Medidas de Posição e Dispersão para Dados Agrupados com Intervalos de Classe

Média Aritmética

Temos dois processos a serem utilizados: O processo longo e o processo breve.

Processo Longo: Para utilizarmos o processo longo convencionamos que todos os valores incluídos em um determinado intervalo de classe coincidem com o seu ponto médio, e determinamos a média aritmética ponderada por meio da fórmula

$$\boxed{\overline{X} = \frac{\sum PM_i \cdot f_i}{\sum f_i}}$$

Exemplo: Considere a distribuição:

i	ESTATURAS (cm)	f_i
1	150 ⊢ 154	4
2	154 ⊢ 158	9
3	158 ⊢ 162	11
4	162 ⊢ 166	8
5	166 ⊢ 170	5
6	170 ⊢ 174	3
		$\sum = 40$

Vamos abrir uma coluna para os pontos médios e outra para os produtos $PM_i \cdot f_i$, então:

i	ESTATURAS (cm)	f_i	PM_i	$PM_i f_i$
1	150 ⊢ 154	4	152	608
2	154 ⊢ 158	9	156	1404
3	158 ⊢ 162	11	160	1760
4	162 ⊢ 166	8	164	1312
5	166 ⊢ 170	5	168	840
6	170 ⊢ 174	3	172	516
		$\sum = 40$		$\sum = 6440$

$$\overline{X} = \frac{\sum PM_i \cdot f_i}{\sum f_i} = \frac{6440}{40} = 161 \text{, logo } \overline{X} = 161\,cm$$

b) Processo Breve: Baseia-se em uma mudança de variável PM_i por outra y, tal que $y_i = \frac{PM_i - PM_0}{h}$, onde PM_0 é o ponto médio da maior frequência.

Com essa mudança de variável a fórmula resulta em $\boxed{\overline{X} = PM_0 + \frac{\left(\sum y_i \cdot f_i\right) \cdot h}{\sum f_i}}$

Na tabela do exemplo anterior $PM_0 = 160$ e como $h = 4$, temos como os novos valores de y_i.

$y_1 = \frac{152-160}{4} = \frac{-8}{4} = -2$ $\quad y_2 = \frac{156-160}{4} = \frac{-4}{4} = -1$ $\quad y_3 = \frac{160-160}{4} = \frac{0}{4} = 0$

$y_4 = \frac{164-160}{4} = \frac{4}{4} = 1$ $\quad y_5 = \frac{168-160}{4} = \frac{8}{4} = 2$ $\quad y_6 = \frac{172-160}{4} = \frac{12}{4} = 3$

Completando a tabela com as colunas correspondentes aos pontos médios, aos valores da nova variável y e aos produtos $y_i f_i$

i	ESTATURAS (cm)	f_i	PM_i	y_i	$y_i f_i$
1	150 ⊢ 154	4	152	-2	-8
2	154 ⊢ 158	9	156	-1	-9
3	158 ⊢ 162	11	160	0	0
4	162 ⊢ 166	8	164	1	8
5	166 ⊢ 170	5	168	2	10
6	170 ⊢ 174	3	172	3	9
		$\sum = 40$			$\sum = 10$

Substituindo os valores na fórmula:

$$\overline{X} = PM_0 + \frac{\left(\sum y_i \cdot f_i\right) \cdot h}{\sum f_i} = 160 + \frac{10 \cdot 4}{40} = 160 + \frac{40}{40} = 160 + 1$$

$$\overline{X} = 161\,cm$$

Resumo:

1ª) Abrimos uma coluna para os valores PM_i

2ª) Escolhermos um dos pontos médios (de prefrência o de maior frequência) para valor de PM_0

3ª) Abrimos uma coluna para os valores de y_i e escrevemos ZERO na linha correspondente à classe onde se encontra o valor de PM_0, a sequência -1, -2, -3, ... logo acima do zero e a sequência 1, 2, 3,... logo abaixo.

4ª) Abirmos uma coluna para os valores do produto $y_i f_i$, conservando os sinais + ou -, e, em seguida, somamos algebricamente esses produtos.

5ª) Aplicamos a fórmula.

Moda

A classe que apresenta a maior frequência é denominada classe modal. Para esse cálculo utilizamos a fórmula de Czuber

$$\boxed{Mo = l_{MO} + \frac{f_{Mo} - f_{Mo}ant}{2 f_{Mo} - \left(f_{Mo}ant + f_{Mo}post\right)} \cdot h}$$

No qual

l_{Mo} é o limite inferior da classe modal

h é a amplitude da classe modal

f_{Mo} é a frequência simples da classe modal

$f_{Mo}ant$ é a frequência simples da classe anterior à classe modal

$f_{Mo}\,post$ é a frequência simples da classe posterior à classe modal

Exemplo: Calcule a moda da distrubuição

i	ESTATURAS (cm)	f_i
1	150 ⊢ 154	4
2	154 ⊢ 158	9
3	158 ⊢ 162	11
4	162 ⊢ 166	8
5	166 ⊢ 170	5
6	170 ⊢ 174	3
		$\sum = 40$

$$Mo = 158 + \frac{11-9}{(11-9)+(11-8)} \cdot 4 = 158 + \frac{2}{2+3} \cdot 4 = 158 + \frac{8}{5} = 158 + 1{,}6 = 159{,}6$$

$$Mo = 159{,}6$$

<u>Mediana</u>

Primeiro determinar a classe na qual se acha a mediana que será a correspondente à frequência acumulada crescente imediatamente superior a $\frac{\sum f_i}{2}$, e utilizamos a fórmula

$$\boxed{Md = l_{Md} + \frac{\left[\frac{\sum f_i}{2} - fac_{Md}(ant)\right].h}{f_{Md}}}$$

Calculando a mediana para o exemplo das alturas temos:

i	ESTATURAS (cm)	f_i	fac_i
1	150 ⊢ 154	4	4
2	154 ⊢ 158	9	13
3	158 ⊢ 162	11	24
4	162 ⊢ 166	8	32
5	166 ⊢ 170	5	37
6	170 ⊢ 174	3	40
		$\sum = 40$	

$$\frac{\sum f_i}{2} = \frac{40}{2} = 20$$

$$Md = 158 + \frac{[20-13] \cdot 4}{11} = 158 + \frac{7 \cdot 4}{11} = 158 + \frac{28}{11} = 158 + 2{,}54 = 160{,}54$$

No caso de existir umaa frequência acumulada exatamente igual a $\frac{\sum f_i}{2}$, a mediana será o limite superior da classe correspondente.
Exemplo:

i	CLASSES	f_i	fac_i
1	0 ⊢ 10	1	1
2	10 ⊢ 20	3	4
3	20 ⊢ 30	9	13
4	30 ⊢ 40	7	20
5	40 ⊢ 50	4	24
6	50 ⊢ 60	2	26
		$\sum = 26$	

$\frac{\sum f_i}{2} = \frac{26}{2} = 13$, logo $Md = L_{Md} = 30$

<u>Desvio Padrão:</u>
a) Processo Longo:

$$\boxed{s = \sqrt{\frac{\sum f_i \cdot PM_i^{\ 2}}{\sum f_i} - \left(\frac{\sum f_i \cdot PM_i}{\sum f_i}\right)^2}}$$

Primeiro abrimos colunas para PM_i, para $f_i PM_i$, para $PM_i^{\ 2}$ e para $f_i \cdot PM_i^{\ 2}$

i	ESTATURAS (cm)	f_i	PM_i	$f_i PM_i$	$PM_i^{\ 2}$	$f_i PM_i^{\ 2}$
1	150 ⊢ 154	4	152	608	23104	92416
2	154 ⊢ 158	9	156	1404	24336	219024
3	158 ⊢ 162	11	160	1760	25600	281600
4	162 ⊢ 166	8	164	1312	26896	215168
5	166 ⊢ 170	5	168	840	28224	141120
6	170 ⊢ 174	3	172	516	29584	88752
		$\sum = 40$		$\sum = 6440$		$\sum = 1.038.080$

$$s = \sqrt{\frac{1.038.080}{40} - \left(\frac{6440}{40}\right)^2} = \sqrt{25.952 - 25.921} = \sqrt{31} = 5{,}567$$

$s = 5{,}57cm$

b) Processo Breve

Baseado na mudança de variável x por outra $y_i = \frac{PM_i - PM_0}{h}$ e abrem-se as colunas correspondentes para o uso da fórmula.

$$\boxed{s = h\sqrt{\frac{\sum f_i \cdot y_i^2}{\sum f_i} - \left(\frac{\sum f_i \cdot y_i}{\sum f_i}\right)^2}}$$

i	ESTATURAS (cm)	f_i	PM_i	y_i	$y_i\,f_i$	$f_i y_i^2$
1	150 ⊢ 154	4	152	-2	-8	16
2	154 ⊢ 158	9	156	-1	-9	9
3	158 ⊢ 162	11	160	0	0	0
4	162 ⊢ 166	8	164	1	8	8
5	166 ⊢ 170	5	168	2	10	20
6	170 ⊢ 174	3	172	3	9	27
		$\sum = 40$			$\sum = 10$	$\sum = 80$

$$s = 4\sqrt{\frac{80}{40} - \left(\frac{10}{40}\right)^2} = 4\sqrt{2 - 0{,}0625} = 4\sqrt{1{,}9375} = 4 \cdot 1{,}3919 = 5{,}5676$$

$$s = 5{,}57cm$$

Resumo:

1ª) Abrimos uma coluna para os valores PM_i

2ª) Escolhermos um dos pontos médios (de prefrência o de maior frequência) para valor de PM_0

3ª) Abrimos uma coluna para os valores de y_i e escrevemos ZERO na linha correspondente à classe onde se encontra o valor de PM_0, a sequência -1, -2, -3, ... logo acima do zero e a sequência 1, 2, 3,... logo abaixo.

4ª) Abrimos uma coluna para os valores do produto $f_i y_i$, conservando os sinais + ou -, e, em seguida, somamos algebricamente esses produtos.

5ª) Abrimos uma coluna para os valores do produto $f_i y_i^2$, obtidos multiplicando dada $f_i y_i$ pelo seu respectivo y_i e em seguida somamos esses produtos

6ª) Aplicamos a fórmula.