Aula 05 - Medidas de Posição

Medida de posição

Medida de posição é o valor ao redor do qual os dados se distribuem.

Principais:

- média;
- mediana;
- moda.

μ = média populacional

- **Média aritmética**:
  A média aritmética é dada por:

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n},$$

em que n corresponde ao tamanho da amostra e x_i ao i-ésimo valor observado.

Exemplo:

Com o objetivo de avaliar a produção de leite, em kg, foram observadas as produções médias diárias de 10 produtores rurais atendidos por um plano governamental, cujos valores são apresentados a seguir:

9,80	9,90	9,95	10,00	8,78
9,90	9,34	10,34	11,75	15,00

$\bar{x} = \frac{\sum x_i}{n}$ $\bar{x} = \frac{104,76}{10}$

$\sum x_i = 104,76$

$n = 10$

$\bar{x} = 10,476$ kg

Conferindo...

A média observada de produção de leite é dada por:

$$\bar{x} = \frac{9,80 + 9,90 + \ldots + 15,00}{10} = \frac{104,76}{10} = 10,48\text{kg}.$$

Observações:

Média sem considerarmos o maior valor observado (15,00):

$$\bar{x} = \frac{9,80 + 9,90 + \ldots + 11,75}{9} = \frac{89,76}{9} = 9,97 \text{ kg}.$$

- A média é bastante afetada por valores extremos;
- não deve ser utilizada quando a distribuição dos dados é assimétrica.

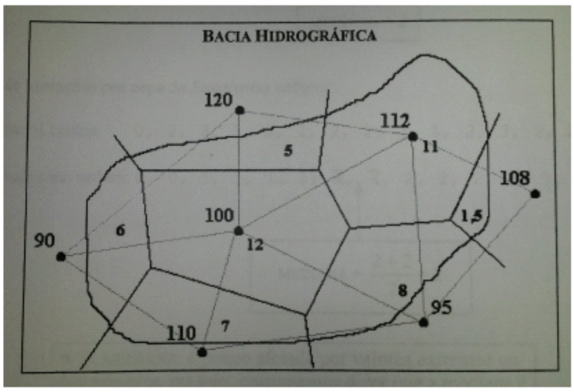
Média Ponderada

A média ponderada dos números $x_1, x_2, \ldots, x_n$, com pesos $p_1, p_2, \ldots, p_n$, representada por $\bar{x}_p$, é definida por

$$\bar{x}_p = \frac{p_1x_1 + p_2x_2 + \ldots + p_nx_n}{p_1 + p_2 + \ldots + p_n} = \frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i}$$

Exemplo: Precipitação média em uma bacia higrográfica

Precipitação (mm) X_i	Área do polígono (km^2) p_i	x_i/p_i
90	6	15
110	7	15,7
120	5	24
100	12	8,3
112	11	10,2
95	8	11,9
108	1,5	72
Total	50,5	



$\bar{x}_p =$ 104,24 mm

Total	50,5

$\bar{x}_p = 104,24$ mm

$$\bar{x}_p = \frac{90.6+110.7+120.5+100.12+112.11+95.8+108.15}{50,5}$$

$$\bar{x}_p = \frac{5264}{50,5} = 104,24$$

observações

$$\bar{x} = \frac{N_1+N_2+N_3}{3}$$

$$\bar{x} = \frac{1+2+3+4+5}{5}$$

$$\bar{x}_p = \frac{N_1 \cdot 5 + N_2 \cdot 6 + N_3 \cdot 8}{19}$$

Dados agrupados em tabelas de frequências

$$\sum f_i \cdot x_i = \frac{5.25+8.12+10.13}{50}$$

X_i	f_i
x_1	f_1
x_2	f_2
$\vdots$	$\vdots$
x_k	f_k
Total	n

X_i	f_i
5	25
8	12
10	13
	50

$$n = \sum_{i=1}^{k} f_i$$

$$\hat{\mu} = \bar{x} = \frac{f_1x_1 + f_2x_2 + \ldots + f_kx_k}{f_1 + f_2 + \ldots + f_k}.$$

Exemplo: Em um estudo realizado para avaliar o número de insetos capturados durante um determinado período, foram utilizadas 50 armadilhas. Os resultados são apresentados na Tabela a seguir.

Tabela 2: Distribuição de frequências para o conjunto de dados de insetos capturados

X_i	f_i
0	1
1	8
2	13
3	20
4	4
5	4
Total	50

$$\bar{x} = \frac{0.1 + 1.8 + 2.13 + 3.20 + 4.4 + 5.4}{50} = \frac{130}{50}$$

$\bar{x} = 2,6$ insetos

- **Dados agrupados em tabelas de classes de frequências**

X_i	m_i	f_i
$c_1 \vdash c_2$	m_1	f_1
$c_2 \vdash c_3$	m_2	f_2
$\vdots$	$\vdots$	$\vdots$
$c_k \vdash c_{k+1}$	m_k	f_k
Total		n

$$m_1 = \frac{c_1 + c_2}{2}$$

$\vdots$

$$m_i = \frac{c_i + c_{i+1}}{2}$$

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{k} m_i f_i}{\sum_{i=1}^{k} f_i}.$$

x

Exemplo: A distribuição de frequências dos pesos dos colmos de cana-de-açúcar, considerando-se uma amostra de tamanho 50, é apresentada a seguir:

$(10,46 + 11,55) \div 2$

Tabela 3: Distribuição de frequências dos pesos dos colmos de cana-de-açúcar

X_i	m_i	f_i	$m_i \cdot f_i$
10,46 ⊢ 11,55	11,005	2	22,01
11,55 ⊢ 12,64	12,095	7	84,665
12,64 ⊢ 13,73	13,185	4	52,74
13,73 ⊢ 14,82	14,275	8	114,2
14,82 ⊢ 15,91	15,365	7	107,555
15,91 ⊢ 17,00	16,455	10	164,55
17,00 ⊢ 18,09	17,545	7	122,815
18,09 ⊢ 19,18	18,635	5	93,175
Total		50	$\Sigma = 761,71$

$$\bar{x} = \frac{\Sigma\, m_i \cdot f_i}{n} = \frac{761,71}{50} = 15,2342$$

$\bar{x} = 15,23$

- **Mediana**
  A mediana é o valor que ocupa a posição central em um conjunto de dados ordenado em ordem crescente (Rol).

Logo,

$$\text{Md} = \begin{cases} x_{[\frac{n+1}{2}]}, & \text{se } n \text{ for ímpar} \\ \frac{x_{[n/2]}+x_{[n/2+1]}}{2}, & \text{se } n \text{ for par} \end{cases}$$

$$\frac{X_3 + X_4}{2} = Md$$

$n = 6$ (par)

X_1	X_2	X_3	X_4	X_5	X_6
55,3	55,4	55,8	58,3	59,6	80,2

$$\frac{55{,}8 + 58{,}3}{2} = 57{,}05 \text{ mediana}$$

$n = 7$ (ímpar)

X_1	X_2	X_3	X_4	X_5	X_6	X_7
55,3	55,4	55,8	58,3	58,3	59,6	80,2

$$Md = X_4$$

Observação: A Mediana é pouco afetada por valores extremos ou discrepantes!

Exemplo: Para os valores observados de produção média diária de leite, tem-se o seguinte rol:

8,78	9,34	9,80	9,90	9,90	9,95	10,00	10,34	11,75	15,00
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}

Como n é par...

$n = 10$

$$Md = \frac{X_5 + X_6}{2} = \frac{9{,}90 + 9{,}95}{2}$$

$$\text{Md} = \begin{cases} x_{[\frac{n+1}{2}]}, & \text{se } n \text{ for ímpar} \\ \frac{x_{[n/2]}+x_{[n/2+1]}}{2}, & \text{se } n \text{ for par} \end{cases}$$

$$\text{Md} = 9{,}925\ \text{Kg}$$

Conferindo...

$$\begin{aligned} \text{Md} &= \frac{x_{[5]} + x_{[6]}}{2} = \frac{9,90 + 9,95}{2} = \frac{19,85}{2} \\ &= 9,925\text{kg} \end{aligned}$$

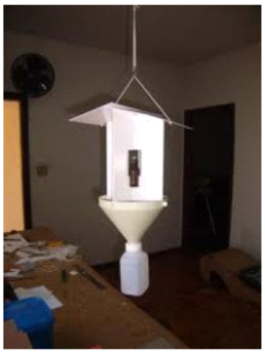
- **Dados agrupados em tabelas de frequências**

Para o exemplo realizado com insetos, para o qual foi observado o número insetos capturados por armadilha.

Tabela 3: Distribuição de frequências para o conjunto de dados número de insetos capturados

X_i	f_i	f'_i	F_i	F'_i
0	1	0,02	1	0,02
1	8	0,16	9	0,18
2	13	0,26	22	0,44
3	20	0,40	42	0,84
4	4	0,08	46	0,92
5	4	0,08	50	1,00
Total	50	1,00	—	—

Md = 3



Observar $F'_i \geq 0,50$

- **Dados agrupados em tabelas de classes de frequências**

Para o exemplo referente ao peso de colmos de cana-de-açúcar, tem-se:

Tabela 3: Distribuição de frequências dos pesos dos colmos

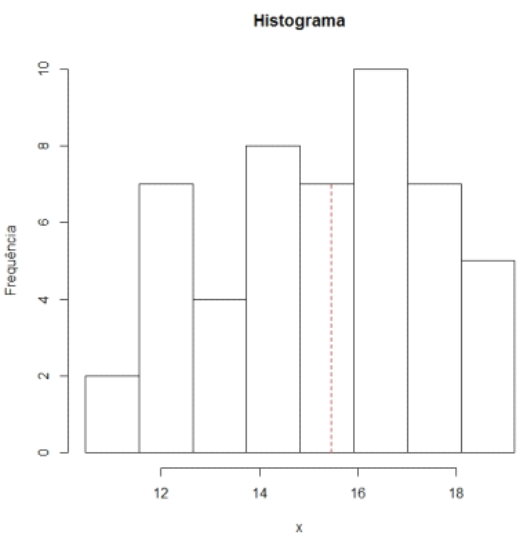
X_i	m_i	f_i	f'_i	F_i	F'_i
10,46 ⊢ 11,55	11,01	2	0,04	2	0,04
11,55 ⊢ 12,64	12,10	7	0,14	9	0,18
12,64 ⊢ 13,73	13,19	4	0,08	13	0,26
13,73 ⊢ 14,82	14,28	8	0,16	21	0,42
14,82 ⊢ 15,91	15,37	7	0,14	28	0,56
15,91 ⊢ 17,00	16,46	10	0,20	38	0,76

classe mediana

$$\begin{cases} 15,91 - 14,82 & -- \quad 0,14 \end{cases}$$

13,73 ⊢ 14,82	14,28	8	0,16	21	0,42
14,82 ⊢ 15,91	15,37	7	0,14	28	0,56
15,91 ⊢ 17,00	16,46	10	0,20	38	0,76
17,00 ⊢ 18,09	17,55	7	0,14	45	0,90
18,09 ⊢ 19,18	18,64	5	0,10	50	1,00
Total		50	1,00		

$$\begin{cases} 15,91 - 14,82 & -- \quad 0,14 \\ \text{Md} - 14,82 & -- \quad 0,08 \end{cases}$$



$14,82 \vdash 15,91$

$0,42 \quad 0,50 \quad 0,56$

$\neq 0,08 \quad \neq 0,06$

$\overbrace{15,91 - 14,82}^{1,09}$ — $0,14$

$\text{Md} - 14,82$ — $0,08$

$0,14(\text{Md} - 14,82) = 0,08 \cdot 1,09$

$0,14\,\text{Md} - 2,0748 = 0,0872$

$\text{Md} = \dfrac{0,0872 + 2,0748}{0,14} = 15,44$

$\text{Md} = 15,44\text{g}$

Conferindo...

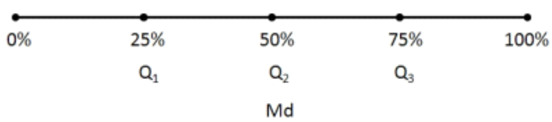
$$\text{Md} = \frac{15,91 - 14,82}{0,14} \times 0,08 + 14,82$$

$$= 15,44 \text{ g.}$$

Quartis e Percentis

- **Quartil**: generalização da mediana.

Quartil $\Rightarrow$ 4 partes



Ex: $n = 60$

$Q_3 = \dfrac{X_{45} + X_{46}}{2}$

$np = 60 \cdot 0,75$

$np = 45$

- **Percentil de ordem** $100p$

$P_{75} = Q_3 = \frac{X_{(45)} \ldots}{2}$

- **Percentil de ordem** $100p$

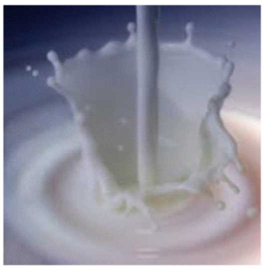
$$P_{100p} = \begin{cases} \frac{x_{[np]}+x_{[np+1]}}{2}, & \text{se } np \text{ for inteiro} \\ x_{[\text{int}(np)+1]}, & \text{se } np \text{ for não inteiro} \end{cases}$$

Ex.: $n = 67$

$P_{75} = Q_3 = X_{51}$

Exemplo: Para os valores observados de produção média diária de leite, tem-se:

$np = 67 \cdot 0{,}75 = 50{,}25$ $\quad 50+1 = 51$



$$P_{100p} = \begin{cases} \frac{x_{[np]}+x_{[np+1]}}{2}, & \text{se } np \text{ for inteiro} \\ x_{[\text{int}(np)+1]}, & \text{se } np \text{ for não inteiro} \end{cases}$$

8,78	9,34	9,80	9,90	9,90
9,95	10,00	10,34	11,75	15,00

$np = 10 \cdot 0{,}75 = 7{,}5$

$X_{7+1} = X_8$

$$n = 10$$

Obter $P_{75} = Q_3$ e P_{20}

$np = 10 \cdot 0{,}2 = 2$

$Q_3 = P_{75} = X_8 = 10{,}34$

$P_{20} = \frac{X_2 + X_3}{2} = \frac{9{,}34 + 9{,}80}{2} = 9{,}57$

Conferindo...

$P_{75} \quad \Rightarrow \quad np = 10 \times 0,75 = 7,5$

$$\begin{aligned} P_{75} = Q_3 &= x_{[\text{int}(7{,}5)+1]} \\ &= x_{[7+1]} = x_{[8]} \\ &= 10,34\text{kg} \end{aligned}$$

$P_{20} \quad \Rightarrow \quad np = 10 \times 0,20 = 2$

$$\begin{aligned} P_{20} &= \frac{x_{[2]}+x_{[3]}}{2} \\ &= \frac{9{,}34+9{,}80}{2} \\ &= 9,57\text{kg} \end{aligned}$$

- Dados agrupados em tabelas de frequências

Tabela 3: Distribuição de frequências para o conjunto de dados número de insetos capturados

X_i	f_i	f_i'	F_i	F_i'
0	1	0,02	1	0,02
1	8	0,16	9	0,18
2	13	0,26	22	0,44
3	20	0,40	42	0,84
4	4	0,08	46	0,92
5	4	0,08	50	1,00
Total	50	1,00		

Obter $P_{25} = Q_1$, P_{90} e $P_{97,5}$

Observar $F_i' \geq 0,25$

$Q_1 = P_{25} = 2$

Observar $F_i' \geq 0,90$

$P_{90} = 4$

Observar $F_i' \geq 0,975$

$P_{97,5} = 5$

Conferindo...

$P_{25} = Q_1 = 2$

$P_{90} = 4$

$P_{97,5} = 5$

- Dados agrupados em tabelas de classes de frequências

Para o exemplo referente ao peso de colmos de cana-de-açúcar, tem-se:

Tabela 3: Distribuição de frequências dos pesos dos colmos de cana-de-açúcar

X_i	m_i	f_i	f_i'	F_i	F_i'
10,46 ⊢ 11,55	11,01	2	0,04	2	0,04
11,55 ⊢ 12,64	12,10	7	0,14	9	0,18
12,64 ⊢ 13,73	13,19	4	0,08	13	0,26
13,73 ⊢ 14,82	14,28	8	0,16	21	0,42
14,82 ⊢ 15,91	15,37	7	0,14	28	0,56
15,91 ⊢ 17,00	16,46	10	0,20	38	0,76
17,00 ⊢ 18,09	17,55	7	0,14	45	0,90
18,09 ⊢ 19,18	18,64	5	0,10	50	1,00

0,08

$\to 0,2 - 0,18 = 0,02$

0,75 − 0,56 = 0,19

Obter P_{20} e $P_{75} = Q_3$

17,00 ⊢ 18,09	17,55	7	0,14	45	0,90
18,09 ⊢ 19,18	18,64	5	0,10	50	1,00
Total		50			

q,3. Obter P_{20} e $P_{75} = Q_3$

$\approx 0{,}19$

$$\begin{cases} \underbrace{13{,}73 - 12{,}64}_{1{,}09} & 0{,}08 \\ P_{20} - 12{,}64 & 0{,}02 \end{cases}$$

$$0{,}08\,(P_{20} - 12{,}64) = 0{,}02 \cdot 1{,}09$$

$$0{,}08\,P_{20} - 1{,}0112 = 0{,}0218$$

$$P_{20} = \frac{0{,}0218 + 1{,}0112}{0{,}08}$$

$$P_{20} \approx 12{,}91 \text{ kg}$$

$$\begin{cases} \underbrace{17 - 15{,}91}_{1{,}09} & 0{,}20 \\ P_{75} - 15{,}91 & 0{,}19 \end{cases}$$

$$0{,}20\,(P_{75} - 15{,}91) = 0{,}19 \cdot 1{,}09$$

$$0{,}20\,P_{75} - 3{,}182 = 0{,}2071$$

$$P_{75} = \frac{0{,}2071 + 3{,}182}{0{,}2}$$

$$P_{75} = 16{,}945 \text{ kg}$$

Conferindo...

$$\begin{cases} 13,73 - 12,64 & -- & 0,08 \\ P_{20} - 12,64 & -- & 0,02 \end{cases}$$

$$P_{20} = \frac{13{,}73 - 12{,}64}{0{,}08} \times 0,02 + 12,64$$
$$P_{20} = 12,91\text{g}.$$

$$\begin{cases} 17,00 - 15,91 & -- & 0,20 \\ P_{75} - 15,91 & -- & 0,19 \end{cases}$$

$$P_{75} = \frac{17{,}00 - 15{,}91}{0{,}20} \times 0,19 + 15,91$$
$$P_{20} = 16,95\text{g}.$$

- **Moda**: corresponde ao valor observado de maior frequência
  - Também pode ser obtida para variáveis qualitativas

Exemplo: Para os valores observados de produção média diária de leite, tem-se:

8,78	9,34	9,80	9,90	9,90
9,95	10,00	10,34	11,75	15,00

Mo = 9,90 kg

Conferindo...

Mo $= 9,90$ kg

- **Dados agrupados em tabelas de frequências**

Para o exemplo realizado com insetos

Tabela 3: Distribuição de frequências para o conjunto de dados número de insetos capturados

X_i	f_i	f_i'	F_i	F_i'
0	1	0,02	1	0,02
1	8	0,16	9	0,18
2	13	0,26	22	0,44
3	20	0,40	42	0,84
4	4	0,08	46	0,92
5	4	0,08	50	1,00
Total	50	1,00		

Mo = 3 insetos

Conferindo...

Mo $= 3$ insetos

- **Dados agrupados em tabelas de classes de frequências**

- Moda bruta: ponto médio da classe com maior frequência

- Moda: método de Czuber $\Rightarrow$ semelhança de triângulos

Para o exemplo referente ao peso dos colmos de cana-de-açúcar, tem-se:

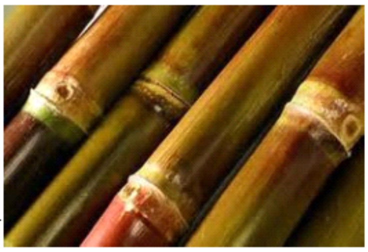
Tabela 3: Distribuição de frequências dos pesos dos colmos de cana-de-açúcar

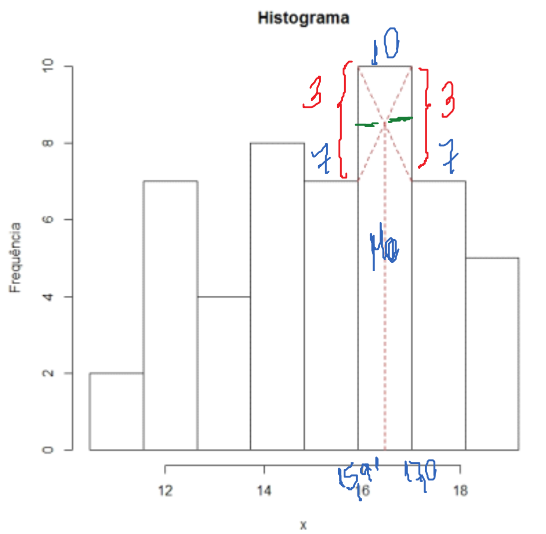
Mo Bruta:

$$Mo = \frac{15,91+17}{2}$$

$$Mo = 16,46$$

X_i	m_i	f_i	f_i'	F_i	F_i'
10,46 ⊢ 11,55	11,01	2	0,04	2	0,04
11,55 ⊢ 12,64	12,10	7	0,14	9	0,18
12,64 ⊢ 13,73	13,19	4	0,08	13	0,26
13,73 ⊢ 14,82	14,28	8	0,16	21	0,42
14,82 ⊢ 15,91	15,37	7	0,14	28	0,56
15,91 ⊢ 17,00	16,46	10	0,20	38	0,76
17,00 ⊢ 18,09	17,55	7	0,14	45	0,90
18,09 ⊢ 19,18	18,64	5	0,10	50	1,00
Total		50	1,00		





$$3 \;—\; 3$$
$$Mo - 15,91 \;—\; 17,0 - Mo$$

$$3(Mo - 15,91) = 3(17 - Mo)$$

$$3Mo - 47,73 = 51 - 3Mo$$

$$6Mo = 51 + 47,73$$

$$Mo = \frac{98,73}{6} = 16,46$$

$$Mo = 16,46$$

$$\begin{cases} 10 - 7 & -- \quad 10 - 7 \\ \text{Mo} - 15,91 & -- \quad 17 - \text{Mo} \end{cases}$$

Conferindo...

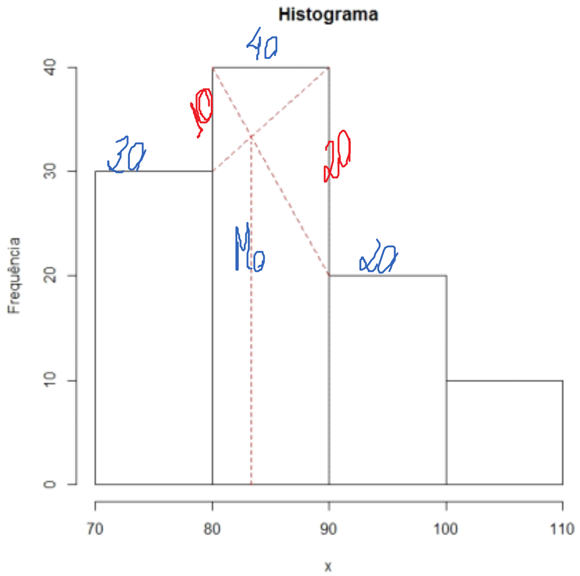
Moda bruta:

$$\begin{aligned} \text{Mo} &= \frac{17,00 + 15,91}{2} \\ &= 16,46 \text{ g} \end{aligned}$$

Moda por Czuber:

$$\begin{aligned} 17,00 - \text{Mo} &= \text{Mo} - 15,91 \\ \text{Mo} &= 16,46\text{g} \end{aligned}$$

Supondo o seguinte histograma para uma variável X qualquer.



$$\text{Mo} = LI_{Mo} + \frac{\Delta_1}{\Delta_1 + \Delta_2} h,$$

$\Delta_1 = 40 - 30 = 10$

$\Delta_2 = 40 - 20 = 20$

$h = 10$

$Mo = 80 + \frac{10}{10+20} \cdot 10$

$Mo = 83,33$

Classe correspondente à moda:
$80 \vdash 90$

ou

$10 \text{ —— } 20$

$Mo - 80 \text{ —— } 90 - Mo$

$20(Mo - 80) = 10(90 - Mo)$

$20Mo - 1600 = 900 - 10Mo$

$Mo \text{ bruta} = \frac{80+90}{2} = 85$

$$20Mo - 1600 = 900 - 10Mo$$

$$20Mo + 10Mo = 900 + 1600$$

$$Mo = \frac{2500}{30} \qquad Mo = 83,33$$

Conferindo...

Moda bruta:

Moda bruta: Mo $= 85$

Moda por Czuber:

$$\begin{cases} 40 - 30 & -- & 40 - 20 \\ \text{Mo} - 80 & -- & 90 - \text{Mo} \end{cases}$$

$$(40-30)(90-\text{Mo}) = (40-20)(\text{Mo}-80)$$

$$\text{Mo} = \frac{(40-30) \times 90 + (40-20) \times 80}{(40-30) + (40-20)}$$
$$= 83,33$$

Ou... Pela fórmula

Fórmula:

$$\text{Mo} = LI_{Mo} + \frac{\Delta_1}{\Delta_1 + \Delta_2} h,$$

em que:

- LI_{Mo} é o limite inferior da classe modal,
- Δ_1 é a diferença de frequência entre a classe modal e a anterior,
- Δ_2 é a diferença de frequência entre a classe modal e a seguinte,
- h é a amplitude do intervalo de classe.