

Observação: A resolução completa das questões inclui a justificação do raciocínio utilizado e a apresentação dos cálculos efectuados.

(13.0) **I.** Seja \mathcal{E} uma experiência aleatória e A um acontecimento associado a \mathcal{E} de probabilidade p ($p \in]0, 1[$). Sejam X a variável aleatória real \mathbb{I}_A e (X_1, X_2, \dots, X_n) uma amostra aleatória de dimensão n de X .

1. Indique o modelo estatístico associado a (X_1, X_2, \dots, X_n) .
2. Obtenha um estimador da máxima verosimilhança, T_n , de p e prove que é único.
3. Conclua que T_n é um estimador de p convergente, precisando os modos de convergência. Interprete o resultado obtido.
4. Obtenha um estimador do desvio padrão de X convergente quase certamente.
5. Deduza a lei limite da v.a.r.

$$Z_n = \sqrt{n} \frac{T_n - p}{\sqrt{T_n(1 - T_n)}}.$$

6. Com o objectivo de avaliar se determinado referendo a uma população poderá vir a ser considerado significativo, pretende-se inferir a correspondente proporção de abstenção, p , a partir duma amostra de n eleitores aleatoriamente escolhidos.
 - a) Obtenha um valor para a dimensão, n , da amostra a observar que lhe permita garantir, com uma confiança de 95%, uma estimação intervalar para p de precisão quando muito igual a 0.03.
 - b) Sabendo que foram inquiridos 5000 eleitores dos quais 2450 declararam vir a abster-se, teste, ao nível de significância 0.05, as hipóteses

$$H_0 : p = 0.5 \text{ contra } H_1 : p < 0.5.$$

- c) Com base na mesma amostra, que decisão tomaria ao nível de significância 0.01? Justifique convenientemente a sua resposta.

(7.0) **II.** Para estudar a influência de um preparado de insulina, \mathbf{x} , na redução do nível de açúcar, Y , foi feita uma experiência em 30 cobaias saudáveis, tendo-se registado para cada valor x_n de \mathbf{x} o correspondente valor do nível de açúcar y_n de Y . Foi analisada, por meio do *software* estatístico SPSS, a existência de uma relação linear da forma $Y_n = ax_n + b + U_n$ ($n = 1, 2, \dots, 30$), com (U_n) variáveis aleatórias independentes e identicamente distribuídas com uma lei normal centrada de desvio padrão σ ($\sigma > 0$, desconhecido). A análise descritiva dos dados conduziu aos seguintes resultados:

Descriptive Statistics			
	Mean	Std. Deviation	N
nível	47.37	10.287	30
insulina	35.53	8.874	30

Correlations

		nível	insulina
Pearson Correlation	nível	1.000	.877
	insulina	.877	1.000
Sig. (1-tailed)	nível	.	.000
	insulina	.000	.
N	nível	30	30
	insulina	30	30

1. Qual a média e a covariância empírica da amostra observada? Que pode concluir sobre a existência de uma dependência funcional linear entre as variáveis em estudo?

Nos quadros abaixo encontram-se os resultados correspondentes à regressão linear efectuada sobre os dados.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		
	B	Std. Error	Beta			Lower Bound	Upper Bound	
1	(Constant)	11.247	3.854		2.918	.007	3.352	19.143
	insulina	1.016	.105	.877	9.649	.000	.801	1.232

a. Dependent Variable: nivel

One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
N		30
Normal Parameters ^{a,b}	Mean	.0000000
	Std. Deviation	4.94642100
Most Extreme Differences	Absolute	.095
	Positive	.095
	Negative	-.065
Kolmogorov-Smirnov Z		.522
Asymp. Sig. (2-tailed)		.948

a. Test distribution is Normal.

b. Calculated from data.

- Indique a relação linear estimada e tire conclusões sobre a adequação, aos dados observados, do modelo estocástico considerado.
- Indique intervalos de confiança para os coeficientes do modelo. Permitirá a amostra concluir que o coeficiente a é significativamente diferente de zero?
- Mostre que $V = \frac{1}{\sigma^2} \sum_{i=1}^{30} U_i^2$ é uma variável fulcral para o modelo estatístico em estudo.
- Com base nos resultados obtidos e sabendo ainda que a soma dos quadrados dos resíduos da estimação é igual a 709.545, deduza um intervalo real que contenha σ com 98% de confiança.