Departamento de Matemática

Secção de Estatística e Aplicações

Probabilidades e Estatística (I)

Exame de 1ª Época / 2º Teste — 2º semestre – 2003/04

Duração: 3 horas / 1 hora e 30 minutos — 24/06/2004 – 9 horas

- Se pretende fazer o **exame** deve resolver **todos os grupos**.
- Se pretende fazer o **2º teste** deve resolver **apenas os grupos III e IV**. Nesse caso as cotações passam a ser o dobro das indicadas.
- Justifique convenientemente **todas as respostas!**

Grupo I — 7.0 valores

1. Segundo um grupo de especialistas, as fundações de um determinado arranha-céus podem falhar devido essencialmente a duas causas. Seja C_i = "Fundações falham devida à causa i", $i = 1, 2$. Tendo em conta que $P(C_1) = 0.001$, $P(C_2) = 0.008$ e $P(C_1 | C_2) = 0.1$, determine:

   (a) A probabilidade de as fundações falharem. (1.5)

   (b) A probabilidade de as fundações falharem devido à causa C_2 mas não à causa C_1. (1.0)

2. Mostre que dois acontecimentos mutuamente exclusivos só podem ser independentes se pelo menos um deles for um acontecimento impossível ou quase impossível. (1.0)

   Nota: Diz-se que um acontecimento $A \neq \emptyset$ é quase impossível se $P(A) = 0$.

3. Dois dados perfeitos são lançados simultaneamente 120 vezes. Considerando que os lançamentos são independentes:

   (a) Mostre que a probabilidade de se obter soma de pontos igual a sete num lançamento dos dois dados é $1/6$. (0.5)

   (b) Escreva a expressão exacta da probabilidade de saírem pelo menos 17 vezes soma de pontos igual a sete nos 120 lançamentos. (1.0)

   (c) Calcule aproximadamente, com e sem correcção de continuidade, a probabilidade pedida na alínea (b). Comente. (2.0)

Grupo II — 3.0 valores

Considere o par aleatório (X, Y) onde X e Y representam, respectivamente, a concentração de antibiótico e de fungos numa cultura de microorganismos (em percentagem). Admita que a função densidade de probabilidade conjunta deste par aleatório é

$$f_{X,Y}(x, y) = \begin{cases} x + y, & 0 < x < 1,\ 0 < y < 1 \\ 0, & \text{caso contrário} \end{cases}$$

(a) Determine as funções densidade de probabilidade marginais de X e de Y. (1.0)

(b) Calcule o valor do coeficiente de correlação entre X e Y. Comente o resultado obtido. (2.0)

Grupo III **6.0 valores**

1. Um engenheiro, ligado a problemas de tráfego, mediu 200 vezes o tempo (em *segundos*) entre passagens sucessivas de viaturas por uma passadeira, situada na proximidade de uma escola. A amostra obtida possui média igual a 9.08 segundos.

   (a) Admitindo que o tempo entre duas passagens sucessivas de viaturas tem distribuição exponencial de parâmetro λ, deduza a estimativa de máxima verosimilhança de λ. Obtenha, justificando, a estimativa de máxima verosimilhança da probabilidade de o tempo entre duas passagens sucessivas exceder 10 segundos. (2.0)

   (b) Um outro especialista em problemas de tráfego é da opinião que se deva instalar um semáforo caso o tempo esperado entre duas passagens sucessivas de viaturas seja inferior ou igual a 10 segundos. Efectue um teste de hipóteses que auxilie na tomada de decisão sobre a instalação ou não do semáforo, ao nível de significância de 5%. (1.5)

2. Num estudo prévio à instalação de uma central eólica em certa região foram colocadas diversas questões acerca da velocidade mínima do vento nessa região (em Km/h). Um metereologista suspeita que esta velocidade possua função de distribuição

$$F_X(x) = 1 - \exp\left[-\exp\left(\frac{x-30}{4.1}\right)\right], x \in I\!R\,.$$

   (a) De modo a averiguar a adequação deste modelo o metereologista pretende recolher n observações diárias dessa variável e organizá-las em k classes com frequência esperada comum e igual a $E_i = n/k$.
   Prove que nesta situação a estatística do teste de ajustamento do qui-quadrado, $\sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$, se simplifica e pode escrever-se do seguinte modo: $\frac{k}{n}\sum_{i=1}^{k} O_i^2 - n$. (0.5)

   (b) O metereologista organizou, de acordo com o referido na alínea anterior, um conjunto de 100 observações na seguinte tabela de frequências absolutas observadas (2.0)

Classe	$(-\infty, 23.85]$	$(23.85, 27.25]$	$(27.25, 29.64]$	$(29.64, 31.95]$	$(31.95, +\infty)$
Frequência	23	24	19	20	14

   Teste, ao nível de significância de 5%, a adequação do modelo referido. Calcule um intervalo para o valor$-p$ e comente.

Grupo IV **4.0 valores**

Um astrónomo resolveu estudar a relação entre a distância e a velocidade de recessão entre nebulosas. Com esse objectivo registou para 24 nebulosas as distâncias a partir da terra (x, em *megaparsecs*) e as respectivas velocidades de recessão (y, em Km/s), tendo obtido:

$$\sum_{i=1}^{24} x_i = 21.873,\ \sum_{i=1}^{24} y_i = 8955,\ \sum_{i=1}^{24} x_i^2 = 29.5178,\ \sum_{i=1}^{24} y_i^2 = 6511425,\ \sum_{i=1}^{24} x_i y_i = 12513.7,$$

com mínimo$(x_1, \ldots, x_{24}) = 0.032$ e o máximo$(x_1, \ldots, x_{24}) = 2$.

Nota: 1 *parsec* = 3.26 *anos luz*.

Assuma que o modelo de regressão linear simples $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ é apropriado.

(a) Obtenha um intervalo de confiança a 90% para o declive da recta de regressão e use esse resultado para verificar se parece ser significativo a existência de uma relação linear entre a velocidade de recessão e a distância. Enuncie as hipóteses de trabalho. (2.0)

(b) Obtenha um intervalo de confiança a 90% para o valor esperado da velocidade de recessão de uma nebulosa a uma distância da terra de 0.55 *megaparsecs*. (1.0)

(c) Calcule o coeficiente de determinação. Comente o valor obtido. (1.0)