

- Em todas as suas respostas defina e apresente, com clareza, a **terminologia** (acontecimentos e variáveis aleatórias), a **metodologia** e as **hipóteses subjacentes aos diferentes passos da sua análise**.
- Nos testes de hipóteses que realizar deve apresentar também o **valor de prova** e uma **conclusão** no contexto do problema.
- Todos os problemas devem ser respondidos em **folhas diferentes**.

1. (4 VALORES) Existe um teste de triagem para o cancro na de próstata que analisa o nível de PSA no sangue. Este teste dá resultado positivo ou negativo, em função do nível de PSA encontrado no sangue do doente. Há uma série de razões, para além do cancro na próstata, que levam a que um homem possa ter níveis elevados de PSA. Além disso, muitos tipos de cancro na próstata desenvolvem-se tão lentamente que nunca chegam a ser um problema. Atualmente não existe nenhum teste capaz de distinguir os diferentes tipos de cancro da próstata e o uso do teste ao PSA é controverso porque é difícil quantificar os valores limite. Também são difíceis de quantificar os danos físicos e psicológicos causados por resultados falso-positivos. Um teste é considerado um **falso-positivo** se, para um doente que não tem um tipo perigoso de cancro da próstata, dá resultado positivo. Um teste é considerado um **verdadeiro-positivo** se, para um doente com um tipo perigoso de cancro na próstata, dá resultado positivo. De modo análogo definem-se falso-negativo e verdadeiro-negativo. A palavra “doente” foi utilizada para designar um homem a quem foi feito o teste do PSA, independentemente de ter, ou não, um cancro na próstata ou qualquer outra doença. Considere que:

- A probabilidade de um tipo perigoso de cancro na próstata entre homens acima de 50 anos é de 0.05%;
- A probabilidade de **verdadeiro-positivos** é de 90%;
- A probabilidade de **falso-positivos** é de 1%.

- a) Calcule a probabilidade de um “doente” ter efetivamente um tipo perigoso de cancro da próstata, sabendo que o teste ao PSA deu resultado positivo.

Solução: $P(\text{Cancro perigoso} \mid \text{Teste positivo}) = 0.0431$

- b) Calcule a probabilidade de o teste ao PSA dar resultado positivo, quando efetuado por um homem com mais de 50 anos.

Solução: $P(\text{Teste positivo}) = 0.0105$

- c) Calcule a probabilidade de em 25 testes ao PSA, realizados a 25 homens com mais de 50 anos, haver pelo menos 2 com resultado positivo.

Solução: $P(Y \geq 2) = 0.0279$

2. (5 valores) Registos históricos da linha de enchimento automático da *Vitaluso* mostram que o volume (em litros) de um garrafão de água segue uma distribuição Normal, com valor esperado $\mu = 4.95 \text{ l}$ e desvio padrão $\sigma = 0.04 \text{ l}$. Um garrafão é considerado conforme as especificações se o volume de água estiver entre 4.88 l e 5.05 l. A produção é realizada em lotes de 200 garrafões e, em cada lote, é selecionada aleatoriamente uma amostra de 15 garrafões para inspeção. Um lote é rejeitado se na amostra forem detetados pelo menos 2 garrafões fora das especificações. O lucro líquido obtido por lote de 200 garrafões depende do número de garrafões não conformes encontrados na amostra inspecionada. Assim, caso todos os garrafões estejam conformes o lucro é 15 €, caso exista apenas um garrafão não conforme o lucro é 5 € e caso existam dois ou mais garrafões não conformes o lucro é -20 € (i.e., passa a ser um prejuízo).

- a) Calcule a probabilidade de um garrafão, selecionado aleatoriamente, não estar conforme as especificações em termos de volume de água.

Solução: 4.63%

- b) Calcule a probabilidade de um lote de 200 garrafões ser rejeitado.

Solução: 15.11% (dupla extração aleatória $\rightarrow B(N = 15, p = 4.63\%)$)

- c) Qual deverá ser o tamanho máximo da amostra a inspecionar de forma a garantir que a probabilidade de rejeição seja inferior a 5%.

Solução: amostra de tamanho 8

- d) Calcule o valor esperado e o desvio padrão do lucro obtido com um lote de 200 garrafões.

Solução: $\mu = 6.14\text{€}$ $\sigma = 11.93\text{€}$

- e) Calcule a probabilidade de a soma dos lucros obtidos com 50 lotes de 200 garrafões ser negativo, i.e., ser um prejuízo.

Solução: 0.0138%

(dado o tamanho da amostra, 50, o TLC garante que: $S = L_1 + \dots + L_{50} \sim N(50 \times 6.14, 50 \times 11.93^2)$)

3. (4 valores). O grupo PRESS detém um conjunto de jornais em suporte físico e em suporte digital relacionados com diferentes temáticas (e.g., economia, desporto, cultura). O grupo vendeu durante o ano passado em média 6 500 assinaturas digitais por mês.

Para averiguar se o volume de jornais digitais vendidos tem aumentado de acordo com a tendência atual de mercado, o grupo recolheu informação sobre as vendas de jornais nos últimos 6 meses.

	Jan-18	Fev-18	Mar-18	Abr-18	Mai-18	Jun-18
# jornais digitais vendidos	8 000	6 000	7 200	9 000	6 500	7 000

$\bar{X} = 7367$

$s = 973$

- a) Existe uma tendência de crescimento no número de jornais vendidos pelo grupo PRESS no último semestre face ao ano transato? Teste ao nível de significância de 5%.

Solução: $ET=2.18$ | $VC=2.015$ (Rejeita H_0)

- b) Determine a probabilidade do erro tipo II efetuado na alínea anterior, admitindo que o valor esperado do número de jornais digitais vendidos pelo grupo PRESS no último semestre é 7 250.

Solução: $X_{crítico}=7300.4$ | Erro tipo II =55%

- c) Ajude o grupo PRESS a perceber se o valor esperado das vendas dos jornais digitais é superior ao valor esperado das vendas dos jornais em papel. Sabe-se que o valor esperado das vendas dos jornais em papel nos últimos 6 meses é 7 000 e o desvio padrão é 700. Construa o intervalo de confiança a 95%.

Solução: Teste à razão de variâncias: $ET=1.93$ | $VC=5.0$ (não rejeita H_0)

I.C. a diferença de valores esperados (variâncias iguais; amostras de pequena dimensão),

IC aberto à direita $\rightarrow [1338.6; +\infty[$

4. (4 VALORES) Na 1ª fase das provas de acesso ao ensino superior em 2017, foram selecionados para um inquérito 100 estudantes de escolas secundárias do distrito do Porto que se candidataram ao Mestrado Integrado em Engenharia Informática e Computação (MIEIC) na FEUP. Para 86 destes estudantes, este curso foi a sua primeira opção de colocação.

- a) Com base na amostra recolhida em 2017, o Departamento de Comunicação da FEUP está a considerar publicitar a seguinte frase: "O MIEIC foi a primeira opção de mais de 85% dos alunos que se candidataram a este curso". Concorda com a frase? Recorra a um teste de hipótese, com $\alpha=5\%$, para justificar a sua resposta.

Solução: $H_0: p=0.85$ e $H_1: p>0.85$; $\alpha=5\%$; $VC=1.645$; $ET=0.280$; $pvalue=38.97\%$ → não rejeitar H_0 (teste inconclusivo)

- b) Com base na amostra de 2017 e utilizando 95% de confiança, que valor recomendaria que fosse usado na frase do Departamento de Comunicação, apresentada na alínea anterior, em vez de "85%"?

Solução: Intervalo de Confiança a 95% para p : $[0.803, +\infty[$ → poderia ser usado o valor 80%

- c) Na 1ª fase das provas de acesso de 2016, o mesmo inquérito revelou que 62 em 80 estudantes se tinham candidatado ao MIEIC como primeira opção. Compare, com base nas amostras de 2016 e 2017, se é possível afirmar que a proporção de alunos candidatos ao MIEIC em primeira opção aumentou?

Solução: $H_0: p_{17} - p_{16}=0$ e $H_1: p_{17} - p_{16}>0$; $\alpha=5\%$; $VC=1.645$; $\hat{p}_0=0.822$; $ET=1.482$; $pvalue=6.91\%$ → não rejeitar H_0 (teste inconclusivo)

5. (3 VALORES) Numa empresa de vendas online, o número de encomendas de um determinado produto segue uma distribuição de Poisson. Até ao final do ano passado, o número de encomendas do produto era, em média, de 4.5 encomendas por dia.

- a) A empresa mantém um *stock* diário de 8 unidades do produto. Calcule a probabilidade de, num dia, ocorrer uma rutura de *stock*.

Solução: $P(\text{rutura de stock}) = P(Y > 8) = 4.03\%$

- b) No início deste ano, a empresa decidiu publicitar o produto num jornal de grande tiragem. Entre Março e Maio (**92 dias**) foram efetuadas 442 encomendas, o que dá, em média, 4.8 encomendas por dia. Teste, ao nível de significância de 5%, se o número médio de encomendas diário do produto aumentou, depois de este ter sido publicitado no jornal.

Solução: teste a λ que, face ao tamanho da amostra (92), pode ser usado o teste aproximado baseado na Normal ($\bar{Y} \sim N(\lambda, \lambda/N)$); $H_0: \mu = \lambda = 4.2$ e $H_1: \mu = \lambda > 4.2$; $\alpha = 5\%$; $VC = 1.65$; $ET = 1.36$; $pvalue = 8.74\%$ → Teste inconclusivo

- c) Recentemente, a empresa decidiu investir numa campanha publicitária em diversas redes sociais. Dada a dimensão do investimento associado a esta campanha, a empresa espera que o número de encomendas diárias do produto passe a ser superior a 6.2 encomendas. Nos **5 dias** seguintes após o fim da campanha foram efetuadas 40 encomendas, o que dá, em média, 8 encomendas por dia. Teste, ao nível de significância de 5%, se a campanha teve o efeito esperado.

*Solução: teste **exato** a λ que, face ao tamanho da amostra (5), não pode ser usado o teste aproximado baseado na Normal; ($Y' = Y_1 + \dots + Y_5 \sim \text{Poisson}(\lambda' = 5 \times 6.2 = 31.0)$); $H_0: \lambda' = 31$ e $H_1: \lambda' > 31$; $\alpha = 5\%$; $ET = 40$; $pvalue = 4.87\%$ → rejeitar H_0*