

Instruções:

- É obrigatória a apresentação de um documento de identificação.
- Não se aceitam provas ou questões escritas a lápis.
- Não pode responder a diferentes questões numa mesma folha de resposta.
- O abandono da sala só poderá efetuar-se decorrida uma hora a partir do início da prova e implica a entrega da mesma.
- É permitida a utilização individual de máquina de calcular, a consulta de uma folha A4 manuscrita pelo aluno (em suporte papel e não são permitidas fotocópias de folhas manuscritas) e a consulta das tabelas fornecidas pelos docentes. Caso consultem outros documentos, por exemplo através da calculadora, a prova será anulada.
- Não é permitido o manuseamento ou exibição de telemóveis durante a prova.
- **Justifique convenientemente todas as respostas.**

Questões:

1. Foi recolhida a seguinte amostra referente ao peso, em kg, de 24 pessoas:

65	63	67	64	68	62	70	66	68	67	69	71
68	66	68	65	69	66	68	65	71	67	68	70

- [2.0] (a) Organize a amostra numa tabela de frequências, construindo classes com recurso à regra de Sturges.
- (b) A primeira linha da tabela refere-se ao peso de 12 pais (X) quando estes completaram 20 anos e a segunda linha da tabela refere-se ao peso dos respetivos 12 filhos (Y) quando estes completaram 20 anos. A informação da tabela encontra-se resumida nas seguintes quantidades:

$$\sum_{i=1}^{12} x_i = 800; \quad \sum_{i=1}^{12} x_i^2 = 53418; \quad \sum_{i=1}^{12} y_i = 811; \quad \sum_{i=1}^{12} y_i^2 = 54849; \quad \sum_{i=1}^{12} x_i y_i = 54107$$

- [1.0] i. Qual o conjunto de dados que apresenta maior dispersão, o peso dos pais aos vinte anos ou peso dos filhos aos vinte anos?
- [2.5] ii. Pretende-se prever o peso de um filho aos vinte anos quando o pai aos 20 anos pesava 98kg. Defina a variável independente e variável dependente, justifique se o modelo de regressão linear é adequado, calcule a reta de regressão, efetue a previsão pretendida e comente os resultados.
- [1.5] iii. Suponha que o peso tem distribuição Normal. Se pretender construir um intervalo de confiança para o peso médio dos filhos aos vinte anos, cuja margem de erro seja de 1 kg, qual deveria ser o grau de confiança do intervalo?
- [1.5] iv. Suponha que o peso tem distribuição Normal. Através de um intervalo de confiança a 95%, podemos concluir que existem evidências de que o desvio padrão do peso dos filhos aos vinte anos é de 2kg?

2. Suponha que o número de vezes em que uma dada máquina é desligada, num dia em que trabalhe 10 horas, tem distribuição de Poisson de variância igual a 2.
- [1.5] (a) Qual a probabilidade da máquina não ter sido desligada, num dia em que só trabalhou 8 horas?
- [1.5] (b) Qual a probabilidade da máquina estar a trabalhar durante pelo menos 7 horas seguidas sem ser desligada?
- [1.0] (c) Considerando que o tempo, em horas, que o operador da máquina trabalha ininterruptamente por dia é uma variável aleatória Y com distribuição uniforme no intervalo $[2, 5]$. Calcule a variância de $W = \frac{1-Y}{2}$.
3. Um estudo sobre congestionamento de tráfego numa dada artéria de uma cidade, revelou que o tempo, em minutos, necessário para o seu percurso em hora de ponta é uma variável aleatória X com distribuição normal e com desvio padrão igual a 3 minutos.
- [1.5] (a) Sabendo que 15.87% dos veículos ligeiros demoram mais de 13 minutos no percurso da referida artéria em hora de ponta, calcule o valor esperado de X .
- [2.0] (b) Suponha que o tempo médio para o percurso da referida artéria em hora de ponta é de 10 minutos. Considere um grupo de 15 veículos ligeiros que percorreram o troço em questão em hora de ponta e suponha que os tempos de percurso são independentes de veículo para veículo. Determine a probabilidade do tempo total de percurso dos 15 veículos ultrapassar duas horas.
4. Um biólogo marinho está a estudar uma espécie de pinguins e está interessado em saber se é o tamanho do macho que leva à formação de casais nestas aves. A variável utilizada para o efeito é o comprimento de um osso da perna que, a partir de estudos anteriores, é tida como um bom indicador da variável em causa. As medidas estão em milímetros arredondadas às décimas e os resultados obtidos em 8 casais de pinguins foram os seguintes:

casal	1	2	3	4	5	6	7	8
macho	17.1	18.5	19.7	16.2	21.3	19.6	17.2	20.1
fêmea	16.5	17.4	17.3	16.8	19.5	18.3	17.5	20.4

- [2.0] (a) Considere que as observações são provenientes de populações normais. Recorrendo a um teste de hipóteses paramétrico e para um nível de significância de 1%, pode afirmar que, em média, o tamanho do pinguim macho pode ser considerado significativamente superior ao tamanho do pinguim fêmea?
- [2.0] (b) Recorreu-se a um teste de hipóteses não paramétrico e obteve-se o seguinte resultado:

```
wilcoxon signed rank test with continuity correction
data: amostra$femea and amostra$macho
V = 6.5, p-value = 0.123
alternative hypothesis: true location shift is not equal to 0
```

Com base neste resultado e para um nível de significância de 5%, que conclusão pode tirar sobre os tamanhos dos casais destas aves? Diga, justificando, que teste não paramétrico foi efetuado, quais as hipóteses testadas e tome a decisão com base na região crítica.

Fim