

## Introdução à Probabilidade e Estatística

### Ficha N°1: Estatística Descritiva

Para as licenciaturas em: Eng. Civil, Eng. das Energias Renováveis, Eng. Geológica,  
Eng. Informática e Eng. Mecatrónica

**2º semestre 2014/15 — 2h Teóricas + 2h Práticas**

Docentes: Patrícia Filipe e Ana Isabel Santos

**Nota:** Alguns dos exercícios propostos nesta ficha foram retirados dos manuais constantes na bibliografia (ver programa da cadeira).

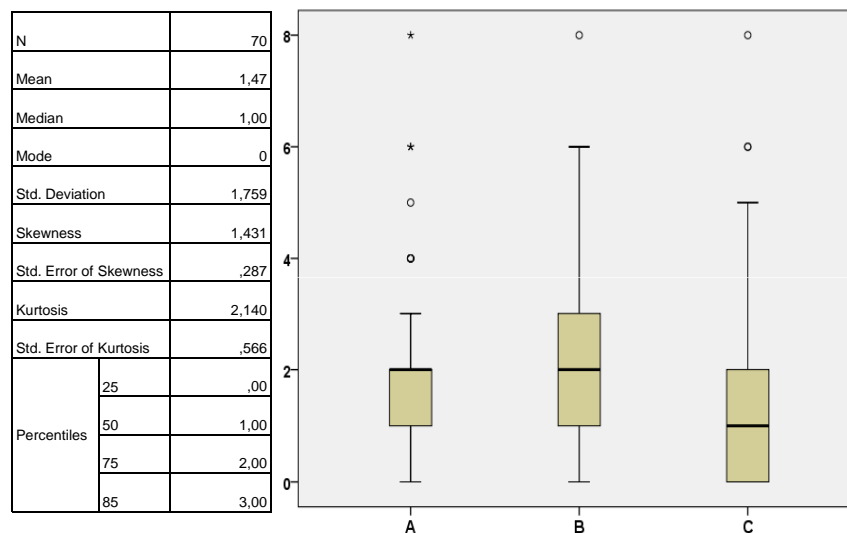
1. (1ª Frequência 2013) O número de chamadas telefónicas recebidas, por minuto, numa determinada central telefónica foi registado durante um período de 50 minutos, observando-se os seguintes valores:

Descriptives		Nº de chamadas	Frequência
N	50	0	6
Median	<b>C</b>	1	20
Variance	1,452	2	<b>A</b>
Range	<b>D</b>	3	10
Minimum	0	4	5
Sum	88		<b>B</b>

- (a) Determine **A**, **B**, **C** e **D** e complete a tabela de frequências.
- (b) Determine e interprete a média, a moda e o desvio-padrão do número de chamadas recebidas por minuto;
- (c) Complete as seguintes frases:
- Em 18% dos minutos foram recebidas \_\_ chamadas;
  - Em 30% dos minutos foram recebidas mais de \_\_ chamadas;
  - Em \_\_ minutos foi recebida no máximo 1 chamada.

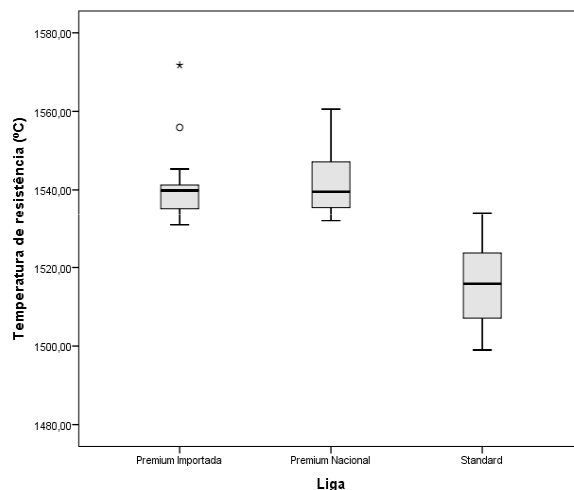
- (d) Classifique a forma da distribuição dos dados quanto à assimetria e ao achatamento.

2. (**2ª Frequência/Exame 2013**) O responsável técnico de uma empresa pretende averiguar qual a origem das interrupções diárias do sistema. Entre outros dados, recolheu informação do número de interrupções diárias do sistema por avaria mecânica, ao longo de 70 dias. Alguns dos resultados obtidos foram os seguintes:



- (a) Interprete a média, a moda, a mediana e o desvio-padrão dos dados.
- (b) Indique qual das boxplot (**A**, **B** ou **C**) corresponde à representação dos dados observados.
- (c) Diga, justificando, se a seguinte afirmação é verdadeira: “A percentagem de dias com mais de 4 interrupções do sistema por avaria mecânica é inferior a 15%”
- (d) Classifique a distribuição dos dados quanto ao tipo de assimetria e achatamento. Justifique a sua resposta.
3. (**Exame de recurso 2013**) Um fabricante da indústria cerâmica pretende determinar se duas novas ligas *premium*, uma nacional e uma importada, possuem uma resistência ao calor superior à da liga *standard* já utilizada. Para tal foram realizados testes em que, para 20 fornadas de cada tipo de liga, se registou a temperatura máxima de resistência ao calor. Apresentam-se abaixo alguns dos resultados obtidos com a análise dos dados realizada com recurso ao *software* SPSS.

Descriptives				
	Liga		Statistic	Std. Error
Temperatura de resistência (°C)	Premium Importada	Mean	1540,77	2,04
		Median	1539,85	
		Std. Deviation	9,10	
		Maximum	1571,77	
		Range	40,90	
		Skewness	2,38	,51
	Premium Nacional	Kurtosis	6,96	,99
		Mean	1541,48	1,68
		Median	1539,55	
		Variance	56,15	
		Minimum	1531,93	
		Maximum	1560,58	
	Standard	Skewness	0,94	,51
		Kurtosis	0,63	,99
		Mean	1516,05	2,26
		Median	1515,86	
		Std. Deviation	10,09	
		Minimum	1498,99	
		Range	35,06	
		Skewness	0,13	,51
		Kurtosis	-0,78	,99



- (a) Interprete os valores da média, mediana e desvio-padrão observados para a temperatura de resistência da liga *premium* nacional.
- (b) O que pode dizer acerca da representatividade da temperatura de resistência média de cada uma das ligas?
- (c) O que sugere a representação gráfica acerca da forma (assimetria e achatamento) da distribuição das amostras correspondentes às três ligas?
4. Na sequência do exercício anterior, são disponibilizados os seguintes dados da temperatura máxima de resistência ao calor registados nos testes das 20 fornadas da liga *premium* nacional:

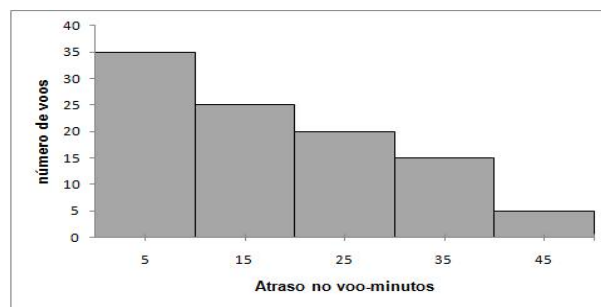
Temperatura	$f_i$
[1531.93; 1537.66[	0.30
[1537.66; 1543.39[	0.40
[1543.39; 1549.12[	0.15
[1549.12; 1554.85[	0.10
[1554.85; 1560.58]	0.05

Com base nestes resultados,

- (a) Calcule os valores da média, moda, mediana e desvio-padrão.
- (b) Determine e interprete o 1º quartil e o percentil 80.

- (c) Através do cálculo de medidas adequadas, classifique a forma da distribuição dos dados quanto à assimetria e ao achatamento.

5. (**1ª Frequência 2011**) Um estudo sobre os atrasos nos 100 voos europeus durante um Verão, realizado em determinado aeroporto, conduziu à seguinte representação gráfica:



- (a) Construa a tabela de frequências a partir da informação fornecida pelo gráfico.  
 (b) Calcule a variância referentes aos minutos de atraso dos voos.  
 (c) O que poderá dizer quanto à assimetria dos dados? O que é que esta característica nos diz sobre os dados?
6. Foi feito um estudo de modo a avaliar quantas vezes durante uma tarde de estudo (5 horas) os alunos enviam e/ou recebem mensagens sms. Para tal, recolheu-se uma amostra aleatória de 100 alunos inscritos na disciplina de Introdução à Probabilidade e Estatística (IPE). O resultado deste estudo apresenta-se na tabela seguinte:

Nº de sms enviados	0	1	2	3	4	5	6	7	8
frequências relativas	0,01	0,2	0,25	0,03	0,28	0,12	0,05	0,04	0,02

- (a) Construa a tabela de frequências associada a estes dados.  
 (b) Represente graficamente (usando a frequência relativa) estes dados.  
 (c) Determine, para este conjunto de dados,:
- a média;
  - a moda e mediana
  - o desvio-padrão.
  - o 1º quartil.
  - o percentil 75.
  - o 8º decil.
- (d) Através do cálculo do grau de assimetria de Bowley o que poderá dizer quanto à assimetria dos dados?
- (e) Qual a proporção de alunos que enviam
- pelo menos 4 sms?
  - entre 2 a 5 sms?

7. **(Exame de época normal 2009).** Devido não só ao aumento dos preços dos preços combustíveis mas também a preocupações ambientalistas, muitos portugueses têm optado por transformar os seus automóveis e instalar GPL (este contribui para ajudar na preservação do meio ambiente, visto ser um combustível muito menos poluente, quando comparado com outros combustíveis). Uma grande empresa de instalação de GPL em veículos automóveis, com 1000 funcionários, pretende medir o tempo gasto na instalação de GPL. Para tal seleccionou uma amostra de 200 funcionários e registou os tempos gastos por cada um na instalação de GPL num veículo automóvel. Os resultados obtidos foram os seguintes:

Tempo (em horas)	$X'_i$	$F_i$
[3, 10[	6,5	0,015
[10, 12[	11	0,03
[12, 15[	13,5	0,07
[15, 20[	17,5	0,245
[20, 30[	25	0,695
[30, 60[	45	0,945
[60, 120]	90	1

- (a) Complete a tabela de frequências associada a estes dados.  
 (b) Represente, como achar mais conveniente, esta colecção de dados.  
 (c) Determine a média e o desvio-padrão associados a esta colecção de dados.  
 (d) Determine ainda a classe modal e a classe mediana.
8. O encarregado do controlo de qualidade de uma fábrica de relógios digitais observou durante 200 dias amostras de 5 relógios, tendo registado o número de relógios defeituosos por amostra:

N.º de defeituosos, por amostra	0	1	2	3	4	5
Frequências	53	68	44	17	16	2

- (a) Construa a tabela de frequências associada a este conjunto de dados.  
 (b) Calcule e interprete a média, a mediana, a moda e o desvio-padrão do número de relógios defeituosos, por amostra.  
 (c) Determine os 1º, 2º e 3º quartis. Interprete cada um dos valores obtidos.  
 (d) Determine e interprete P10 e P75 (percentis 10% e 75%).  
 (e) Numa outra fábrica de relógios (B) obteve-se uma média igual a 2,702 relógios defeituosos por amostra e uma variância igual a 1,6. Qual das duas fábricas apresenta resultados mais homogêneos?
9. Numa empresa recolheu-se uma amostra aleatória relativa à produção de energia eléctrica em KW/h em dois tipos de geradores, I e II. Admita que a produção de energia de cada gerador segue uma distribuição normal e que  $\sigma_I^2 = \sigma_{II}^2$ . Os resultados amostrais foram os seguintes:

Gerador tipo I (n=27)	15.01	3.81	2.74	16.82	14.30	13.45	8.75
	9.40	16.84	17.21	2.74	4.91	5.05	9.72
	9.02	12.31	14.10	9.64	10.21	10.34	9.04
	5.02	10.59	11.91	9.44	7.21	11.07	
Gerador tipo II (n=23)	10.87	8.07	10.31	11.08	10.84	6.34	10.05
	9.37	8.94	8.78	15.01	6.93	15.91	13.45
	6.84	9.37	10.04	10.94	2.04	16.89	14.04
	4.32	10.71					

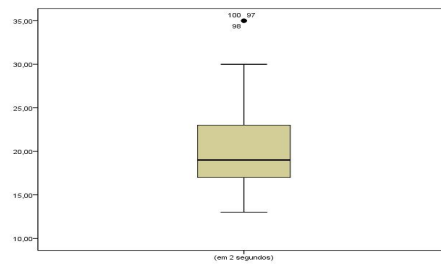
- (a) Complete a tabela de frequências associada à amostra do gerador I.

Classes	$n_i$
	6
	1
	11
	5
	4
	27

- (b) Relativamente ao gerador I, determine a média considerando os dados agrupados e não agrupados. Compare os valores obtidos.
- (c) Compare, através de uma caixa-de-bigodes, estas duas colecções de dados.
- (d) Calcule, para os dados não agrupados:
- duas medidas de localização e uma medida de dispersão que lhe pareçam ser as mais adequadas para ambos os conjunto de dados. Interprete estes valores.
  - Determine, apenas para o gerador I, o quantil 85% e interprete este valor.
10. De modo a estudar a performance dos atletas de alta competição do Eborense Futebol Clube foram seleccionados, ao acaso, 100 atletas e registado o número de metros percorridos em 2 segundos. Os resultados obtidos apresentam-se na seguinte tabela de frequências:

$X'_i$	$F_i$
14,6	0,24
17,8	0,53
21	0,71
24,2	0,89
27,4	0,92
30,6	0,96
33,8	1

- (a) Complete a tabela de frequências associada a estes dados.
- (b) Dada a representação abaixo, indique quais as medidas de localização e as medidas de dispersão adequadas para caracterizar esta colecção de dados.



(c) O que pode dizer sobre a forma (apenas em termos de assimetria) desta colecção de dados?

11. Tal como no caso do exercício anterior, pretendia estudar-se a performance dos atletas, agora não profissionais, do Eborense Futebol Clube. Para tal, foram seleccionados, ao acaso, 25 atletas e registado o número de metros percorridos em 5 segundos. Os resultados obtidos encontram-se resumidos no quadro seguinte:

N	Valid	25
	Missing	0
Mean		60,6400
Median		65,0000
Mode		57,00a
Std. Deviation		16,88411
Variance		285,073
Skewness		-,660
Std. Error of Skewness		,464
Kurtosis		-,078
Std. Error of Kurtosis		,902
Range		68,00
Minimum		20,00
Maximum		88,00
Percentiles	10	35,6000
	25	49,0000
	30	54,0000
	50	65,0000
	70	71,2000
	75	72,5000

a. Multiple modes exist. The smallest value is shown

- (a) Determine o coeficiente de Bowley e interprete-o.
- (b) O treinador destes atletas afirma que 30% dos atletas efectuem mais de  $x$  metros em 5 segundos. De quantos metros está o treinador a falar?
12. (1ª Frequência 2006). De modo a estudar a rapidez com que os alunos resolviam problemas de estatística, foi recolhida aleatoriamente uma amostra de 25 estudantes de várias licenciaturas (com a mesma matéria e os mesmos professores), aos quais foi apresentado um exercício para resolverem. Os dados que se seguem dizem respeito ao tempo (em minutos) que os alunos levaram a resolver o exercício proposto;

26	17	37	38	33	29	63	43	47	19	49	41	
21	53	59	42	60	58	40	48	44	54	47	31	32

- (a) Construa a tabela de frequências associada a estes dados.
- (b) Construa o histograma e o polígono de frequências correspondente à distribuição de frequências relativas.
- (c) Determine a média e o desvio-padrão deste conjunto de dados.

- (d) Através do coeficiente de assimetria ( $G_1$ ) o que poderá dizer quanto à assimetria dos dados?

13. Foi feito um inquérito a um grupo de compradores de 40 carros novos para determinar quantas reparações ou substituições de peças foram feitas durante o primeiro ano de utilização dos carros. Obtiveram-se os seguintes resultados:

1	4	1	2	2	3	3	2	1	2
3	2	3	1	0	1	2	7	4	3
5	1	2	4	2	1	3	1	0	1
2	1	1	3	1	0	4	2	3	5

- (a) Construa um quadro de distribuição de frequências absolutas.  
 (b) Calcule as frequências relativas.  
 (c) Construa um gráfico para as frequências absolutas.  
 (d) Calcule as frequências acumuladas.  
 (e) Construa um gráfico de frequências acumuladas.
14. Num estudo para analisar a capacidade de germinação de certo tipo de cereal foram semeadas cinco sementes em cada um dos vasos dum conjunto de vasos iguais, contendo o mesmo tipo de solo, e registou-se o número de sementes germinadas. Os resultados obtidos foram os seguintes:

N.º de sementes germinadas por vaso	0	1	2	3	4	5
Nº de vasos	16	32	89	137	98	25

- (a) Calcule a média, a mediana e a moda do número de sementes germinadas.  
 (b) Represente graficamente os resultados.  
 (c) Calcule a proporção de vasos com mais de três sementes germinadas.  
 (d) Determine os 1º, 2º e 3º quantis. Interprete cada um dos valores obtidos.  
 (e) Determine e interprete P10 e P75 (percentis 10% e 75%).  
 (f) Calcule o desvio padrão e o coeficiente de variação.
15. Os dados que se seguem referem-se ao comprimento total (em cm) de uma colecção de dados de achigãs de uma barragem:

29.9	40.2	37.8	19.7	30.0	29.7	19.4	39.2	24.7	20.4
19.1	34.7	33.5	18.3	19.4	27.3	38.2	16.2	36.8	33.1
41.4	13.6	32.2	24.3	19.1	37.4	23.8	33.3	31.6	20.1
17.2	13.3	37.7	12.6	39.6	24.6	18.6	18.0	33.7	38.2

- (a) Ordene os dados e calcule a média, mediana, desvio padrão, quartis e o quantil de ordem  $2/3$ . Encontre um valor tal que 70% dos peixes observados tenham comprimentos superiores a esse valor.



- (b) Faça um agrupamento dos dados em classes, de forma conveniente, e represente-os graficamente.
- (c) Calcule a média e variância para os dados classificados. Compare estes valores aproximados com os correspondentes valores exactos obtidos em (a). Indique a classe modal e interprete o seu significado.
16. O departamento de pessoal de uma certa firma fez um levantamento dos salários dos 120 funcionários do sector administrativo, obtendo os seguintes resultados:

Faixa Salarial (Nº de salários mínimos)	Frequência Relativa
$[0, 2[$	0,25
$[2, 4[$	0,40
$[4, 6[$	0,20
$[6, 10[$	0,15

- (a) Esboce a representação gráfica apropriada a este conjunto de dados.
- (b) Calcule a média e o desvio padrão.
- (c) Se for concedido um aumento de 100% para todos os funcionários haverá alteração na média? E na variância? Justifique convenientemente a sua resposta.
- (d) Se for concedido um abono de 2 salários mínimos para todos os 120 funcionários, haverá alteração na média? Justifique.
17. Foram medidas a altura e o peso de um grupo de homens e de um grupo de mulheres. O quadro seguinte contém a média e o desvio padrão dos dados observados:

Grupos	Altura (cm)		Peso (Kg)	
	$\bar{x}$	$s$	$\bar{x}$	$s$
Homens	173	3,2	71	3,2
Mulheres	158	3,2	54	2,7

- (a) Calcule o coeficiente de variação da altura das mulheres. Suponha que a altura é medida em milímetros. Qual o novo coeficiente de variação? Justifique a resposta.
- (b) Compare a dispersão da altura e do peso dentro de cada grupo e entre os dois grupos de indivíduos.
18. No quadro seguinte indicam-se os preços dum bem alimentar (em unidades monetárias) praticado durante 12 meses consecutivos e as quantidades vendidas.

Preço	110	90	80	76	74	71	70	65	63	60	55	50
Vendas	55	70	90	100	90	105	80	110	125	115	130	131

- a)* Represente graficamente a informação disponibilizada.
- b)* Através da análise gráfica, parece-lhe existir relação linear entre o preço e as quantidades vendidas observadas?
- c)* Calcule e interprete o valor do coeficiente de correlação amostral.

**Docentes:** Dulce Gomes, Patrícia Filipe, Telma Santos e Ana Isabel Santos