

- 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**.
- Deve também mostrar explicitamente o raciocínio, apresentando as fórmulas que usar e mostrando como se utilizam no caso concreto substituindo os valores (nas fórmulas com somatórios basta mostrar a primeira e a última parcela). Não é necessário mostrar cálculos detalhados.
- Nos **testes de hipóteses** que realizar deve apresentar também o **valor de prova**, mesmo que aproximado, e uma **conclusão no contexto do enunciado**.
- Deve responder a **cada problema em folhas diferentes**.

1. (7.5 VALORES) Um componente é produzido em grandes quantidades. Dados históricos mostram que a percentagem de peças defeituosas produzidas é de 1.5%, que se admite ser constante. Os componentes são agrupados aleatoriamente em lotes de 200 unidades para posteriormente serem expedidos para os clientes (i.e, cada lote tem 200 unidades). Diariamente são produzidos 50 lotes do referido componente. De cada lote é seleccionada aleatoriamente uma amostra de 15 componentes para inspeção. Caso sejam encontrados dois ou mais componentes defeituosos na amostra todo o lote é rejeitado.
- a) Calcule a probabilidade de:
- i. um lote ser rejeitado.
  - ii. em 10 lotes seleccionados aleatoriamente haver pelo menos dois rejeitados.
- b) Calcule a probabilidade de duas peças consecutivas serem defeituosas.
- c) Calcule a probabilidade de o número total de peças defeituosas produzidas num dia ser superior a 125.
2. (7.5 VALORES) Na empresa FAZTUDO existem dois turnos de operários, A e B, que executam uma determinada tarefa. O supervisor de produção acredita que o turno A tem uma produtividade média superior em 1 unidade ao turno B, o que estes negam. Para averiguar quem fala verdade foi decidido efetuar um teste em que ficou estabelecido que, se na experiência a realizar, o valor da diferença de médias for inferior a um valor limite correspondente a um nível de significância de 5%, rejeita-se a posição defendida pelo supervisor de produção. Na tabela seguinte apresentam-se valores relativos a 2 amostras independentes de um indicador de produtividade para cada um dos 2 turnos.

Dimensão da Amostra	Media Amostral ( $\bar{X}$ )	Desvio Padrão Amostral (S)
$N_A = 51$	24.29	1.71
$N_B = 51$	23.97	1.13

- a) Calcule o intervalo de confiança para indicador de produtividade do turno A.
- b) Qual deverá ser a conclusão a tirar perante os dados fornecidos a um nível de significância de 5%? Defina convenientemente o teste de hipóteses a efetuar e determine o valor de prova.
- c) Suponha que, de facto, a diferença entre as produtividades médias dos dois turnos é de 0.32 unidades. Determine a nova dimensão da amostra a recolher para o indicador de produtividade dos operários do turno B de modo que a potência do teste que se irá efetuar seja, no mínimo, igual a 40%. (NOTA: o turno A mantém a mesma amostra).

3. (5 VALORES) Uma empresa de distribuição de eletricidade dedica grande parte da sua atividade na manutenção e/ou reparação do seu parque de transformadores de alta para média tensão. A nível de políticas de manutenção, a empresa pode escolher utilizar um de três níveis de condição (C1, C2 e C3), em que o nível C1 corresponde a uma política mais conservadora onde se privilegiam ações de reparação e de substituição antecipada de equipamentos e o nível C3 corresponde a uma política mais otimista onde a reparação e substituição de equipamentos tende a ser feito só em último caso. Dada a relevância diferenciada de cada transformador para o negócio, nem todos são sujeitos ao mesmo nível de condição. De forma geral, 5% dos transformadores são mantidos com o nível C1, 20% com o C2 e os restantes com o C3. No entanto, sabe-se que existe uma probabilidade de o transformador avariar antes de o mesmo atingir um determinado nível de condição. Com base no histórico de dados da empresa sabe-se que:

- a. 1% dos transformadores avariaram antes de atingir o nível de condição C1;
- b. 3% dos transformadores avariaram antes de atingir o nível de condição C2;
- c. 7% dos transformadores avariaram antes de atingir o nível de condição C3;

- a) Caso a empresa selecione aleatoriamente um transformador, qual a probabilidade de este não avariar?
- b) Calcule a probabilidade de numa seleção aleatória de 20 transformadores com o nível de política de manutenção C3, a probabilidade de ocorrer no máximo 3 avarias.

Recentemente, a empresa de distribuição de eletricidade tem vindo a desenvolver uma ferramenta de *data mining* que permite prever a condição dos transformadores com bases nos dados históricos de manutenção. Mais concretamente, a empresa aplica a técnica de regressão linear multivariada sobre os dados referente às concentrações dos gases ( $H_2$ ,  $CH_4$ ,  $C_2H_2$ ,  $C_2H_4$ ,  $C_2H_6$ ,  $CO$  e  $CO_2$ ) que se encontram presentes no óleo dos transformadores, para quantificar a condição dos mesmos. O impacto de cada um dos gases na previsão da condição dos transformadores é apresentado na tabela seguinte:

Variável	Coefficiente ( $\beta_i$ )	Coefficiente Padronizado ( $\hat{\beta}_i$ )	Valor de prova
$H_2$	-1,82	-0,476	0,049
$CH_4$	1,950	0,823	0,012
$C_2H_2$	0,376	0,052	0,261
$C_2H_4$	0,839	0,713	0,023
$C_2H_6$	0,375	0,642	0,035
$CO$	0,187	0,253	0,075
$CO_2$	-0,344	-0,951	0,001

- c) Admitindo um nível de significância de 5%, indique quais as variáveis que têm um peso estatisticamente significativo na previsão da condição dos transformadores. Justifique devidamente a sua resposta.
- d) Indique as três variáveis com mais impacto na previsão da condição dos transformadores. Justifique devidamente a sua resposta.