Departamento de Matemática da Universidade de Coimbra

Teste de Estatística

Duração: 1h30m 26–04–2006

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

1. Considere uma variável estatística (v.e.) bidimensional (X, Y) e sejam $(x_1, y_1), ..., (x_N, y_N)$ N observações de tal variável.

a) Defina coeficiente de correlação, r, entre X e Y e indique uma condição sobre a nuvem de pontos $\{(x_1, y_1), ..., (x_N, y_N)\}$ que assegure a sua existência.

b) Sob a condição referida na alínea anterior, prove que $|r| = 1$ se e só se os pontos observados $(x_i, y_i), \forall i \in \{1, ..., N\}$, pertencem à recta de equação

$$y = \overline{y} + r\frac{s_Y}{s_X}(x - \overline{x}),$$

onde $\overline{x}$, $\overline{y}$, s_X e s_Y designam, respectivamente, as médias e os desvios padrão das v.e. X e Y.

(**Sugestão:** Tenha em conta que $\displaystyle\frac{1}{N}\sum_{i=1}^{N}\left[(y_i - \overline{y}) - r\frac{s_Y}{s_X}(x_i - \overline{x})\right]^2 \geq 0$).

2. Para estudar o efeito da qualidade do ar na água de um lago fizeram-se, em diferentes períodos, medições do pH da água, descrito por uma variável aleatória real (v.a.r.) Y, e do índice de qualidade do ar, descrito por uma v.a.r. X (valores crescentes de X indicam aumento de poluição). Pretende-se aproximar a relação existente entre Y e X por uma relação linear, aplicando o método dos mínimos quadrados aos 6 valores daquelas medições representados no quadro seguinte:

índice X	40	50	60	10	15	85
pH Y	4.5	4.1	4.0	6.0	6.1	3.2

a) Calcule, justificando convenientemente, uma estimativa cêntrica do valor médio do pH da água e do índice de qualidade do ar, isto é, do par $(E(X), E(Y))$.

b) Calcule o coeficiente de correlação da amostra observada. Que pode afirmar sobre o sentido da variação do pH da água e do índice de qualidade do ar?

c) Determine, a partir da amostra observada, a recta de regressão de Y sobre X e estime, a partir dela, o pH da água do referido lago se for de 20 o índice de qualidade do ar.

v.s.f.f.

3. Num processo de controlo de qualidade de determinado tipo de material regista-se o número de peças desse material que são produzidas até ocorrer a primeira com defeito. Sabendo que as peças são executadas independentemente umas das outras e sempre nas mesmas condições, a variável aleatória real (v.a.r.) X que descreve aquele número aleatório é discreta de função de probabilidade

$$f(x) = p(1-p)^{x-1}\mathbb{I}_{\mathbb{N}}(x)\,, \text{ com } p \in \,]0,1[\, \text{ desconhecido.}$$

Além disso, $E(X) = \frac{1}{p}$ e $V(X) = \frac{1-p}{p^2}$.

Com o objectivo de estimar p, considere uma amostra aleatória de dimensão n de X, $(X_1, X_2, ..., X_n)$.

a) Obtenha o estimador da máxima verosimilhança de p, T_n, e verique que T_n é também o estimador dos momentos de p.

b) Mostre que T_n é um estimador de p quase certamente convergente.

c) Deduza o estimador dos momentos e da máxima verosimilhança do número médio de peças do referido material que são produzidas até ocorrer a primeira com defeito.

d) Calcule o erro quadrático médio do estimador.obtido na alínea anterior.
