Departamento de Matemática da Universidade de Coimbra

Exame de Estatística

Duração: 2h 30m 05-09-08

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

1. Seja $(X_1, X_2, \dots, X_n)$ uma amostra de uma variável aleatória real (v.a.r.) X de valor médio m e variância σ^2 ($m \in \mathbb{R}$ e $\sigma > 0$). Sejam $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ e $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$, respectivamente, a média e a variância da amostra. Mostre que $(\overline{X}_n, S_n)$ é um estimador quase certamente convergente de (m, σ).

2. Com o objectivo de estimar a relação entre a cilindrada do motor de um automóvel (em cm^3), X, e o correspondente consumo de gasolina em estrada (em $litros/100km$), Y, recolheu-se uma amostra de 40 automóveis cujos valores de cilindrada e correspondente consumo de gasolina deram origem aos resultados seguintes, obtidos através do *software* estatístico SPSS.

Correlations

		consumo	cilindrada
Pearson Correlation	consumo	1,000	,987
	cilindrada	,987	1,000

Coefficients(a)

Model		Unstandardized Coefficients	
		B	Std. Error
1	(Constant)	1,963	,111
	cilindrada	,003	,000

a Dependent Variable: consumo

a) Indique o coeficiente de correlação da amostra observada e interprete-o.

b) Obtenha uma estimativa para o consumo (em $litros/100km$) de um automóvel com cilindrada igual a 1560 cm^3.

c) Os estudos de natureza inferencial envolvendo as variáveis em estudo requerem, classicamente, o pressuposto de que as diferenças entre os valores observados da variável Y e os correspondentes valores obtidos a partir da recta de regressão de Y sobre X, ditos resíduos, constituam uma amostra de uma v.a.r. seguindo uma lei normal centrada. Para a amostra observada, os valores de tais resíduos, que se encontram resumidos no quadro seguinte, apresentam média nula e desvio padrão igual a 0.14.

Resíduos	$]-0.4, -0.15]$	$]-0.15, 0]$	$]0, 0.15]$	$]0.15, 0.3]$
Frequências	4	17	13	6

Ao nível de significância 0.05, pode considerar que o referido pressuposto de normalidade é válido?

d) Admitindo a veracidade da hipótese testada na alínea anterior, determine um intervalo real que contenha o desvio padrão dos resíduos com 95% de confiança.

v.p.f.

3. O tempo de vida de determinado tipo de componentes electrónicas é descrito por uma v.a.r. X de lei $\mathcal{L}(\theta, 2)$, absolutamente contínua, de densidade

$$f_\theta(x) = \frac{1}{\theta}\,\mathrm{e}^{\frac{1}{\theta}(2-x)}\,\mathbb{1}_{]2,+\infty[}(x),$$

onde θ é um parâmetro positivo, desconhecido. Nestas condições, o tempo de vida médio de tais componentes é $\theta + 2$ *u.t.* (unidades de tempo) e o correspondente desvio padrão é θ *u.t.*. Seja $(X_1, X_2, \ldots, X_n)$ uma amostra aleatória de X.

a) Com o objectivo de estimar θ, consideram-se os estimadores

$$T_n = n\left(\min_{1\leq i\leq n} X_i - 2\right) \quad \text{e} \quad Z_n = \bar{X}_n - 2.$$

(i) Verifique que a v.a.r. $M_n = \min\limits_{1\leq i\leq n} X_i$ segue a lei $\mathcal{L}(\frac{\theta}{n}, 2)$.

(ii) Prove que os estimadores T_n e Z_n são cêntricos.

(iii) Compare os estimadores T_n e Z_n relativamente à função de risco quadrático.

b) Obtenha o estimador de máxima verosimilhança de θ.

c) Prove que, para n suficientemente grande, $\bar{X}_n$ segue aproximadamente a lei normal de média $\theta + 2$ e desvio padrão $\frac{\theta}{\sqrt{n}}$.

d) Estudos estatísticos anteriores levaram à conclusão de que o tempo de vida médio das referidas componentes era de $4\,u.t.$. Recentemente, foram introduzidos alguns aperfeiçoamentos no processo de fabrico com o objectivo de aumentar aquele tempo de vida. Para verificar tal facto, recolheu-se uma amostra de tamanho 100 de tais componentes, a qual apresentou uma média de $4.25\,u.t.$.

(i) Verifique que o teste de região crítica da forma

$$\mathcal{C} = \left\{(x_1, x_2, \ldots, x_{100}) \in\,]2, +\infty[^{100} \colon \sum_{i=1}^{100} x_i > k\right\},$$

com $k > 0$, é o mais potente ao seu nível para testar $H_0 : E(X) = 4$ contra $H_1 : E(X) = 4.5$ e indique a decisão a que conduz a amostra observada, ao nível de significância 0.05.

(ii) Perante a decisão tomada, qual a probabilidade de errar?

Cotação

1. 2.0 valores
2. 7.0 valores
3. 11.0 valores