Escola Secundária de Jácome Ratton

Medidas de Dispersão

Ano Lectivo 2010/2011

Matemática Aplicada às Ciências Sociais

Na Escola Secundária do Sucesso anualmente é premiado o aluno que tiver melhor média nas suas classificações nas diferentes disciplinas. No ano lectivo 2009/10, os dois melhores alunos obtiveram as seguintes classificações:

Nuno	18	16	17	20	18	19	18	18
Leonor	18	17	15	18	20	20	18	18

a) Qual é a média das classificações de cada aluno? É possível indicar quem é o melhor aluno tendo como base apenas a média?

Nuno:

$$\bar{x} = \frac{18+16+17+20+18+19+18+18}{8} \Leftrightarrow \bar{x} = \frac{144}{8} \Leftrightarrow \bar{x} = 18$$

Leonor:

$$\bar{x} = \frac{18+17+15+18+20+20+18+18}{8} \Leftrightarrow \bar{x} = \frac{144}{8} \Leftrightarrow \bar{x} = 18$$

Como a média das classificações dos dois alunos é 18 não é possível indicar quem é o melhor aluno.

Nuno	18	16	17	20	18	19	18	18
Leonor	18	17	15	18	20	20	18	18

b) Determine a moda, a mediana, o 1º e o 3º quartil para cada conjunto de dados.

Nuno:

$M_o = 18$

16 17 18 18 18 18 19 20

$\tilde{x} = 18$

$Q_1 = 17{,}5$

$Q_3 = 18{,}5$

Leonor:

$M_o = 18$

15 17 18 18 18 18 20 20

$\tilde{x} = 18$

$Q_1 = 17{,}5$

$Q_3 = 19$

Nuno	18	16	17	20	18	19	18	18
Leonor	18	17	15	18	20	20	18	18

c) Calcule, para cada distribuição, a amplitude do conjunto de dados.

Nuno:

$$20-16=4$$

Leonor:

$$20-15=5$$

Amplitude de um conjunto de dados:

Chama-se amplitude de um conjunto de dados à diferença entre o máximo e o mínimo desse conjunto de dados.

Nota:

- A amplitude não deve ser utilizada para comparar a variabilidade de várias amostras, excepto se tiverem a mesma dimensão.
- Medida **pouco resistente** à presença uma observação muito alta ou muito baixa (*outliers*).

Nuno	18	16	17	20	18	19	18	18
Leonor	18	17	15	18	20	20	18	18

d) Para cada distribuição, determine a diferença entre o 3º e o 1º quartil.

Nuno:

$Q_1 = 17{,}5$ $\quad$ $Q_3 = 18{,}5$ $\quad$ $Q_3 - Q_1 = 18{,}5 - 17{,}5 = 1$

Leonor:

$Q_1 = 17{,}5$ $\quad$ $Q_3 = 19$ $\quad$ $Q_3 - Q_1 = 19 - 17{,}5 = 1{,}5$

Amplitude interquartis:

Chama-se amplitude interquartis à diferença entre o 3º e o 1º quartil, isto é, igual a $Q_3 - Q_1$.

Nota:

- Uma amplitude interquartis nula não significa que não exista variabilidade. Mas, se as observações forem todas iguais a amplitude interquartis é nula.

- Medida **resistente** relativamente à existência de *outliers*, porque é definida à custa dos quartis que são medidas resistentes;

Nuno	18	16	17	20	18	19	18	18
Leonor	18	17	15	18	20	20	18	18

e) As medidas de localização determinadas nas alíneas a) e b) dão alguma indicação relativamente à dispersão dos dados? E as medidas calculadas nas alínea c) e d)?

As medidas de localização não caracterizam a variabilidade dos dados, enquanto que as medidas calculadas nas alíneas c) e d) ajudam a descrever essa dispersão.

Nuno	18	16	17	20	18	19	18	18
Leonor	18	17	15	18	20	20	18	18

f) O conjunto de professores, a quem competia escolher o vencedor, decidiu escolher o aluno que tivesse obtido classificações mais concentradas em torno da média para isso baseou-se na tabela que se segue. Copie a tabela e complete-a.

$$\bar{x} = 18$$

Classificações do Nuno					Classificações da Leonor				
x_i	n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$n_i \times (x_i - \bar{x})^2$	x_i	n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$n_i \times (x_i - \bar{x})^2$
16	1	-2	$(-2)^2=4$	1x4=4	15	1	-3	9	9
17	1	-1	$(-1)^2=1$	1x1=1	17	1	-1	1	1
18	4	0	$0^2=0$	4x0=0	18	4	0	0	0
19	1	1	$1^2=1$	1x1=1	20	2	2	4	8
20	1	2	$2^2=4$	1x4=4	---	---	---	-----	------
Total	8	0	10	10	Total	8	-2	14	18

Variância:

A variância é a razão entre a soma dos quadrados dos desvios e o número de dados menos um.

$$s^2 = \frac{\left(x_1 - \overline{x}\right)^2 + \left(x_2 - \overline{x}\right)^2 + \cdots + \left(x_n - \overline{x}\right)^2}{n-1} = \frac{\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2}{n-1}$$

Se os k valores diferentes da variável têm ni como frequências absolutas então:

$$s^2 = \frac{n_1\left(x_1 - \overline{x}\right)^2 + n_2\left(x_2 - \overline{x}\right)^2 + \cdots + n_k\left(x_k - \overline{x}\right)^2}{n-1} = \frac{\sum_{i=1}^{k} n_i \left(x_i - \overline{x}\right)^2}{n-1}$$

Nota:

- A variância permite avaliar o grau de dispersão dos valores da variável em relação à média, quanto maior for o seu valor maior será a dispersão.
- **Mas** a sua unidade é o quadrado da unidade das observações, o que pode levantar problemas de interpretação.

Classificações do Nuno					Classificações da Leonor				
x_i	n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$n_i \times (x_i - \bar{x})^2$	x_i	n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$n_i \times (x_i - \bar{x})^2$
16	1	-2	$(-2)^2=4$	1x4=4	15	1	-3	9	9
17	1	-1	$(-1)^2=1$	1x1=1	17	1	-1	1	1
18	4	0	$0^2=0$	4x0=0	18	4	0	0	0
19	1	1	$1^2=1$	1x1=1	20	2	2	4	8
20	1	2	$2^2=4$	1x4=4	---	---	---	-----	------
Total	8	0	10	10	**Total**	8	-2	14	18

g) Determine a variância das classificações do Nuno e das classificações da Leonor.

Nuno:

$$s^2 = \frac{10}{8-1} = \frac{10}{7} \approx 1{,}43$$

Leonor:

$$s^2 = \frac{18}{8-1} = \frac{18}{7} \approx 2{,}57$$

Desvio-Padrão:

O desvio-padrão é a raiz quadrada positiva da variância.

$$s = \sqrt{\frac{\left(x_1 - \bar{x}\right)^2 + \left(x_2 - \bar{x}\right)^2 + \cdots + \left(x_n - \bar{x}\right)^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}}$$

Se os k valores diferentes da variável têm *ni* como frequências absolutas então:

$$s = \sqrt{\frac{n_1\left(x_1 - \bar{x}\right)^2 + n_2\left(x_2 - \bar{x}\right)^2 + \cdots + n_k\left(x_k - \bar{x}\right)^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{k} n_i\left(x_i - \bar{x}\right)^2}{n-1}}$$

Nota:

- O desvio-padrão é sempre não negativo e quanto maior for o seu valor maior será a dispersão dos dados em relação à média.
- Se o desvio-padrão é igual a zero é porque não existe variabilidade, isto é, os dados são todos iguais.

Classificações do Nuno					Classificações da Leonor				
x_i	n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$n_i \times (x_i - \bar{x})^2$	x_i	n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$n_i \times (x_i - \bar{x})^2$
16	1	-2	$(-2)^2=4$	1x4=4	15	1	-3	9	9
17	1	-1	$(-1)^2=1$	1x1=1	17	1	-1	1	1
18	4	0	$0^2=0$	4x0=0	18	4	0	0	0
19	1	1	$1^2=1$	1x1=1	20	2	2	4	8
20	1	2	$2^2=4$	1x4=4	---	---	---	-----	------
Total	8	0	10	10	Total	8	-2	14	18

h) Calcule o desvio-padrão das classificações do Nuno e das classificações da Leonor.

Nuno:

$$s = \sqrt{\frac{10}{8-1}} = \sqrt{\frac{10}{7}} \approx 1{,}20$$

Leonor:

$$s = \sqrt{\frac{18}{8-1}} = \sqrt{\frac{18}{7}} \approx 1{,}60$$

i) De acordo com os critérios do conjunto de professores qual dos alunos será o vencedor? Justifique.

O vencedor será o Nuno porque as suas classificações estão menos dispersas relativamente à média.

Influência da alteração de dados no Desvio-Padrão

- Ao adicionar a cada dado uma constante, o desvio-padrão da distribuição não se altera.

- Ao multiplicar cada dado por uma constante, diferente de zero, o desvio-padrão vem multiplicado por essa constante.