# Probabilidades e Estatística

**LEAN, LEE, LEIC-A, LEIC-T, LEMat, MEBiol, MEBiom, MEEC, MEFT, MEMec, MEQ**

1º semestre – 2016/2017
12/01/2017 – **9:00**

Duração: 90 minutos **2º teste**

**Justifique convenientemente todas as respostas!**

**Grupo I** 10 valores

**1.** Um fabricante garante que o desvio (em milímetros), em relação à norma, das peças que produz é uma variável aleatória $X$ com distribuição normal de valor esperado $\mu$ e variância unitária. Ao analisar uma amostra aleatória de 10 peças por ele fornecidas, verificou-se que $\sum_{i=1}^{10} x_i = 0.2$ onde $x_i$ designa o desvio (em milímetros) da peça $i = 1,\ldots,10$.

(a) Determine a estimativa de máxima verosimilhança de $\mu$. (3.0)

$\mathscr{L}(\mu; x_1,\ldots,x_{10}) \overset{iid}{=} \prod_{i=1}^{10} f_X(x_i;\lambda) = \prod_{i=1}^{10} (2\pi)^{-1/2} e^{-\frac{(x_i-\mu)^2}{2}} = (2\pi)^{-10/2} e^{-\frac{1}{2}\sum_{i=1}^{10}(x_i-\mu)^2}$

$\log\mathscr{L}(\mu; x_1,\ldots,x_{10}) = -5\log(2\pi) - \frac{1}{2}\sum_{i=1}^{10}(x_i-\mu)^2$ (diferenciável em ordem a $\mu$ em $I\!R$)

$\frac{d\log\mathscr{L}(\mu;x_1,\ldots,x_{10})}{d\mu} = 0 \iff \sum_{i=1}^{10}(x_i-\mu) = 0 \iff \sum_{i=1}^{10} x_i - 10\mu \iff \mu = \bar{x}$

$\frac{d^2\log\mathscr{L}(\mu;x_1,\ldots,x_{10})}{d\mu^2} = -10 < 0,\ \forall \mu \in I\!R.$

$\therefore \hat{\mu}_{MV} = \bar{X}.$

Com a amostra observada temos a estimativa $\hat{\mu}_{MV} = 0.02$.

(b) Determine o erro quadrático médio do estimador $\bar{X}$ do parâmetro $\mu$. (2.0)

$E[\bar{X}] = E\left[\frac{\sum_{i=1}^n X_i}{n}\right] = \frac{\sum_{i=1}^n E[X_i]}{n} = E[X] = \mu$

$Var[\bar{X}] = Var\left[\frac{\sum_{i=1}^n X_i}{n}\right] = \frac{\sum_{i=1}^n Var[X_i]}{n^2} = \frac{Var[X]}{n} = \frac{1}{n}$

$EQM[\bar{X}] = Var[\bar{X}] + \left(E[\bar{X}]-\mu\right)^2 = \frac{1}{n} = 0.1.$

**2.** Para comparar dois tipos de materiais cerâmicos usados em pavimentos, instalaram-se mosaicos do tipo I e do tipo II, num corredor de uma universidade. Sejam $X_1$ e $X_2$ as variáveis aleatórias que medem (numa escala conveniente) os desgastes ao fim de dois anos em mosaicos dos tipos I e II, respetivamente. Tendo sido selecionados ao acaso 81 mosaicos do tipo I e 71 do tipo II e medidos os respetivos desgastes ao fim de 2 anos, obteve-se: $\bar{x}_1 = 290$, $\bar{x}_2 = 321$, $s_1 = 12$ e $s_2 = 14$. Assuma que $X_1$ e $X_2$ são variáveis aleatórias com distribuição normal.

(a) Construa um intervalo de confiança a 90% para o desgaste médio de mosaicos do tipo I. (2.5)

Seja $T = \frac{\bar{X}_1-\mu_1}{\sqrt{\frac{S_1^2}{n}}} \sim t_{(n-1)}$.

e $a = F^{-1}_{t_{(80)}}(0.95) = 1.664$

$P(-a \le T \le a) = 0.95 \iff P\left(\bar{X}_1 - a\sqrt{\frac{S_1^2}{80}} \le \mu_1 \le \bar{X}_1 + a\sqrt{\frac{S_1^2}{80}}\right) = 0.90$

$\text{IAC}_{0.90}(\mu_1) = \left[\bar{X}_1 - a\sqrt{\frac{S_1^2}{80}}, \bar{X}_1 + a\sqrt{\frac{S_1^2}{80}}\right]$

$\text{IC}_{0.90}(\mu_1) = [287.78, 292.22]$

(b) Teste, ao nível de significância de 1%, a hipótese de igualdade entre os desgastes médios dos dois tipos de pavimento. Admita que as variâncias dos desgastes em ambos os casos são iguais. (2.5)

Pretende-se testar $H_0: \mu_1 - \mu_2 = 0$ contra $H_1: \mu_1 - \mu_2 \neq 0$.

Seja $Z = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim t_{(n_1+n_2-2)}$. Sob $H_0$ obtemos a estatística do teste,

$Z_0 = \frac{(\bar{X}_1-\bar{X}_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim t_{(n_1+n_2-2)}.$

Para $\alpha = 0.01$ deve rejeitar-se $H_0$ se $|Z_0| > F^{-1}_{t_{(150)}}(0.995) = 2.609$.

Com $z_0 \approx -14.7$ pertence à região de rejeição então $H_0$ é rejeitada para $\alpha = 0.01$.

| **Grupo II** | 10 valores |
|---|---|

**1.** Um processo de laboração contínua causa interrupções na produção. A observação do número de interrupções por semana, $X$, durante o último ano (52 semanas) conduziu aos seguintes resultados:

| Nº interrupções | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Nº semanas | 12 | 13 | 14 | 13 |

Teste a hipótese de $X$ ter uma distribuição uniforme discreta entre 0 e 3. Decida com base no valor-$p$. (4.0)

Pretende-se testar $H_0 : X \sim U(\{0,\ldots,3\})$ contra $H_1 : X \not\sim U(\{0,\ldots,3\})$.
Seja $p_i^0 = P(X = i \mid H_0) = 1/4$, $i = 0,\ldots,3$.

| i | $o_i$ | $p_i^0$ | $e_i = np_i^0$ |
|---|---|---|---|
| 0 | 12 | 0.25 | 13 |
| 1 | 13 | 0.25 | 13 |
| 2 | 14 | 0.25 | 13 |
| 3 | 13 | 0.25 | 13 |
| | $n = 52$ | | |

Como todas as classes têm uma frequência esperada superior a 5 não é necessário agrupar classes ($k = 4$) e, não havendo qualquer parâmetro estimado ($\beta = 0$), a estatística de teste é $Q_0 = \sum_{i=0}^{3} \frac{(O_i - E_i)^2}{E_i} \underset{H_0}{\overset{a}{\sim}} \chi^2_{(3)}$.
Tem-se $q_0 \approx 0.154$ e valor$-p = P(Q_0 > q_0 \mid H_0) = 1 - F_{\chi^2_{(3)}}(0.154) \approx 0.985$. Deve-se rejeitar $H_0$ para níveis de significância $\geq 0.985$ e não rejeitar no caso contrário. Para os níveis de significância usuais os dados não fornecem evidência para a rejeição de $H_0$.

**2.** Num ensaio clínico aplicaram-se a 10 pacientes determinadas doses de medicação antialérgica. O registo do tempo de ação do fármaco, $Y$, em função da dose de medicação, $x$, conduziu aos seguintes resultados:

| $x_i$ (em mg) | 3.0 | 3.0 | 4.0 | 5.0 | 6.0 | 6.0 | 7.0 | 8.0 | 8.0 | 9.0 |
|---|---|---|---|---|---|---|---|---|---|---|
| $y_i$ (em horas) | 9.1 | 5.5 | 12.3 | 9.2 | 14.2 | 16.8 | 22.0 | 18.3 | 24.5 | 22.7 |

$\sum_{i=1}^{10} x_i = 59.0$, $\sum_{i=1}^{10} x_i^2 = 389.00$, $\sum_{i=1}^{10} y_i = 154.0$, $\sum_{i=1}^{10} y_i^2 = 2\,767.30$, $\sum_{i=1}^{10} x_i y_i = 1\,025.70$.

(a) Obtenha as estimativas dos mínimos quadrados dos coeficientes da reta de regressão. (2.0)

$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = 2.8632$ e $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = -1.4923$

(b) Após ter enunciado as hipóteses de trabalho que entender mais convenientes, construa um intervalo de confiança a 95% para o declive da recta de regressão. (3.0)

Admite-se que as variáveis $Y_i = Y \mid x = x_i$ são não correlacionadas e que $Y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$ para $i = 1,\ldots,10$.
Sejam $T = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2 - 10\bar{x}^2}}} \sim t_{(8)}$ e $a = F^{-1}_{t_{(8)}}(0.975) = 2.306$

$P(-a \leq T \leq a) = 0.95 \iff P\left(\hat{\beta}_1 - a\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2 - 10\bar{x}^2}} \leq \beta_1 \leq \hat{\beta}_1 + a\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2 - 10\bar{x}^2}}\right) = 0.95$

$\mathrm{IAC}_{0.95}(\beta_1) = \left[\hat{\beta}_1 - 2.306\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2 - 10\bar{x}^2}}, \hat{\beta}_1 + 2.306\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2 - 10\bar{x}^2}}\right]$
$\mathrm{IC}_{0.95}(\beta_1) = [1.8721, 3.8541]$

(c) Estime o valor esperado do tempo de ação do fármaco para uma dose de 7.0 mg. (1.0)

Pretende-se estimar $E[Y|x = 7.0] = \beta_0 + 7.0\beta_1$.
$\hat{E}[Y|x = 7.0] = \hat{\beta}_0 + 7\hat{\beta}_1 = 18.554$