

Introdução à Probabilidade e Estatística



Universidade de Évora

Departamento de Matemática

Ano lectivo 2015/16

Ana Isabel Santos

Estatística Descritiva

Aula 2

Exercício 1.1 :

Considere os resultados finais de Estatística de 20 estudantes de uma Universidade:

9 14 12 8 14 12 16 16 8 14 11 12 12 11 11 18 14 18 15 15

- a)** Os dados em estudo são de tipo qualitativo ou quantitativo?
- b)** Construa a tabela de frequências associada a estes dados.
- c)** Represente graficamente a informação disponibilizada.
- d)** Determine e interprete a média, a moda e a mediana.
- e)** Calcule a variância e o desvio padrão.
- f)** Calcule e interprete o valor do percentil 15 e do 8º decil.
- g)** Indique a amplitude da amostra e a amplitude interquartil.
- h)** Determine e interprete o coeficiente de variação.
- i)** Estude a distribuição dos dados quanto à assimetria e ao achatamento.
- j)** Apresente os dados numa caixa de bigodes.

Medidas de Dispersão

Coeficiente de dispersão:

$$CD = \frac{S}{X}, \quad 0 \leq CD \leq 1;$$

Coeficiente de variação:

$$CV = \frac{S}{X} \times 100\%, \quad 0 \leq CV \leq 100;$$

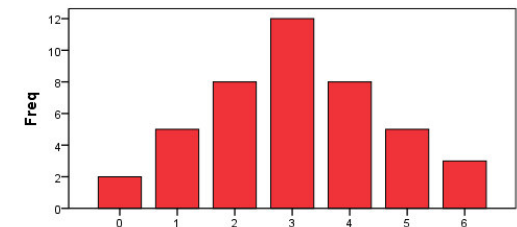
Para valores de *CV* inferiores a 50%, a média será tanto mais representativa dos valores da amostra quanto menor for o valor deste coeficiente. Valores de *CV* superiores a 50% indicam uma baixa representatividade da média.

Medidas de Assimetria (Skewness)

A distribuição dos dados classifica-se, quanto à assimetria, em:

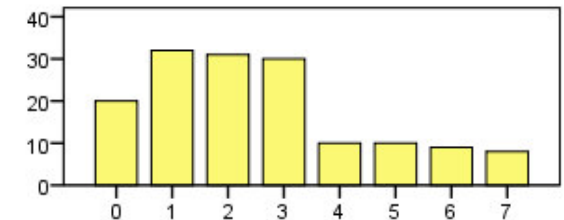
➤ **Simétrica**

$$\bar{X} = M_e = M_0$$



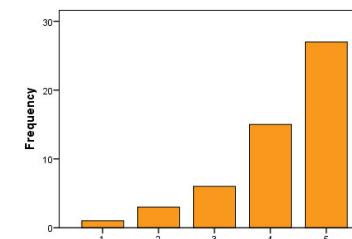
➤ **Assimétrica Positiva ou enviesada à esquerda**

$$\bar{X} > M_e > M_0$$



➤ **Assimétrica Negativa ou enviesada à direita**

$$\bar{X} < M_e < M_0$$



Medidas de Assimetria (Skewness)

➤ Grau de Assimetria de Pearson:

$$G_P = \frac{\bar{X} - M_0}{S}, \quad -3 < G_P < 3$$

➤ Grau de Assimetria de Bowley:

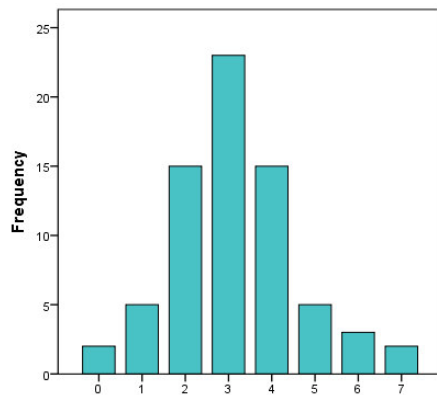
$$G_B = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}, \quad -1 < G_B < 1$$

O tipo de assimetria é determinado pelo sinal de G_P ou de G_B

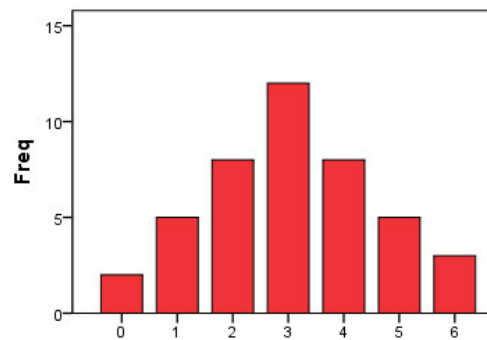
> 0 – **Assimétrica+**; $= 0$ – **Simétrica**; < 0 – **Assimétrica-**

Medidas de Achatamento (Kurtosis)

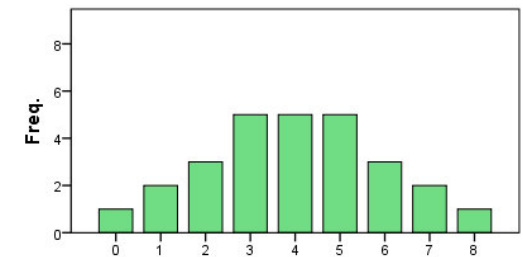
A distribuição dos dados classifica-se, quanto ao achatamento, em:



Leptocúrtica



Mesocúrtica



Platicúrtica

Coeficiente Percentil de Achatamento:

$$K = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})} \begin{cases} < 0.263 & \text{Distribuição Leptocúrtica} \\ = 0.263 & \text{Distribuição Mesocúrtica} \\ > 0.263 & \text{Distribuição Platicúrtica} \end{cases}$$

Assimetria e Achatamento no SPSS

Coeficiente de assimetria do SPSS:

$$g_{SPSS} = \frac{Skewness}{Std. \text{ error of Skewness}}.$$

- ★ Se $g_{SPSS} < -1,96$, então assume-se que a distribuição é **Assimétrica Negativa**;
- ★ Se $|g_{SPSS}| < 1,96$, então assume-se que a distribuição é **Simétrica**;
- ★ Se $g_{SPSS} > 1,96$, então assume-se que a distribuição é **Assimétrica Positiva**.

Coeficiente de achatamento do SPSS:

$$K_{SPSS} = \frac{Kurtosis}{Std. \text{ error of Kurtosis}}.$$

- ★ Se $K_{SPSS} < -1,96$, então assume-se que a distribuição é **Platicúrtica**;
- ★ Se $|K_{SPSS}| < 1,96$, então assume-se que a distribuição é **Mesocúrtica**;
- ★ Se $K_{SPSS} > 1,96$, então assume-se que a distribuição é **Leptocúrtica**.

Exercício 1.1 : *outputs do SPSS*

Notas finais de Estatística

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 8	2	10,0	10,0	10,0
9	1	5,0	5,0	15,0
11	3	15,0	15,0	30,0
12	4	20,0	20,0	50,0
14	4	20,0	20,0	70,0
15	2	10,0	10,0	80,0
16	2	10,0	10,0	90,0
18	2	10,0	10,0	100,0
Total	20	100,0	100,0	

Statistics

Notas finais de Estatística

N	Valid	20
	Missing	0
Mean		13,00
Median		13,00
Mode		12 ^a
Std. Deviation		2,920
Variance		8,526
Skewness		-,042
Std. Error of Skewness		,512
Kurtosis		-,552
Std. Error of Kurtosis		,992
Minimum		8
Maximum		18
Percentiles	15	9,30
	25	11,00
	50	13,00
	75	15,00
	80	15,80

a. Multiple modes exist. The smallest value is shown

Outliers

Um **Outlier** é um valor que se afasta de modo evidente do centro da distribuição.

X_i é um **outlier moderado** se ultrapassa uma das barreiras **moderadas**:

$$X_i < Q_1 - 1.5 IQ \quad \text{ou} \quad X_i > Q_3 + 1.5 IQ$$

X_i é um **outlier severo** se ultrapassa uma das barreiras **severas**:

$$X_i < Q_1 - 3 IQ \quad \text{ou} \quad X_i > Q_3 + 3 IQ$$

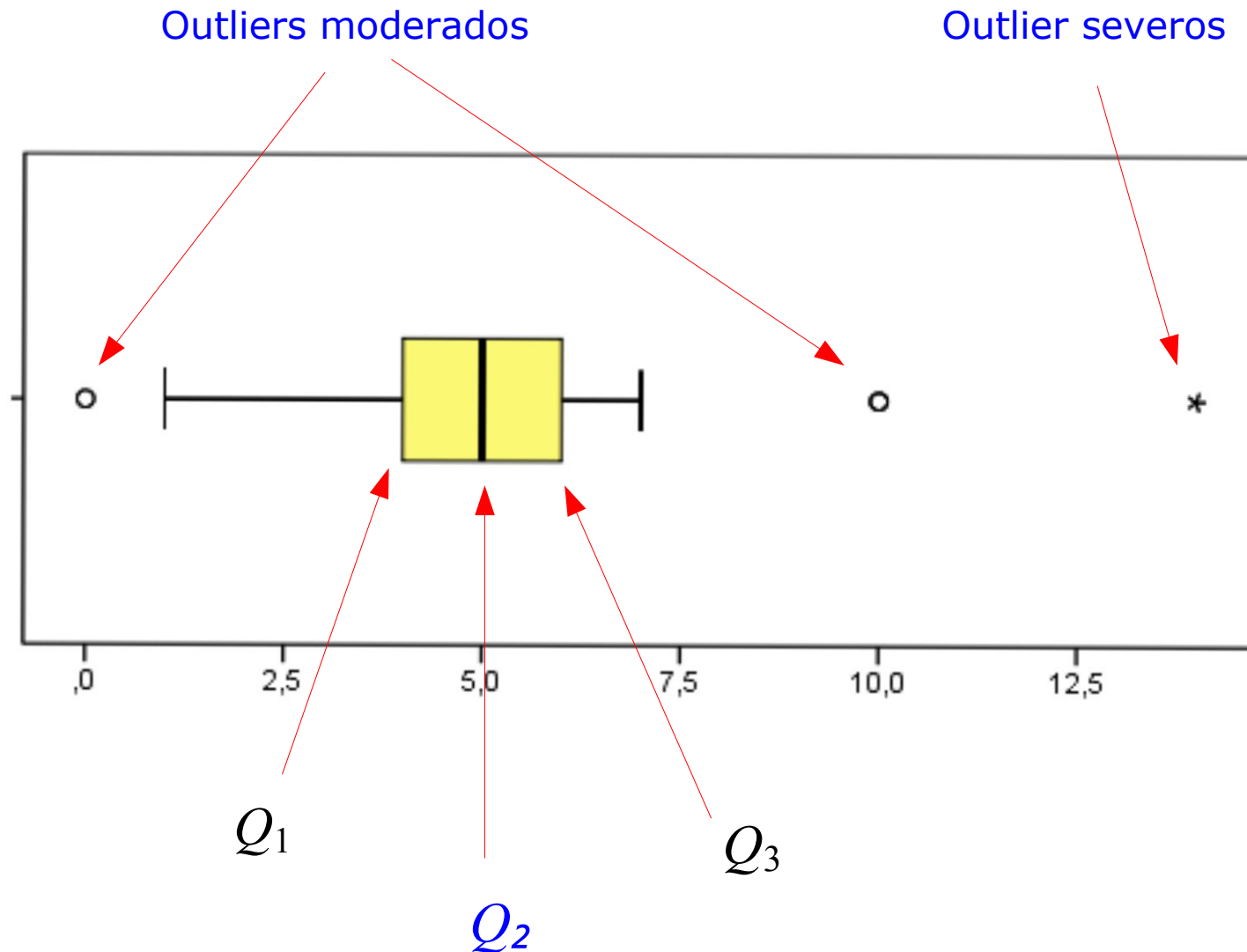
Caixa de bigodes, diagrama de extremos e quartis ou Boxplot

Caixa de Bigodes: apresenta algumas das principais características descritivas de um certo conjunto de dados, numa imagem compacta. Encontram-se representados Q_1 , Q_2 , Q_3 , IQ , menor valor não outlier, o maior valor não outlier e os outliers.

Fornece uma boa visualização da variabilidade dos dados e do tipo da assimetria e achatamento da distribuição.

Nota: Nas boxplot do SPSS os outliers moderados são assinalados com um "O" e os outliers severos com um "★".

Caixa de bigodes, diagrama de extremos e quartis ou Boxplot



Exercício 1.1 : *Resolução no SPSS*

➤ No OUTPUT do SPSS

GRAPHS

LEGACY DIALOGS

BOX-PLOT

SIMPLE

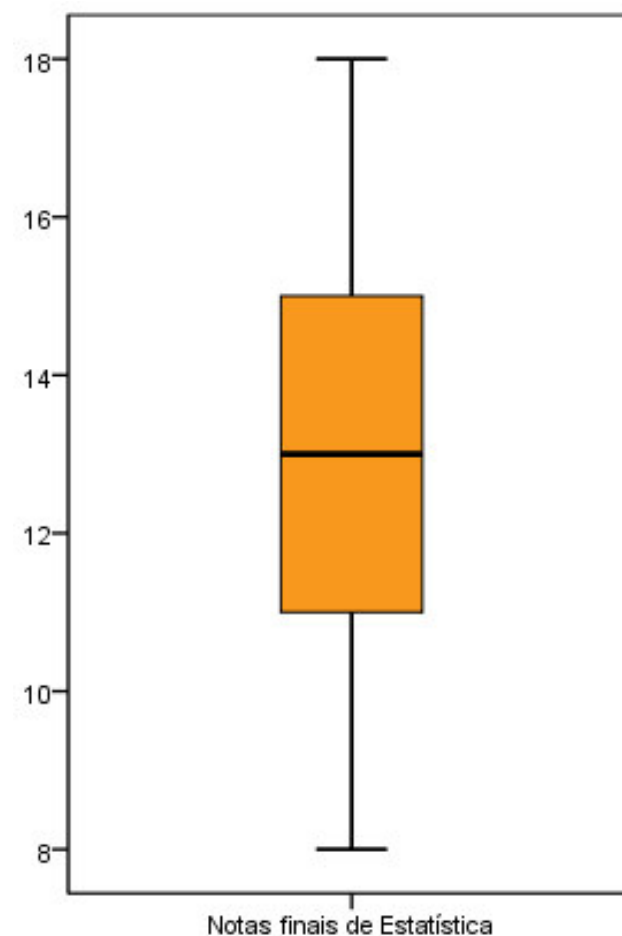
DATA IN CHART ARE

SUMMARIES OF SEPARATE VARIABLES

DEFINE

selecionar a var (notas_fin) para boxes represent

OK



Exercício 3:

Recolheu-se a seguinte informação diária referente à humidade relativa (em %) e à temperatura máxima (em °C) na determinada estação meteorológica no mês de agosto de 2014:

Statistics

Áreas		Humidade	Temperatura
N	Valid	31	31
	Missing	0	0
Mean		53,224	23,252
Median		55,273	21,997
Variance		169,278	32,464
Skewness		-0,867	0,948
Std. Error of Skewness		0,421	0,421
Kurtosis		0,172	-0,4108
Std. Error of Kurtosis		0,821	0,821
Minimum		22,187	15,895
Maximum		71,215	35,335
Percentiles	10	33,137	17,763
	25	45,246	18,804
	75	62,148	24,613
	90	69,092	32,748