2º teste (9/6/2011, código B)

Grupo I

1. seja X = 'v.a. nº de fluxos de tráfego num intervalo de tempo'

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

a) seja $\underset{\sim}{X} = (X_1, \ldots, X_n)$ a.a. de dimensão n de X

Dedução do Estimador de M.V. de λ

- função de verosimilhança duma amostra $\underset{\sim}{x} = (x_1, \ldots, x_n)$

$$L(\lambda|\underset{\sim}{x}) \underset{\text{i.i.d.}}{=} \prod_{i=1}^{n} P(X=x_i) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} \quad \text{se } \forall x_i > 0, \lambda > 0$$

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

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

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

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

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

o Estimador de M.V. de λ é $\hat{\lambda} = \bar{X}$.

b) $\underset{\sim}{x} = (x_1, \ldots, x_{200})$: $\sum_{i=1}^{200} x_i = 1314 \Rightarrow \bar{x} = 6.57$

Estimativa de M.V. da $P(X=0) = \hat{P}(X=0) = ?$

$$\hat{P}(X=0) \underset{\uparrow}{=} e^{-\hat{\lambda}} = e^{-6.57} \simeq 0.00144$$

pela prop. da Invariância

2. Seja $X =$ 'tempo de vida da bateria de lítio (mh)'

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

Amostra $\underset{\sim}{x} = (x_1, \dots, x_{10})$: $\bar{x} = 24$

a) Pretende-se testar:

$H_0: \mu = 27$ vs $H_1: \mu \neq 27$ ao nível de significância $\alpha = 5\%$

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

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

$T|_{H_0} = T_0 = \dfrac{\bar{X} - 27}{\frac{1}{\sqrt{10}}} \sim N(0,1)$

Para o nível de significância $\alpha = 5\%$ rejeitamos Ho se

$|T_0| > z_{0.975} = \Phi^{-1}(0.975) = 1.96$

$R.R_{T_0} =]-\infty, -1.96[\cup]1.96; +\infty[$

valor observado da estatística de teste : $t_0 = \dfrac{24-27}{\frac{1}{\sqrt{10}}} \simeq -9.487$

Decisão: Como $t_0 \in R.R.$, para $\alpha = 5\%$, rejeitamos Ho, ie, com base nesta amostra parece que a duração esperada destas baterias não é igual a 27mh.

b) $P(\text{'rej Ho'} / \mu = 25) = ?$

Rej. Ho se $|T_0| > 1.96 \iff \left| \dfrac{\bar{X} - 27}{\frac{1}{\sqrt{10}}} \right| > 1.96 \iff \bar{X} < -\dfrac{1.96}{\sqrt{10}} + 27 \ \vee\ \bar{X} > \dfrac{1.96}{\sqrt{10}} + 27 \iff$

$\bar{X} < 26.380 \ \vee\ \bar{X} > 27.620$

$RR_{\bar{X}} =]-\infty, 26.380[\cup]27.620, +\infty[$

Assim, $P((\bar{X} < 26.380 \vee \bar{X} > 27.620) | \mu = 25) \underset{\text{mut. excl.}}{=} \Phi\left(\dfrac{26.380 - 25}{\frac{1}{\sqrt{10}}}\right) + 1 - \Phi\left(\dfrac{27.620 - 25}{\frac{1}{\sqrt{10}}}\right) \simeq$

$\simeq \Phi(4.36) + 1 - \Phi(8.29) \simeq 1$

Grupo II

1. Seja X = 'comprimento do ovo de tartaruga (cm)'

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

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

a) Calculo de E_1 e E_3

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

Assim, $E_1 = 100\times p_1^0$ e $p_1^0 = P(X\in C_1|H_0) = P(X\leq 2.5|H_0) = G_X(2.5) =$
$= \frac{(2.5)^2}{25} = 0.25$

$\therefore E_1 = 25$

$E_3 = 100\times p_3^0$ e $p_3^0 = P(X\in C_3|H_0) = P(5/\sqrt{2}<X\leq\sqrt{75/4}) =$
$= G_X\left(\sqrt{\frac{75}{4}}\right) - G_X\left(\frac{5}{\sqrt{2}}\right) = 0.75 - 0.5 = 0.25$

$\therefore E_3 = 25$

b) Pretende-se testar $H_0: X\sim G_X(x)$ vs $H_1: X\not\sim G_X(x)$

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

Calculo do valor observado da estatística

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

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

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

$$= \begin{cases} 0.6063, \text{ ie, não rejeitamos } H_0 \text{ para } \alpha<60{,}63\% & \text{(máq. calcular)} \\ 0.6<1-F_{\chi^2_{(3)}}(1.84)<0.7, \text{ ie, para } \alpha\geq 70\% \text{ rej } H_0 \text{ e para } \alpha\leq 60\% \text{ ñ rej } H_0 & \text{(tabelas)} \end{cases}$$

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

2. Seja x = 'percentagem de gordura alimentar'
Y = 'v.a. aumento do perímetro abdominal (cms)'

Amostra de dimensão 30: $\sum_{i=1}^{30} x_i = 229.2$; $\sum_{i=1}^{30} x_i^2 = 1781.58$; $\sum_{i=1}^{30} y_i = 87.90$; $\sum_{i=1}^{30} y_i^2 = 263.13$

$\sum_{i=1}^{30} x_i y_i = 683.55$

Modelo de R.L.S.: $Y = \beta_0 + \beta_1 x + \varepsilon$, com $E(\varepsilon) = 0$, $var(\varepsilon) = \sigma^2$ e $cov(\varepsilon_i, \varepsilon_j) = 0 \quad \forall i \neq j$; $i,j = 1, \ldots, 30$

a) Pretende-se $I.C._{99\%}(\beta_1) = ?$

v. fulcral: $T = \dfrac{\hat{\beta}_1 - \beta_1}{\sqrt{\dfrac{\hat{\sigma}^2}{\sum_i x_i^2 - 30\bar{x}^2}}} \sim t_{(28)}$

$P(a \leq T \leq b) = 0{,}99$, escolhendo o I.C. simétrico tem que:

$a = -b = t_{0.995(28)} = F^{-1}_{t_{(28)}}(0.995) = 2.763$

Assim, $P\left(-2.763 \leq \dfrac{\hat{\beta}_1 - \beta_1}{\underbrace{\sqrt{\dfrac{\hat{\sigma}^2}{\sum_i x_i^2 - 30\bar{x}^2}}}_{\to D}} \leq 2.763\right) = 0{,}99$

Sendo o I.C.A. da forma:

$P(\hat{\beta}_1 - 2.763 \times D \leq \beta_1 \leq \hat{\beta}_1 + 2.763 \times D) = 0.99$

$I.C.A._{99\%}(\beta_1) = [\hat{\beta}_1 \pm 2.763 \times D]$; concretização do I.C.A.:

$\hat{\beta}_1 = \dfrac{\sum_{i=1}^{30} x_i y_i - 30\bar{x}\bar{y}}{\sum_{i=1}^{30} x_i^2 - 30\bar{x}^2} \simeq 0{,}393349$; $d \simeq 0{,}031833$

$\hat{\sigma}^2 = \dfrac{1}{28}\left[\sum_{i=1}^{30} y_i^2 - 30\bar{y}^2 - \hat{\beta}_1^2\left(\sum_i x_i^2 - 30\bar{x}^2\right)\right] \simeq 0.0309$

$\therefore \; I.C._{99\%}(\beta_1) = [0{,}3054\,;\,0{,}4813]$

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 $IC_{99\%}(\beta_1)$.

Uma vez que $0 \notin IC_{99\%}(\beta_1)$, para $\alpha = 1\%$, rejeitamos H_0, ie, com base nesta amostra parece que este modelo de RLS é adequado a estes dados.