

Data: 22 de abril de 2023

Duração: 2 horas

Nome: _____

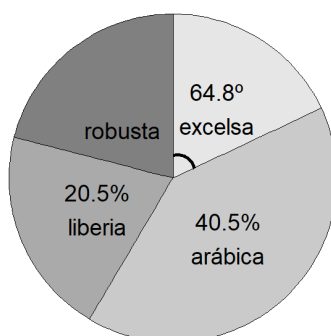
Número: _____

Instruções:

- As questões devem ser respondidas em folha de teste 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 teste não pode responder a diferentes questões numa mesma folha de resposta. Coloque o seu nome e número em todas as folhas de testes 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 teste para obter os dados e no fim do teste para enviar o script (ficheiro com extensão .R).
- Durante o teste 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 decorrido 60 minutos a partir do início da prova e implica a entrega da mesma.
- **Justifique convenientemente todas as respostas na folha de teste da ESTSetúbal** e recorra ao software R apenas para efetuar os cálculos necessários.

Questões:

1. Considere o seguinte gráfico circular referente ao registo da preferência em relação ao sabor do café dos 4 principais tipos de grãos de café, manifestada num inquérito feito a 1000 consumidores.



- [1.5] (a) Construa a tabela de frequências só com as frequências absolutas e relativas e calcule as medidas de localização e dispersão que achar adequadas. Justifique a sua resposta.
- (b) No ficheiro `cafe.txt`, disponível no Moodle, são apresentados os dados recolhidos referentes à produção de café no ano de 2019 em alguns dos países com plantações de café. Nesse ficheiro encontram-se os seguintes campos:
- country = nome do país
 - bags = número de sacos de café de 60 kg (em milhares)
 - month = mês em que é feita a colheita (4 = abril, 6 = junho, 10 = outubro)
 - price = preço pago aos produtores (em centimos de dolar por tonelada)

- [1.0] i. Caso seja possível, identifique a População e a Amostra indicando as suas dimensões, a unidade estatística, as variáveis estatísticas e os dados estatísticos, classificando-os.
- [0.5] ii. Qual o país que produziu mais em 2019? Qual o país que foi melhor pago em 2019?
- [1.5] iii. Construa a tabela de frequências completa relativa ao mês em que é feita a colheita e represente-a graficamente recorrendo às frequências relativas.
- [1.0] iv. Compare, recorrendo ao diagrama de extremos e quartis, o preço pago aos produtores de acordo com o mês em que é feita a colheita. Comente os resultados.
- [1.0] v. Recorrendo às medidas de dispersão adequadas, compare quanto à dispersão os dados referentes ao número de sacos de café de 60kg e o preço pago aos produtores. Comente os resultados.
- [1.5] vi. Em relação ao preço pago aos produtores, construa classes recorrendo à regra de Sturges, comece as classes com o mínimo dos dados e considere classes abertas à direita e fechadas à esquerda. Com base nas classes definidas, organize os dados numa tabela de frequências e represente-a graficamente considerando as frequências absolutas.

2. A variável aleatória X representa o diâmetro, em milímetros, dos parafusos produzidos numa determinada linha de montagem e tem a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} x & , 0 < x < 1 \\ \frac{x}{8} & , 1 \leq x < 3 \\ 0 & , \text{caso contrário} \end{cases}.$$

- [1.5] (a) Calcule a função de distribuição da variável aleatória X .
- [1.5] (b) Calcule $P\left(X \leq \frac{5}{3} \mid \frac{1}{4} < X < \frac{8}{3}\right)$.
- [1.0] (c) Qual o diâmetro máximo que deve constar nas especificações destes parafusos de modo a que apenas 5% dos parafusos tenha diâmetro superior ao diâmetro máximo.
- (d) De acordo com dados históricos da produção, sabe-se que o número de parafusos defeituosos por dia de produção (8 horas) segue uma distribuição de Poisson com média de 1.2 parafusos defeituosos.
- [1.0] i. Qual é a probabilidade de, em 10 dias de produção, serem encontrados 20 parafusos defeituosos?
- [1.5] ii. Num dia de produção, qual a probabilidade de passarem pelo menos 6 horas entre o tempo de produção de dois parafusos defeituosos?

3. Num estabelecimento, que vende materiais de construção, sabe-se que as vendas diárias de areia, em toneladas, têm um comportamento aleatório, traduzido por uma distribuição normal, de média 2 toneladas.

- [1.5] (a) O responsável pelo estabelecimento afirma que em 95% dos dias as vendas de areia ultrapassam os 1500 quilos. Calcule o valor do desvio padrão das vendas diárias de areia que torna verdadeira esta afirmação.
- (b) Suponha que a variância das vendas diárias de areia é de 0.04 toneladas².
- [1.5] i. Sabendo que numa manhã o estabelecimento já vendeu pelo menos uma tonelada de areia, qual a probabilidade de, nesse dia, as vendas serem superiores a 2.5 toneladas?
- [1.5] ii. Determine a probabilidade de, em 20 dias úteis de um mês, haver pelo menos 5 dias com vendas que não atingem a média diária.
- [1.0] (c) Suponha que Y é uma variável aleatória que representa o número de camiões utilizados diariamente para transporte da areia e que assume os valores 0, 2 e 3 com probabilidades 0.1, 0.3 e 0.6, respectivamente. Sabendo que o custo diário (em unidades monetárias) de cada camião é dado pela variável aleatória $W = 3Y^2 - 1$, calcule qual é o custo médio diário com os camiões.