2	6,5	1	2,7	-	-	-	-	2	5,0	1	2,3
TOTAL												
12   18 anos	22	10,5	14	6,4	25	8,2	5	10,4	47	9,1	19	7,1
18   30 anos	132	63,2	153	69,9	199	65,0	29	60,4	331	64,3	182	68,2
30   45 anos	48	23,0	47	21,4	65	21,2	11	22,9	113	21,9	58	21,7
45	7	3,3	5	2,3	17	5,6	3	6,3	24	4,7	8	3,0



### 5.4.5 Gráficos

A representação dos dados com elementos geométricos permite uma descrição imediata do fenômeno. Representa uma forma atrativa e expressiva, uma vez que facilita a visão do conjunto com apenas uma olhada, e possibilita ver o abstrato com facilidade.

A representação gráfica apresenta algumas limitações (Ander-Egg, 1978: 268):

- a. não pode representar tantos dados como um quadro ou tabela estatística;
- b. não permite a apreciação de detalhes;
- c. não pode dar valores exatos;
- d. requer maior tempo em sua execução do que os quadros ou tabelas;
- e. presta-se a deformações, pelas escalas utilizadas."

Há inúmeros tipos de gráficos estatísticos, mas todos eles podem formar dois grupos:

1. *Gráficos Informativos* (ou de Informação) – cujo objetivo é dar ao leitor ou ao investigador um conhecimento da situação real, atual, do problema estudado ou de interesse. Devem ser feitos com cuidado, de modo que o desenho impressione bem, tenha algo de atraente. Todavia, esse cuidado não pode ser exagerado, a ponto de prejudicar o observador na apreensão dos dados.
2. *Gráficos Analíticos* (ou de Análise) – cujo objetivo é, além de fornecer informações, oferecer ao pesquisador elementos de interpretação, cálculo, inferências e previsões.

Os gráficos devem conter o mínimo de construções e ser simples. Principais gráficos de informação:

#### *De Base Matemática*

1. Lineares:
  - a. retilíneos;
  - b. curvilíneos.
2. De superfície:
  - a. retangulares (barras ou colunas);
  - b. circulares;
    - de setores;
    - de círculos concêntricos;

- de ordenada polar;
  - de gráficos em espiral;
  - c. triangulares;
  - d. quadrangulares.
3. Estereométricos:
    - a. cúbicos;
    - b. prismáticos;
    - c. piramidais.

#### *De Base não Matemática*

4. Cartogramas:
  - a. mapas;
  - b. cartas.
5. Pictogramas.
6. Organogramas.
7. Livres ou especiais.

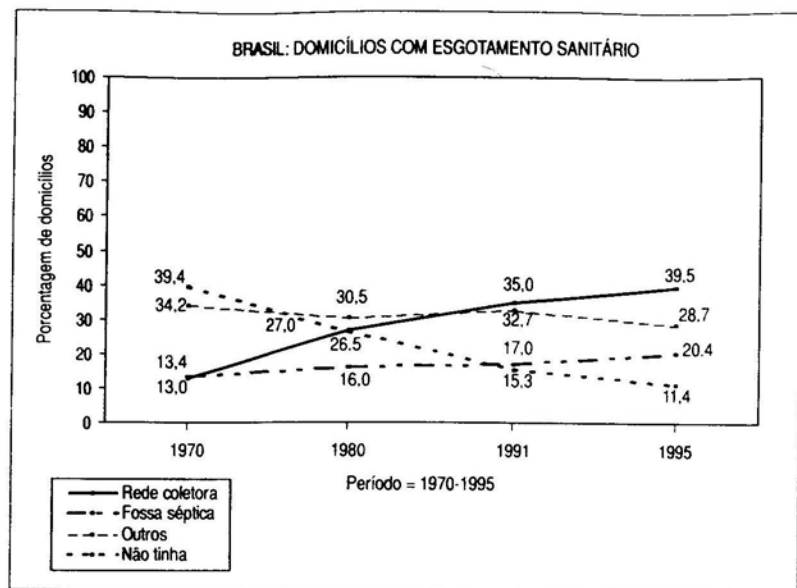
### 5.4.5.1 DE BASE MATEMÁTICA

#### *a. Gráficos lineares*

O diagrama linear é um tipo de gráfico muito simples e empregado com grande frequência. Representa alterações quantitativas sob a forma de uma linha reta ou curva, que avança pelo quadrilátero. É o tipo mais eficaz para representar as séries em marcha.

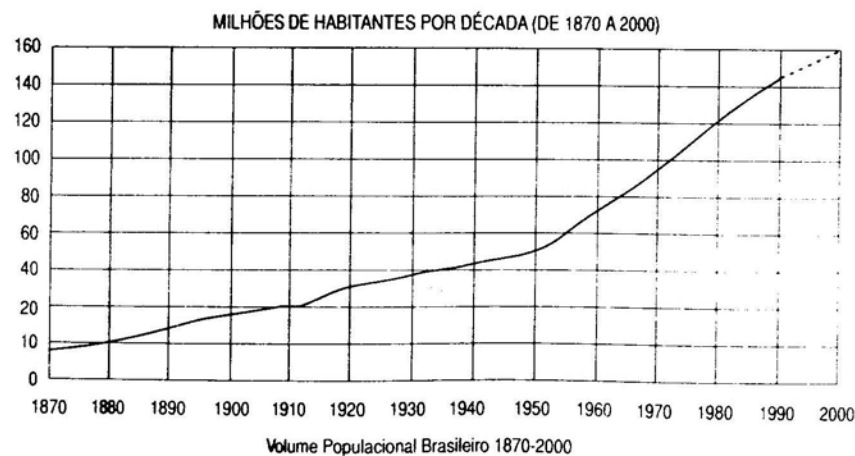
O gráfico linear apresenta uma série de variações e é largamente empregado ao lado do gráfico de colunas.

*Exemplo:*



FONTE: IBGE, PNAD, 1995.

Gráfico linear retilíneo.



FONTE: IBGE.

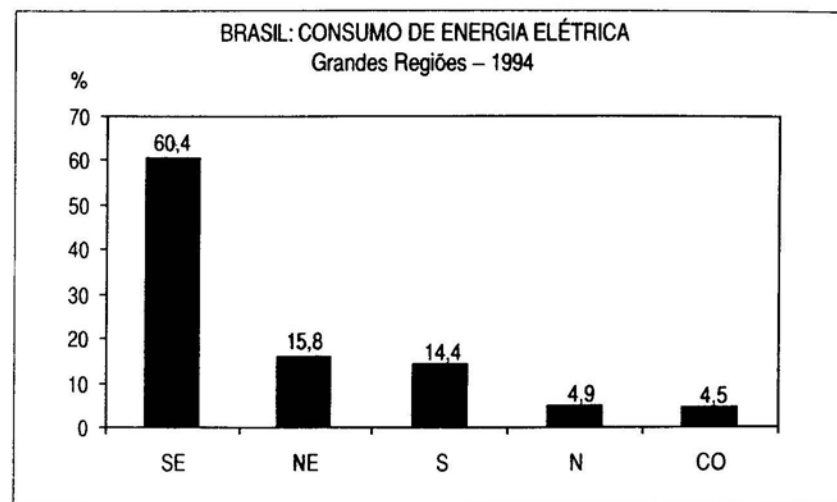
Gráfico linear curvilíneo.

## b. Gráficos de superfície

1. *Retangulares*. Também denominados de barras ou colunas, são representações formadas por retângulos alongados, de base assentada sobre uma linha horizontal ou vertical. No primeiro caso, têm-se gráficos de colunas; no segundo, os de barra.

As barras são escolhidas arbitrariamente e as alturas são proporcionais aos valores ou dados da série respectiva.

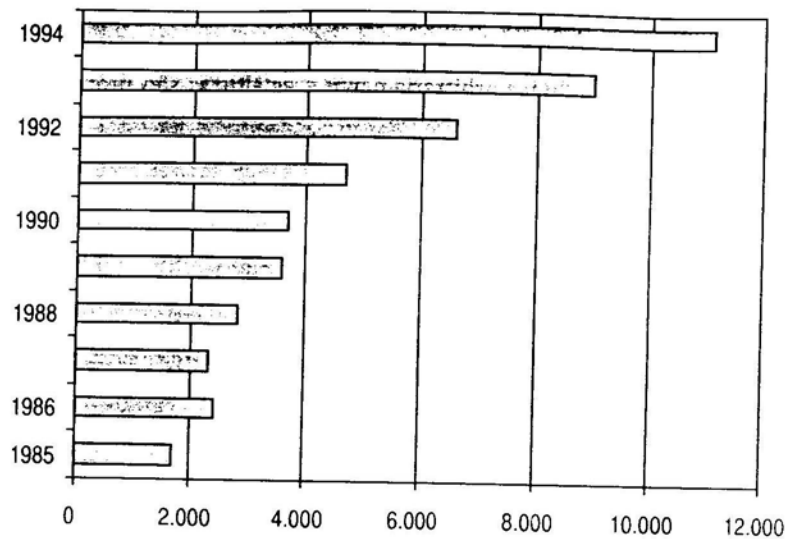
Quando as legendas forem muito extensas, convém usar o gráfico de barras.



FONTE: IBGE, Anuário Estatístico do Brasil, 1994.

Gráfico em colunas.

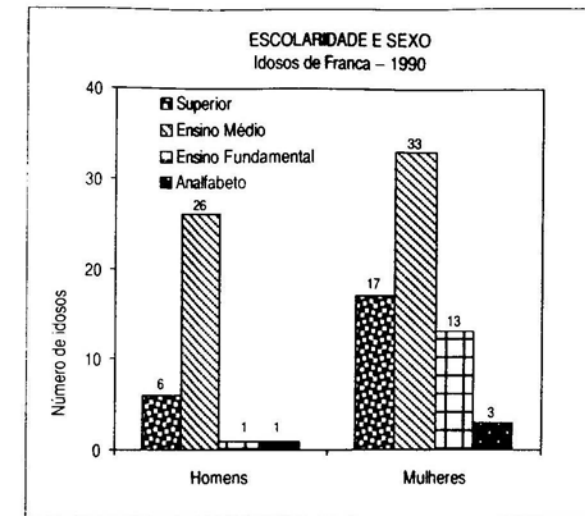
BRASIL – MERCOSUL: TOTAL DO COMÉRCIO  
(milhões de dólares)



FONTE: BRANDÃO, Antônio Salazar, PEREIRA, Lia Valls (Orgs.). *Mercosul: perspectivas da integração*. FGV, 1995.

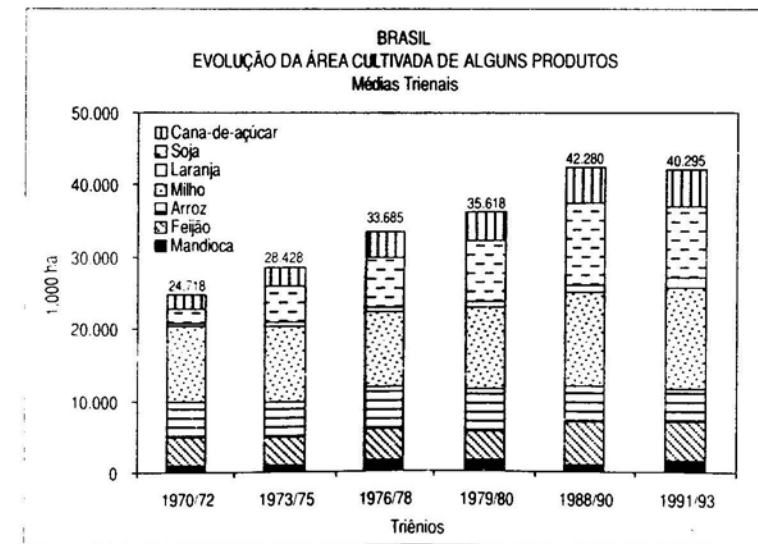
Gráfico de barras.

“O histograma é uma modalidade do diagrama retangular, no qual se representa, por um retângulo, cada classe da série, determinando a frequência de classe respectiva à altura do retângulo. No caso em que as classes sejam desiguais, deve-se introduzir uma correção na altura do retângulo correspondente” (Ander-Egg, 1978:281). Consiste, portanto, em colunas ou barras duplas que se dispõem sem espaço entre si ou com alguma separação, dependendo da exatidão, clareza e estética do conjunto.



FONTE: DI GIANNI, Victalina Maria Pereira. *A convivência social do idoso francano*. Tese de Mestrado, 1990.

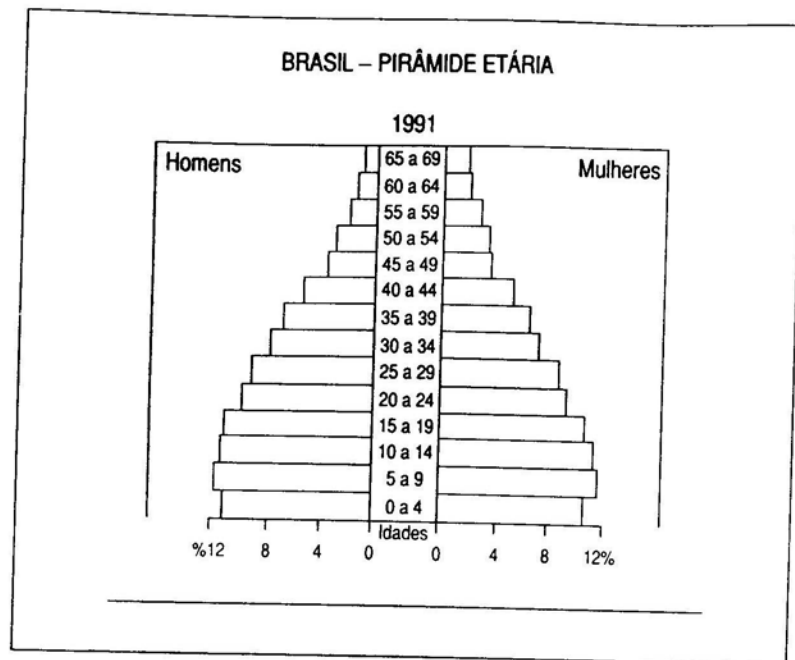
Histograma.



FONTE: MELLO, Fernando Homem de. *O problema alimentar no Brasil*. São Paulo: Paz e Terra, 1983. IBGE, Anuários Estatísticos do Brasil, 1992 a 1995.

Histograma.

De entre os diagramas retangulares, as *pirâmides de idade* constituem uma modalidade muito usada.



FONTE: IBGE

2. **Circulares.** Dos gráficos circulares, o mais utilizado, em área, é o de *setores*, que presta para confrontar as partes integrantes de um total. Os valores são dispostos num círculo, onde o total equivale a uma amplitude de 360°; a área do círculo será dividida em setores proporcionais aos acontecimentos que se quer representar. A operação matemática consiste em dividir 360° em setores proporcionais aos valores. Tem-se assim o número de graus para cada valor.

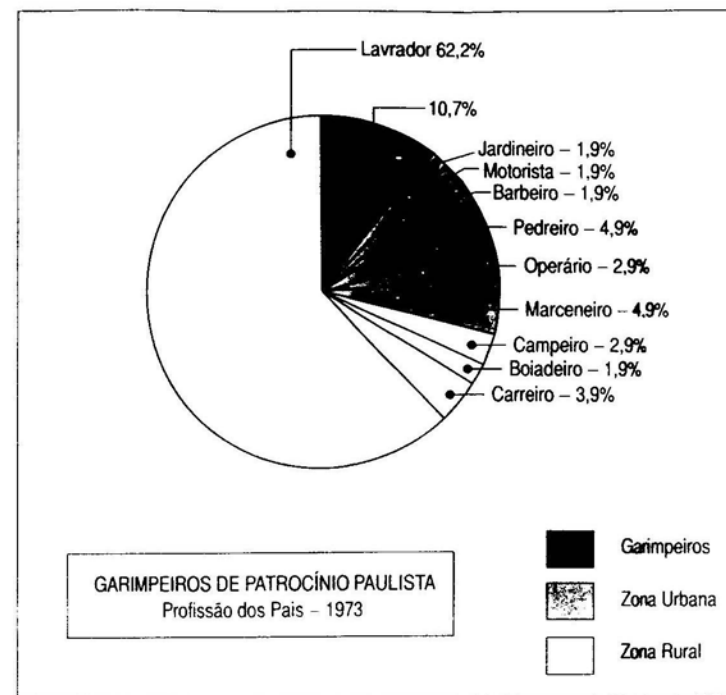
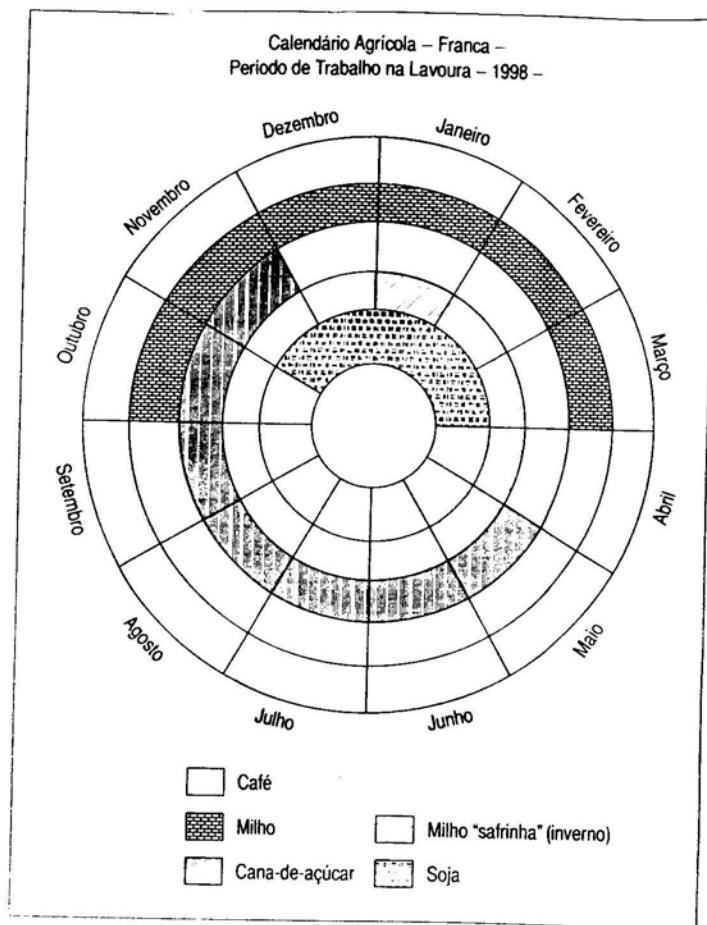


Gráfico de setores.

Organizado por Marina de Andrade Marconi – 1973 – Desenho de Valdete.

O gráfico de *círculos concêntricos* é de grande utilidade para representar um conjunto de fenômenos em épocas diferentes. Às vezes, é difícil de confeccionar, quando as amplitudes são pouco desiguais.



FONTE: Casa da Lavoura, Franca, 1998.

Gráfico de círculos concêntricos.

Os diagramas com base em ordenadas polares são mais empregados nos processos dinâmicos, de caráter cíclico, de período bem definido. As divisões são marcadas mediante círculos concêntricos, cada um com um valor determinado: 10, 20, 30, 40 etc. A representação estatística é feita por meio de um polígono irregular (linha poligonal fechada), que indica as variações no tempo ou no espaço (7:60). São *simples* quando representam apenas um fenômeno; *compostos* quando se consideram dois ou mais.

Exemplo: Extração de diamantes em um ano, em P. P.

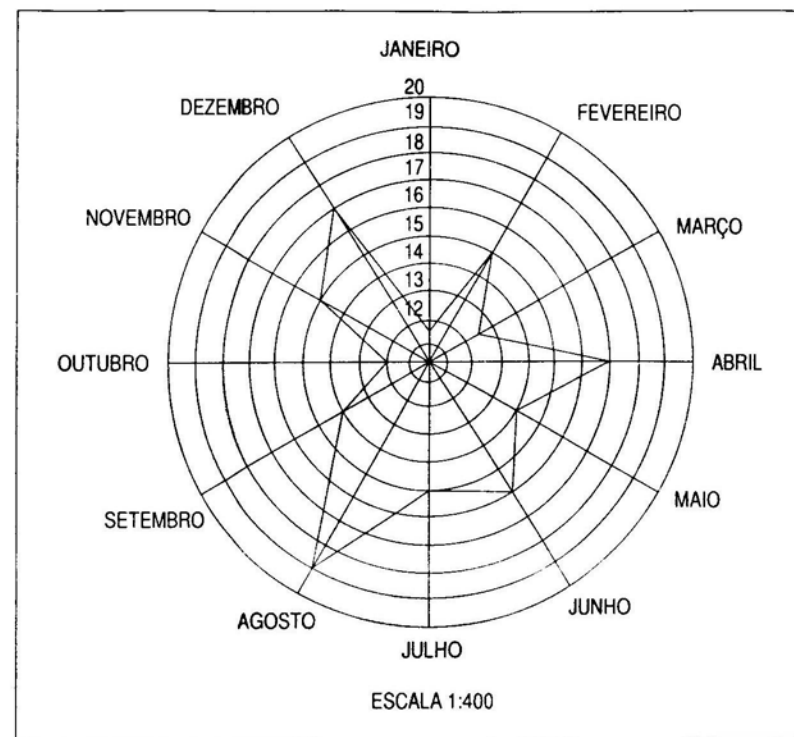
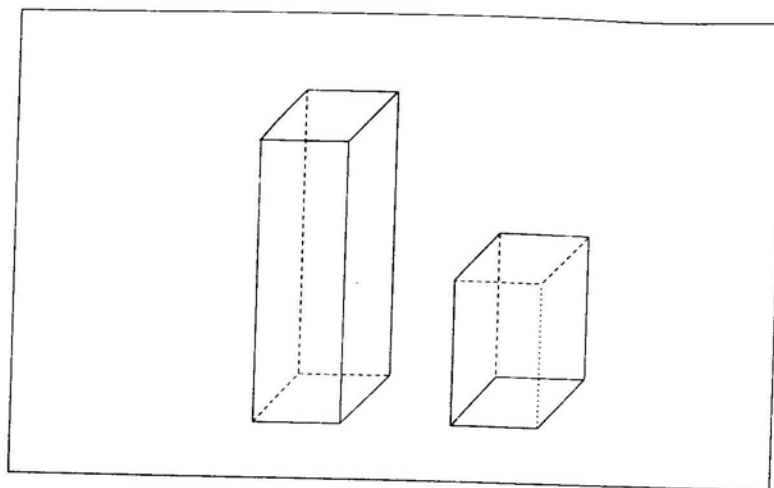


Gráfico polar.

### c. Gráficos estereométricos

Utilizados quando se deseja representar fenômenos com duas ou mais variáveis, mediante o emprego de figuras geométricas, principalmente prismas ou cubos; o menos utilizado é o de pirâmide. Em geral, representam fenômenos medidos em unidades cúbicas:  $m^3$ ,  $l^3$ , t. etc. Medem valores e suas variações no tempo e/ou no espaço.

Exemplo: Produção de feijão de um Estado, no período das águas (20 t.) e no período da seca (10 t.).



FONTE: Casa da lavoura de franca, 1998.

Gráfico estereométrico.

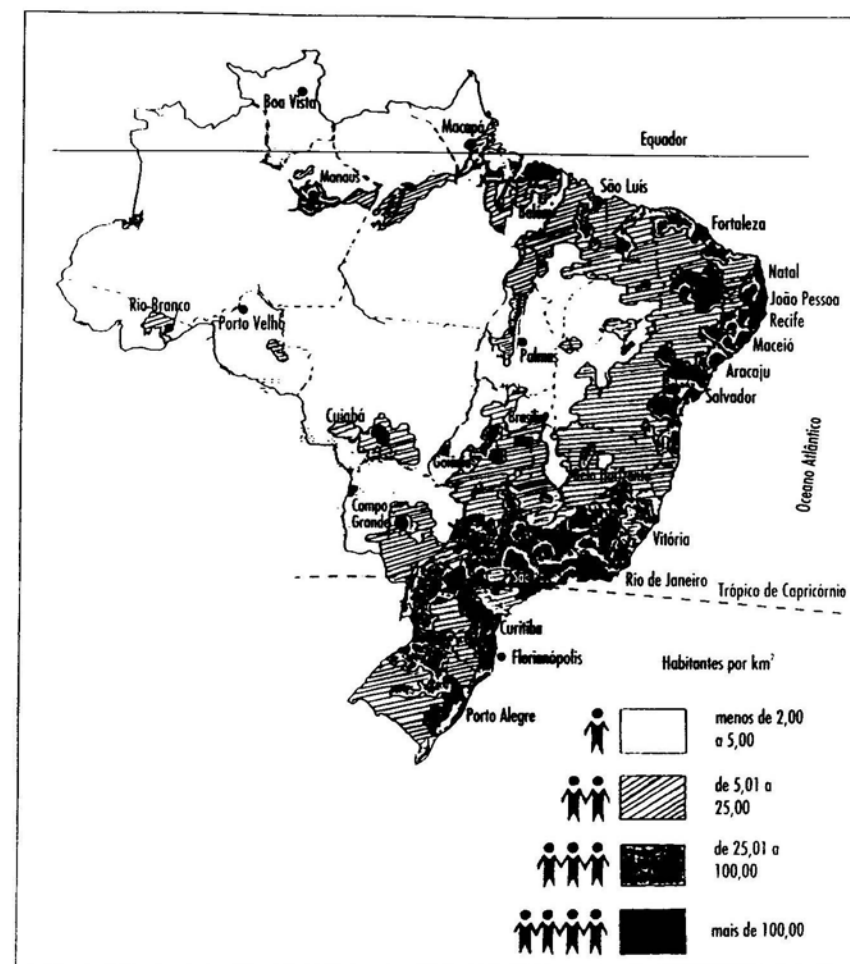
#### 5.4.5.2 DE BASE NÃO MATEMÁTICA

##### a. Cartogramas

Constituem uma associação entre mapas geográficos e as representações propriamente estatísticas. São gráficos estabelecidos sobre mapas. Mesmo não havendo base matemática, devem ser confeccionados com precisão em relação ao fenômeno representado.

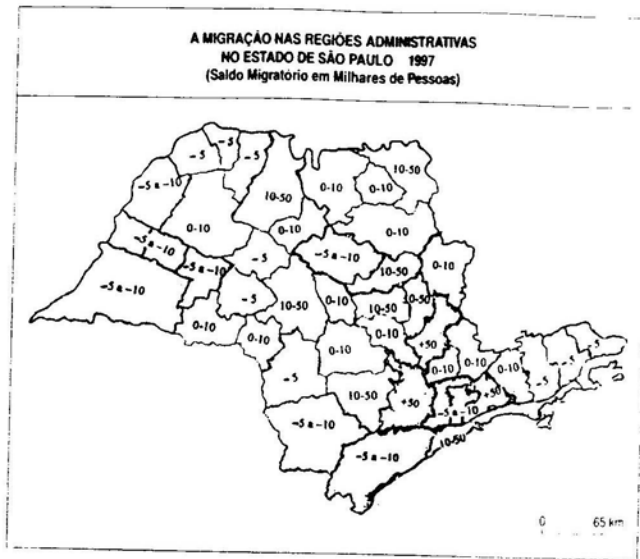
Há vários tipos de cartogramas (Hofmann, 1974:25):

- De densidade.* Quando as regiões do mapa se apresentam diferenciadas por cores ou traços mais ou menos espaçados.
- Ponteados.* A densidade progressiva das variáveis está representada, no mapa, por pontos mais ou menos numerosos, ou seja, por círculos cuja superfície seja proporcional ao número a ser representado.
- De gráficos.* Em cada região do mapa se situa um pequeno gráfico correspondendo à região, que pode ser de colunas, barras, circulares, de superfície, figurativo etc.
- Cifrados.* Em lugar de um pequeno gráfico, nada impede de se escrever um número por região, se não houver mais de duas magnitudes a representar, segundo seu valor por região.



FONTE: IBGE.

Densidade da população em 1991.

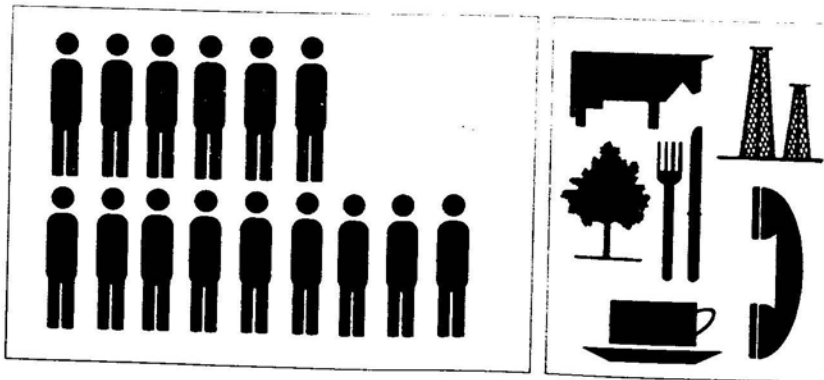


FONTE: Roberto de Toledo, in *Folha de São Paulo*, 9 mar. 1997, p. 3-6

Cartograma cifrado.

### b. Pictograma

Esse gráfico nada mais é do que uma forma artística dos gráficos de barras ou colunas. As figuras têm tamanho proporcional ao valor atribuído na série. As figuras (isótipos) representam um fenômeno, explicado pela própria natureza da figura. “Os pictogramas só devem ser utilizados para fazer comparações e não para apresentar números isolados” (Ander-Egg, 1978:320).

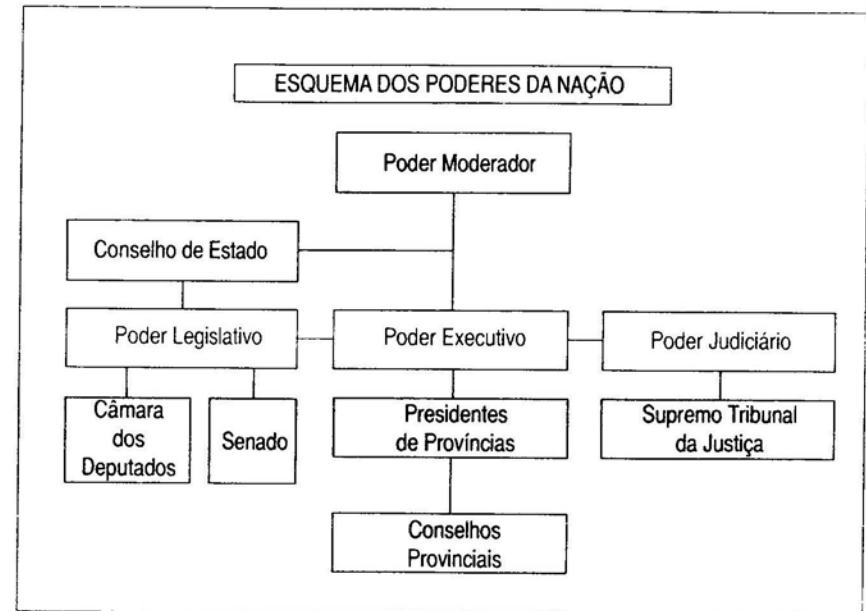


Pictogramas.

### c. Organograma

Não são gráficos no sentido próprio da palavra, porque não servem para comparar quantidades, mas para representar esquematicamente os diversos órgãos de uma administração: comercial, governamental, educacional, industrial etc.

Brasil: Constituição de 1824.



Organograma.

### d. Livres ou Especiais

São gráficos que escapam a toda regra ou norma estatística para sua construção. Para que tenham validade, é preciso observar duas regras: precisão e clareza.

#### 5.4.6 Regras para a Construção de Gráficos

Embora haja outras, as regras gerais para a construção de gráficos, apontadas por Ander-Egg (1978:303), são as seguintes:

- a. o título deve ser claro, conciso, preciso e figurar dentro dos limites da trama;
- b. todo gráfico deve ter uma legenda, para explicar e esclarecer o fenômeno representado;
- c. não se deve representar grande número de componentes, mesmo que o fenômeno seja composto, a fim de evitar confusões;
- d. escolher sempre o sistema mais adequado ao tipo de fenômeno que se deseja representar;
- e. o quadro estatístico, com os dados numéricos equivalentes, deve acompanhar o gráfico;
- f. devem-se utilizar cores contrastantes nas representações dos fenômenos compostos, para destacar o dado e favorecer a interpretação;
- g. se a apresentação gráfica abranger uma superfície muito grande, deve-se truncá-la, a fim de evitar tamanhos descomunais;
- h. a escolha da escala deve ser efetuada de maneira que as diferentes intensidades constituam valores perfeitamente adaptáveis ao gráfico."

### 5.5 OS TESTES DE HIPÓTESES COMO INSTRUMENTAL DE VALIDAÇÃO DA INTERPRETAÇÃO (*Estatística Inferencial*)

Como se viu, a apresentação e a organização descritiva adequada dos dados obtidos constituem-se em notável contribuição estatística, quando procedem do estudioso de ciências sociais. Em última instância, consistem na formulação atualizada de preocupação talvez tão antiga quanto a Humanidade, qual seja, a sistematização dos resultados dos esforços empregados.

Outras possibilidades de apresentação descritiva de dados são as medidas de posição, de dispersão, comparação de frequências e taxas.

No entanto, modernamente, o recurso mais relevante que a estatística nos fornece é o procedimento inferencial. Também conhecido como Testes de Hipóteses, consiste no instrumental metodológico que permite ao pesquisador, por exemplo, apreciar sobre a validade de expandir seus dados para amplas generalizações ou, ao contrário, verificar se esses são extremamente valiosos por diferirem do que se conhece até então a esse respeito.

Podem ser o caso, por exemplo, de o investigador, ao longo de um estudo de uma população de pré-escolares, coletar os respectivos dados antropométricos (peso e altura, entre outros), visando esclarecer tratar-se crianças com crescimento deficiente ou, pelo contrário, representativas das de nossa cidade.

Uma seqüência de passos deverá ser seguida, começando pelo conhecimento de dados populacionais com os quais se pretende fazer a comparação. Em um estudo realizado por Gonçalves (1979:64), a condução de tal questão resultou nos dados sumariados na Tabela I.

TABELA I

VALORES MÉDIOS DOS DADOS ANTROPOMÉTRICOS DOS ALUNOS ESTUDADOS E DO GRUPO DE CONTROLE (PARA A MESMA IDADE E SEXO).

Valor Antropométrico	Sexo	Grupo Estudado		Grupo de Controle	
		Média	Desvio-padrão	Média	Desvio-padrão
Peso	Masculino	24,77	± 3,53	23,56	± 3,29
	Feminino	24,19	± 6,34	23,53	± 3,59
Altura sentada	Masculino	67,80	± 3,12	67,00	± 2,70
	Feminino	66,90	± 2,94	66,53	± 2,89
Altura tronco-cefálica	Masculino	125,50	± 6,13	122,56	± 5,3
	Feminino	123,80	± 5,45	122,62	± 5,6

Componentes do grupo estudado: 57 elementos (masculino); 47 elementos (feminino).

Componentes do grupo de controle: 380 elementos (masculino); 340 elementos (feminino).

A simples inspeção da tabela mostra-nos uma diferença de valores antropométricos médios entre ambos os grupos.

A pergunta que se coloca é, portanto, "as diferenças observadas correm por conta de fatores não relevantes, ou realmente refletem realidades diferentes (digamos, crianças com crescimento normal e com crescimento comprometido)?" Aqui, entretanto, há que se recorrer a um teste estatístico.

Nessa altura, diante de todas estas percepções, iniciam-se as várias fases de sua aplicação. Já a questão enunciada pode ser formulada em termos estatísticos no sentido de se pretender verificar se as diferenças observadas são não relevantes ou efetivamente significativas. Simbolicamente teríamos:

$$H_0 : \bar{x}_a = \bar{x}_b \text{ ou}$$

$$H_a : \bar{x}_a \neq \bar{x}_b$$

em que  $\bar{x}_a$  e  $\bar{x}_b$  são as médias de cada uma das amostras.

Isso porque todo pesquisador trabalha com uma hipótese que pretende comprovar (hipótese alternativa,  $H_a$  ou  $H_1$ ). Para isso, deve contrapô-la ao já conhecido, por um imperativo científico ( $H_0$ ).



Uma vez assim definidas as hipóteses e situado o problema a ser tratado, cabe decidir qual o instrumental estatístico para resolvê-lo (eis a função primordial do estatístico).

No caso, por se tratar de comparação entre médias, o indicado é o teste *t* de Student.

### 5.5.1 Teste *t* de Student

A fórmula deste teste consiste em:

$$t = \frac{\bar{x}_a - \bar{x}_b}{\sqrt{\frac{(N_a + N_b) [(N_a - 1) V_a + (N_b - 1) V_b]}{N_a N_b (N_a + N_b - 2)}}$$

onde as representações com índice *a* são referentes a uma amostra e as com índice *b* a outra amostra e  $\bar{X}$ , *N* e *V*, respectivamente, média, tamanho da amostra e variância. Por sua vez (*N* + *N* - 2), para o caso, é o grau de liberdade.

“Via de regra, os graus de liberdade de uma estatística, objeto de cálculo, indicam o número de fatores (a partir dos quais a estatística é calculada) que podem ser alterados independentemente, sem alterar o valor da estatística. Suponhamos, por exemplo, calcular a soma de três números:

$$x + y + z = S$$

Se se fixasse o valor de *S*, poder-se-ia atribuir valor a (diga-se) *x* e *y*, mas uma vez determinados *x* e *y*, estará determinado *z*. Ou seja, se a soma é fixada, e todas, menos uma das variáveis são determinadas, também esta última variável estará determinada. Portanto, todas as variáveis menos uma são ‘livres’. Neste caso, portanto, há dois graus de liberdade. Em geral, se há *n* variáveis, e uma equação que define a estatística, há *n* - 1 graus de liberdade” (Ackoff, 1967:246-247).

A variância é uma medida de dispersão. Muitos são os meios de medir a dispersão. Um deles seria tomar a soma dos desvios da média. No entanto, a soma dos desvios da média é sempre igual a zero. Por exemplo, a média aritmética de 2, 4, 6, 8, 10 e 12 é 7; os desvios em relação à média são -5, -3, -1, 1, 3 e 5; o total destes desvios é igual a zero. Tal dificuldade pode ser contornada tornando positivo o sinal de todos os desvios, ao elevá-los ao quadrado. A média desses quadrados é exatamente a variância.

Voltando à Tabela I, calcular-se *t* para cada valor antropométrico. Tome-se como exemplo, para efeitos de cálculo, o peso para alunos do sexo masculino. Ter-se-ia, então:

$$t = \frac{24,77 - 23,56}{\frac{437 [(56 \times 12,46) + (379 \times 10,82)]}{57 \times 380 \times (57 + 380 - 2)}}$$

$$t = \frac{1,21}{\frac{437 [697,76 + 4.100,78]}{9.422.100}}$$

$$t = \frac{1,21}{\sqrt{\frac{437 \times 4.789,54}{9.422.100}}} = \frac{1,21}{\sqrt{\frac{2.096.961,9}{9.422.100}}}$$

$$t = \frac{1,21}{\sqrt{0,2225}} = \frac{1,21}{0,47} = 2,574$$

Da mesma forma obter-se-ão, para os demais valores antropométricos, 1,064; 0,205; 0,822; 3,818 e 1,395, respectivamente.

Que significam tais números? Eles são, a seguir, comparados com um valor de *t* acima do qual se rejeita  $H_0$  e se aceita  $H_a$  e abaixo do qual se procede contrariamente. Tal valor assim tão relevante, tão crítico, é fornecido facilmente pelos manuais correntes de estatística, nos quais ele é detectado a partir das entradas chamadas grau de liberdade e  $\alpha$  (alfa). Daí ter-se mencionado, ao apresentar a fórmula estatística, este primeiro elemento. Já o alfa, ou erro de tipo I, consiste no risco que o pesquisador dispõe a correr, de rejeitar  $H_0$  quando esta é verdadeira. Habitualmente chega até a 5%, excepcionalmente a 10%, quando se trata de experimento extremamente caro, pioneiro ou preliminar.

Essa probabilidade, ou risco, denominado  $\alpha$ , é geralmente especificada antes da extração de qualquer amostra de modo que os resultados obtidos não influenciam a escolha.

De fato, para  $\alpha = 5\%$  e 385 graus de liberdade, obtém-se um *t* crítico de 1,960, indicando, assim, que a diferença significativa observada entre ambas as amostras se refere ao peso e à altura tronco-cefálica das meninas. Portanto, dos seis valores apresentados, apenas esses dois diferem significativamente entre ambos os grupos analisados, contrariamente às eventuais conclusões aventadas pela simples inspeção da Tabela I.

**Expressão geral** - A fórmula mais comum de Student, onde:

$$t = \frac{\bar{X} - u}{S_s} = \sqrt{N - 1}$$

$\gamma$  = Grau de liberdade

$\gamma$  = *N* - 1

### 5.5.2 Teste de $\chi^2$ (Qui Quadrado)

Quando se lida com diferenças de proporções entre duas amostras, os padrões de raciocínio e procedimento são os mesmos, com a diferença, apenas, de que se aplica o teste adequado,  $\chi^2$  (qui quadrado). É o caso, por exemplo, da Tabela II, em que o pesquisador, ao descrever pela primeira vez geneticamente uma doença humana, interessa-se em aplicar a técnica dos dermatóglifos e, portanto, torna-se-lhe fundamental discernir se as diferenças entre as proporções observadas, nos padrões digitais de ambas as populações (doentes e sãs), são casuais ou associadas a peculiaridades biológicas do grupo de afetados (Gonçalves, 1977:46).

TABELA II

FREQÜÊNCIA DOS VÁRIOS PADRÕES DIGITAIS DOS ELEMENTOS VIVOS, AFETADOS PELA SÍNDROME DE MOUNER-KUHN

Padrões	Grupo estudado		Grupo de controle		$\chi^2$
	N	%	N	%	
Arco	13	13,0	60	3,4	30,09*
Presilha radial	2	2,0	97	5,5	
Presilha ulnar	68	68,0	1.096	61,9	
Verticilos	17	17,0	517	29,2	
Total	100	100,0	1.770	100,0	

Sendo a fórmula do  $\chi^2$ :

$$\chi^2 = \frac{\sum (o - e)^2}{e}$$

onde  $\Sigma$  = somatória (soma),

$o$  = cada um dos valores observados,

$e$  = cada um dos valores esperados,

os valores esperados são obtidos da seguinte forma: para cada um dos números de ambos os grupos (estudado e de controle) atribui-se uma representação literal; temos, portanto:

Padrões	Grupo estudado	Grupo de controle	Total
Arco	$a_o$	$e_o$	$a_o + e_o$
Presilha radial	$b_o$	$f_o$	$b_o + f_o$
Presilha ulnar	$c_o$	$g_o$	$c_o + g_o$
Verticilos	$d_o$	$h_o$	$d_o + h_o$
Total	$a_o + b_o + c_o + d_o$	$e_o + f_o + g_o + h_o$	$n$

A seguir calcula-se o valor esperado para cada casela (a, b, ..., h):

$$a_e = \frac{(a_o + b_o + c_o + d_o) \times (a_o + e_o)}{n}$$

$$a_e = \frac{(13 + 2 + 68 + 17) \times (13 + 60)}{1.870} = \frac{(100) \times (73)}{1.870} = \frac{7.300}{1.870} = 3,9$$

Do mesmo modo obteríamos:  $b_e = 62,3$ ;  $c_e = 5,3$ ;  $d_e = 28,5$ ;  $e_e = 69,1$ ;  $f_e = 1.101,7$ ;  $g_e = 93,7$ ;  $h_e = 505,5$ .

Como tabela auxiliar para o cálculo, constrói-se a que segue:

	e	$o - e$	$(o - e)^2$	$(o - e)^2/e$
a	3,9	9,1	82,81	21,23
b	62,3	5,7	32,49	0,52
c	5,3	3,3	10,89	2,05
d	28,5	11,5	132,25	4,64
e	69,1	9,1	82,81	1,20
f	1.101,7	5,7	32,49	0,02
g	93,7	3,3	10,49	0,17
h	505,5	11,5	132,25	0,26
				$\Sigma 30,09$

Portanto o valor  $\chi^2$  é 30,09, com 7 graus de liberdade (número de caselas: 8 - menos 1, igual a 7).

Da mesma forma que se procedeu para o teste  $t$  de Student, o valor obtido para  $\chi^2$  é comparado com um valor de  $\chi^2$ , acima do qual se rejeita  $H_o$  e se aceita  $H_a$  e abaixo do qual se procede de forma contrária. Tal valor ( $\chi^2$  crítico) também é fornecido facilmente pelos livros de estatística.

No caso, o  $\chi^2$  crítico para 7 graus de liberdade e  $\alpha = 5\%$  é 14,6. Portanto, o valor de  $\chi^2$  obtido foi superior ao crítico, isto é, foi significativo (daí ser representado com \*) e, conseqüentemente, rejeita-se  $H_o$ : há diferenças efetivas entre as proporções dos padrões digitais existentes na população e os encontrados no

grupo estudado de afetados. A partir daí o problema excede o campo da estatística e cabe ao pesquisador, mercê de sua formação específica, aprofundar-se no estudo das possíveis peculiaridades biológicas envolvidas em tal resultado.

Como se viu na aplicação, os cálculos e a interpretação desta estatística são bastante simples e não requerem formação matemática profunda. Realmente, a dificuldade reside em qual teste indicar, diante de cada situação em particular. Há que se verificar qual a distribuição do fenômeno estudado, se a amostra em questão merece correção ou não em função de seu tamanho e, entre outras preocupações, se o que se está testando é uma diferença entre médias ou entre proporções. Daí a necessidade de o investigador, especialista em ciências humanas, estar em contato com o colega de formação e experiência estatística.

Esta seqüência de procedimentos, no teste estatístico de hipóteses, aqui apresentada de forma tão descritiva, foi sistematizada por Montenegro (1981: 1.579-1.589), de forma bastante operativa, para ser aplicada pelo estudante passo a passo. Desse modo, ter-se-iam, no teste de hipóteses, oito passos a serem observados:

- a. identificação dos dados do problema, ou seja, identificar com qual média, desvio-padrão e tamanho amostral se está trabalhando;
- b. localização do problema: discriminar se está lidando com duas amostras ou uma amostra e uma população; se se trata de pequenas ou de grandes amostras; se a questão envolve teste de médias ou de proporção, por exemplo;
- c. estabelecimento das hipóteses ( $H_0$  e  $H_a$ );
- d. escolha de fórmula. No caso, para o teste de médias escolheu-se  $t$  de Student e, para proporções,  $\chi^2$ ;
- e. cálculos;
- f. determinação da estatística crítica;
- g. avaliação das hipóteses diante do objeto das estatísticas calculadas e das críticas;
- h. conclusão. isto é, enunciado da hipótese aceita, em termos do problema inicial.

Uma última observação pertinente à questão em foco é que, nesta era da tecnologia em que se vive, já se dispõe de um número bastante grande e diversificado de calculadoras, de gabinete, de mesa ou mesmo de bolso, as quais já se encontram programadas para realizar tais estatísticas, bastando para o aluno obtê-las, após a leitura dos respectivos manuais de instrução, e saber operar adequadamente seus botões. No entanto, obviamente, é desejável que o aluno conheça pelo menos as bases apresentadas, para saber o sentido do que está fazendo: não apenas carregar pedras, mas construir a catedral!

### 5.5.3 A Pesquisa Social e a Estatística

Essa discussão sobre os testes de hipóteses leva, de forma bastante direta, à questão da aplicação da estatística às ciências sociais. Habitualmente o cientista social admite que a estatística se constitui de um instrumento adicional no arsenal de que dispõe para abordar a realidade empírica que estuda e sobre a qual atua, chegando até a reconhecer não lhe ser lícito, como profissional da pesquisa, ignorar ou desprezar um recurso já conhecido como adequado.

De fato, isto é profundamente verdadeiro. Com a propriedade que lhe é peculiar, diz mestre Carneiro Leão Ribeiro que a estatística gera sua própria demanda, isto é, tal instrumental, eficiente e refinado, embora eventualmente de aparência difícil, estimula a projeção de estudos de outra forma inviáveis. Isso sem se referir à sua natureza de evidente agente de mudança em favor da quantificação, ao imprimir maior exatidão na descrição e apreciação dos fenômenos.

Assim, as reservas do cientista social a esse respeito situam-se sobretudo no nível operativo. Na prática, tende a admiti-la no máximo quando se lida com massa muito grande de dados, na qual o nível de complexidade envolvido bloqueia uma primeira percepção globalizadora das questões em pauta. No entanto, a estatística pode auxiliá-lo ainda mais em tais situações, ao prevenir tal desgaste, na medida em que oferece as decorrências da teoria da amostragem para se poder lidar apenas com pequeno número de casos que seja legitimamente representativo do todo.

De início, já a linguagem simbólica e literal, em muitas situações, lhe causa estranheza. Adicionalmente, sua formação e experiência no trato da pesquisa já lhe filtram certo *insight* das tendências descritivas do fato em estudo, de sorte que a estatística vem situar-se como esoterismo a demonstrar o óbvio. Finalmente, tal descrença o leva a um comportamento conceitual e axiológico de rejeição. Em síntese, se tudo que o estatístico propôs der certo, ainda assim terá sido uma fonte adicional de recursos, talvez não tão necessária!

Estas visões são completamente compreensíveis e o esforço consiste em relacionar ambos os profissionais pela complementaridade e não pelo antagonismo, pelo entendimento e não por uma conversa de surdos. Um é homem do todo, o outro do detalhe. O trabalho de campo de um se completa pela montagem e análise de gabinete. Em linguagem administrativa, é a articulação do coordenador com o assessor. É da atuação conjugada de ambos que resulta o trabalho final, a refletir o concurso de experiências convergentes.

Mais detalhadamente, o planejamento da amostragem é o primeiro encontro de ambos. A consolidação, a tabulação, a ordenação e a apresentação dos dados sobre a sociedade se beneficiam igualmente desta atuação a dois, pela montagem, como se viu, racional e clara das tabelas, gráficos e medidas de tendência central e de dispersão cujo resultado final exige um conhecimento básico de suas propriedades e indicações. Já abandonando a fase eminentemente estatístico-descritiva, vê-se que o cientista social é impelido, em sua bus-

ca pela verdade, a cogitar da procedência das generalizações. Então é aí que se lhe pode propiciar a possibilidade do uso de testes de hipóteses. Realmente, torna-se bastante difícil discriminar qual das duas contribuições da estatística às ciências sociais é mais relevante, se a descritiva ou a inferencial.

Tudo isso, porém, deve ser entendido e vivido realisticamente. A supervalorização dos números e seu uso excessivo e a sofisticação são riscos de que se precisa ser alertado constantemente. A estatística é um instrumento e não um fim em si mesmo, ainda quando indispensável.

### LITERATURA RECOMENDADA

ACKOFF, Russell L. *Planejamento de pesquisa social*. 2. ed. São Paulo: EPU: Edusp, 1967. Capítulo 5.

AZEVEDO, Amílcar Gomes; CAMPOS, Paulo H. B. *Estatística básica*. 3. ed. Rio de Janeiro: Livros Técnicos e Científicos, 1978. Capítulos 3, 4, 5, 6 e 7.

BELTRÃO, Pedro Calderon. *Demografia: ciência da população, análise e teoria*. Porto Alegre: Sulina, 1972. Capítulos 8, 9 e 10.

BERQUÓ, Elza Salvatori et al. *Bioestatística*. São Paulo: EPU, 1980. Capítulos 10, 13 e 14.

COUTO, Gracília M. Almeida. *Iniciação à estatística*. Rio de Janeiro: Reper, s.d.

GATTI, Bernadete A.; FERES, Nagib Lima. *Estatística básica para ciências humanas*. São Paulo: Alfa-Omega, 1975. Capítulos 6 e 7.

GOODE, William J., HATT, Paulo K. *Métodos em pesquisa social*. 3. ed. São Paulo: Nacional, 1969. Capítulos 19 e 20.

HOFMANN, Abraham. *Los gráficos en la gestión*. Barcelona: Técnicos, 1974. Capítulo Generalidades.

KAPLAN, Abraham. *A conduta na pesquisa: metodologia para as ciências do comportamento*. São Paulo: EPU: Edusp, 1975. Capítulo 6.

KIRSTEN, José Tiacci et al. *Estatística para as ciências sociais: teoria e aplicações*. São Paulo: Saraiva, 1980. Capítulos 3, 4, 8 e 9.

MOREIRA, José dos Santos. *Elementos de estatística*. São Paulo: Atlas, 1979. Capítulos 3, 4 e 6.

PHILLIPS, Bernard S. *Pesquisa social: estratégias e táticas*. Rio de Janeiro: Agir, 1974. Capítulo 14.

RUMMEL, J. Francis. *Introdução aos procedimentos de pesquisa em educação*. 3. ed. Porto Alegre: Globo, 1977. Apêndice B e Capítulo 10.

# 6

## TRABALHOS CIENTÍFICOS

### 6.1 TRABALHOS CIENTÍFICOS

Os trabalhos científicos devem ser elaborados de acordo com normas preestabelecidas e com os fins a que se destinam. Serem inéditos ou originais e não só contribuir para a ampliação de conhecimentos ou a compreensão de certos problemas, mas também servirem de modelo ou oferecerem subsídios para outros trabalhos.

Para Salvador (1980:11), os trabalhos científicos originais devem permitir a outro pesquisador, baseado nas informações dadas:

- a. reproduzir as experiências e obter os resultados descritos, com a mesma precisão e sem ultrapassar a margem de erro indicada pelo autor;
- b. repetir as observações e julgar as conclusões do autor;
- c. verificar a exatidão das análises e deduções que permitiram ao autor chegar às conclusões."

Rey (1978:29) aponta como trabalhos científicos:

- a. *Observações ou descrições originais* de fenômenos naturais, espécies novas, estruturas e funções, mutações e variações, dados ecológicos etc.
- b. *Trabalhos experimentais* cobrindo os mais variados campos e representando uma das férteis modalidades de investigação, por submeter o fenômeno estudado às condições controladas da experiência.

- c. *Trabalhos teóricos de análise ou síntese de conhecimentos, levando à produção de conceitos novos por via indutiva ou dedutiva; apresentação de hipóteses, teorias etc.*"

Os trabalhos científicos podem ser realizados com base em fontes de informações primárias ou secundárias e elaborados de várias formas, de acordo com a metodologia e com os objetivos propostos.

Neste capítulo, os trabalhos científicos abordados englobam: relatórios de pesquisa, monografias, dissertações, teses, artigo científico e resenha crítica.

## 6.2 RELATÓRIOS

O relatório é a parte final da pesquisa, devendo ser considerado também como "o ato culminante do pensar reflexivo", afirma Whitney (1958:365).

Seu objetivo consiste em "dar às pessoas interessadas o resultado completo do estudo, com pormenores suficientes e dispostos de modo a permitir que o leitor compreenda os dados e determine para si a validade das conclusões" (Goode e Hatt, 1969:456). É, portanto, a exposição de um estudo, via de regra original, e de suas conclusões.

Os relatórios de pesquisa "dependem das condições e dos objetivos fixados no projeto de estudo e do tipo de dados colhidos na investigação" (Hirano, 1979:108). Diferem uns dos outros em relação a forma, estilo, extensão e ao leitor a que se destinam: público em geral ou especializado, instituição ou entidades patrocinadoras da pesquisa etc.

É mais do que a apresentação dos dados coletados, pois tem um propósito: comunicar os resultados da pesquisa em toda sua dimensão, apresentando fatos, dados, procedimentos utilizados, resultados obtidos, análise, chegando a certas conclusões e recomendações.

Para esse tipo de redação é indispensável um roteiro, um esquema básico formal e a observância de certas normas. "Ao se desenvolver o esquema, deve-se conciliar um plano cronológico com o lógico, ciente de que a melhor maneira de relatar é seguir a seqüência natural da pesquisa efetuada" (Salomon, 1999:238).

### 6.2.1 Estrutura do Relatório

Geralmente os relatórios, curtos ou longos, seguem um sistema, cujos elementos constituem uma estrutura básica.

O esquema proposto, baseado em Best (1972:209-217) e Ander-Egg (1978:321-326), com algumas alterações, obedece à seguinte ordem:

#### *Seção Preliminar*

1. Capa (nº de série, título, autor, entidade responsável, data).
2. Folha de rosto (título, autor, entidade responsável, data).
3. Agradecimentos (se houver).
4. Prefácio (dispensável).
5. *Abstract*/Resumo.
6. Relação das tabelas e dos gráficos (se existirem).
7. Sumário (só omitido em relatórios curtos).

#### *Corpo do Relatório*

1. Introdução:
  - a. Explicitação da pesquisa realizada.
  - b. Significado da pesquisa.
  - c. Objeto investigado.
  - d. Aspectos teóricos.
  - e. Definições operacionais utilizadas.
2. Revisão da bibliografia relacionada com os temas ou análise das pesquisas efetuadas.
3. Esquema da Investigação:
  - a. Procedimentos empregados.
  - b. Fontes dos dados.
  - c. Metodologia e utilização.
4. Apresentação, análise e interpretação dos dados:
  - a. Os dados.
  - b. Análise dos dados.
  - c. Interpretação dos dados.
5. Resumo e Conclusões:
  - a. Principais descobertas e conclusões.
  - b. Sugestões para pesquisas posteriores.
6. Recomendações.

#### *Seção de Referências*

1. Anexos ou Apêndice:
  - a. Tabelas e quadros.
  - b. Gráficos.
  - c. Figuras.
  - d. Material suplementar (questionários, glossários etc.).
2. Bibliografia.



### 6.2.1.1 SEÇÃO PRELIMINAR

A seção preliminar é uma espécie de apresentação geral do trabalho; abrange questões formais e convencionais. Divide-se em 7 itens:

- a. *Capa*. A primeira página ou capa do relatório de pesquisa, embora com formatos diferentes de uma instituição para outra, deve incluir: número de série, título completo, autor(es), nome da entidade responsável e data da apresentação. Anotações específicas referentes à entidade patrocinadora são colocadas no reverso da capa.
- b. *Folha de rosto*. Nesta parte, merece especial atenção o título: deve ser conciso, claro, breve e indicar os propósitos da pesquisa. Recomenda-se um título com um mínimo de palavras, ser atraente, de modo que desperte o interesse pela leitura do relatório, e expressar, tanto quanto possível, o âmbito da pesquisa. Depois do título vem o nome do(s) autor(es) do relatório e da entidade patrocinadora. Os títulos dos capítulos devem ser datilografados com todas as letras maiúsculas e os subtítulos (grifados) somente com a inicial em maiúscula.
- c. *Agradecimentos*. Os nomes das pessoas que, de uma forma ou de outra, contribuíram na pesquisa devem constar na página de agradecimentos, expressos de maneira simples e sóbria.
- d. *Prefácio*. Habitualmente não é incluído no relatório, a menos que maiores explicações sejam necessárias sobre o histórico, o âmbito do trabalho e a metodologia empregada. Quando aparece, deve ser assinado ou trazer as iniciais do autor.
- e. *Abstract* ou *Resumo*. Apresentação concisa e freqüentemente seletiva do texto; deve dar relevo aos elementos de maior interesse e importância. Em geral, é redigido no final, após o término do trabalho.
- f. *Sumário* ou *Índice*. Deve ser suficientemente analítico, oferecendo ao leitor uma visão global do estudo realizado. Inclui todos os títulos principais do trabalho e os subtítulos e deve seguir uma ordem progressiva, acompanhando o relatório. A tendência atual é colocar o Sumário, que é mais detalhado, no início e o Índice no final do trabalho.
- g. *Relação das Tabelas e Gráficos*. Se o trabalho incluir listas de tabelas e de gráficos, deve aparecer uma página para cada lista.

### 6.2.1.2 CORPO DO RELATÓRIO

É a parte em que o autor apresenta suas idéias, fatos e provas. Contém o material coletado, elaborado, analisado, interpretado. A forma de apresenta-

ção deve ser objetiva, clara e concisa, levando-se em consideração o tema central, o desenvolvimento lógico e a seqüência dos passos.

#### a. Introdução

É a apresentação do trabalho. Devem ser incluídos os motivos da realização da pesquisa, sua importância, caráter e delimitação, indicando também os objetivos da pesquisa. Abrange:

1. *Explicitação da pesquisa realizada*. Exposição clara sobre a natureza do problema focalizado, juntamente com as questões específicas relacionadas com ele. Cada divisão principal do problema deve ser apresentada em um capítulo.
2. *Significado da pesquisa*. Explicações sucintas, mas suficientes, que demonstrem a relevância da pesquisa e a razão pela qual foi levada em consideração.
3. *Objeto investigado*. Especificação do tema geral em torno do qual a pesquisa foi realizada; justificativa da escolha, indicando também lacunas no conhecimento científico.
4. *Aspectos teóricos*. Referências à teoria de base na qual o estudo se apoiou.
5. *Definições operacionais utilizadas*. Definição cuidadosa dos termos importantes utilizados na pesquisa, a fim de que o leitor possa compreender os conceitos sob os quais a pesquisa se desenvolveu.

#### b. Revisão bibliográfica

Análise dos dados levantados em fontes secundárias. Refere-se às informações ligadas ao estudo e ao resumo das conclusões mais importantes.

Bibliografias pertinentes ao tema da pesquisa não devem ser omitidas, assim como trabalhos que nada tenham a ver com o assunto devem ser excluídos.

#### c. Esquema da investigação

Relaciona-se com o desenho da investigação. Engloba:

1. *Procedimentos empregados*. Descrição dos procedimentos utilizados, indicando, da melhor maneira, como se realizou a pesquisa.
2. *Fontes de dados*. Identificação dos tipos de fontes onde foram coletados os dados, segundo sua procedência.
3. *Metodologia e sua utilização*. Explicação da metodologia empregada na obtenção dos dados, relatando, em pormenores, as próprias experiências e observações e indicando o manejo dos instrumentos

empregados. Justificativa da escolha tanto dos métodos quanto das técnicas utilizados e o grau de precisão e validade dos instrumentos.

#### d. Apresentação, análise e interpretação dos dados

Parte principal do corpo do relatório, que descreve o desenvolvimento do trabalho e os resultados obtidos. Abrange:

1. *Dados.* Apresentação dos resultados, passo a passo, indicando os elementos mais importantes, delineando as fases do estudo e demonstrando o que a pesquisa realmente apurou.
2. *Análise dos dados.* Análise crítica dos dados, tentando explicar o fenômeno e as relações existentes entre ele e alguns fatores antecedentes ou independentes, valendo-se de processos matemáticos e estatísticos. A comprovação ou refutação da hipótese se faz por meio da análise.
3. *Interpretação dos dados.* Discussão dos resultados encontrados e confronto com os obtidos por outros estudiosos, apontando pontos mais importantes e realçando determinados setores. Indicação dos vínculos de tais resultados com os objetivos propostos pela hipótese, incorporando-se num sistema teórico-prático.

#### e. Resumo e conclusões

Exposição sucinta do pensamento do autor ou do conteúdo do trabalho. Engloba:

1. *Principais achados e conclusões.* Apresentação sumária dos principais achados e explicitação clara e concisa dos resultados finais, de preferência ordenados por itens. "A conclusão confirma total ou parcialmente a hipótese ou hipóteses colocadas na introdução" (Rehfeldt, 1980:49).
2. *Sugestões para pesquisas posteriores.* Indicação de problemas secundários não abordados, mas prometedores, e sugestões para futuros estudos no mesmo campo.

#### f. Recomendações

Em se tratando de pesquisa aplicada, devem-se fazer recomendações para que outros interessados possam valer-se das informações ou repetir as experiências e observações.

#### 6.2.1.3 SEÇÃO DE REFERÊNCIAS

Esta seção abrange dois aspectos: apêndice e bibliografia.

- a. *Anexos e Apêndices.* Inclusão de tabelas, quadros, gráficos, inventários, ilustrações ou figuras, formulários ou questionários, glossários, documentos e notas explicativas longas, usadas no estudo, e todo material suplementar que não é necessário à compreensão das informações mas que ilustra o conteúdo da pesquisa. Denomina-se Apêndice quando é da própria autoria e Anexo, de outro autor.
- b. *Bibliografia.* Citação completa de todas as obras utilizadas no trabalho, colocando-as em ordem alfabética.

#### 6.2.1.4 REDAÇÃO E ESTILO

A representação científica de um trabalho deve ser o mais didática possível e ter um estilo simples, claro, preciso e objetivo. Alguns autores ainda acrescentam estilo criador.

A apresentação do trabalho requer linguagem perfeita em relação às regras gramaticais, evitando-se não só o vocabulário popular, vulgar, mas também o pomposo.

Se uma das finalidades da redação é a objetividade, o trabalho deve ter caráter impessoal, abolindo-se os pronomes pessoais como *eu, nós, vocês* etc.

*Exemplo:*

Em vez de "eu escolhi", dizer "foi escolhido".

Na descrição dos procedimentos realizados na pesquisa, usa-se o verbo no passado. Abreviações no texto do trabalho devem ser evitadas, salvo em notas de rodapé e nas bibliografias e tabelas.

"O bom estilo requer uma apresentação orgânica e inteligente do material", afirma Ander-Egg (1978:325).

Para Barras (1979:92), "um bom estilo depende de inteligência, imaginação e bom gosto de quem escreve; depende da sinceridade, da modéstia, planejamento cuidadoso e atenção para com as regras de redação científica".

#### 6.2.2 Relatório Progressivo

Os relatórios progressivos, oferecidos sobretudo a órgãos patrocinadores de pesquisa (Fapesp, Capes, CNPq etc.), precisam ser apresentados sistematicamente em períodos curtos, estabelecidos pela entidades. Sendo assim, o rela-

tório progressivo dá conta, explicitamente, do que já foi realizado num período e o que se pretende fazer no período subsequente.

*Exemplo: Plano de Trabalho*

*O Artesanato Feminino (em X lugar)*

*1º ano*

1. Levantamento de bibliografia atinente ao artesanato, dado que a pesquisa se propõe a investigar o artesanato.
2. Levantamento das artesãs da região.
3. Exame da bibliografia metodológica.
4. Desenvolvimento da pesquisa de campo: aplicação de formulários e de entrevistas; seleção de elementos para histórias de vida.

*2º ano*

1. Estudos das diferentes técnicas de trabalho e do instrumental utilizado pelas artesãs.
2. Sistematização e interpretação dos dados coletados sobre as artesãs.
3. Análise da dinâmica cultural na qual se integra a artesã:
  - a. Tradição x industrialização.
  - b. Subemprego, ocupação e/ou atividade valorizada.
  - c. Lazer ou necessidade.
4. Análise e interpretação dos dados obtidos na pesquisa de campo.
5. Redação provisória dos resultados obtidos.
  - a. Revisão da redação.
  - b. Redação definitiva.
  - c. Seleção das ilustrações.

Este plano pode ser subdividido, de acordo com o órgão patrocinador, em etapas de 6 meses e até de menos. Os relatórios progressivos obedecem à seqüência do plano.

### 6.3 MONOGRAFIA

São numerosos os conceitos sobre monografia encontrados em obras cujos autores abordam este assunto, dos quais foram selecionados alguns.

#### 6.3.1 Conceitos

Asti Vera (1979:164) define monografia como sendo o “tratamento escrito de um tema específico” e Salomon (1999:254), como o “tratamento escrito de um tema específico que resulte de interpretação científica com o escopo de apresentar uma contribuição relevante ou original e pessoal à ciência”.

Descrição ou tratado especial de determinada parte de uma ciência qualquer, dissertação ou trabalho escrito que trata especialmente de determinado ponto da ciência, da arte, da história etc. ou “trabalho sistemático e completo sobre um assunto particular, usualmente pormenorizado no tratamento, mas não extenso em alcance” (American Library Association) são outros conceitos.

Trata-se, portanto, de um estudo sobre um tema específico ou particular, com suficiente valor representativo e que obedece a rigorosa metodologia. Investiga determinado assunto não só em profundidade, mas também em todos os seus ângulos e aspectos, dependendo dos fins a que se destina.

Tem como base a escolha de uma unidade ou elemento social, sob duas circunstâncias: (1) ser suficientemente representativo de um todo cujas características se analisa; (2) ser capaz de reunir os elementos constitutivos de um sistema social ou de refletir as incidências e fenômenos de caráter autenticamente coletivo.

Para Rehfeldt (1980:9), a monografia “é degrau rumo à pesquisa mais ampla, além de possibilitar ao futuro professor condições de habilitar-se para o treinamento dos jovens”. É uma consequência da investigação científica, que exige tratamento reflexivo.

#### 6.3.2 Características

Analisando-se os diferentes conceitos, pode-se observar que a monografia apresenta algumas características:

- a. trabalho escrito, sistemático e completo;
- b. tema específico ou particular de uma ciência ou parte dela;
- c. estudo pormenorizado e exaustivo, abordando vários aspectos e ângulos do caso;
- d. tratamento extenso em profundidade, mas não em alcance (nesse caso, é limitado);
- e. metodologia específica;
- f. contribuição importante, original e pessoal para a ciência.

A característica essencial não é a extensão, como querem alguns autores, mas o caráter do trabalho (tratamento de um tema delimitado) e a qualidade



da tarefa, isto é, o nível da pesquisa, que está intimamente ligado aos objetivos propostos para a sua elaboração.

A monografia implica originalidade, mas até certo ponto, uma vez que é impossível obter total novidade em um trabalho; isto é relativo, pois a ciência, sendo acumulativa, está sujeita a contínuas revisões.

### 6.3.3 Estrutura da Monografia

Os trabalhos científicos, em geral, apresentam a mesma estrutura: introdução, desenvolvimento e conclusão.

Pode haver diferenças quanto ao material, o enfoque dado, a utilização desse ou daquele método, dessa ou daquela técnica, mas não em relação à forma ou à estrutura.

- a. **Introdução.** Formulação clara e simples do tema da investigação; é a apresentação sintética da questão, importância da metodologia e rápida referência a trabalhos anteriores, realizados sobre o mesmo assunto.
- b. **Desenvolvimento.** Fundamentação lógica do trabalho de pesquisa, cuja finalidade é expor e demonstrar.  
No desenvolvimento, podem-se levar em consideração três fases ou estágios: explicação, discussão e demonstração.
  - **Explicação** “é o ato pelo qual se faz explícito o implícito, claro o escuro, simples o complexo” (Asti Vera, 1979:169).  
Explicar é apresentar o sentido de uma noção, é analisar e compreender, procurando suprimir o ambíguo ou o obscuro.
  - **Discussão** é o exame, a argumentação e a explicação da pesquisa: explica, discute, fundamenta e enuncia as proposições.
  - **Demonstração** é a dedução lógica do trabalho; implica o exercício do raciocínio. Demonstra que as proposições, para atingirem o objetivo formal do trabalho e não se afastarem do tema, devem obedecer a uma seqüência lógica.
- c. **Conclusão.** Fase final do trabalho de pesquisa, mas não somente um fim. Como a introdução e o desenvolvimento, possui uma estrutura própria.

A conclusão consiste no resumo completo, mas sintetizado, da argumentação dos dados e dos exemplos constantes das duas primeiras partes do trabalho. Da conclusão devem constar a relação existente entre as diferentes partes da argumentação e a união das idéias e, ainda, conter o fecho da introdução ou síntese de toda reflexão.

### 6.3.4 Tipos de Monografias

Os estudantes, ao longo de suas carreiras, precisam apresentar uma série de trabalhos que se diferenciam uns dos outros quanto ao nível de escolaridade e quanto ao conteúdo. Via de regra, para o término do curso de graduação, os estudantes têm o compromisso de elaborar um trabalho baseado, geralmente, em fontes bibliográficas, que não precisa ser extenso nem muito específico. À medida que ascendem na carreira universitária, esses trabalhos vão exigindo maior embasamento, mais reflexão, mais amplitude e criatividade.

Alguns autores, apesar de darem o nome genérico de monografia a todos os trabalhos científicos, diferenciam uns dos outros de acordo com o nível da pesquisa, a profundidade e a finalidade do estudo, a metodologia utilizada e a originalidade do tema e das conclusões.

Dessa maneira, podem-se distinguir três tipos: monografia, dissertação e tese, que obedecem a esta ordem ascendente em relação à originalidade, à profundidade e à extensão. Há os que incluem, nesta relação, a memória científica, que ora se aproxima da monografia apresentada no final do curso de graduação (memória recapitulativa), ora da dissertação de mestrado (memória científica original) e até mesmo da tese de doutoramento.

Salomon (1999:254) classifica a monografia em duas categorias:

- a. “*Lato*. Todo o trabalho científico de ‘primeira mão’, que resulte da investigação científica”. Inclui nesse item as dissertações, as exercitações, as tesinas, certos relatórios de pesquisa, informes científicos ou técnicos e as memórias científicas.
- b. “*Estrito*. Quando se identifica com a tese” (ver conceito de monografia).

Há os que apresentam outra divisão:

- a. *Monografias escolares* ou trabalhos de caráter didático, apresentados ao final de um curso específico, elaborados por alunos iniciantes na autêntica monografia, ou de “iniciação à pesquisa e como preparação de seminários” (Salvador, 1980:32). Também chamados trabalhos de média divulgação, porque baseados em dados de segunda mão.
- b. *Monografias científicas*. Trabalhos científicos apresentados ao final do curso de mestrado, com o propósito de obter o título de mestre.

### 6.3.5 Escolha do Tema

Na escolha do tema, o estudante poderá tomar a iniciativa, selecionando um assunto ou problema de trabalho, de acordo com suas preferências, eviden-

ciadas durante o curso de graduação. Pode aceitar o tema indicado pelo professor ou escolher um tópico, constante de uma relação oferecida pelo orientador, tendo sempre em vista o seu interesse.

O tema geral de um estudo também “pode ser sugerido por alguma vantagem prática ou interesse científico ou intelectual em benefício dos conhecimentos sobre certa situação particular”, afirma Seltiz (1965:33-34).

Escolhido o tema, a primeira coisa a fazer é procurar conhecer o que a ciência atual sabe sobre ele, para não cair no erro de apresentar como novo o que já é conhecido há tempos, de demonstrar o óbvio ou de preocupar-se em demasia com detalhes sem grande importância, desnecessários ao estudo.

Este trabalho prévio abrange três aspectos:

- a. orientação geral sobre a matéria que vai ser desenvolvida;
- b. conhecimento da bibliografia pertinente;
- c. reunião, seleção e ordenação do material levantado.

A bibliografia relacionada com o estudo muitas vezes é indicada pelo próprio professor e/ou orientador. Nesse caso, o estudante tem à sua disposição o material necessário ao seu trabalho.

Outros pontos importantes a serem considerados: relevância do assunto, áreas controversas ou obscuras, natureza e extensão da contribuição.

No conhecimento da bibliografia faz-se necessário consultar, ler e fichar os estudos já realizados sobre o tema, com espírito crítico, valendo-se da literatura especializada, a partir dos trabalhos mais gerais e indo a seguir para os estudos mais específicos.

Quanto ao assunto escolhido, devem-se ainda observar algumas qualidades importantes:

- a. ser proporcional (em suas partes);
- b. ter valor científico;
- c. não ser extenso demais ou muito restrito;
- d. ser claro e bem delineado.

As monografias referentes ao grau de conclusão de estudante universitário não podem ser consideradas verdadeiros trabalhos de pesquisa (para o qual os estudantes não estão ainda capacitados, salvo raras exceções), mas estudos iniciais de pesquisa.

O trabalho de investigação – teórico ou prático, bibliográfico ou de campo – dá oportunidade ao estudante para explorar determinado tema ou problema, levando-o a um estudo com maior ou menor profundidade e/ou extensão. Possibilita o desenvolvimento de sua capacidade de coletar, organizar e relatar in-

formações obtidas e, mais, de analisar e até de interpretar os dados de maneira lógica e apresentar conclusões.

## 6.4 DISSERTAÇÃO

A dissertação é “um estudo teórico, de natureza reflexiva, que consiste na ordenação de idéias sobre um determinado tema” (Salvador, 1980:35), “aplicação de uma teoria já existente, para analisar determinado problema” (Rehfeldt, 1980:62), ou “trabalho feito nos moldes da tese, com a peculiaridade de ser ainda uma tese inicial ou em miniatura” (Salomon, 1999:263).

A dissertação é, portanto, um tipo de trabalho científico apresentado no final do curso de pós-graduação, visando obter o título de mestre. Requer defesa de tese.

Tem caráter didático, pois se constitui em um treinamento e/ou iniciação à investigação.

Como estudo teórico, de natureza reflexiva, requer sistematização, ordenação e interpretação dos dados. Por ser um estudo formal, exige metodologia própria do trabalho científico.

Situa-se entre a monografia e a tese, porque aborda temas em maior extensão e profundidade do que aquela e é fruto de reflexão e de rigor científico, próprios desta última.

A estrutura e o plano de trabalho da dissertação praticamente são idênticos ao da tese, mas esta se distingue da dissertação pela contribuição significativa na solução de problemas importantes, colaborando para o avanço científico na área em que o estudo se realiza.

Para Salomon (1999:258-262), há dois tipos de dissertação:

- a. *Dissertação monográfica* ou tratamento escrito de assunto específico, com metodologia adequada e de caráter eminentemente didático.
- b. *Dissertação científica* ou tratamento escrito, original, de assunto específico, com metodologia própria e resultante de pesquisa pura ou aplicada.

Para Salvador (1980:35), a dissertação pode ser:

- a. *Expositiva*. Quando reúne e relaciona material obtido de diferentes fontes, expondo o assunto com fidedignidade e demonstrando habilidade não só de levantamento, mas também de organização.
- b. *Argumentativa*. Quando requer interpretação das idéias apresentadas e o posicionamento do pesquisador.

Alguns autores usam os termos tese de mestrado e memória doutoral, opondo-se aos citados anteriormente, mas é menos usual.

A dissertação (tese de mestrado) é de natureza semelhante à tese (memória doutoral) no sentido de que contribui, de modo substancial, na solução de problemas importantes.

Além dos aspectos de qualidade, existem as limitações de tempo, de fundos e de esforços, que geralmente restringem a extensão e a quantidade do estudo, aspectos que não podem deixar de ser considerados em trabalhos desse tipo.

## 6.5 TESE

A tese significa proposição sobre determinado aspecto de qualquer ciência, devendo ser apresentada e defendida publicamente.

É uma monografia científica, escrita, original, sobre um tema específico, cuja contribuição amplia os conhecimentos do tema escolhido. Representa, portanto, um avanço na área científica em que se situa.

São várias, mas não contraditórias, as definições de tese formuladas por diferentes autores.

Tese é “opinião ou posição que alguém sustenta e está preparado para defender” (Barras, 1979:152), proposição clara e terminantemente formulada, em um de seus aspectos, formal ou material, e que se submete à discussão ou prova. “ato culminante do pensar reflexivo” (Whitney, 1958:368).

A tese doutoral, por sua “importância e transcendência, é o projeto de fim de carreira e deve ter uma apresentação cuidadosa, esmerada, já que à primeira vista deve dizer, transmitir, dar uma idéia da importância de seu conteúdo”, afirma Vega (1969:658).

Apresenta o mais alto nível de pesquisa e requer não só exposição e explicação do material coletado, mas, e principalmente, análise e interpretação dos dados.

É um tipo de trabalho científico que levanta, coloca e soluciona problemas; argumenta e apresenta razões, baseadas em evidências dos fatos, com o objetivo de provar se as hipóteses levantadas são falsas ou verdadeiras.

A tese pode ser considerada como um teste de conhecimento para o candidato, que deve demonstrar capacidade de imaginação, de criatividade e habilidade não só para relatar o trabalho, mas também para apresentar soluções para determinado problema.

Na elaboração da tese deve-se fazer uso do raciocínio lógico para se chegar a conclusões válidas. A argumentação tem em vista justificar e persuadir, objetivando convencer o leitor em relação a determinadas idéias ou posições.

O objetivo básico da tese é a argumentação, e o imediato é o de colaborar na solução de dado problema.

Pode resultar de estudo teórico ou de pesquisa de campo, de trabalho em laboratório ou experimental. Utiliza, como todo trabalho científico, um ou mais métodos e técnicas, usados concomitantemente ou separados. Possui a mesma estrutura da monografia ou da dissertação: introdução, desenvolvimento e conclusão, acrescidos das preliminares e das ilustrações. Todavia, distingue-se das outras no que concerne à profundidade, originalidade, extensão e objetividade.

### 6.5.1 Partes da Tese

#### A. Preliminares

1. Folha de rosto
2. Página de aprovação
3. Agradecimentos
4. Lista das tabelas
5. Sumário

#### B. Corpo da tese

1. Introdução (proposição)
2. Desenvolvimento (demonstração), geralmente dividido em Capítulos  
O problema  
Revisão da Bibliografia  
Procedimentos Metodológicos  
Apresentação e Análise dos Dados
3. Conclusões e Recomendações

#### C. Parte referencial

1. Apêndices e Anexos
2. Referências Bibliográficas

## 6.6 ARTIGOS CIENTÍFICOS

Os artigos científicos, para publicações em revistas ou periódicos, são pequenos estudos que tratam de uma questão verdadeiramente científica, mas não chegam a constituir-se em matéria de um livro.

Distinguem-se dos diferentes tipos de monografias, abordados anteriormente, pela sua reduzida dimensão e pelo conteúdo.

Têm a mesma estrutura exigida para trabalhos científicos, ou seja: introdução, desenvolvimento e conclusão.

A *introdução* apresenta o assunto, o objetivo e a metodologia empregada; o *desenvolvimento*, por ser o núcleo do trabalho, consiste na exposição, explicação e demonstração da matéria; a *conclusão* expõe resumidamente os resultados.

O conteúdo abrange os mais variados aspectos e, em geral, apresenta temas e/ou abordagens novos, atuais, diferentes. Pode:

- a. versar sobre um estudo pessoal, uma descoberta ou dar um enfoque contrário ao já conhecido;
- b. oferecer soluções para questões controvertidas;
- c. levar ao conhecimento do público intelectual, ou especializado no assunto, idéias novas, para sondagem de opiniões ou atualização de informes;
- d. abordar aspectos secundários levantados em alguma pesquisa, mas que não serão utilizados nela.

O estabelecimento de um esquema para expor, de maneira lógica e sistemática, os diferentes itens do assunto evita repetições ou omissões ao longo da dissertação.

O público a que se destina o artigo também deve ser levado em consideração; isto pode ser mais ou menos previsto, conhecendo-se de antemão a natureza da revista: científica, didática, de divulgação etc.

O estilo deve ser claro, conciso, objetivo e a linguagem correta, precisa, coerente e simples. Adjetivos supérfluos, rodeios, repetições ou explicações inúteis devem ser evitados, assim como a forma excessivamente compacta, que pode prejudicar a compreensão do texto.

O título também merece atenção: precisa corresponder de maneira adequada ao conteúdo.

## 6.7 RESENHA CRÍTICA

Resenha crítica é a apresentação do conteúdo de uma obra. Consiste na leitura, no resumo e na crítica, formulando o resenhista um conceito sobre o valor do livro.

A resenha, em geral, é feita por cientistas que, além do conhecimento sobre o assunto, têm capacidade de juízo crítico. Também pode ser feita por estudantes; neste caso, como um exercício de compreensão e crítica. Para iniciar-se nesse tipo de trabalho, a maneira mais prática seria começar por resenhas de capítulos.

O resenhista deve resumir o assunto e ressaltar as falhas, sem entrar em muitos pormenores, pois o seu objetivo é informar. Deve ter competência no assunto, mas isso não lhe dá o direito de fazer juízo de valor ou de deturpar o pensamento do autor.

Mesmo não fazendo parte dos trabalhos científicos de primeiro nível, a resenha crítica apresenta uma estrutura ou roteiro:

1. *Referência bibliográfica*  
Autor(es)  
Título (subtítulo)  
Imprentas (local da edição, editora, data)  
Número de páginas  
Ilustrações (tabelas, gráficos, fotos etc.)
2. *Credenciais do autor*  
Informações gerais sobre o autor  
Autoridade no campo científico  
Quem fez o estudo?  
Quando? Por quê? Em que local?
3. *Conhecimento*  
Resumo detalhado das idéias principais  
Do que trata a obra? O que diz?  
Tem alguma característica especial?  
Como foi abordado o assunto?  
Exige conhecimentos prévios para entendê-lo?
4. *Conclusão do autor*  
O autor faz conclusões? (ou não?)  
Onde foram colocadas? (final do livro, dos capítulos)  
Quais foram?
5. *Quadro de referência do autor*  
Modelo teórico  
Que teoria serviu de embasamento?  
Qual o método utilizado?
6. *Apreciação*
  - a. Julgamento da obra  
Científica, didática, de divulgação  
Como se situa o autor em relação:  
– às escolas ou correntes científicas, filosóficas, culturais?

- às circunstâncias culturais, sociais, econômicas, históricas etc.?
- b. Mérito da obra:
  - Qual a contribuição dada?
  - Idéias verdadeiras, originais, criativas?
  - Conhecimentos novos, amplos, abordagem diferente?
- c. Estilo
  - Conciso, objetivo, simples?
  - Claro, preciso, coerente?
  - Linguagem correta?
  - Ou o contrário?
- d. Forma
  - Lógica, sistematizada?
  - Há originalidade e equilíbrio na disposição das partes?
- e. Indicação da obra
  - A quem é dirigida: grande público, especialistas, estudantes?

**Observação:** Quando se tratar de resenhas realizadas por estudantes, sob orientação de um professor, podem-se solicitar mais outros tópicos, para avaliação:

Qual o sistema utilizado para fazer o trabalho?

Quantas vezes leu? Leu tudo de uma vez?

Adquiriu conhecimentos? Reforçou conhecimentos anteriores?

## LITERATURA RECOMENDADA

ANDER-EGG, Ezequiel. *Introducción a las técnicas de investigación social: para trabajadores sociales*. 7. ed. Buenos Aires: Humanitas, 1978. Quarta parte, Capítulo 26.

BARRAS, Robert. *Os cientistas precisam escrever: guia de redação para cientistas, engenheiros e estudantes*. São Paulo: T. A. Queiroz: Edusp, 1979. Capítulos 12 e 13.

BASTOS, Lília da Rocha et al. *Manual para a elaboração de projetos e relatórios de pesquisa, teses e dissertações*. Rio de Janeiro: Zahar, 1979.

CASTRO, Cláudio de Moura. *Estrutura e apresentação de publicações científicas*. São Paulo: McGraw-Hill do Brasil, 1978b.

MANZO, Abelardo J. *Manual para la preparación de monografías: una guía para presentar informes y tesis*. Buenos Aires: Humanitas, 1971. Capítulos 8 e 10.

MARTINS, Joel; CELANI, M. Antonieta Alba. *Subsídio para redação ou teses de mestrado e doutoramento*. 2. ed. São Paulo: Cortez & Moraes, 1979.

NOGUEIRA, Oracy. *Pesquisa social: introdução às suas técnicas*. São Paulo: Nacional: Edusp, 1968. Capítulo 16.

PARDINAS, Felipe. *Metodología y técnica de investigación en ciencias sociales*. 2. ed. México: Siglo Veintiuno, 1977. Capítulos 7 e 8.

REY, Luís. *Como redigir trabalhos científicos*. São Paulo: Edgard Blücher, 1978. Capítulos 3 e 4.

REHFELDT, Gládis Knak. *Monografia e tese: guia prático*. Porto Alegre: Sulina, 1980.

RUMMEL, J. Francis. *Introdução aos procedimentos de pesquisa em educação*. 3. ed. Porto Alegre: Globo, 1977. Capítulo 11 e Apêndice A.

SALOMON, Délcio Vieira. *Como fazer uma monografia: elementos de metodologia do trabalho científico*. 9. ed. São Paulo: Martins Fontes, 1999. Segunda parte, Capítulos 1 e 3, Terceira parte, Capítulos 1, 2 e 4.

SALVADOR, Ângelo Domingos. *Métodos e técnicas de pesquisa bibliográfica: elaboração de trabalhos científicos*. 8. ed. Porto Alegre: Sulina, 1980. Introdução.

SELLTIZ, C. et al. *Métodos de pesquisa nas relações sociais*. São Paulo: Herder, 1965. Capítulo 12.

SPINA, Segismundo. *Normas gerais para os trabalhos de grau: um breviário para o estudante de pós-graduação*. São Paulo: Fernando Pessoa, 1974.

VEGA, Javier Lasso de la. *Manual de Documentación*. Barcelona: Labor, 1969. Sexta parte, Capítulos 34, 35 e 36.

WHITNEY, Frederick L. *Elementos de investigación*. Barcelona: Omega, 1958. Capítulo 16.