Probabilidades e Estatística

LEAN, LEE, LEGI, LEMat, LETI, LMAC, MEAmb, MEAer, MEBiol, MEBiom, MEEC, MEFT, MEQ

DM DEPARTAMENTO DE MATEMÁTICA TÉCNICO LISBOA

2º semestre – 2015/2016
09/06/2016 – **09:00**

Duração: 90 minutos **2º teste A**

Justifique convenientemente todas as respostas!

Grupo I	10 valores

1. Tem-se assumido que o impacto hidrodinâmico, X, do casco de um navio sobre uma onda em determinada região do globo possui distribuição exponencial de parâmetro λ desconhecido, com valores medidos em unidades apropriadas. Tendo por base uma amostra aleatória $(X_1, X_2, \ldots, X_{10})$ de X:

(a) Determine a estimativa de máxima verosimilhança do parâmetro λ para uma realização $(x_1, x_2, \ldots, x_{10})$ da amostra aleatória tal que $\sum_{i=1}^{10} x_i = 19.39$. (3.0)

- **V.a. de interesse**
  X = impacto hidrodinâmico do casco de um navio sobre uma onda
- **Distribuição**
  $X \sim \text{exponencial}(\lambda)$
- **F.d.p. de X**
  $$f_X(x) = \begin{cases} \lambda\, e^{-\lambda x}, & x > 0 \\ 0, & \text{c.c.} \end{cases}$$
- **Parâmetro desconhecido**
  $\lambda, \quad \lambda > 0$
- **Amostra**
  $\underline{x} = (x_1, \ldots, x_n)$ amostra de dimensão n proveniente da população X
- **Obtenção da estimativa de MV de λ**

  Passo 1 — Função de verosimilhança
  $$\begin{aligned} L(\lambda|\underline{x}) &= f_{\underline{X}}(\underline{x}) \overset{X_i\, indep}{=} \prod_{i=1}^{n} f_{X_i}(x_i) \overset{X_i \sim X}{=} \prod_{i=1}^{n} f_X(x_i) \\ &= \prod_{i=1}^{n} \left(\lambda\, e^{-\lambda x_i}\right) \\ &= \lambda^n\, e^{-\lambda \sum_{i=1}^{n} x_i}, \quad \lambda > 0 \end{aligned}$$

  Passo 2 — Função de log-verosimilhança
  $$\ln L(\lambda|\underline{x}) = n \ln(\lambda) - \lambda \sum_{i=1}^{n} x_i$$

  Passo 3 — Maximização
  A estimativa de MV de λ é doravante representada por $\hat{\lambda}$ e atente-se que $\hat{\lambda} = argmax_\lambda L(\lambda|\underline{x}) = argmax_\lambda \ln L(\lambda|\underline{x})$. Neste caso em particular:
  $$\hat{\lambda} \quad : \quad \begin{cases} \left.\dfrac{d \ln L(\lambda|\underline{x})}{d\lambda}\right|_{\lambda=\hat{\lambda}} = 0 & \text{(ponto de estacionaridade)} \\ \left.\dfrac{d^2 \ln L(\lambda|\underline{x})}{d\lambda^2}\right|_{\lambda=\hat{\lambda}} < 0 & \text{(ponto de máximo)} \end{cases}$$
  $$\begin{cases} \frac{n}{\hat{\lambda}} - \sum_{i=1}^{n} x_i = 0 \\ -\frac{n}{\hat{\lambda}^2} < 0 \end{cases}$$
  $$\begin{cases} \hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} \\ \text{Proposição verdadeira já que } \sum_{i=1}^{n} x_i > 0. \end{cases}$$

- **Passo 4 — Estimativa de MV de λ**

$$\begin{aligned}\hat{\lambda} &= \frac{n}{\sum_{i=1}^{n} x_i}\\ &\overset{n=10}{=} \frac{10}{19.39}\\ &\simeq 0.515729.\end{aligned}$$

(b) Tendo em vista a estimação do valor esperado de X, compare a eficiência do estimador $\bar{X}$ relativamente ao estimador $T = \frac{X_1+X_{10}}{2}$. (1.5)

- **Parâmetro desconhecido**
$\mu = E(X)$

- **Estimador de $\mu = E(X)$**
$\bar{X}$

- **Erro quadrático médio de $\bar{X}$**

$$\begin{aligned}EQM_{\mu}(\bar{X}) &= V(\bar{X}) + \left[bias_{\mu}(\bar{X})\right]^2\\ &= V(\bar{X}) + \left[E(\bar{X}) - \mu\right]^2\\ &\overset{X_i \underset{i.i.d.}{\sim} X}{=} \frac{V(X)}{n} + [E(X) - E(X)]^2\\ &= \frac{V(X)}{n} \qquad \text{[onde } V(X) = 1/\lambda^2 \text{ e } n = 10.]\end{aligned}$$

- **Outro estimador de $\mu = E(X)$**
$T = \frac{X_1+X_{10}}{2}$

- **Erro quadrático médio de T**

$$\begin{aligned}EQM_{\mu}(T) &= V(T) + \left[bias_{\mu}(T)\right]^2\\ &= V(T) + \left[E(T) - \mu\right]^2\\ &= V\left(\frac{X_1+X_{10}}{2}\right) + \left[E\left(\frac{X_1+X_{10}}{2}\right) - E(X)\right]^2\\ &\overset{X_i \underset{i.i.d.}{\sim} X}{=} \frac{2\,V(X)}{4} + \left[\frac{2\,E(X)}{2} - E(X)\right]^2\\ &= \frac{V(X)}{2}\end{aligned}$$

- **Eficiência do estimador $\bar{X}$ relativamente ao estimador $T = \frac{X_1+X_{10}}{2}$**

$$\begin{aligned}e_{\mu}(\bar{X}, T) &= \frac{EQM_{\mu}(T)}{EQM_{\mu}(\bar{X})}\\ &\overset{n=10}{=} \frac{\frac{V(X)}{2}}{\frac{V(X)}{10}}\\ &= 5\end{aligned}$$

- **Comentário**
Tendo em conta que $e_{\mu}(\bar{X}, T) = 5 > 1$ (i.e., $EQM_{\mu}(T) > EQM_{\mu}(\bar{X})$) pode afirmar-se que $\bar{X}$ é um estimador mais eficiente que $T = \frac{X_1+X_{10}}{2}$ no que respeita à estimação de $\mu = E(X)$.

2. Para estudar o consumo de combustível em automóveis do modelo 1 (resp. 2), considerou-se a variável aleatória X_1 (resp. X_2) denotando a quilometragem efetuada por litro (km/litro) de combustível por um automóvel do modelo 1 (resp. 2) escolhido ao acaso. Tendo selecionado ao acaso 8 automóveis do modelo 1 e 9 automóveis do modelo 2 e registado as respetivas quilometragens efetuadas por litro de combustível, observaram-se os seguintes resultados:

$$\sum_{i=1}^{8} x_{1i} = 194.7, \qquad \sum_{i=1}^{8} x_{1i}^2 = 4\,743.69, \qquad \sum_{i=1}^{9} x_{2i} = 280.9, \qquad \sum_{i=1}^{9} x_{2i}^2 = 8\,954.45$$

Assumindo que as quilometragens efetuadas por litro de combustível em automóveis dos modelos 1 e 2 possuem distribuição normal com igual variância:

(a) Obtenha um intervalo de confiança a 90% para a variância da kilometragem efetuada por litro de combustível em automóveis do modelo 2. (2.5)

- **V.a. de interesse**
  X_2 = kilometragem efetuada por litro de combustível em automóveis do modelo 2
- **Situação**
  $X_2 \sim \text{normal}(\mu_2, \sigma_2^2)$
  μ_2 desconhecido
  σ_2^2 DESCONHECIDO
- **Obtenção do IC para σ_2^2**
  Passo 1 — Selecção da v.a. fulcral para σ_2^2

$$Z = \frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2_{(n_2-1)}$$

uma vez que é suposto determinar um IC para a variância de uma população normal, com valor esperado desconhecido.

Passo 2 — Obtenção dos quantis de probabilidade
Ao ter-se em consideração que $n_2 = 9$ e $(1-\alpha) \times 100\% = 90\%$, far-se-á uso dos quantis

$$(a_\alpha, b_\alpha) \quad : \quad \begin{cases} P(a_\alpha \leq Z \leq b_\alpha) = 1-\alpha \\ P(Z < a_\alpha) = P(Z > b_\alpha) = \alpha/2. \end{cases}$$

$$\begin{cases} a_\alpha = F^{-1}_{\chi^2_{(n_2-1)}}(\alpha/2) = F^{-1}_{\chi^2_{(8)}}(0.05) \stackrel{tabela/calc.}{=} 2.733 \\ b_\alpha = F^{-1}_{\chi^2_{(n_2-1)}}(1-\alpha/2) = F^{-1}_{\chi^2_{(8)}}(0.95) \stackrel{tabela/calc.}{=} 15.51. \end{cases}$$

Passo 3 — Inversão da desigualdade $a_\alpha \leq Z \leq b_\alpha$

$$P(a_\alpha \leq Z \leq b_\alpha) = 1-\alpha$$

$$P\left[a_\alpha \leq \tfrac{(n_2-1)S_2^2}{\sigma_2^2} \leq b_\alpha\right] = 1-\alpha$$

$$P\left[\tfrac{1}{b_\alpha} \leq \tfrac{\sigma_2^2}{(n_2-1)S_2^2} \leq \tfrac{1}{a_\alpha}\right] = 1-\alpha$$

$$P\left[\tfrac{(n_2-1)S^2}{b_\alpha} \leq \sigma_2^2 \leq \tfrac{(n_2-1)S_2^2}{a_\alpha}\right] = 1-\alpha$$

Passo 4 — Concretização
Atendendo ao par de quantis acima e ao facto de

$$\begin{aligned} s_2^2 &= \frac{1}{n_2-1}\left[\sum_{i=1}^{n_2} x_{2i}^2 - n_2\,(\bar{x}_2)^2\right] \\ &= \frac{1}{9-1}\left[8954.45 - 9 \times (280.9/9)^2\right] \\ &= 23.406(1) \end{aligned}$$

$$IC_{(1-\alpha)\times 100\%}(\sigma_2^2) = \left[\frac{(n_2-1)\,s_2^2}{F^{-1}_{\chi^2_{(n_2-1)}}(1-\alpha/2)}, \frac{(n_2-1)\,s_2^2}{F^{-1}_{\chi^2_{(n_2-1)}}(\alpha/2)}\right],$$

segue-se:

$$\begin{aligned} IC_{90\%}(\sigma_2^2) &= \left[\frac{(9-1) \times 23.406(1)}{15.51}, \frac{(9-1) \times 23.406(1)}{2.733}\right] \\ &\simeq [12.0728, 68.5149]. \end{aligned}$$

(b) Teste ao nível de significância de 10% a hipótese de igualdade dos valores esperados das (3.0) quilometragens efetuadas por litro de combustível em automóveis dos modelos 1 e 2.

- **V.a. de interesse**
  X_i = quilometragens efetuadas por litro de combustível em automóveis do modelo i, $i = 1,2$
- **Situação**
  $X_1 \sim \text{Normal}(\mu_1, \sigma_1^2) \perp\!\!\!\perp X_2 \sim \text{Normal}(\mu_1, \sigma_2^2)$
  $(\mu_1 - \mu_2)$ DESCONHECIDO
  σ_1^2 e σ_2^2 desconhecidos, no entanto, assume-se que são IGUAIS: $\sigma_1^2 = \sigma_2^2 = \sigma^2$
  $n_1 = 8 \leq 30$ ou $n_1 = 9 \leq 30$
- **Hipóteses**
  $H_0 : \mu_1 - \mu_2 = \mu_0 = 0$
  $H_1 : \mu_1 - \mu_2 \neq \mu_0$
- **Nível de significância**
  $\alpha_0 = 10\%$
- **Estatística de teste**

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_0}{\sqrt{\frac{(n_1-1)\,S_1^2 + (n_2-1)\,S_2^2}{n_1+n_2-2} \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim_{H_0} t_{(n_1+n_2-2)}$$

  dado que se pretende efectuar um teste sobre a diferença de valores esperados de duas populações normais independentes, com variâncias desconhecidas mas que se assume serem iguais.
- **Região de rejeição de** H_0 (para valores da estatística de teste)
  Estamos a lidar com um teste bilateral ($H_1 : \mu_1 - \mu_2 \neq \mu_0$), logo a região de rejeição de H_0 (para valores da estatística de teste) é $W = (-\infty, -c) \cup (c, +\infty)$, onde $c : P(\text{Rejeitar } H_0 | H_0) = \alpha_0$, i.e.,

$$\begin{aligned} c \quad : \quad & P(T \in W \mid H_0) = \alpha_0 \\ & 2 \times \left[1 - F_{t_{(n_1+n_2-2)}}(c)\right] = \alpha_0 \\ & c = F^{-1}_{t_{(n_1+n_2-2)}}(1 - \alpha_0/2) \\ & c = F^{-1}_{t_{(15)}}(0.95) \\ & c \overset{tabela/calc.}{=} 1.753. \end{aligned}$$

- **Decisão**
  Atendendo a que
  - $n_1 = 8$
    $\bar{x}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} x_{1i} = \frac{194.7}{8} = 24.3375$
    $s_1^2 = \frac{1}{n_1-1}\left[\left(\sum_{i=1}^{n_1} x_{1i}^2\right) - n_1\,(\bar{x}_1)^2\right] = \frac{1}{8-1}\left(4743.69 - 8 \times 24.3375^2\right) = \frac{5.17875}{7} = 0.7398$
  - $n_2 = 9$
    $\bar{x}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} x_{2i} = \frac{280.9}{9} = 31.2(1)$
    $s_2^2 \overset{a}{=} 23.406(1)$,

  o valor observado da estatística de teste é igual a

$$\begin{aligned} t \quad & = \quad \frac{(\bar{x}_1 - \bar{x}_2) - \mu_0}{\sqrt{\frac{(n_1-1)\,s_1^2 + (n_2-1)\,s_2^2}{n_1+n_2-2} \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \\ & = \quad \frac{(24.3375 - 31.2(1)) - 0}{\sqrt{\frac{(8-1)\times 0.7398 + (9-1)\times 23.406(1)}{8+9-2}\left(\frac{1}{8} + \frac{1}{9}\right)}} \\ & \simeq \quad -3.9491. \end{aligned}$$

  Como $t \simeq -3.9491 \in W = (-\infty, -1.753) \cup (1.753, +\infty)$, devemos rejeitar H_0 ao n.s. $\alpha_0 = 10\%$ [ou a qualquer n.s. superior a $\alpha_0 = 10\%$].

Grupo II	10 valores

1. Numa dada eleição para a Presidência da República concorrem 3 candidatos: A, B e C. Suponha que dos 2000 eleitores inquiridos numa sondagem aleatória: 1000 são apoiantes do candidato A, 600 preferem o candidato B e 400 preferem o candidato C. (4.0)

Vários analistas sustentam que, entre os eleitores, a base de apoio do candidato A é dupla da base de apoio do candidato B e tripla da base de apoio do candidato C. Averigúe, aplicando um teste apropriado, se a opinião dos analistas é consistente com os resultados da sondagem. Decida com base no valor-p.

- **V.a. de interesse e f.p.**
  X = candidato apoiado pelo leitor inquirido
  $$p_i = \begin{cases} P(X=i), & i = A,B,C \\ 0, & \text{caso contrário} \end{cases}$$

- **Hipóteses**
  $H_0 : p_i = p_i^0 \ (i = A,B,C)$ onde
  $$\begin{cases} p_A^0 = 2p_B^0 = 3p_C^0 \\ p_A^0 + \frac{p_A^0}{2} + \frac{p_A^0}{3} = 1 \end{cases} \Leftrightarrow \begin{cases} - \\ \frac{(6+3+2)\,p_A^0}{6} = 1 \end{cases} \Leftrightarrow \begin{cases} p_A^0 = \frac{6}{11} \\ p_B^0 = \frac{3}{11} \\ p_C^0 = \frac{2}{11} \end{cases}$$
  $H_1 : p_i \neq p_i^0$, para algum i

- **Estatística de teste**
  $$T = \sum_{i=A,B,C} \frac{(O_i - E_i)^2}{E_i} \ \overset{a}{\sim}_{H_0} \ \chi^2_{(k-\beta-1)},$$
  onde:

  k = No. de classes = 3 (candidadtos)

  O_i = Frequência absoluta observável da classe i

  E_i = Frequência absoluta esperada, sob H_0, da classe i

  β = No. de parâmetros a estimar = 0 [dado que a distribuição conjecturada em H_0 está completamente especificada, i.e., H_0 é uma hipótese simples.]

- **Região de rejeição de** H_0 (para valores de T)
  Ao efectuarmos um teste de ajustamento do qui-quadrado a região de rejeição de H_0 é um intervalo à direita $W = (c, +\infty)$.

- **Cálculo das frequências absolutas esperadas sob** H_0
  As frequências absolutas esperadas sob H_0 são dadas por $E_i = n \times p_i^0$ $(i = A,B,C)$ e iguais a

  $E_A = (1000 + 600 + 400) \times \frac{6}{11} = 1090.(90)$

  $E_B = 2000 \times \frac{3}{11} = 545.(45)$

  $E_C = 2000 \times \frac{3}{11} = 363.(63)$.

  [Importa notar que não é necessário fazer qualquer agrupamento de classes uma vez que em pelo menos 80% das classes se verifica $E_i \geq 5$ e $E_i \geq 1$ para todo o i.]

- **Decisão (com base no valor-p)**
  No cálculo do valor observado da estatística de teste convém recorrer à seguinte tabela auxiliar:

i	Freq. abs. obs. o_i	Freq. abs. esper. sob H_0 $E_i = n \times p_i^0$	Parcelas valor obs. estat. teste $\frac{(o_i-E_i)^2}{E_i}$
A	1000	1090.(90)	$\frac{[1000-1090.(90)]^2}{1090.(90)} = 7.(57)$
B	600	545.(45)	$\frac{[600-545.(45)]^2}{545.(45)} = 5.(45)$
C	400	363.(63)	$\frac{[400-363.(63)]^2}{363.(63)} = 3.(63)$
	$\sum_{i=A,B,C} o_i = n = 2000$	$\sum_{i=A,B,C} E_i = n = 2000$	$t = \sum_{i=A,B,C} \frac{(o_i-E_i)^2}{E_i} = 16.(6)$

Assim, temos

$$\begin{aligned} t &= \sum_{i=A,B,C} \frac{(o_i - E_i)^2}{E_i} \\ &= 16.(6). \end{aligned}$$

Uma vez que a região de rejeição de H_0 é para este teste um intervalo à direita temos:

$$\begin{aligned} valor - p &= P(T > t \mid H_0) \\ &= P[T > 16.(6) \mid H_0] \\ &\simeq 1 - F_{\chi^2_{(3-1-0)}}[16.(6)] \\ &\overset{calc.}{\simeq} 0.00024. \end{aligned}$$

Consequentemente, é suposto:

- não rejeitar H_0 a qualquer n.s. $\alpha_0 \leq 0.024\%$;
- rejeitar H_0 a qualquer n.s. $\alpha_0 > 0.024\%$, pelo que a opinião dos analistas não é consistente com os dados a qualquer dos níveis usuais de significância (1%, 5% e 10%).

[Em alternativa, poderíamos recorrer às tabelas de quantis da distribuição do qui-quadrado com 2 graus de liberdade e adiantar um intervalo para o *p-value*:

$$\begin{aligned} F^{-1}_{\chi^2_{(2)}}(0.9995) = 15.20 &< t = 16.(6) \\ 0.9995 &< F_{\chi^2_{(2)}}[16.(6)] \\ 1 - F_{\chi^2_{(2)}}[16.(6)] &< 1 - 0.9995 \\ valor - p &< 0.0005. \end{aligned}$$

Logo o intervalo para o *valor-p* é $(0, 0.0005)$ e devemos:

- rejeitar H_0 a qualquer n.s. $\alpha_0 \geq 0.05\%$, pelo que a opinião dos analistas não é consistente com os dados a qualquer dos níveis usuais de significância (1%, 5% e 10%).]

2. Para descrever a relação existente entre o volume de uma massa de um gás ideal clássico e a respetiva pressão, registaram-se 10 valores do logaritmo de base 10 do volume, x (com o volume medido em polegadas ao quadrado), e os correspondentes valores experimentais do logaritmo de base 10 da pressão, Y (com a pressão medida em psi). Pretendendo avaliar-se a validade do modelo de regressão linear simples para descrever a relação existente entre o logaritmo da pressão do gás e o logaritmo do seu volume, efetuaram-se os seguintes cálculos:

$$\sum_{i=1}^{10} x_i = 19.4, \quad \sum_{i=1}^{10} x_i^2 = 38.06, \quad \sum_{i=1}^{10} y_i = 14.8, \quad \sum_{i=1}^{10} y_i^2 = 22.76, \quad \sum_{i=1}^{10} x_i y_i = 28.12$$

(a) Obtenha as estimativas de mínimos quadrados dos parâmetros da recta de regressão linear simples de Y em x e interprete o significado do sinal da estimativa do parâmetro β_1 do modelo. (2.0)

- **Estimativas de β_0 e β_1**
  Dado que
  - $n = 10$
  - $\sum_{i=1}^{n} x_i = 19.4$
    $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{19.4}{10} = 1.94$
    $\sum_{i=1}^{n} x_i^2 = 38.06$
    $\sum_{i=1}^{n} x_i^2 - n\,(\bar{x})^2 = 38.06 - 10 \times 1.94^2 = 0.424$
  - $\sum_{i=1}^{n} y_i = 14.8$
    $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{14.8}{10} = 1.48$
    $\sum_{i=1}^{n} y_i^2 = 22.76$
    $\sum_{i=1}^{n} y_i^2 - n\,(\bar{y})^2 = 22.76 - 10 \times 1.48^2 = 0.856$
  - $\sum_{i=1}^{n} x_i y_i = 28.12$
    $\sum_{i=1}^{n} x_i y_i - n\,\bar{x}\,\bar{y} = 28.12 - 10 \times 1.94 \times 1.48 = -0.592,$

  as estimativas de β_1 e β_0 são, para este modelo de RLS, iguais a:

$$\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_{i=1}^{n} x_i y_i - n\,\bar{x}\,\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\,(\bar{x})^2} \\
&= \frac{-0.592}{0.424} \\
&\simeq -1.396226 \\
\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \times \bar{x} \\
&\simeq 1.48 - (-1.396226) \times 1.94 \\
&= 4.188678
\end{aligned}$$

- **Interpretação do sinal da estimativa de β_1**
  $\hat{\beta}_1 \simeq -1.396226$
  Como o sinal de $\hat{\beta}_1$ é negativo, espera-se que um aumento no logaritmo de base 10 do volume do gás provoque uma DIMINUIÇÃO no valor esperado do logaritmo de base 10 da pressão.

(b) Indicando as hipóteses de trabalho convenientes, obtenha um intervalo de confiança a 95% para o parâmetro β_1 do modelo de regressão linear simples de Y em x. O que pode concluir sobre a significância do modelo de regressão ao nível de significância de 5%? (4.0)

- **Hipóteses de trabalho**
  $\epsilon_i \stackrel{i.i.d.}{\sim} \text{Normal}(0,\sigma^2),\ i = 1,\ldots,n$
  β_0, β_1 e σ^2 DESCONHECIDOS

- **Obtenção do IC para β_1**
  Passo 1 — Selecção da v.a. fulcral para β_1

$$Z = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2}}} \sim t_{(n-2)}$$

  Passo 2 — Obtenção dos quantis de probabilidade
  Neste caso $n = 10$ e $(1-\alpha) \times 100\% = 95\%$, logo usaremos os quantis de probabilidade

$$(a_\alpha, b_\alpha) \quad : \quad \begin{cases} P(a_\alpha \le Z \le b_\alpha) = 1-\alpha \\ P(Z < a_\alpha) = P(Z > b_\alpha) = \alpha/2. \end{cases}$$

$$\begin{cases} a_\alpha = -F^{-1}_{t_{(n-2)}}(1-\alpha/2) = -F^{-1}_{t_{(8)}}(0.975) \stackrel{tabela/calc.}{=} -2.306 \\ b_\alpha = F^{-1}_{t_{(n-2)}}(1-\alpha/2) = F^{-1}_{t_{(8)}}(0.975) \stackrel{tabela/calc.}{=} 2.306. \end{cases}$$

Passo 3 — Inversão da desigualdade $a_\alpha \leq Z \leq b_\alpha$

$P\,(a_\alpha \leq Z \leq b_\alpha) = 1-\alpha$

...

$$P\left[\hat\beta_1 - F^{-1}_{t_{(n-2)}}(1-\alpha/2)\times\sqrt{\frac{\hat\sigma^2}{\sum_{i=1}^n x_i^2 - n\,\bar x^2}} \leq \beta_1 \leq \hat\beta_1 + F^{-1}_{t_{(n-2)}}(1-\alpha/2)\times\sqrt{\frac{\hat\sigma^2}{\sum_{i=1}^n x_i^2 - n\,\bar x^2}}\right] = 1-\alpha.$$

Passo 4 — Concretização

Atente-se que

$$\begin{aligned}
\hat\sigma^2 &= \frac{1}{n-2}\left[\left(\sum_{i=1}^n y_i^2 - n\,\bar y^2\right) - \left(\hat\beta_1\right)^2\left(\sum_{i=1}^n x_i^2 - n\,\bar x^2\right)\right] \\
&= \frac{1}{10-2}\left[0.856 - (-1.396226)^2 \times 0.424\right] \\
&= 0.003679 \\
IC_{(1-\alpha)\times 100\%}(\beta_1) &= \left[\hat\beta_1 \pm F^{-1}_{t_{(n-2)}}(1-\alpha/2)\times\sqrt{\frac{\hat\sigma^2}{\sum_{i=1}^n x_i^2 - n\,\bar x^2}}\right].
\end{aligned}$$

Logo

$$\begin{aligned}
IC_{95\%}(\beta_1) &\simeq \left[-1.396226 \pm 2.306 \times \sqrt{\frac{0.003679}{0.424}}\right] \\
&\simeq [-1.396226 \pm 0.214812] \\
&= [-1.611038, -1.181414].
\end{aligned}$$

- **Teste de significância do modelo de RLS**

  Hipóteses

  $H_0 : \beta_1 = \beta_{1,0} = 0$

  $H_1 : \beta_1 \neq \beta_{1,0}$

  N.s.

  $\alpha_0 = 0.05$

  Decisão

  Invocando a relação entre intervalos de confiança e testes de hipóteses, podemos adiantar que:

  – o valor conjecturado para β_1 em H_0 é

  $$\begin{aligned}
  \beta_{1,0} &= 0 \\
  &\not\in IC_{95\%}(\beta_1) = [-1.611038, -1.181414];
  \end{aligned}$$

  – assim sendo, a hipótese $H_0 : \beta_1 = \beta_{1,0} = 0$ deve ser rejeitada ao nível de significância $\alpha = 5\%$ [ou a qualquer n.s. $\alpha_0 > 5\%$].