Probabilidade e Estatística

LEBiol, LEIC-A, LEQ; LEIC-T

2º Semestre – 2023/2024
02/07/2024
13:00–15:00

Duração: **120** minutos

Exame Época Recurso – A

- As questões de escolha múltipla têm uma e apenas uma resposta correta. Caso assinale uma resposta errada ou mais de uma resposta, serão descontados 0.5 valores.
- As respostas às questões de escolha múltipla têm de ser assinaladas na folha de instruções previamente distribuída.
- Para as restantes questões, terá de justificar convenientemente as suas respostas.
- Só pode sair da sala uma hora após o início do exame, devendo nesse caso entregar a sua prova ou desistir da mesma.

Pergunta 1 — 2 valores

Uma fábrica produz três tipos de produtos, dos quais 30% são do tipo A, 50% do tipo B e 20% do tipo C. Alguns destes produtos têm defeitos, tendo-se verificado que 5% (respetivamente, 3%) dos produtos do tipo A (respectivamente, B) têm defeitos. Da produção total desta fábrica, 3.5% dos produtos têm defeitos.

Qual é a probabilidade de um produto do tipo C ter defeitos?

- **Quadro de acontecimentos e probabilidades**

Acontecimento	Probabilidade
A = produto ser do tipo A	$P(A) = 0.3$
B = produto ser do tipo B	$P(B) = 0.5$
C = produto ser do tipo C	$P(C) = 0.2$
D = produto ter defeitos	$P(D) = 0.035$
	$P(D \mid A) = 0.05$
	$P(D \mid B) = 0.03$

- **Prob. pedida**

  Ao aplicar a lei das probabilidades totais, obtemos

$$
\begin{aligned}
P(D) &= P(D \mid A) \times P(A) + P(D \mid B) \times P(B) + P(D \mid C) \times P(C) \\
0.035 &= 0.05 \times 0.3 + 0.03 \times 0.5 + P(D \mid C) \times 0.20 \\
P(D \mid C) &= \frac{0.035 - 0.05 \times 0.30 - 0.03 \times 0.50}{0.20} = 0.025
\end{aligned}
$$

Pergunta 2 — 2 valores

Uma engenheira dispõe de k ($k > 1$) amostras de material biológico entre as quais uma está contaminada e por identificar. Seja X o número total de tentativas até à detecção da amostra contaminada, admitindo que a seleção de amostras é efetuada aleatoriamente e com reposição.

Qual é o valor de k tendo em conta que $E(X^2) = 45$?

A: 4 **B**: 5 **C**: 6 **D**: 7

- **V.a. de interesse; distribuição**

  X = no. total de tentativas até à detecção da amostra contaminada (selec. casual com reposição)

  $X \sim \text{geométrica}(p = 1/k)$[, pois X representa o número de provas de Bernoulli independentes e com probabilidade de sucesso comum $p = \frac{1 \text{ amostra contaminada}}{k \text{ amostras}}$].

- **Obtenção de k**

$$k>1 \quad : \quad E(X^2)=45 \quad \Leftrightarrow \quad V(X)+E^2(X)=45 \quad \Leftrightarrow \quad \frac{1-p}{p^2}+\left(\frac{1}{p}\right)^2=45 \qquad \text{(ver formulário)}$$

$$\frac{1-1/k}{(1/k)^2}+\left(\frac{1}{1/k}\right)^2=45 \quad \Leftrightarrow \quad k(k-1)+k^2=45$$

$$k=\frac{-(-1)\pm\sqrt{(-1)^2-4\times 2\times(-25)}}{2\times 2}=\frac{1\pm\sqrt{361}}{4}=\frac{1\pm 19}{4}=5 \text{ ou } -4.5 \quad \Leftrightarrow \quad k=5.$$

Pergunta 3	2 valores

O tempo de reparação (X, em horas) de uma peça eletrónica é tal que $\ln(X)$ tem distribuição normal, com valor esperado μ e variância σ^2.

Considerando $\mu=1$ e $\sigma=0.1$, calcule a probabilidade de o tempo de reparação ser inferior a três horas, sabendo que a reparação foi iniciada há mais de duas horas.

- **V.a.**

  $X=$ tempo de reparação da peça eletrónica

- **Distribuição de** $\ln(X)$

  $\ln(X)\sim \text{normal}(\mu,\sigma^2)$

- **Probabilidade pedida**

$$\begin{aligned}
P(X<3\mid X>2) &= \frac{P(X<3,X>2)}{P(X>2)}=\frac{P(2<X<3)}{1-P(X\le 2)}=\frac{P[\ln(2)<\ln(X)<\ln(3)]}{1-P[\ln(X)\le\ln(2)]}\\
&= [\frac{P\left[\frac{\ln(2)-\mu}{\sigma}<Z=\frac{\ln(X)-\mu}{\sigma}<\frac{\ln(3)-\mu}{\sigma}\right]}{1-P\left[Z<\frac{\ln(2)-\mu}{\sigma}\right)}]\\
&\simeq \frac{\Phi(0.99)-\Phi(-3.07)}{1-\Phi(-3.07)}=\frac{\Phi(0.99)-[1-\Phi(3.07)]}{1-[1-\Phi(3.07)]}\overset{tabelas/calc.}{=}\frac{0.8389-(1-0.998930)}{1-(1-0.998930)}\\
&\simeq 0.838727.
\end{aligned}$$

Pergunta 4	2 valores

De acordo com registos de um laboratório de química, que está dividido em duas zonas, A e B, a função de probabilidade conjunta do número mensal de frascos de ácido clorídrico gastos na zona A do laboratório (X) e do número mensal de frascos de ácido clorídrico gastos na zona B do laboratório (Y) é dada pela tabela à direita.

	Y		
X	0	1	2
0	0.01	0.1	0.07
1	0.15	0.3	0.15
2	0.09	0.1	0.03

Determine a variância do número total de frascos de ácido clorídrico gastos mensalmente no laboratório.

- **Par aleatório**

  (X,Y)

  $X=$ número mensal de frascos de ácido clorídrico gastos na zona A do laboratório

  $Y=$ número mensal de frascos de ácido clorídrico gastos na zona B do laboratório

- **V.a. de interesse e variância pedida**

  $T=X+Y=$ número mensal de frascos de ácido clorídrico gastos no laboratório

  $V(T)=V(X+Y)=V(X)+V(Y)+2\,cov(X,Y)$

  Assim sendo, são necessários alguns cálculos auxiliares que envolverão a f.p. conjunta de (X,Y) e as f.p. marginais de X e Y dadas por $P(X=x)=\sum_y P(X=x,Y=y)$ e $P(Y=y)=\sum_x P(X=x,Y=y)$.

X	Y 0	1	2	$P(X=x)$
0	0.01	0.1	0.07	0.18
1	0.15	0.3	0.15	0.60
2	0.09	0.1	0.03	0.22
$P(Y=y)$	0.25	0.5	0.25	1

- **Valor esperado e variância de** X

$$\begin{aligned} E(X) &= \sum_{x=0}^{2} x \times P(X=x) = 1 \times 0.6 + 2 \times 0.22 = 1.04 \\ V(X) &= E(X^2) - E^2(X) = \sum_{x=0}^{2} x^2\, P(X=x) - E^2(X) = \left(1^2 \times 0.6 + 2^2 \times 0.22\right) - 1.04^2 = 1.48 - 1.0816 \\ &= 0.3984 \end{aligned}$$

- **Valor esperado e variância de** Y

$$\begin{aligned} E(Y) &= \sum_{y=0}^{2} x \times P(Y=y) = 1 \times 0.5 + 2 \times 0.25 = 1 \\ V(Y) &= E(Y^2) - E^2(Y) = \sum_{y=0}^{2} y^2\, P(Y=y) - E^2(Y) = \left(1^2 \times 0.5 + 2^2 \times 0.25\right) - 1^2 = 1.5 - 1 = 0.5 \end{aligned}$$

- **Valor esperado de** XY

$$E(XY) = \sum_{x=0}^{2}\sum_{y=0}^{2} xy \times P(X=x, Y=y) = 1 \times 1 \times 0.3 + 1 \times 2 \times 0.15 + 2 \times 1 \times 0.1 + 2 \times 2 \times 0.03 = 0.92$$

- **Covariância entre** X **e** Y

$$cov(X,Y) = E(XY) - E(X) \times E(Y) = 0.92 - 1.04 \times 1 = -0.12$$

- **Variância pedida** (cont.)

$$V(X+Y) = V(X) + V(Y) + 2\,cov(X,Y) = 0.3984 + 0.5 + 2 \times (-0.12) = 0.6584.$$

Pergunta 5 2 valores

O número de erupções vulcânicas anuais na zona da Islândia denominada *Zona Vulcânica Oriental* é uma variável aleatória com distribuição de Poisson com valor esperado 0.173. O número de erupções vulcânicas por ano no resto do país é também uma variável aleatória com distribuição de Poisson, independente da anterior, mas com valor esperado 0.027.

Calcule a probabilidade do número total de erupções vulcânicas na Islândia, durante os próximos 5 anos, ser de pelo menos 3, admitindo também a independência na atividade vulcânica ao longo dos anos.

- **V.a.**

  X_1 = número anual de erupções vulcânicas na Zona Vulcânica Oriental

  $X_1 \sim \text{Poisson}(\lambda)$, onde λ : $E(X_1) = 0.173 \Leftrightarrow \lambda = 0.173$

  Y_1 = número anual de erupções vulcânicas fora da Zona Vulcânica Oriental

  $Y_1 \sim \text{Poisson}(\lambda)$, onde λ : $E(Y_1) = 0.027 \Leftrightarrow \lambda = 0.027$

- **V.a. de interesse**

  X_5 = número de erupções vulcânicas na Zona Vulcânica Oriental em 5 anos

  Y_5 = número de erupções vulcânicas fora da Zona Vulcânica Oriental em 5 anos

  T_5 = número de erupções vulcânicas na Islândia em 5 anos = $X_5 + Y_5$

- **Distribuições**

  Pela propriedade reprodutiva da distribuição de Poisson (admitindo que são válidas as necessárias condições de independência), temos:

$$\begin{aligned} X_5 &\sim \text{Poisson}(0.173\times 5 = 0.865);\\ Y_5 &\sim \text{Poisson}(0.027\times 5 = 0.135);\\ T_5 &= X_5+Y_5 \sim \text{Poisson}(0.865+0.135=1). \end{aligned}$$

- **F.p. de T_5**

  $P(T_5=x)=\frac{e^{-1}\times 1^x}{x!}=\frac{e^{-1}}{x!},\quad x\in\mathbb{N}_0$

- **Probabilidade pedida**

$$P(T_5\geq 3) = 1-P(T_5\leq 2)=1-F_{Poisson(1)}(2)=1-\sum_{x=0}^{2}\frac{e^{-1}}{x!} \overset{tabelas/calc.}{\simeq} 1-0.9197=0.0803.$$

Pergunta 6	2 valores

Numa produção em série, o número de chips consecutivos que passam no controlo de qualidade é uma variável aleatória, X, com função de probabilidade dada por

$$P(X=x)=\exp\left[-\left(\frac{x}{20}\right)^{\alpha}\right]-\exp\left[-\left(\frac{x+1}{20}\right)^{\alpha}\right],\ x=0,1,2,\ldots,$$

onde α é uma constante desconhecida que toma valores em $\{4,5,6\}$. A concretização de uma amostra aleatória de dimensão 4 proveniente de X conduziu a $x_1=22$, $x_2=22$, $x_3=10$ e $x_4=21$.

Obtenha a estimativa de máxima verosimilhança de α.

- **V.a. de interesse**

  X = número de chips consecutivos que passam no controlo de qualidade

- **Parâmetro desconhecido e espaço paramétrico**

  $\alpha,\quad \Theta=\{4,5,6\}$

- **Amostra**

  $\underline{x}=(22,22,10,21)$, amostra de dimensão 4 proveniente da população X.

- **Obtenção da estimativa de MV de α**

  Será representada por $\hat{\alpha}$ e $L(\hat{\alpha}\mid\underline{x})=\max_{\alpha\in\Theta}L(\alpha\mid\underline{x})$, onde $L(\alpha\mid\underline{x})$ representa a f. de verosimilhança.

  Função de verosimilhança

$$\begin{aligned} L(\alpha\mid\underline{x}) &= P(\underline{X}=\underline{x}) \overset{X_i\ indep}{=} \prod_{i=1}^{4}P(X_i=x_i) \overset{X_i\sim X}{=} \prod_{i=1}^{4}P(X=x_i)\\ &= \left[e^{-(22/20)^{\alpha}}-e^{-(23/20)^{\alpha}}\right]^2\times\left[e^{-(10/20)^{\alpha}}-e^{-(11/20)^{\alpha}}\right]\times\left[e^{-(21/20)^{\alpha}}-e^{-(22/20)^{\alpha}}\right] \end{aligned}$$

  Maximização e concretização

  [Como Θ contém apenas três valores, a estimativa de MV de α obtém-se calculando os vários valores de $L(\alpha\mid\underline{x})$, para $\alpha\in\Theta$, e identificando o ponto de máximo — i.e., faz-se por pesquisa ponto por ponto.]

α	$L(\alpha\mid\underline{x})$
4	5.763801×10^{-6}
5	6.321893×10^{-6}
6	5.473500×10^{-6}

A inspecção da tabela anterior leva a concluir que a função de verosimilhança atinge o máximo quando $\alpha = 5$, pelo que $\hat{\alpha} = 5$.

Pergunta 7	2 valores

Uma engenheira considera que o número de avarias por turno de operação, em determinada unidade fabril, é uma variável aleatória X com valor esperado μ e variância σ^2 desconhecidos. A seleção casual de 100 registos de tal número conduziu a uma média e variância corrigida amostrais iguais a $\bar{x} = 1.07$ e $s^2 = 1.0355$.

Indique um intervalo de confiança aproximado a 98% para μ.

A: [0.8291, 1.3109] **B**: [0.8610, 1.2790] **C**: [0.8573, 1.2827] **D**: [0.8333, 1.3067]

- **V.a. de interesse**

  $X =$ número de avarias por turno de operação

- **Situação**

  X com distribuição arbitrária

  $\mu = E(X)$ DESCONHECIDO

  $\sigma^2 = V(X)$ desconhecida

- **Obtenção de IC aproximado para μ**

  Passo 1 - Seleção da v.a. fulcral para μ

  $$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} \stackrel{a}{\sim} \text{normal}(0,1)$$

  Passo 2 — Obtenção dos quantis de probabilidade

  Como $(1-\alpha) \times 100\% = 98\% \Leftrightarrow \alpha = 0.02$, lidaremos com os quantis seguintes:

  $$\begin{cases} a_\alpha = \Phi^{-1}(\alpha/2) = -\Phi^{-1}(1-\alpha/2) = -\Phi^{-1}(0.99) \stackrel{tabela/calc.}{=} -2.3263 \\ b_\alpha = \Phi^{-1}(1-\alpha/2) = \Phi^{-1}(0.99) = 2.3263. \end{cases}$$

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

  $$P(a_\alpha \le Z \le b_\alpha) \simeq 1-\alpha$$

  $$P\left[a_\alpha \le \frac{\bar{X}-\mu}{S/\sqrt{n}} \le b_\alpha\right] \simeq 1-\alpha$$

  $$P\left[\bar{X} - b_\alpha \times S/\sqrt{n} \le \mu \le \bar{X} - a_\alpha \times S/\sqrt{n}\right] \simeq 1-\alpha$$

  Passo 4 — Concretização

  A expressão geral do intervalo aproximado de confiança a $(1-\alpha)\times 100\%$ para μ é

  $$IC_{(1-\alpha)\times 100\%}(\mu) \simeq \left[\bar{x} \pm \Phi^{-1}(1-\alpha/2) \times s/\sqrt{n}\right].$$

  Ao termos em conta que $n = 100$, $\alpha = 0.02$, $\Phi^{-1}(1-\alpha/2) = 2.3263$ $\bar{x} = 1.07$ e $s^2 = 1.0355$,

  $$\begin{aligned} IC_{98\%}(p) &\simeq \left[1.07 - 2.3263 \times \sqrt{1.0355/100},\, 1.07 + 2.3263 \times \sqrt{1.0355/100}\right] \\ &\simeq [0.8333, 1.3067]. \end{aligned}$$

Pergunta 8	2 valores

Um engenheiro químico afirma que as medições (X) associadas a um método para determinação do pH, para soluções neutras, seguem uma distribuição normal com desvio padrão 0.5. Com o objectivo de validar essa informação, realizaram-se 10 medições independentes do pH em soluções neutras, tendo-se obtido os seguintes resultados: $\sum_{i=1}^{10} x_i = 70.4$, $\sum_{i=1}^{10} x_i^2 = 501.76$.

Averigue se os dados suportam a informação anunciada relativamente ao desvio padrão de X, testando $H_0: \sigma = 0.5$ contra $H_1: \sigma \neq 0.5$ ao nível de significância de 10%.

- **V.a. de interesse**

  X = resultado da determinação do pH

- **Situação**

  $X \sim \text{normal}(\mu, \sigma^2)$

  μ desconhecido

  σ DESCONHECIDO

- **Hipóteses**

  $H_0: \sigma = \sigma_0 = 0.5$ vs. $H_1: \sigma \neq \sigma_0$

- **N.s.**

  $\alpha_0 = 10\%$

- **Estatística de teste**

$$\frac{(n-1)\,S^2}{\sigma_0^2} \sim_{H_0} \chi^2_{(n-1)}$$

- **Região de rejeição de H_0 (para valores da estatística de teste)**

  O teste é bilateral ($H_1: \sigma \neq \sigma_0$), logo a região de rejeição de H_0 é do tipo $W = (0, a) \cup (b, +\infty)$, onde

$$\begin{aligned} a &= F^{-1}_{\chi^2_{(n-1)}}(\alpha/2) = F^{-1}_{\chi^2_{(9)}}(0.05) \stackrel{tabelas/calc.}{=} 3.325 \\ b &= F^{-1}_{\chi^2_{(n-1)}}(1-\alpha/2) = F^{-1}_{\chi^2_{(9)}}(0.95) \stackrel{tabelas/calc.}{=} 16.92. \end{aligned}$$

- **Decisão**

  Uma vez que

$$\begin{aligned} \bar{x} &= \frac{\sum_{i=1}^n x_i}{n} = \frac{70.4}{10} = 7.04 \\ s^2 &= \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n-1} = \frac{501.76 - 10 \times 7.04^2}{10-1} = 0.6827, \end{aligned}$$

  o valor observado da estatística de teste é igual a

$$t = \frac{(n-1)\,s^2}{\sigma_0^2} = \frac{(10-1) \times 0.6827}{0.5^2} \simeq 24.58$$

  Como $t \in W = (0, 3.325) \cup (16.92, +\infty)$, deve rejeitar-se H_0 ao nível de significância em causa, $\alpha = 10\%$, ou seja, há evidência para afirmar (a este nível) que os dados não suportam a informação anunciada relativamente ao desvio padrão de X.

Pergunta 9	2 valores

Um engenheiro pretende averiguar se a variação diária da taxa de câmbio entre duas moedas segue uma distribuição de Laplace com função de densidade de probabilidade $f_0(x) = \frac{1}{2} e^{-|x|}$, $x \in \mathbb{R}$ (hipótese H_0).

Calcule o valor-p aproximado (ou um intervalo para o valor-p aproximado) do teste de ajustamento do qui-quadrado, atendendo à tabela de frequências seguinte baseada em 100 registos casuais e da qual constam as absolutas esperadas sob H_0 (arredondadas à unidade):

Classe	$(-\infty,0]$	$(0,0.2]$	$(0.2,0.4]$	$(0.4,0.6]$	$(0.6,+\infty)$
Frequência absoluta observada	42	12	6	5	35
Frequência absoluta esperada sob H_0	50	9	7	6	28

Qual deverá ser a conclusão do engenheiro ao nível de significância de 7.5%?

- **V.a. de interesse**

  X = variação diária da taxa de câmbio entre duas moedas

- **Hipóteses**

  $H_0: f_X(x) = f_0(x) = f_0(x) = \frac{1}{2}e^{-|x|}, \quad x \in \mathbb{R}$

  $H_1: f_X(x) \neq f_0(x), \quad$ para algum $x \in \mathbb{R}$

- **Estatística de teste**

$$T = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \overset{a}{\sim}_{H_0} \chi^2_{(k-1)}$$

  onde: k = número de classes = 5; O_i = frequência absoluta observável da classe i; E_i = frequência absoluta esperada, sob H_0, da classe i.

- **Região de rejeição de H_0** (para valores de T)

  Lidamos com um teste de ajustamento, donde a região de rejeição de H_0 seja o intervalo à direita $W = (c, +\infty)$.

- **Decisão (com base no valor-p aproximado)**

i	Classe i	Freq. abs. obs. o_i	Freq. abs. esp. sob H_0 E_i	Parcelas valor obs. estat. teste $\frac{(o_i-E_i)^2}{E_i}$
1	$(-\infty,0]$	42	50	$\frac{(42-50)^2}{50} \simeq 1.28$
2	$(0,0.2]$	12	9	1.00
3	$(0.2,0.4]$	6	7	0.14
4	$(0.4,0.6]$	5	6	0.17
5	$(0.6,+\infty)$	35	28	1.75
		$\sum_{i=1}^k o_i = n = 100$	$\sum_{i=1}^k E_i = n = 100$	$t = \sum_{i=1}^k \frac{(o_i-E_i)^2}{E_i} \simeq 4.34$

  Atendendo a que a região de rejeição de H_0 é um intervalo à direita num teste de ajustamento do qui-quadrado, temos

$$valor-p \quad = \quad P(T > t =| H_0) \simeq 1 - F_{\chi^2_{(5-1)}}(4.34) \overset{calc.}{\simeq} 0.361943.$$

  Assim sendo, devemos decidir pela:

  - não rejeição de H_0 a qualquer n.s. $\alpha_0 \leq 36.1943\%$, nomeadamente aos n.u.s. de 1%, 5% e 10%, assim como ao n.s. de 7.5%[;
  - rejeição de H_0 a qualquer n.s. $\alpha_0 > 36.1943\%$].

  [Alternativamente, a consulta das tabelas dos quantis de probabilidade da distribuição do qui-quadrado com $k-1=4$ graus de liberdade permitem concluir que

$$F^{-1}_{\chi^2_{(5-1)}}(0.60) = 4.045 \quad < \quad 4.34 < 4.878 = F^{-1}_{\chi^2_{(5-1)}}(0.70)$$
$$0.30 = 1 - 0.70 \quad < \quad 1 - F_{\chi^2_{(5-1)}}(4.34) < 1 - 0.60 = 0.40.$$

  Logo, devemos:

– não rejeitar de H_0 a qualquer n.s. $\alpha_0 \leq 30\%$, nomeadamente aos n.u.s. de 1%, 5% e 10%, bem como ao n.s. de 7.5%[;
– rejeitar H_0 a qualquer n.s. $\alpha_0 \geq 40\%$].

Pergunta 10 2 valores

Num estudo para avaliar o efeito da percentagem de um novo aditivo (x) no número de octanas da gasolina (Y) realizaram-se 6 ensaios em relação aos quais foram calculadas as seguintes quantidades: $\sum_{i=1}^{6} x_i = 21$, $\sum_{i=1}^{6} x_i^2 = 91$, $\sum_{i=1}^{6} y_i = 496.8$, $\sum_{i=1}^{6} y_i^2 = 41149.12$, $\sum_{i=1}^{6} x_i y_i = 1754.3$.

Admitindo a validade do modelo de regressão linear simples, $Y = \beta_0 + \beta_1 x + \epsilon$, e as hipóteses de trabalho convenientes, obtenha um intervalo de confiança a 95% para o parâmetro β_0.

- [**Modelo de RLS e hipóteses de trabalho**

  Y = número de octanas da gasolina (v.a. resposta)

  x = percentagem de um novo aditivo (variável explicativa)

  $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \dots, n \quad \text{com } \epsilon_i \overset{i.i.d.}{\sim} \text{normal}(0, \sigma^2), \quad i = 1, \dots, n$]

- **Obtenção do IC a** 95% **para** β_0

  Passo 1 — V.a. fulcral para β_0

$$Z = \frac{\hat{\beta}_0 - \beta_0}{se(\hat{\beta}_0)} \sim t_{(n-2)}, \quad \text{com} \quad se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2 \times \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2}\right]}$$

  Passo 2 — Quantis de probabilidade

$$\begin{cases} a_\alpha = -b_\alpha \\ b_\alpha = F^{-1}_{t_{(n-2)}}(1-\alpha/2) = F^{-1}_{t_{(6-2)}}(1-0.05/2) \overset{tabela/calc.}{=} 2.776 \end{cases}$$

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

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

$$P\left[a_\alpha \leq \frac{\hat{\beta}_0 - \beta_0}{se(\hat{\beta}_0)} \leq b_\alpha\right] = 1 - \alpha$$

$$P\left[\hat{\beta}_0 - b_\alpha \times se(\hat{\beta}_0) \leq \beta_0 \leq \hat{\beta}_0 - a_\alpha \times se(\hat{\beta}_0)\right] = 1 - \alpha$$

  Passo 4 — Concretização

  Temos

  - $n = 6$
  - $\sum_{i=1}^{n} x_i = 21$

    $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{21}{6} = 3.5$

    $\sum_{i=1}^{n} x_i^2 = 91$

    $\sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2 = 91 - 6 \times 3.5^2 = 17.5$
  - $\sum_{i=1}^{n} y_i = 496.8$,

    $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{496.8}{6} = 82.8$ $\sum_{i=1}^{n} y_i^2 = 41149.12$

    $\sum_{i=1}^{n} y_i^2 - n\,\bar{y}^2 = 41149.12 - 6 \times 82.8^2 = 14.08$
  - $\sum_{i=1}^{n} x_i y_i - n\,\bar{x}\,\bar{y} = 1754.3 - 6 \times 3.5 \times 82.8 = 15.5$

- $\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - n\,\bar{x}\,\bar{y}}{\sum_{i=1}^n x_i^2 - n\,\bar{x}^2} = \frac{15.5}{17.5} \simeq 0.885714$
- $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \times \bar{x} \simeq 82.8 - \frac{15.5}{17.5} \times 3.5 = 79.7$,

bem como

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n-2}\left[\left(\sum_{i=1}^n y_i^2 - n\,\bar{y}^2\right) - \left(\hat{\beta}_1\right)^2\left(\sum_{i=1}^n x_i^2 - n\,\bar{x}^2\right)\right] \\
&\simeq \frac{1}{6-2}\left(14.08 - 0.8857143^2 \times 17.5\right) \\
&\simeq 0.087857.
\end{aligned}$$

Além disso, a expressão geral do $IC_{(1-\alpha)\times 100\%}\left(\beta_0\right)$ é dada por $\left[\hat{\beta}_0 \pm b_\alpha \times se\left(\hat{\beta}_0\right)\right]$. Donde,

$$\begin{aligned}
IC_{95\%}(\beta_0) &= \left[79.7 \pm 2.776 \times \sqrt{0.08785703 \times \left(\frac{1}{6} + \frac{3.5^2}{17.5}\right)}\right] \\
&\simeq [78.934,\ 80.466].
\end{aligned}$$