

Data: 11 de julho de 2023

Duração: 2 horas e 30 minutos

Nome: \_\_\_\_\_

Número: \_\_\_\_\_

### Instruções:

- As questões devem ser respondidas em folha de exame da ESTSetúbal (não se aceitam questões escritas a lápis) e os cálculos devem ser entregues num script do R (ficheiro com extensão .R) que deve ser entregue via Moodle.
- Na folha de exame não pode responder a diferentes questões numa mesma folha de resposta. Coloque o seu nome e número em todas as folhas de exame e no enunciado.
- Não é permitido partilhar ficheiros durante a prova nem utilizar o email ou outras formas de comunicação.
- Só é permitido aceder ao Moodle no início do exame para obter os dados e no fim do exame para enviar o script (ficheiro com extensão .R).
- Durante o exame o seu computador deverá estar sempre no R ou RStudio, não é permitido aceder a outros programas.
- É permitida a consulta do caderno/apontamentos e dos slides das aulas em versão papel.
- O abandono da sala só poderá efetuar-se decorridos 60 minutos a partir do início da prova e implica a entrega da mesma.
- **Justifique convenientemente todas as respostas na folha de exame da ESTSetúbal** e recorra ao software R apenas para efetuar os cálculos necessários.

### Questões:

1. Um grande hipermercado pretende estudar a preferência dos consumidores relativamente aos cereais de pequeno almoço americanos que comercializa, tendo em conta as componentes nutricionais de cada um e a localização nas prateleiras. No ficheiro cereal.txt, disponível no Moodle, são apresentados os dados recolhidos no estudo e no qual se encontram os seguintes campos:
  - name = nome do cereal de pequeno almoço
  - calories = calorias por dose do cereal (em quilocalorias)
  - carbo = hidratos de carbono por dose do cereal (em gramas)
  - sugars = quantidade de açúcar no cereal (em gramas)
  - shelf = prateleira do hipermercado onde se encontra o cereal  
(1 = mais baixa, 2 = prateleira do meio, 3 = prateleira mais alta)
  - client = consumidor a quem se destina o cereal (I = infantil, A = adulto)
  - rating = pontuação atribuída ao cereal
- [1.5] (a) Construa a tabela de frequências completa relativamente à prateleira em que os cereais se encontram e calcule as medidas de localização e dispersão adequadas. Justifique a sua resposta.
- [1.5] (b) Com base num histograma, classifique os dados referentes aos hidratos de carbono quanto à simetria. Na construção do histograma considere 7 classes, a primeira classe a começar no mínimo dos dados, amplitude por classe de 3.5 gramas e classes fechadas à direita e abertas à esquerda.

- [2.0] (c) A Direção Geral de Saúde suspeita que os cereais de pequeno almoço destinados ao consumidor infantil contêm, em média, mais açúcar do que os destinados a adultos. Com base nos dados disponibilizados e recorrendo a um teste de hipóteses paramétrico, verifique se esta suspeita é verdadeira. Considere um nível de significância de 4% e tome a sua decisão com base na região crítica.
- [1.5] (d) Por uma questão de marketing, os responsáveis desta cadeia de hipermercados pretendem que, na prateleira mais alta, a percentagem de embalagens destinadas ao consumidor infantil seja de 10%. Através de um intervalo de confiança, com 96% de confiança, verifique se este hipermercado está a cumprir essa orientação.
- [1.5] (e) Com base nos dados que dispõe, diga, justificando, a partir de que nível de significância se pode considerar que existe associação entre a prateleira onde se encontra a embalagem de cereais e o consumidor a quem se destina.
- [2.0] (f) Pretende-se estimar, através da regressão linear simples, a pontuação de um cereal de pequeno almoço que apresente 125 quilocalorias por dose. Avalie, justificando, se o modelo de regressão linear é adequado, defina a variável independente e dependente, calcule a reta de regressão e, com base nessa reta, efetue a previsão pretendida. Comente o resultado obtido.

2. O responsável da loja I-TEL de uma grande superfície comercial está a analisar a estratégia de mercado que está a ser aplicada. Seja  $X$  a variável aleatória discreta, que representa a procura diária de um dado modelo de *smartphone* na loja I-TEL, com a seguinte função de probabilidade:

$x$	0	1	2	3
$f(x)$	0.25	0.125	0.5	0.125

- [1.5] (a) Num dia em que já houve procura de pelo menos um daqueles *smartphones*, qual a probabilidade da procura ser inferior a 3?
- [1.0] (b) Determine o valor esperado de  $Y = -3X + 2$ .
- [1.5] (c) O tempo entre entradas consecutivas de clientes na loja I-TEL é uma variável aleatória Exponencial com valor esperado 15 minutos. Determine a probabilidade de, em 3 horas, entrarem pelo menos 10 clientes na loja.
- (d) O tempo de produção de um lote de *smartphones* é uma variável aleatória normal de média 50 dias e desvio padrão 5 dias.
- [1.0] i. A empresa tem de estabelecer um prazo de entrega para um lote cuja produção vai ser iniciada agora. Que prazo deve ser indicado ao cliente de forma a ser cumprido em 95% das vezes?
- [2.0] ii. Foram encomendados 10 lotes de *smartphones*. Qual a probabilidade de apenas 3 lotes demorarem mais de 60 dias em produção?

3. Para avaliar a viabilidade de instalação de uma central de energia eólica na localização A, efetuou-se um estudo sobre o comportamento do vento, tendo-se recolhido aleatoriamente 11 medições da velocidade do vento (em nós) em diferentes alturas do ano, cujos valores foram:

55, 61, 41, 42, 43, 28, 37, 40, 42, 48, 25

Admite-se que a velocidade do vento tem um comportamento normal.

- [1.5] (a) Obteve-se o seguinte intervalo de confiança para o desvio padrão da velocidade do vento:

$$]7.6878; 16.5710[.$$

Determine o grau de confiança associado a esse intervalo.

- [1.5] (b) Interprete e teste as seguintes hipóteses:

$$H_0 : \mu = 40 \quad \text{contra} \quad H_1 : \mu = 38.$$

Considere um nível de significância de 1%.