

Data: 16 de julho de 2021

Duração: 2 horas e 30 minutos

Instruções:

- Não se aceitam provas ou questões escritas a lápis.
- Não pode responder a diferentes questões numa mesma folha de resposta.
- É permitida a utilização individual de máquina de calcular, a consulta de tabelas e a consulta de uma folha A4 manuscrita pelo aluno (não são permitidas fotocópias de folhas manuscritas nem a consulta de outros documentos através da máquina de calcular).
- Não é permitido o manuseamento ou exibição de telemóveis durante a prova.
- O abandono da sala só poderá efetuar-se decorrida uma hora a partir do início da prova e implica a entrega da mesma.
- **Justifique convenientemente todas as respostas.**

Questões:

1. Uma empresa que se dedica à área do desporto pretende estudar a relação entre a durabilidade, em anos, das suas sapatilhas de desporto e o seu preço de venda. Para tal recolheu a seguinte informação:

sapatilha	1	2	3	4	5	6	7	8	9	10
$X = \text{durabilidade (em anos)}$	1	4	4	3	5	3	3	3	6	5
$Y = \text{preço de venda (em euros)}$	17.99	12.99	56.99	53.99	25.99	16.99	43.99	9.99	16.99	51.99

$$\sum_{i=1}^{10} x_i = 37, \quad \sum_{i=1}^{10} x_i^2 = 155, \quad \sum_{i=1}^{10} x_i y_i = 1164.63, \quad \sum_{i=1}^{10} y_i = 307.9, \quad \sum_{i=1}^{10} y_i^2 = 12645.841$$

- [1.5] (a) Construa a tabela de frequências completa da durabilidade das sapatilhas de desporto.
- [1.5] (b) Em relação à durabilidade das sapatilhas de desporto, calcule as medidas de localização central e, com base nessas medidas classifique os dados quanto à simetria.
- [1.0] (c) Qual o conjunto de dados que apresenta maior dispersão relativa, a durabilidade ou o preço de venda?
- [1.5] (d) Como foi referido pretende-se estudar a relação entre a durabilidade das sapatilhas e o seu preço. Acha que o modelo de regressão linear simples é o adequado? Justifique a sua resposta.
2. Uma loja de material de desporto tem para venda produtos de uma determinada marca. Seja X a variável aleatória que representa o número de produtos dessa marca vendidos diariamente.
- [1.5] (a) Acerca desta variável aleatória sabe-se:
- o domínio é $\{0, 1, 2, 3\}$ e os valores 1 e 2 são igualmente prováveis;
 - em 20% dos dias não se vendem produtos dessa marca e em 15% dos dias vendem-se mais de 2 produtos dessa marca.
- Escreva a função de probabilidade e a função de distribuição da variável aleatória X .

(b) Afinal sabe-se que a variável aleatória X segue uma distribuição Uniforme com o domínio indicado na alínea anterior.

[1.5] i. Determine a probabilidade de num dia venderem menos de dois produtos dessa marca, sabendo que nesse dia já venderam produtos dessa marca.

[1.0] ii. Calcule a variância da variável aleatória $Y = 1 - 3X$.

3. O tempo, em horas, que um atleta de alta competição em período de estágio dorme por noite é uma variável aleatória com distribuição normal de média 8 horas e desvio padrão 0.6 horas.

[1.0] (a) Determine o número mínimo de horas de sono que um atleta dorme 33% das noites.

[1.5] (b) Qual a probabilidade de pelo menos 4 atletas, de um grupo de 15 atletas que estão em estágio, dormirem no máximo 7 horas por noite?

[1.5] (c) Qual a probabilidade de um atleta dormir em média mais de 8.5 horas por noite, num estágio de 15 dias?

(d) Para comparar a eficiência de dois métodos de treino, 24 atletas de alta competição foram divididos aleatoriamente em dois grupos. Cada grupo foi treinado de acordo com um método diferente. Os resultados no fim do treino foram os seguintes (numa escala de 0 a 100 e quanto mais alto, melhor o treino):

método 1	$n_1 = 13$	$\bar{x}_1 = 74.5$	$s_1^2 = 82.6$
método 2	$n_2 = 11$	$\bar{x}_2 = 71.8$	$s_2^2 = 112.6$

Suponha que as populações são normais.

[1.5] i. Obteve-se o seguinte intervalo de confiança para comparar os desvios padrão dos resultados dos treinos:

$$]0.5021, 1.4203[.$$

Qual o grau de confiança utilizado no cálculo deste intervalo? E qual a conclusão que pode tirar deste intervalo?

[1.5] ii. Com base num intervalo de confiança a 90% acha que, em média, os métodos de treino podem ser considerados iguais?

4. Um importante fabricante de equipamento desportivo pretende lançar no mercado da alta competição em natação um novo tipo de calções que, alegadamente, permitem reduzir os tempos obtidos em cada prova. Com o objetivo de verificar a vantagem do novo tipo de calções decidiu fazer uma experiência com base numa amostra emparelhada. Para tal seis nadadores fizeram a prova de 100 metros livres duas vezes, uma com o novo tipo de calções e outra com calções tradicionais e os resultados obtidos, em segundos, apresentam-se na tabela seguinte:

Nadador	1	2	3	4	5	6
Novo tipo de calções	53.94	55.86	49.83	58.89	55.90	57.01
Calções tradicionais	53.97	55.88	49.85	58.90	55.90	56.95

[1.5] (a) Recorra ao teste não paramétrico adequado para testar, a um nível de significância de 5%, se o novo tipo de calções permite obter melhores resultados. Tome a decisão com base na região crítica. Justifique a escolha do teste.

[2.0] (b) Suponha que as observações são provenientes de populações normais. Recorra ao teste paramétrico adequado para testar, a um nível de significância de 5%, se há evidência estatística que, em média, o tipo de calções influencia os resultados. Tome a decisão com base no valor-p.

Fim do Exame