Departamento de Matemática

Probabilidades e Estatística

MEC + LEGM + LET + LEIC-A

Segundo teste — 1º semestre – 2010/2011

Duração: 90 minutos — 08/01/2011 – 11:00

▼ Justifique convenientemente todas as respostas ▼

Grupo I — 5 valores

1. Considere que o tempo que um servidor demora a responder a pedidos dos clientes, em milisegundos (ms), é uma variável aleatória com distribuição exponencial de parâmetro $\lambda \in \mathbb{R}^+$.

(a) Com base numa amostra aleatória de tamanho n, $X_1, \ldots, X_n$, obtenha o estimador de máxima verosimilhança do parâmetro λ. (1.0)

$\mathcal{L}(\lambda; x_1, \ldots, x_n) = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$, $\frac{d \log \mathcal{L}}{d\lambda} = 0 \Leftrightarrow \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0 \Leftrightarrow \lambda = \bar{x}^{-1}$ e $\frac{d^2 \log \mathcal{L}}{d\lambda^2} = -\frac{n}{\lambda^2} < 0,\ \forall \lambda > 0, \forall n \in \mathbb{N}$.

$\therefore \hat{\lambda}_{MV} = \bar{X}^{-1}$.

(b) Compare a eficiência de $\bar{X}$ com a do estimador $T = \frac{X_1 + (n-1)X_n}{n}$ na estimação de $1/\lambda$, para $n > 2$. (1.5)

$E[\bar{X}] = E[T] = 1/\lambda$, $Var[\bar{X}] = 1/n\lambda^2$, $Var[T] = \dfrac{(n-1)^2+1}{(n\lambda)^2}$.

$e(T, \bar{X}) = \frac{(n-1)^2+1}{n} > 1,\ \forall n > 2 \Rightarrow \bar{X}$ é mais eficiente que T na estimação de $E(X) = 1/\lambda$.

Numa amostra casual de 40 observações obteve-se um tempo total de 1737.18 ms.

(c) Calcule a estimativa de máxima verosimilhança da percentagem de pedidos que são respondidos em menos de 100 ms. (1.0)

Sendo $\phi = P(X < 100) = 1 - e^{-100\lambda}$, tem-se $\hat{\phi}_{MV} = 1 - e^{-100\bar{X}^{-1}}$, pela invariância dos estimadores de máxima verosimilhança. Na amostra observada tem-se $\bar{x} = 1737.18/40 \approx 43.43$ e logo $\hat{\phi}_{MV} = 1 - e^{-2.303} \approx 0.9$.

(d) Determine um intervalo de confiança para λ a um nível de confiança aproximado de 96%. (1.5)

Pelo T.L.C. tem-se que $Z = \frac{\bar{X} - E(\bar{X})}{\sqrt{Var(\bar{X})}} = \frac{\bar{X} - E(X)}{\sqrt{Var(X)/n}} = \frac{\bar{X} - 1/\lambda}{\sqrt{1/n\lambda^2}} = \sqrt{n}(\lambda\bar{X} - 1) \overset{a}{\sim} N(0,1)$.

$P(-2.0537 < Z < 2.0537) \approx 0.96 \Leftrightarrow P\left(\bar{X}^{-1}\left(1 - \frac{2.0537}{\sqrt{n}}\right) < \lambda < \bar{X}^{-1}\left(1 + \frac{2.0537}{\sqrt{n}}\right)\right) \approx 0.96$.

$\therefore IC_{\approx 0.96}(\lambda) =]0.016, 0.031[$.

Grupo II — 5 valores

1. Com o intuito de avaliar a solidez de um subsolo rochoso para efeitos de construção, recolheram-se *in situ* amostras de rocha a partir das quais se prepararam provetes cilíndricos por meio de uma máquina de corte. O uso diferenciado desta máquina conduziu a provetes de dois tipos distintos, rotulados de A e B. Tais provetes são submetidos a um teste em laboratório para a determinação da sua resistência à compressão X (em kNm^{-2}). Estudos prévios mostraram que esta variável aleatória pode ser modelada por uma distribuição Normal com valor médio denotado por μ_A ou μ_B, consoante o tipo de provete, e com um desvio padrão de 11 kNm^{-2} para qualquer tipo de provete.

Uma 1ª série de ensaios realizados sobre 10 provetes do tipo A originou valores de X sumariados em $\sum_{i=1}^{10} x_{iA} = 606.0$ kNm^{-2}. Uma 2ª série de ensaios em 6 provetes do tipo B, independente da primeira, conduziu ao valor $\sum_{i=1}^{6} x_{iB} = 343.8$ kNm^{-2}.

(a) Teste a conjectura de igualdade dos valores médios da resistência à compressão dos dois tipos de provetes com base nos dados indicados. (1.5)

Pretende-se testar $H_0 : \mu_A = \mu_B$ contra $H_1 : \mu_A \neq \mu_B$. Uma vez que as variâncias de ambas as populações são conhecidas, recorre-se à variável fulcral $Z = \frac{\bar{X}_A - \bar{X}_B - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \sim N(0,1)$. A estatística de teste é então $Z_0 = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{121}{10} + \frac{121}{6}}} \overset{H_0}{\sim} N(0,1)$ e $z_0 = \frac{60.6 - 57.3}{5.68} \approx 0.581$.
Como se trata de um teste bilateral o valor-p é dado por $P(|Z_0| > 0.581) = 2P(Z_0 > 0.581) = 2(1 - \Phi(0.581)) = 2(1 - 0.7194) \approx 0.5612$. Uma vez que este valor está acima da gama de valores habituais para o nível de significância de um teste de hipóteses, não se deve rejeitar H_0, ou seja, os dados não fornecem evidência significativa para se rejeitar a conjectura de igualdade dos valores médios da resistência à compressão dos dois tipos de provetes.

(b) Qual a probabilidade de o teste de hipóteses anterior conduzir à não rejeição incorrecta da hipótese nula ao nível de significância de 2% quando se sabe que $\mu_A = \mu_B + 4$? (1.0)

Para um nível de significância de 2%, H_0 deve ser rejeitada se $|Z_0| > \Phi^{-1}(0.99) = 2.326$. Quando $\mu_A - \mu_B = 4$ tem-se que $Z^* = \frac{\bar{X}_A - \bar{X}_B - 4}{5.68} \sim N(0,1)$. Então, $P(\text{Não rejeitar } H_0 \mid \mu_A - \mu_B = 4) = P(|Z_0| \leq 2.326 \mid Z^* \sim N(0,1)) = P(-3.030 \leq Z^* \leq 1.622) = \Phi(1.622) - \Phi(-3.030) \approx 0.95$.

2. O número diário de avarias registadas num dado equipamento fabril é uma variável aleatória cuja distribuição se pode caracterizar pelas seguintes probabilidades: $P(X = 0) = 0.7$, $P(X = 1) = 0.2$ e $P(X \geq 2) = 0.1$. (2.5)

Para se analisar a adequabilidade deste modelo seleccionaram-se ao acaso 200 dias de laboração da fábrica nos quais se verificou terem ocorrido 2 ou mais avarias por dia em 30 dias, uma avaria por dia em 45 dias e nenhuma nos restantes dias.

Será que a conclusão a extrair destes dados sobre a adequabilidade do modelo conjecturado varia com o valor que se está disposto a tolerar para a probabilidade de se rejeitar incorrectamente tal conjectura?

Considerem-se as classes $C_1 = \{0\}$, $C_2 = \{1\}$ e $C_3 = \{2, \ldots\}$ e seja $p_i = P(X \in C_i)$, $i = 1, 2, 3$. Pretende-se testar a hipótese $H_0 : p_1 = 0.7,\ p_2 = 0.2,\ p_3 = 0.1$, recorrendo à estatística de Pearson

$$Q_0 = \sum_{i=1}^{3} \frac{(O_i - E_i)^2}{E_i} \overset{a}{\sim} \chi^2_{(2)}.$$

Classes	O_i	$E_i = np_i^0$	$(O_i - E_i)^2/E_i$
C_1	125	140	1.607
C_2	45	40	0.625
C_3	30	20	5.000
	$n = 200$	200	$q_0 = 7.232$

Valor-$p = P(Q_0 \geq 7.232) = 1 - F_{\chi^2_{(2)}}(7.232) \approx 1 - 0.973 = 0.027$.
A hipótese H_0 deve ser rejeitada para $\alpha > 0.027$ e não deve ser rejeitada no caso contrário. Como o valor-p se encontra no intervalo de valores usuais para o nível de significância de um teste de hipóteses, a decisão vai naturalmente depender do nível de significância que for fixado.