

# Escola Superior de Tecnologia e Gestão de Viseu

## 1ª Frequência de Estatística

Gestão de Empresas e Gestão de Empresas - PL 2º Ano

10 de novembro de 2018

Duração: 2h00min

- O tempo que os automóveis levam a abastecer é uma variável aleatória normal,  $X$ , de média 4 minutos e desvio padrão 1 minuto.
  - Qual a percentagem de automóveis abastecidos em menos de 3 minutos?
  - Qual é o tempo mínimo de abastecimento dos 2.5% automobilistas mais rápidos?
  - Em 3 automóveis, qual a probabilidade de um deles efetuar o abastecimento em menos de 4 minutos?
  - Foram escolhidos aleatoriamente 35 automóveis. Qual a probabilidade do tempo total de abastecimento destes 35 automóveis estar entre os 139 minutos e os 142 minutos?
- A central telefónica do INEM de uma grande cidade recebe chamadas, umas genuínas e outras falsas, isto é, correspondentes ou não a verdadeiros acidentes. Sabe-se que 2% das chamadas recebidas pela central são falsas. Destas, 20% são efetuadas durante o período da manhã, 40% durante o período da tarde e as restantes à noite. Das chamadas genuínas recebidas na central, 30% são feitas durante a manhã. Determine a probabilidade de uma chamada ser efetuada durante a manhã.
- O número de automóveis de certa marca vendidos, em cada semana, por um stand de automóveis é uma variável aleatória  $X$ , com a seguinte função de probabilidade:

$x$	0	1	2	3
$f(x)=P(X=x)$	$a$	0.4	0.15	$B$

- Sabendo que a esperança de  $X$  é igual a 1.15, determine o valor de  $a$  e de  $b$ .
  - Sabendo que numa determinada semana o número de automóveis vendidos foi superior à média, qual a probabilidade de terem sido vendidos menos de 3 automóveis?
- O tempo (em minutos) que um trabalhador da companhia A demora em cada viagem entre a sua casa e o escritório, é uma variável aleatória  $X$  com função densidade de probabilidade:

$$f(x) = \begin{cases} 0 & , \text{ se } x < 20 \\ \frac{1}{4} & , \text{ se } 20 \leq x < 24 \\ 0 & , \text{ se } x \geq 24 \end{cases}$$

Determine a probabilidade do trabalhador demorar menos de 21 minutos.

5. Efetuou-se um inquérito junto das famílias de um determinado bairro, cujos resultados obtidos foram os seguintes:

Satisfação no casamento	Rendimento familiar			Total
	Baixo (RB)	Médio (RM)	Elevado (RE)	
Elevada (SE)	190	217	44	<b>451</b>
Reduzida (SR)	60	143	56	<b>259</b>
<b>Total</b>	<b>250</b>	<b>360</b>	<b>100</b>	<b>710</b>

- a) Escolhida uma família desse bairro ao acaso, determine a probabilidade de a satisfação no casamento ser elevada.
- b) Escolhida uma família de rendimento familiar elevado, qual a probabilidade de a satisfação no casamento ser reduzida?
- c) Os acontecimentos “rendimento familiar baixo” e “satisfação no casamento elevada” são independentes? Justifique.
6. O número de acidentes de aviação, na companhia aérea A, durante a época de verão segue distribuição de Poisson. Sabe-se que os aviões da companhia A têm como número médio de acidentes na referida época o valor 0.1. Calcule a probabilidade de o total de acidentes dos aviões da companhia A em três épocas de verão consecutivas não exceder 2, admitindo que o número de acidentes é independente de época para época.
7. No seguimento da obrigatoriedade de equipar automóveis movidos a combustível diesel com filtro de partículas para redução efetiva da poluição, surgiram reservas quanto à utilização deste novo tipo de filtros, nomeadamente sobre o potencial aumento do nível de ruído emitido. A indústria produtora destes filtros veio defender que estes não implicam mais ruído do que os catalisadores usuais para automóveis a diesel. No sentido de estudar este problema, uma empresa do ramo automóvel recolheu dados relativos a uma amostra de automóveis. Usando o SPSS obteve-se o seguinte output:

## Estatísticas

Nível de ruído (decibéis)		
N	Válido	36
	Omisso	0
Média		810,14
Mediana		820,00
Moda		820
Desvio Padrão		29,216
Variância		853,552
Assimetria		-,892
Erro padrão da assimetria		,393
Amplitude		95
Mínimo		760
Máximo		855
Percentis	10	763,50
	25	777,50
	30	785,50
	50	820,00
	70	825,00
	75	832,50
	90	845,00

Para o nível de ruído complete as afirmações.

- (a) Nesta amostra registaram-se observações relativas a \_\_\_\_\_ veículos, sendo o nível médio de ruído desses veículos de \_\_\_\_\_ decibéis. No entanto, o valor mais elevado observado para o nível de ruído foi de \_\_\_\_\_ decibéis.
- (b) Em termos globais, pode afirmar-se que \_\_\_\_\_% dos veículos observados apresentam um nível de ruído que não ultrapassa os 832,50 decibéis, mas registaram-se 18 veículos com nível de ruído superior ou igual a \_\_\_\_\_ decibéis.
- (c) \_\_\_\_\_% dos veículos apresentam valores de ruído superiores a 785,50 decibéis, mas 70 % apresentam níveis inferiores ou iguais a \_\_\_\_\_ decibéis.
- (d) Os veículos observados apresentam valores do nível de ruído \_\_\_\_\_ heterogêneos, como se pode verificar pelo valor da amplitude, que foi \_\_\_\_\_, ou ainda pelo valor observado 29,216 para o \_\_\_\_\_.
- (e) A distribuição dos níveis de ruído apresenta uma assimetria \_\_\_\_\_, como se pode verificar por \_\_\_\_\_. Isto permite-nos concluir que o mais frequente na amostra foram veículos com níveis de ruído mais \_\_\_\_\_, embora em menor número, registaram-se também alguns veículos com níveis de ruído mais \_\_\_\_\_.

## 2.º Frequência (19-12-2018)

1. No quadro seguinte figuram os resultados de um inquérito sobre a concordância com o conteúdo de um determinado Decreto-Lei sobre o orçamento, conforme o género da pessoa inquirida:

	Nº de pessoas inquiridas	Nº de respostas concordantes
Homens	200	75
Mulheres	300	120

- (a) Usando um teste de hipóteses e um nível de significância de 0,05, averigue se a proporção de homens que concorda com o Decreto-Lei é superior a 35%.
- (b) Determine o p-value referente ao teste realizado na alínea anterior. Explique o seu significado.
- (c) Com base num intervalo, com 95% de confiança, é possível concluir que a proporção de homens concordantes com o decreto de lei é inferior à proporção correspondente para as mulheres?
- (d) Os resultados do inquérito permitem concluir que a concordância com o decreto é independente do género? Use, para o efeito, um nível de significância de 0,05.
- (e) Que outro teste de hipóteses poderia usar para dar resposta à alínea anterior? Indique as hipóteses em teste e a estatística de teste. De entre os dois testes que indicou qual o mais adequado? Justifique.
2. Um treinador pretende saber qual o número ótimo de dias semanais de treino para os seus atletas. Para tal mediu a performance de três grupos de atletas (característica quantitativa) separados consoante o número de dias de treino semanal: um, dois e três dias. Através da Análise de Variância pretende-se averiguar se existem diferenças entre as performances dos 3 grupos ( $\alpha=0.05$ ). Note que entende-se que quanto maior a performance melhor o desempenho do atleta.

**Quadro A: Tests of Normality**

Número de dias de treino	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
1	,224	21	,200*	,869	21	,223
2	,249	21	,200*	,865	21	,206
3	,135	21	,200*	,984	21	,969

a. Lilliefors Significance Correction

\*. This is a lower bound of the true significance.

**Quadro B: Test of Homogeneity of Variances**

Levene Statistic	df1	df2	Sig.
4,637	2	60	,13

**Quadro C: ANOVA**

Performance

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	2525,691	2	1262,846	<b>(A)</b>	,000
Within Groups	.....	....	.....		
Total	8500,415	62			

### Quadro D Multiple Comparisons

Dependent Variable: Performance

Tukey HSD

(I) Número de dias de treino		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-9,98789*	3,23759	,009	-17,7789	-2,1969
	3	-15,69947*	3,23759	,000	-23,4905	-7,9085
2	1	9,98789*	3,23759	,009	2,1969	17,7789
	3	-5,71158	3,23759	,191	-13,5026	2,0794
3	1	15,69947*	3,23759	,000	7,9085	23,4905
	2	5,71158	3,23759	,191	-2,0794	13,5026

\*. The mean difference is significant at the **0.05** level.

- Indique quais dos seguintes são pressupostos subjacentes à aplicação da ANOVA?
- Averigue se a amostra recolhida para os atletas que treinam uma vez por semana provém de uma população normal? Indique as hipóteses em teste, o p-value e conclua.
- Recorrendo ao p-value e ao teste relativo ao Quadro B que conclusões pode tirar.
- Determine o valor de **(A)** do Quadro C, indicando o cálculo.
- Que conclusão pode tirar do teste relativo ao Quadro C?
- Das conclusões seguintes, indique as verdadeiras (V) e as falsas (F), assinalando com uma cruz a opção correta. Considere um nível de significância de 0.05.

(Cada resposta certa=0.5 valores, cada resposta errada = -0.1, ausência de resposta=0)

	V	F
Há evidência para afirmar que os grupos 2 e 3 ultrapassam o grupo 1 em termos de performance média.	<input type="checkbox"/>	<input type="checkbox"/>
Há evidência para afirmar que o grupo 3 tem, em média, uma maior performance do que o grupo 2, mas não se pode dizer o mesmo relativamente ao grupo 1.	<input type="checkbox"/>	<input type="checkbox"/>
Entre o grupo 1 e os grupos 2 e 3, não há evidência de haver diferenças nas performances médias.	<input type="checkbox"/>	<input type="checkbox"/>

- Um fabricante de pneus pretende comparar, através de ensaios piloto, dois métodos de produção dos pneus. Seleccionados 10 e 8 pneus produzidos, respetivamente segundo o primeiro e segundo métodos, resolve-se testá-los. Os pneus da primeira amostra foram testados numa zona A, os da segunda numa zona B, tendo-se obtido (em unidades de 100km):

	Zona	N	Mean	Std. Deviation	Std. Error Mean
Duração	A	10	61,40	4,300	1,360
	B	8	59,88	3,563	1,260

**Independent Samples Test**

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
Duração	Equal variances assumed	,203	,658	,805	16	,433
	Equal variances not assumed			,823	15,961	,423

Sabe-se de estudos anteriores que a duração de um pneu varia segundo uma distribuição normal, em que o valor esperado é eventualmente influenciável pelo método de produção, e cujo desvio padrão é suscetível de ser fortemente afetado pelas características da zona onde se procede a rodagem.

- a) Usando um intervalo de confiança a 90% diga se será que se pode admitir que a duração esperada de um pneu na Zona B não excede 7000km?
- b) Por dois processos diferentes, teste a igualdade das variâncias. Considere um nível de significância de 0.05.
- c) Através de um teste de hipóteses verifique se existem diferenças significativas na duração média dos dois tipos de pneus. Use  $\alpha = 0,05$ . Conclua com base no p-value e com base no valor observado e na região crítica.
- d) Qual é a relação de ordem entre o p-value e o nível de significância quando o valor observado na estatística de teste não pertence à região crítica num qualquer teste de hipóteses?

**Exame (12-01-2019)**

1. Para estudar o número de golos marcados por jogo na 1ª Liga Portuguesa de Futebol Profissional, considerou-se como amostra a época de 2017/2018:

Nº de golos	0	1	2	3	4	5	6	7	8
Nº de jogos	27	67	80	69	29	19	12	2	1

- 1.1 Determine a média, a moda e a mediana do número de golos marcados por jogo, na época de 2017/2018, e interprete os seus valores.
- 1.2 Estude a dispersão do número de golos marcados por jogo, na época de 2017/2018, utilizando três medidas de dispersão.
- 1.3 Com base num intervalo com 95% de confiança comente a afirmação: “Na 1ª Liga Portuguesa de Futebol Profissional, o número médio de golos marcados por jogo é superior a 2”. Para responder a esta questão, admita que para a amostra considerada a média é igual a 2.3 golos e o desvio padrão é igual a 1.7 golos.
- 1.4 Admita que na 1ª Liga Espanhola de Futebol Profissional na época de 2017/2018 (também constituída por 306 jogos) existiram 25 jogos onde não foram marcados golos. Com base num intervalo com 96% de confiança verifique se é possível concluir em qual dos países a proporção de jogos sem golos é superior.
- 1.5 Suponha que o número de golos marcados por jogo segue uma distribuição de Poisson de média igual a 2.3. Qual a probabilidade de, em 4 jogos, terem sido marcados menos de 3 golos?
2. O tempo de funcionamento, em milhões de horas, de um equipamento informático é uma variável aleatória  $X$  com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} 6x - 6x^2, & \text{se } 0 < x < 1; \\ 0, & \text{caso contrário.} \end{cases}$$

- 2.1 Qual a probabilidade do equipamento funcionar entre 0.5 e 1.5 milhões de horas?
- 2.2 Calcule o tempo médio de vida do equipamento e o respetivo desvio padrão, sabendo que  $E(X^2) = 0.3$ .
- 2.3 O peso do equipamento informático em causa segue uma distribuição normal de média 50 gramas e desvio padrão igual a  $k$  gramas. Os equipamentos deste tipo são embalados em caixas contendo 50 equipamentos.
- 2.3.1 Calcule  $k$  de modo a que 9.85% dos equipamentos apresentem peso superior a 51 gramas. (NOTA: Caso não consiga responder a esta questão, nas perguntas seguintes considere que  $k=0.78$  gramas)
- 2.3.2 Determine a percentagem de equipamentos cujo peso varia entre as 48 e as 52 gramas.
- 2.3.3 Qual a probabilidade do peso total das caixas (considerando apenas o peso dos equipamentos informáticos) exceder 2.49 Kg?
- 2.3.4 Qual a probabilidade de, numa caixa, existir no máximo um equipamento com peso superior a 50 gramas?

3. Sabe-se que 30% dos portugueses são funcionários públicos e os restantes são funcionários do sistema privado. Dos funcionários públicos, 50% têm um rendimento mensal superior a 700 euros. A probabilidade de um trabalhador ser funcionário do sistema privado e ganhar, por mês, 700 euros ou menos é igual a 0.7.
- 3.1 Escolhendo ao acaso um trabalhador, qual a probabilidade do seu rendimento mensal ser inferior ou igual a 700 euros?
- 3.2 Escolheu-se ao acaso um trabalhador e verificou-se que este ganhava mais de 700 euros por mês. Qual a probabilidade deste trabalhador ser funcionário público?
4. Num determinado estudo, com o objetivo de averiguar se existe qualquer relação entre a qualidade do ar (bom e razoável) com a zona do país (litoral e interior) analisaram-se 150 zonas, tendo-se obtido os seguintes resultados:

Qualidade do ar \* Zona Crosstabulation

			Zona		Total
			litoral	interior	
Qualidade do ar	Bom	Count	32	23	55
		Expected Count	39,2	15,8	55,0
	Razoável	Count	75	20	95
		Expected Count	67,8	27,2	95,0
Total		Count	107	43	150
		Expected Count	107,0	43,0	150,0

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-Square	7,274	--	,007
Continuity Correction <sup>b</sup>	6,365	--	,012
Likelihood Ratio	7,191	--	,0071
Fisher's Exact Test			
Linear-by-Linear Association	7,296	--	,0072

- 4.1 Enuncie o teste em questão e retire conclusões, utilizando um nível de significância de 1%.
- 4.2 Comente a seguinte afirmação, apresentando cálculos que entender adequados “a associação que existe entre a qualidade do ar e a zona do país é considerada forte”.
5. Pretende-se averiguar se existem diferenças significativas nos comprimentos de uma determinada substância química. O analista aponta o método de medição como fator que pode explicar a existência de tais diferenças. Fizeram-se alguns testes no SPSS, cujos resultados da medição do comprimento da substância entre 3 métodos existentes, M1, M2 e M3, são apresentados nas tabelas abaixo. Sabe-se, ainda, que  $s_1^2 = 39.8$ ;  $s_2^2 = 46.667$ ;  $s_3^2 = 55.467$  e que  $\bar{x}_1 = 29.4$ ;  $\bar{x}_2 = 38.33$ ;  $\bar{x}_3 = 28.67$ .

Testes de Normalidade – Quadro 1

	Método	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Estatística	Df	Sig.	Estatística	Df	Sig.
Comprimento	M1	,260	5	,200*	,890	5	,355
	M2	,169	6	,200*	,933	6	,605
	M3	,185	6	,200*	,908	6	,426

\*. Este é um limite inferior da significância verdadeira.

a. Correlação de Significância de Lilliefors

ANOVA – Quadro 3

Comprimento					
	Soma dos Quadrados	Graus liberdade	Variância	Razão F	Sig.
Entre Grupos	-----	---	-----	----	,057
Dentro dos grupos	669,87	----	-----		
Total	-----	----	-----		



**Teste de Homogeneidade de Variância – Quadro 2**

	Estatística de Levene	df1	df2	Sig.
Comprimento Com base em média	,053	2	14	,949

- 5.1 Qual a relevância dos quadros 1 e 2? Interprete, os resultados que estes quadros fornecem. Utilize um nível de significância de 5%.
- 5.2 Complete o quadro 3 e justifique se pode utilizar os resultados que dele se obtêm.
- 5.3 A partir do quadro 3, formule as hipóteses do teste que permite perceber se o método de medição influencia o comprimento das substâncias. Tire conclusões de duas formas diferentes utilizando um nível de significância de 5%.
- 5.4 Com um nível de significância de 1%, será possível concluir que o comprimento médio de M1 é superior a 29?

**Exame de recurso (05-02-2019) Duração: 2:45 horas**

1. Efetuou-se um estudo sobre o emprego da população de uma região, baseado numa amostra. Os resultados do inquérito foram organizados no quadro seguinte:

	Empregados	Desempregados	total
Mulheres	105	15	120
Homens	122	8	130
Total	227	23	250

- (a) Ao nível de significância de 0.1, o que se pode concluir relativamente à independência entre o sexo e a situação de emprego. Justifique convenientemente a sua resposta.
- (b) Comente a seguinte afirmação, apresentando cálculos que entender adequados “a associação que existe entre o sexo e a situação de emprego é considerada forte”.

Considere agora que os valores indicados no quadro anterior são relativos à população.

- (c) Selecionou-se um indivíduo ao acaso e verificou-se que é mulher, qual a probabilidade de esta se encontrar desempregada?
- (d) Os acontecimentos “o indivíduo selecionado é mulher” e “o indivíduo selecionado está desempregado” são independentes? Justifique convenientemente a sua resposta.
2. Uma empresa, monopolista no mercado de determinado produto, tem produção constante de 90 toneladas por mês e tem conhecimento que a procura mensal é uma variável aleatória com distribuição normal de valor médio 80 toneladas e desvio padrão 10 toneladas.
- (a) Diz-se que a procura mensal é excedentária se a procura mensal exceder a produção mensal da empresa. Calcule a probabilidade de haver procura mensal excedentária?
- (b) Para satisfazer uma eventual procura excedentária a empresa tem possibilidade de importar no princípio do mês um stock de segurança. Qual deve ser esse stock de segurança  $S$ , de forma a que a probabilidade de haver procura insatisfeita baixe para 0.025?
- (c) Calcule a probabilidade de em 36 meses a procura média exceda 82 toneladas?
- (d) Em 6 meses, qual é a probabilidade de em pelo menos um deles existir procura excedentária?
3. Verifica-se que o número de defeitos de um certo tipo, na pintura da chapa exterior de um automóvel, é bem descrito por uma variável aleatória de Poisson de média 0.15 defeitos por  $m^2$ . A área total da chapa do automóvel do modelo *Expe2000* é de  $8 m^2$ .
- (a) Qual é a probabilidade de um *Expe2000* ter pelo menos um desses defeitos?
- (b) Diga, justificando, se será admissível que um defeito deste tipo se possa encontrar numa área de  $0.5 m^2$  da chapa.

4. Realizou-se um estudo para avaliar o tempo de secagem do verniz A, em horas. Para o efeito, recolheu-se uma amostra de tamanho 9 tendo-se obtido os seguintes resultados,

6 5.7 5.8 6.5 7 6.3 5.6 6.1 5

Considere que o tempo de secagem do verniz A tem distribuição normal.

- (a) Calcule a média, a mediana e a variância da amostra.

**Nota:** Caso não tenha resolvido (a) considere que a média da amostra é de 6 horas e a variância é de 0.33.

- (b) Determine um intervalo de confiança para o tempo médio de secagem do verniz do tipo A com uma confiança a 90%. O que pode concluir?
- (c) Comente a seguinte afirmação, justificando convenientemente a sua resposta: “Com uma confiança a 90%, e de forma a garantir que o erro máximo cometido na estimação tempo médio de secagem do verniz A seja inferior a 0.1 o tamanho da amostra deve ser superior a 251.”
- (d) Um dos requisitos importantes num verniz é que a variabilidade do tempo de secagem seja inferior a 0.4 horas<sup>2</sup>. Será que o verniz A cumpre este requisito? (use  $\alpha=0.05$ )
- (e) Um estudo análogo foi realizado sobre um verniz concorrente - verniz B. Foi recolhida uma amostra de tamanho 23 donde se obteve uma média de 5.48 horas e um desvio padrão de 0.26 horas. Sabe-se ainda que o tempo de secagem do verniz B tem distribuição normal. O fabricante do verniz B afirma que o tempo médio de secagem do seu verniz é inferior aos dos seus concorrentes. Ao nível de significância de 0.05, o que pode concluir relativamente à afirmação do fabricante quando compara o verniz A com o B?

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Tempo_secagem	Equal variances assumed	5,987	,020	3,590	30	,001	,52174	,14534
	Equal variances not assumed			2,623	9,287	,027	,52174	,19887

5. Para estudar a imprevisibilidade - variabilidade - do estado de tempo numa dada cidade, escolheu-se, aleatoriamente 41 dias de **2010** e 61 dias de **2018**, registou-se a temperatura e obteve-se uma média de 17.3 °C e desvio-padrão 10 °C para 2010 e uma média de 17.34 °C e desvio-padrão 12 °C para o ano 2018. Será que estes valores justificam que se conclua, com um nível de significância de 5%, que o estado do tempo está mais imprevisível? Considere que a temperatura diária tem distribuição normal em ambos os anos e que são independentes.

6. Pretende-se estudar o consumo de água de um município de acordo com as temperaturas máximas diárias de um determinado mês (26°C, 27°C, 30°C e 35°C). Escolheram-se aleatoriamente 124 residências, 31 para cada grupo registando-se a água consumida, tendo-se obtido consumos médias de 147.7, 155.9, 168.2, 180.4 para os grupos de temperaturas 26°C, 27°C, 30°C e 35°C respetivamente. Da aplicação da ANOVA, obtiveram-se os seguintes resultados:

Fonte de Variação	Soma de Quadrados	graus de liberdade	Variância ( soma média de quadrados)	Razão F
Entre grupos	30566,393	(b)	(e)	(g)
Dentro dos grupos	83394,875	(c)	(f)	
Total	(a)	(d)		

- (a) Indique os pressupostos da ANOVA.
- (b) Complete a tabela ANOVA preenchendo os espaços assinalados pelas letras de (a) a (g).
- (c) Com base na tabela de ANOVA diga se os consumos são significativamente diferentes para diferentes máximas diárias. Justifique, indicando o valor observado e a região crítica. (use  $\alpha=0,05$ ).
- (d) Caso os consumos dependam das temperaturas máximas diárias que técnica estatística usaria para averiguar entre que pares de temperaturas existem diferenças significativas no que diz respeito ao consumo diário?