ESTATÍSTICA I - 2º Ano/Gestão do Desporto 2º sem., 1ª Prova Intercalar 10. 04. 18

1horas. (10 valores)

Nome:____ Nº__________

Espaço reservado para classificações

1 a). (10)	2a.(10)	*3 a.(10)*	4a.(5)	
1 b). (10)	2b.(15)	3 b.(10)	4b.(5)	T:
1 c). (15)		*3 c. (10)*		

Atenção: 1. Devem apresentar na folha de exame a formalização e Justificação dos cálculos efectuados no EXCEL.
2. Devem fazer os cálculos no ficheiro EXCEL em folhas separadas para cada questão.
3. Nas questões de resposta múltipla uma resposta errada desconta 2.5.

1. Considere que o tempo, em segundos, dos atletas de alta competição que correram uma maratona Olímpica masculina pode ser bem modelado por uma variável aleatória com distribuição Normal. O ficheiro EXCEL- folha 1 contém os valores registados para uma amostra de 20 atletas.

   a. Estime os parâmetros da população pelo método dos momentos.

   $X \sim N(\mu, \sigma^2)$, como se tem 2 parâmetros desconhecidos ter-se-á de definir o sistema seguinte:

   $$\begin{cases} E(X) = \bar{X} \\ E(X^2) = \frac{\sum_{i=1}^{20} X_i^2}{n} \end{cases} \Leftrightarrow \begin{cases} \mu = \bar{X} \\ Var(X) + \mu_X^2 = \frac{\sum_{i=1}^{20} X_i^2}{n} \end{cases} \Leftrightarrow \begin{cases} \tilde{\mu} = \bar{X} \\ \tilde{\sigma}_X^2 = \frac{\sum_{i=1}^{20} X_i^2}{n} - \tilde{\mu}_X^2 \end{cases}$$ são

   os estimadores pelo método dos momentos de μ e σ^2 e as estimativas são:

   $\tilde{\mu} = \bar{x} = 8557.95$ e $\tilde{\sigma}_x^2 = 73420972 - 8557.95^{\ 2} = 182464 \Rightarrow \tilde{\sigma}_x$ =427.1581

   b. Estude o enviesamento e consistência do estimador para a média da população encontrado na alínea a).

   $E(\tilde{\mu}) = E(\bar{X}) = \mu$ pelo que o estimador é não enviesado

   $Var(\tilde{\mu}) = Var(\bar{X}) = \frac{\sigma^2}{n}$,

   então $$\begin{cases} \lim_{n\to\infty} E(\tilde{\mu}) = \lim_{n\to\infty} E(\bar{X}) = \lim_{n\to\infty} \mu = \mu \\ \lim_{n\to\infty} Var(\tilde{\mu}) = \lim_{n\to\infty} Var(\bar{X}) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0 \end{cases}$$ logo o estimador é consistente

c. Se considerarmos uma amostra aleatória de tempos de 10 atletas de alta competição que correram esta maratona Olímpica, qual a probabilidade de o tempo mais curto na amostra ser inferior a 8200 segundos?

$$P(min\{X_i\} < 8200) = \{1 - [1 - F_X(8200)]^{10}\} \approx 1 - [1 - 0.201]^{10} = 0.894$$

2. De uma população normal, extraíram-se várias amostras de dimensão $\boldsymbol{n}$**.** Comente a afirmação “ A média da amostra é sempre igual à média da população”. Corrija se necessário

R: A afirmação é falsa porque dada a **variabilidade da amostra**, a média da amostra diferirá de amostra para amostra e muito dificilmente será igual à media da população. A **afirmação verdadeira é que a média da média da amostra é igual à média da população**

3. Com base na amostra, que consta do ficheiro EXCEL- folha 1.
   a. Calcule o intervalo de confiança para a média da população para um nível de confiança de 90%.

   R: Variável fulcral - $\frac{\bar{X}-\mu}{S'/\sqrt{n}} \sim t(n-1)$ $\qquad IC_{\mu}^{90\%} = (8388.5,\ \ 8727{,}4)$

   b. Interprete o resultado obtido.

   R: O tempo médio, em segundos, dos atletas de alta competição que correram uma maratona Olímpica masculina pertencerá ao intervalo $(8388.5,\ \ 8727{,}4)$ **em 90% dos Intervalos de Confiança para a média calculados com informação de amostras de igual dimensão** retiradas desta população.

   c. Explique como e porque varia a amplitude do intervalo de confiança com a dimensão da amostra e o nível de confiança.

   R: Quando aumenta a dimensão da amostra, mantendo constante o nível de confiança, diminui a amplitude do intervalo de confiança porque uma amostra maior tem mais informação o que permite ter o mesmo grau de confiança com um intervalo de menor amplitude.

   Quando aumenta o nível de confiança, mantendo constante a dimensão da amostra, a amplitude do intervalo de confiança aumenta porque só se pode ter maior confiança em que acertamos no “alvo” se ele tiver uma maior dimensão.

4. Pretende-se estimar a proporção de atletas com mais de 25 anos numa população. Para tal recolheu-se uma amostra aleatória de 50 indivíduos cujos valores constam do ficheiro EXCEL- folha 2.
   a. Calcule a estimativa por intervalos de confiança para o parâmetro desejado com um grau de confiança de 90%.

   R: Variável fulcral - $\frac{\bar{X}-\theta}{\sqrt{\frac{\bar{X}(1-\bar{X})}{(n)}}} \dot{\sim} N(0,1)$ $\qquad IC_{\theta}^{90\%} = (0.5283, 0.7517)$

b. Determine a dimensão da amostra necessária para garantir um erro de amostragem máximo de 10% na estimação da proporção pretendida.

Erro de amostragem - $\epsilon = z\alpha/_2 \frac{s'}{\sqrt{n}}$

$$n: z\alpha/_2 \frac{s'}{\sqrt{n}} \leq 0.1 \Leftrightarrow n \geq 63$$

5.

a. Para cada uma das proposições seguintes diga se se trata ou não de uma hipótese estatística, assinalando com um S (sim) ou N (não) no quadrado respectivo:

$\theta \leq 0.6$ S $\quad$ $\bar{x} \leq 2$ N $\quad$ $1.2 < S'^2 \leq 3.5$ N $\quad$ $\bar{X} > 5$ N

[**Nota: a notação acima é a utilizada no âmbito desta unidade curricular**]

b. Se um teste de hipótese simples contra hipótese simples tem potência 0.3 isto significa que: (assinale com uma cruz no quadrado respetivo.

Quando H_1 é verdadeira, a probabilidade de H_0 ser rejeitada é 0.3 ☐

Quando H_0 é falsa, a probabilidade de H_0 ser rejeitada é 0.3 X

Quando H_0 é verdadeira, a probabilidade de H_0 não ser rejeitada é 0.3 ☐

Quando H_1 é verdadeira, a probabilidade de H_0 não ser rejeitada é 0.3 ☐