2º teste (9/6/2011, código A)

Grupo I

1. seja Y = 'v.a. nº de pacotes de rede que chegam a um router num dado intervalo de tempo'

$Y \sim Poi(\lambda)$; $P(Y=y) = \begin{cases} \frac{e^{-\lambda}\lambda^y}{y!} & , y \in \mathbb{N}_0, \lambda > 0 \\ 0 & , c.c. \end{cases}$

a) seja $\underset{\sim}{Y} = (Y_1, \dots, Y_n)$ a.a. de dimensão n de Y

Dedução do Estimador de M.V. de λ

. função de verosimilhança duma amostra $\underset{\sim}{y} = (y_1, \dots, y_n)$

$$L(\lambda|\underset{\sim}{y}) \underset{i.i.d.}{=} \prod_{i=1}^{n} P(Y=y_i) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{y_i}}{y_i!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_i!} \quad \text{se } \forall y_i \in \mathbb{N}_0, \lambda > 0$$

Maximizar $L(\lambda|\underset{\sim}{y}) \Leftrightarrow$ Maximizar $\ln L(\lambda|\underset{\sim}{y})$

$$\ln L(\lambda|\underset{\sim}{y}) = -n\lambda + \sum_{i=1}^{n} y_i \ln\lambda - \sum_{i=1}^{n} \ln y_i!$$

$$\left.\frac{d \ln L(\lambda|\underset{\sim}{y})}{d\lambda}\right|_{\lambda=\hat{\lambda}} = 0 \Leftrightarrow -n + \frac{\sum_{i=1}^{n} y_i}{\hat{\lambda}} = 0 \Leftrightarrow -n\hat{\lambda} + \sum_{i=1}^{n} y_i = 0 \Leftrightarrow \hat{\lambda} = \frac{\sum_{i=1}^{n} y_i}{n} = \bar{y}$$

verificar q̄ $\hat{\lambda}$ é máximo de $\ln L(\lambda|\underset{\sim}{y})$

$$\left.\frac{d^2 \ln L(\lambda|\underset{\sim}{y})}{d\lambda^2}\right|_{\lambda=\hat{\lambda}} = -\frac{\sum_{i=1}^{n} y_i}{\hat{\lambda}^2} < 0 \quad \forall \hat{\lambda} > 0 \quad \therefore \hat{\lambda} = \bar{y} \text{ é a estimativa de M.V. de } \lambda$$

O Estimador de M.V. de λ é $\hat{\lambda} = \bar{Y}$

b) $\underset{\sim}{y} = (y_1, \dots, y_{100})$: $\sum_{i=1}^{100} y_i = 657 \Rightarrow \bar{y} = 6.57$

Estimativa de M.V. da $P(Y \geq 1) = \hat{P}(Y \geq 1) = ?$

$$\hat{P}(Y \geq 1) = 1 - \hat{P}(Y < 1) = 1 - \hat{P}(Y=0) \underset{\downarrow}{=} 1 - e^{-\hat{\lambda}} = 1 - e^{-6.57} \simeq 0.9986$$

pelo Princ. da Invariância

2. Seja X = 'duração da lâmpada fluorescente do televisor LCD (mh)'

$X \sim N(\mu; 1)$

Amostra $\underset{\sim}{x} = (x_1, \dots, x_{10})$: $\bar{x} = 12$

a) Pretende-se testar:

$H_0: \mu = 15$ vs $H_1: \mu \neq 15$ Ao nível de significância $\alpha = 1\%$.

v. fulcral: $T = \dfrac{\bar{X} - \mu}{\frac{1}{\sqrt{10}}} \sim N(0,1)$

Sob a validade de H_0 tem-se a estatística de teste

$$T|H_0 = T_0 = \frac{\bar{X} - 15}{\frac{1}{\sqrt{10}}} \sim N(0,1)$$

Para o nível $\alpha = 1\%$ rejeitamos H_0 se

$$|T_0| > z_{0.995} = \Phi^{-1}(0.995) \simeq 2.5758$$

$$R.R_{T_0} =]-\infty, -2.5758[\cup]2.5758, +\infty[$$

Valor observado da estatística de teste: $t_0 = \dfrac{12-15}{\frac{1}{\sqrt{10}}} \simeq -9.487$

Decisão: Como $t_0 \in R.R.$, para $\alpha = 1\%$, rejeitamos H_0, ie, com base nesta amostra parece que a duração esperada das lâmpadas não é igual a 15 mh.

b) $P(\text{'rej } H_0' \mid \mu = 13) = ?$

rej H_0 se $|T_0| > 2.5758 \Leftrightarrow \left|\dfrac{\bar{X} - 15}{\frac{1}{\sqrt{10}}}\right| > 2.5758 \Leftrightarrow$

$$\Leftrightarrow \bar{X} < -\frac{2.5758}{\sqrt{10}} + 15 \vee \bar{X} > \frac{2.5758}{\sqrt{10}} + 10 \Leftrightarrow \bar{X} < 14.1855 \vee \bar{X} > 15.8145$$

$$R.R_{\bar{X}} =]-\infty, 14.1855[\cup]15.8145; +\infty[$$

Assim $P((\bar{X} < 14.1855 \vee \bar{X} > 15.8145) / \mu = 13) \underset{\text{mut. excl.}}{=} \Phi\left(\dfrac{14.1855 - 13}{\frac{1}{\sqrt{10}}}\right) + 1 - \Phi\left(\dfrac{15.8145 - 13}{\frac{1}{\sqrt{10}}}\right)$

$$\simeq \Phi(3.75) + 1 - \Phi(8.90) \simeq 0.999912 + 1 - 1 \simeq 0.999912$$

Grupo II

1. Seja X='comprimento do ovo de cuco (em mm)'

$$F_0(x)=\begin{cases}0 & , x<0\\ \left(\frac{x}{25}\right)^2 & , 0\le x<25\\ 1 & , x\geq 25\end{cases} \qquad H_0: F_X(x)=F_0(x)\ , x\in\mathbb{R}$$

$\underset{\sim}{x}=(x_1,\dots,x_{100})$ amostra que foi organizada em 4 classes, $C_i, i=1,\dots,4$

a) Calculo de E_3 e E_4

$E_i=E(O_i|H_0)=np_i^0$; onde O_i='v.a. frequência da classe C_i' e
$p_i^0=P(X\in C_i|H_0)$

Assim, $E_3=100\times p_3^0$ e $p_3^0=P(X\in C_3|H_0)=P(25\sqrt{0.5}<X\leq 25\sqrt{0.75})=$
$=F_0(25\sqrt{0.75})-F_0(25\sqrt{0.5})=0.75-0.5=0.25$

$\therefore E_3=25$

$E_4=100\times p_4^0$ e $p_4^0=P(X\in C_4|H_0)=P(25\sqrt{0.75}<X\leq 25)=$
$=F_0(25)-F_0(25\sqrt{0.75})=1-0.75=0.25$

$\therefore E_4=25$

b) Pretende-se testar $H_0: X\sim F_0(x)$ vs $H_1: X\nsim F_0(x)$

Sob a validade de H_0 tem-se a estatistica de teste: $T_0=\sum_{i=1}^{4}\frac{(O_i-E_i)^2}{E_i}\underset{a}{\sim}\chi^2_{(4-\beta-1)}$ porque $\forall E_i>5$, $\beta=0$, $\equiv\chi^2_{(3)}$

Calculo do valor observado da estatistica de teste

$$t_0=\frac{(26-25)^2}{25}+\frac{(23-25)^2}{25}+\frac{(30-25)^2}{25}+\frac{(21-25)^2}{25}=1.84$$

Rejeita-se H_0 se $T_0>c$. Assim, o valor-p $=P$ do teste é:

$P=P(T_0>1.84)=1-F_{\chi^2_{(3)}}(1.84)=$

$\approx$ { 0.6063, ie, não rejeitamos H_0 para $\alpha<60.63\%$. (máq. calculas)

0.6 $<1-F_{\chi^2_{(3)}}(1.84)<0.7$, ie, $\alpha\geq 70\%$ rej. H_0 e para $\alpha\leq 60\%$ ñ rej. H_0 (tabelas)

Assim, a conjectura de que $F_X(x)=F_0(x)$ não é rejeitada aos niveis de significância usuais.

2. Seja x = 'percentagem de gordura ingerida'
Y = 'Aumento do perímetro abdominal (cm)'

Amostra de dimensão 10: $\sum_{i=1}^{10} x_i = 764$; $\sum_{i=1}^{10} x_i^2 = 593.86$; $\sum_{i=1}^{10} y_i = 29.3$; $\sum_{i=1}^{10} y_i^2 = 87.71$

$\sum_{i=1}^{10} x_i y_i = 227.85$

Modelo de R.L.S.: $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \ (i=1,\dots,10)$, com $E(\varepsilon_i)=0$, $var(\varepsilon_i)=\sigma^2$, $\forall i$
$cov(\varepsilon_i, \varepsilon_j) = 0$, $\forall i \neq j$

a) Pretende-se $I.C._{95\%}(\beta_1) = ?$

v. fulcral: $T = \dfrac{\hat{\beta}_1 - \beta_1}{\sqrt{\dfrac{\hat{\sigma}^2}{\sum_{i=1}^{10} x_i^2 - 10\bar{x}^2}}} \sim t_{(8)}$

$P(a \leq T \leq b) = 0.95$, escolhendo o I.C. simétrico tem-se que

$a = -b = t_{0.975(8)} = F^{-1}_{t_{(8)}}(0.975) = 2.306$

Assim, $P\left(-2.306 \leq \dfrac{\hat{\beta}_1 - \beta_1}{\underbrace{\sqrt{\dfrac{\hat{\sigma}^2}{\sum_i x_i^2 - 10\bar{x}^2}}}_{\to D}} \leq 2.306\right) = 0.95$

Sendo o I.C. Aleatório da forma:

$P(\hat{\beta}_1 - 2.306 \times D \leq \beta_1 \leq \hat{\beta}_1 + 2.306 \times D) = 0.95$

$I.C.A_{95\%}(\beta_1) = [\hat{\beta}_1 \pm 2.306 \times D]$

concretização do I.C.A

$\hat{\beta}_1 = \dfrac{\sum_{i=1}^{10} x_i y_i - 10\bar{x}\bar{y}}{\sum_{i=1}^{10} x_i^2 - 10\bar{x}^2} \simeq 0.393349$

$\hat{\sigma}^2 = \frac{1}{8}\left[\sum_{i=1}^{10} y_i^2 - 10\bar{y}^2 - \hat{\beta}_1^2\left(\sum_{i=1}^{10} x_i^2 - 10\bar{x}^2\right)\right] \simeq 0.036049$; $d \simeq 0.059554$

$\therefore IC_{95\%}(\beta_1) = [0.2560; 0.5307]$

b) Se $\beta_1 = 0 \Rightarrow E[Y|x] = \beta_0 + \beta_1 x = \beta_0$ (modelo não é adequado, a regressão não é significativa)

Pretende-se testar $H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$, relacionando com $I.C._{95\%}(\beta_1)$

Uma vez que $0 \notin IC_{95\%}(\beta_1)$, para $\alpha = 5\%$, rejeitamos H_0, ie, com base nesta amostra parece que este modelo de R.L.S é adequado a estes dados.