MEDIDAS ESTATÍSTICAS de dados agrupados (em classes)

Exemplo: Dada a tabela resumo

Tabela 1: Tabela resumida de Distribuição de Frequências

i	Diâmetros (mm)	f_i
1	72,0 ├ 72,4	1
2	72,4 ├ 72,8	7
3	72,8 ├ 73,2	10
4	73,2 ├ 73,6	17
5	73,6 ├ 74,0	13
6	74,0 ├ 74,4	9
7	74,4 ├ 74,8	3
		$\Sigma f_i = 60$

1.1 Média da Amostra ($\bar{x}$)

Para calcular a média deve-se multiplicar cada valor pelo número atribuído à sua importância no conjunto, somar todos os produtos obtidos e dividir o total pela soma dos pesos.

$$\bar{x} = \frac{\Sigma x_i . f_i}{\Sigma f_i}$$

Como $\Sigma f_i = n$.

$$\bar{x} = \frac{\Sigma x_i}{n}$$

onde,

$\bar{x}$ é a média da amostra de n elementos;

n é o números de elementos da amostra;

x_i ($i = 1, 2, 3, \dots, n$) são os valores das n observações.

Resolução: O ideal é criar duas colunas na **Tabela de Distribuição de Freqüência** responsável pelas médias das classe ($\bar{x}_i$) e pelos resultados dos produtos de cada média pelo peso correspondente (**$\bar{x}_i . f_i$**), como mostra a tabela 2.

Tabela 2: Tabela de Distribuição de Frequências com o Cálculo da Média

i	Diâmetros (mm)	f_i	$\overline{x_i}$	$\overline{x_i}.f_i$
1	72,0 ├ 72,4	1	72,2	**72,2**
2	72,4 ├ 72,8	7	72,6	**508,2**
3	72,8 ├ 73,2	10	73,0	**730,0**
4	73,2 ├ 73,6	17	**73,4**	**1247,8**
5	73,6 ├ 74,0	13	73,8	**959,4**
6	74,0 ├ 74,4	9	74,2	**667,8**
7	74,4 ├ 74,8	3	74,6	**223,8**
		$\Sigma f_i = 60$		**$\Sigma x_i.f_i = 4409,2$**

Os cálculos:

$$x_i.f_i = x_1.f_1 = 72,2.1 = 72,2$$

$$x_2.f_2 = 72,6.7 = 508,2$$

$$x_3.f_3 = 73,0.10 = 730,0$$

$$x_4.f_4 = 73,4.17 = 1247,8$$

$$x_5.f_5 = 73,8.13 = 959,4$$

$$x_6.f_6 = 74,2.9 = 667,8$$

$$x_7.f_7 = 74,6.3 = 223,8$$

Assim,

$$\Sigma x_i.f_i = 72,2 + 508,2 + 730,0 + 1247,8 + 959,4 + 667,8 + 223,8 = 4409,2$$

A média será $\bar{x} = \frac{\Sigma x_i}{n} = \frac{4409,2}{60}$ (= 73,48666...) = 73,49 mm

(usando duas casa decimal, pois todos os elementos da amostra têm uma).

1.2 Moda da Amostra (Mo)

A moda é o valor da variável que aprece com mais freqüência, ou seja, o valor que aparece mais vezes.

Numa **Tabela de Distribuição de Freqüência** com perda de informação a classe com maior freqüência é chamada de **classe modal** e o valor da moda é dado pelo ponto médio da classe modal.

No nosso exemplo, Mo = 73,4mm na 4ª classe (modal)

1.3 Mediana da Amostra (Me)

A mediana é o valor que se encontra no centro de um conjunto de valores, estando a amostra ordenada.

Numa **Tabela de Distribuição de Freqüência** com perda de informação a mediana é calculada seguindo os passos:

Posição 1	**72,0 (mín)**
Posição 2	
...	
Posição 59	
Posição 60	**74,8 (máx)**

Salto de 59 (60 – 1)

Queremos dividir este salto em dois de 29,5 (59/2)

Posição 1	**72,0 (mín)**
Posição 2	
...	
Pos. 30,5	**Me**
...	
Posição 59	
Posição 60	**74,8 (máx)**

Salto de 29,5
1 + 29,5 – 30,5

Salto de 29,5
30,5 + 29,5 = 60

Pela tabela 3, com as frequências acumuladas Fi, vemos que 30, 5 está entre 18 e 35

Tabela 3: Tabela de Distribuição de Frequências com o Cálculo da Mediana

i	Diâmetros (mm)	f_i	F_i
1	72,0 ├ 72,4	1	1
2	72,4 ├ 72,8	7	8
3	72,8 ├ 73,2	10	**18**
4	**73,2** ├ 73,6	17	**35**
5	**73,6** ├ 74,0	13	48
6	74,0 ├ 74,4	9	57
7	74,4 ├ 74,8	3	60

Há 18 medidas "antes" de 73,2.

19º → 73,2

Há 35 medidas "antes" de 73,6.

36º → 73,6

Uma média ponderada entre 73,2 e 73,6 nos dará a mediana, pois a sua posição é a 30,5ª que está entre 19ª e 36ª.

Os pesos serão as distâncias entre 30,5 e 19 (11,5) e 30,5 e 36 (5,5).

Como 30,5 está mais perto de 36 que de 19, a mediana deverá estar mais próxima de 73,6 que 73,2. Assim o peso maior será do valor 73,6.

Me = $\frac{73,2 \times 5,5 + 73,6 \times 11,5}{5,5 + 11,5}$ = **73,47**

Esquematicamente:

dist	pos	valor		pesos		
11,5	19	73,2	x	5,5	=	402,6
	30,5	**73,47**				
5,5	36	73,6	x	11,5	=	846,4
17						1249

1.4. Medidas de dispersão.

- **Amplitude Total : Amp = 2,8 (=74,8 – 72,0);**
- **Amplitude Modal: AmpMo = 17** (a maior frequência – 4ª classe)
- **Intervalo interquartil: IQL = $Q_3 - Q_1$** (Calcular Q_1 e Q_3 como Me = Q_2)
- **Desvio Médio: dm =** $\frac{|x1-\bar{x}|.f1+|x2-\bar{x}|.f2+\cdots+|x3-\bar{x}|.fn}{f1+f2+\cdots+fn} = \frac{\sum|\bar{x}i-\bar{x}|.fi}{n}$
- **Desvio Padrão: s =** $\sqrt{\frac{\sum|\bar{x}i-\bar{x}|^2.fi}{n-1}}$
- **Variância da Amostr:** $s^2 = \frac{\sum|\bar{x}i-\bar{x}|^2.fi}{n-1}$
- **Coeficiente de Variação: CV =100** $\frac{s}{\bar{x}}$

Estas medidas são determinadas de maneiras análogas quando não temos todos os dados, mas somente os dados agrupados em classes.