Probabilidades e Estatística

LEE, LEGI, LENO, LETI, LMAC, MEAer, MEAmbi, MEBiol, MEBiom, MEEC, MEFT, MEMat, MEQ

DM
DEPARTAMENTO DE MATEMÁTICA
TÉCNICO LISBOA

1º semestre – 2018/2019
12/06/2019 – **9:00**

Duração: 90 minutos **2º Teste A**

Justifique convenientemente todas as respostas

Grupo I 10 valores

1. Considere que o número de ovos partidos em cada embalagem de 6 ovos vendida por uma cadeia de supermercados é uma variável aleatória X com distribuição binomial com parâmetros $n = 6$ e $p \in [0,1]$.

(a) Obtenha o estimador de máxima verosimilhança do parâmetro p baseado numa amostra aleatória $(X_1, X_2, \ldots, X_m)$ de X. (3.0)

- **V.a. de interesse / distribuição**
  $X =$ número de ovos partidos em embalagem de 6 ovos
  $X \sim \text{binomial}(6, p)$
- **F.p. de** X
  $P(X = x) = \binom{6}{x}\, p^x\, (1-p)^{6-x}, \quad x = 0, 1, \ldots, 6$
- **Parâmetro desconhecido**
  $p, \quad 0 \le p \le 1$
- **Amostra aleatória**
  $\underline{X} = (X_1, \ldots, X_m)$
  $X_i \stackrel{i.i.d}{\sim} X, \quad i = 1, \ldots, m$
- **Amostra**
  $\underline{x} = (x_1, \ldots, x_m)$ amostra de dimensão m proveniente da população X
- **Obtenção da estimativa de MV de** p

  Passo 1 — Função de verosimilhança

$$\begin{aligned}
L(p \mid \underline{x}) &= P(\underline{X} = \underline{x}) \\
&\stackrel{X_i \stackrel{i.i.d}{\sim} X}{=} \prod_{i=1}^{m} P(X = x_i) \\
&= \prod_{i=1}^{m} \left[\binom{6}{x_i} p^{x_i} (1-p)^{6-x_i} \right] \\
&= \left[\prod_{i=1}^{m} \binom{6}{x_i} \right] p^{\sum_{i=1}^{m} x_i} (1-p)^{\sum_{i=1}^{m}(6-x_i)}, \quad 0 \le p \le 1
\end{aligned}$$

  Passo 2 — Função de log-verosimilhança

$$\ln L(p \mid \underline{x}) = \sum_{i=1}^{m} \ln \left[\binom{6}{x_i} \right] + \ln(p) \sum_{i=1}^{m} x_i + \ln(1-p) \sum_{i=1}^{m} (6 - x_i), \quad 0 < p < 1$$

  Passo 3 — Maximização
  A estimativa de MV de p é doravante representada por $\hat{p}$ e

$$\hat{p} \;:\; \begin{cases} \left. \frac{d \ln L(p|\underline{x})}{dp} \right|_{p=\hat{p}} = 0 & \text{(ponto de estacionaridade)} \\ \left. \frac{d^2 \ln L(p|\underline{x})}{dp^2} \right|_{p=\hat{p}} < 0 & \text{(ponto de máximo)} \end{cases}$$

$$\begin{cases} \frac{\sum_{i=1}^{m} x_i}{\hat{p}} - \frac{\sum_{i=1}^{m}(6-x_i)}{1-\hat{p}} = 0 \\ -\frac{\sum_{i=1}^{m} x_i}{\hat{p}^2} - \frac{\sum_{i=1}^{m}(6-x_i)}{(1-\hat{p})^2} < 0 \end{cases}$$

$$\hat{p} \quad : \quad \begin{cases} (1-\hat{p})\sum_{i=1}^m x_i - \hat{p}\sum_{i=1}^m (6-x_i) = 0 \\ \text{Prop. verdadeira já que} \quad \sum_{i=1}^m x_i \in \{0,1,\ldots,6m\}, \quad \sum_{i=1}^m (6-x_i) \in \{6m, 6m-1,\ldots,0\} \end{cases}$$

$$\begin{cases} \sum_{i=1}^m x_i - 6\,m\,\hat{p} = 0 \\ --- \end{cases}$$

$$\begin{cases} \hat{p} = \frac{\sum_{i=1}^m x_i}{6\,m} \qquad \text{[estimativa também válida se } \sum_{i=1}^m x_i = 0 \text{ ou } \sum_{i=1}^m x_i = 6m] \\ --- \end{cases}$$

- **Passo 4 — Estimador de MV de p**

$$\begin{aligned} EMV(p) &= \frac{\sum_{i=1}^m X_i}{6\,m} \\ &= \frac{\bar{X}}{6}. \end{aligned}$$

(b) Com base numa amostra $(x_1,\ldots,x_{200})$ tal que $\sum_{i=1}^{200} x_i = 20$, calcule a estimativa de máxima verosimilhança da probabilidade de uma embalagem de ovos não conter qualquer ovo partido. (1.5)

- **Estimativa de MV de p**

$$\begin{aligned} \hat{p} &= \frac{\sum_{i=1}^m x_i}{6\,m} \\ &= \frac{20}{6\times 200} \\ &= \frac{1}{60} \\ &= 0.01(6) \end{aligned}$$

- **Outro parâmetro desconhecido**
  $h(p) = P(X=0) = (1-p)^6$

- **Estimativa de MV de $h(p)$**
  Ao invocarmos a propriedade de invariância dos estimadores de máxima verosimilhança, obtemos a estimativa de MV de $h(p)$:

$$\begin{aligned} \widehat{h(p)} &= h(\hat{p}) \\ &= (1-\hat{p})^6 \\ &= [1-0.01(6)]^6 \\ &\simeq 0.904075. \end{aligned}$$

2. Uma fábrica de instrumentos de medida efectua estudos da qualidade da sua produção. De anteriores estudos sabe-se que as leituras efetuadas são realizações independentes de uma variável aleatória X e possuem distribuição normal. As leituras feitas em nove instrumentos selecionados ao acaso conduziram a $\bar{x} = 0.995$ e $s = 0.0085$.

(a) Construa um intervalo de confiança a 95% para o valor esperado da leitura dos instrumentos fabricados. (2.5)

- **V.a. de interesse**
  X = leitura na medição do objecto-padrão

- **Situação**
  $X \sim \text{normal}(\mu, \sigma^2)$
  μ DESCONHECIDO
  σ^2 desconhecido

- **Obtenção do IC para μ**

  Passo 1 — Selecção da v.a. fulcral para μ

  $$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{(n-1)}$$

  [dado que é suposto determinar um IC para o valor esperado de uma população normal, com variância desconhecida.]

  Passo 2 — Obtenção dos quantis de probabilidade
  Ao ter-se em consideração que $(1-\alpha) \times 100\% = 95\%$, far-se-á uso dos quantis

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

  $$\begin{cases} a_\alpha = F^{-1}_{t_{(n-1)}}(\alpha/2) = F^{-1}_{t_{(9-1)}}(0.025) = -F^{-1}_{t_{(8)}}(1-0.025) \stackrel{tabela/calc.}{=} -2.306 \\ b_\alpha = F^{-1}_{t_{(n-1)}}(1-\alpha/2) = F^{-1}_{t_{(9-1)}}(0.975) \stackrel{tabela/calc.}{=} 2.306. \end{cases}$$

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

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

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

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

  Passo 4 — Concretização
  Tendo em conta os quantis acima, as concretizações de $\bar{X}$ e S,

  $$\bar{x} = 0.995$$
  $$s = 0.0085,$$

  e o facto de

  $$IC_{(1-\alpha)\times 100\%}(\mu) = \left[\bar{x} - F^{-1}_{t_{(n-1)}}(1-\alpha/2) \times \frac{s}{\sqrt{n}}, \quad \bar{x} + F^{-1}_{t_{(n-1)}}(1-\alpha/2) \times \frac{s}{\sqrt{n}}\right],$$

  segue-se

  $$\begin{aligned} IC_{95\%}(\mu) &= \left[0.995 - 2.306 \times \frac{0.0085}{\sqrt{9}}, \quad 0.995 + 2.306 \times \frac{0.0085}{\sqrt{9}}\right] \\ &= [0.988466,\ 1.001534]. \end{aligned}$$

- **Comentário**
  Atendendo a que o valor-alvo da leitura, 1, pertence ao $IC_{95\%}(\mu)$, podemos afirmar que os dados não apontam para qualquer desvio sistemático na calibração dos instrumentos.

(b) Avalie se os dados não contrariam a hipótese H_0 de que o desvio-padrão das leituras é igual a 0.008 a favor da hipótese H_1 de que tal desvio-padrão é superior a 0.008. Decida com base no valor-p. (3.0)

- **Hipóteses**
  $H_0 : \sigma = \sigma_0 = 0.008$
  $H_1 : \sigma > \sigma_0$

- **Estatística de teste**

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

  [pois pretendemos efectuar teste sobre o desvio-padrão de população normal, com valor esperado desconhecido.]

- **Região de rejeição de H_0** (para valores de T)
  Tratando-se de um teste unilateral superior ($H_1 : \sigma > \sigma_0$), a região de rejeição de H_0 é do tipo $W = (c, +\infty)$.

  Decisão (com base em intervalo para o valor-p)
  O valor observado da estatística de teste é dado por

$$\begin{aligned} t &= \frac{(n-1)s^2}{\sigma_0^2} \\ &= \frac{(9-1)\times 0.0085^2}{0.008^2} \\ &\simeq 9.031250. \end{aligned}$$

  Dado que a região de rejeição deste teste é um intervalo à direita, temos:

$$\begin{aligned} valor-p &= P(T > t \mid H_0) \\ &= 1 - F_{\chi^2_{(n-1)}}(t) \\ &\simeq 1 - F_{\chi^2_{(8)}}(9.031250). \end{aligned}$$

  Recorrendo às tabelas de quantis da distribuição do qui-quadrado podemos adiantar um intervalo para o valor-p:

$$\begin{array}{ccccc} F^{-1}_{\chi^2_{(8)}}(0.60) = 8.351 & < & 9.031250 & < & 9.524 = F^{-1}_{\chi^2_{(8)}}(0.70) \\ 1-0.7 = 0.3 & < & valor-p = 1 - F_{\chi^2_{(8)}}(9.031250) & < & 0.40 = 1 - 0.60. \end{array}$$

  Consequentemente:
  - não devemos rejeitar H_0 a qualquer n.s. $\alpha_0 \leq 30\%$, nomeadamente aos n.u.s. de 1%, 5% e 10%;
  - devemos rejeitar H_0 a qualquer n.s. $\alpha_0 \geq 40\%$.

- **[Decisão (com base no valor-p)**
  Tendo em conta o que se refere acima

$$\begin{aligned} valor-p &\simeq 1 - F_{\chi^2_{(8)}}(9.031250) \\ &\stackrel{calc.}{\simeq} 0.339667. \end{aligned}$$

  Logo, é suposto:
  - não rejeitar H_0 a qualquer n.s. $\alpha_0 \leq 33.9667\%$, nomeadamente aos n.u.s. de 1%, 5% e 10%;
  - rejeitar H_0 a qualquer n.s. $\alpha_0 > 33.9667\%$.]

Grupo II 10 valores

1. Uma investigadora defende a hipótese H_0 de que a variável aleatória X, que representa o número de dias de internamento por doente em certo serviço hospitalar, possui função de probabilidade dada por

$$P(X = x) = \binom{x-1}{3} p^4 (1-p)^{x-4}, \quad x \in \{4, 5, \ldots\},$$

onde $p \in]0,1[$ é desconhecido. Para avaliar esta conjectura, ela recolheu dados referentes a 200 internamentos selecionados ao acaso e agrupou-os em 5 classes na tabela de frequências abaixo:

Classe	$\{4,5\}$	$\{6\}$	$\{7\}$	$\{8\}$	$\{9,10,\ldots\}$
Frequência absoluta observada	63	33	37	24	43
Estimativa da frequência absoluta esperada sob H_0	51.24	e_2	33.35	26.27	e_5

(a) Sabendo que a estimativa de máxima verosimilhança de p é $\hat{p} = 0.55$, obtenha os valores das estimativas e_2 e e_5 (aproximando-as às centésimas). (1.0)

- **V.a. de interesse**
  X = número de dias de internamento por doente em certo serviço hospitalar
- **F.p. conjecturada**
  $P(X=x) = \binom{x-1}{3}\, p^4 (1-p)^{x-4},\ x \in \{4,5,\ldots\},$ onde p é um parâmetro desconhecido em $]0,1[$.
- **Estimativas das frequências absolutas esperadas omissas**
  Atendendo à dimensão da amostra $n = 200$, à f.p. conjecturada, à estimativa de MV facultada $\hat{p} \simeq 0.55$ e à propriedade de invariância dos EMV, temos:

$$\begin{aligned} e_2 &= n \times \widehat{P(X=6)} \\ &= n \times P(X=6 \mid p=\hat{p}) \\ &= 200 \times \binom{6-1}{3} \hat{p}^4 (1-\hat{p})^{6-4} \\ &\simeq 37.06; \\ e_5 &= n - \sum_{i=1}^{4} e_i \\ &\simeq 200-(51.24+37.06+33.35+26.27) \\ &= 52.08. \end{aligned}$$

(b) Teste H_0, ao nível de significância de 5%. (3.0)

- **Hipóteses**
  $H_0 : P(X=x) = \binom{x-1}{3}\, p^4 (1-p)^{x-4},\ x \in \{4,5,\ldots\}$ (p desconhecido)
  $H_1 : \neg H_0$
- **Nível de significância**
  $\alpha_0 = 5\%$
- **Estatística de teste**

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

  onde:

  k = No. de classes = 5

  O_i = Frequência absoluta observável da classe i

  E_i = ESTIMADOR da frequência absoluta esperada, sob H_0, da classe i

  β = No. de parâmetros a estimar = 1 [dado que p é uma probabilidade desconhecida.]
- **Estimativas das frequências absolutas esperadas sob H_0**
  De acordo com a tabela facultada e a alínea (a), as estimativas [de MV] das frequências absolutas esperadas sob H_0 aproximadas às centésimas são: $e_1 \simeq 51.24$; $e_2 \simeq 37.06$; $e_3 \simeq 33.35$; $e_4 \simeq 26.27$; $e_5 \simeq 52.08$.
  [Não é necessário fazer qualquer agrupamento de classes uma vez que em pelo menos 80% das classes se verifica $e_i \geq 5$ e que $e_i \geq 1$ para todo o i. Caso fosse preciso efectuar agrupamento de classes, os valores de k e $c = F^{-1}_{\chi^2_{(k-\beta-1)}}(1-\alpha_0)$ teriam que ser recalculados...]
- **Região de rejeição de H_0** (para valores de T)
  Tratando-se de um teste de ajustamento, a região de rejeição de H_0 é o intervalo à direita $W = (c, +\infty)$, onde

$$\begin{aligned} c &= F^{-1}_{\chi^2_{(k-\beta-1)}}(1-\alpha_0) \\ &= F^{-1}_{\chi^2_{(5-1-1)}}(1-0.05) \\ &\overset{tabela/calc.}{=} 7.815. \end{aligned}$$

- **Decisão**

i	Classe i	Freq. abs. obs. o_i	Estim. freq. abs. esp. sob H_0 e_i	Parcelas valor obs. estat. teste $\frac{(o_i-e_i)^2}{e_i}$
1	$\{4,5\}$	63	51.24	$\frac{(63-51.24)^2}{51.24} \simeq 2.699$
2	$\{6\}$	33	37.06	$\frac{(33-37.06)^2}{37.06} \simeq 0.445$
3	$\{7\}$	37	33.35	0.399
4	$\{8\}$	24	26.27	0.196
5	$\{9,10,\dots\}$	43	52.08	1.583
		$\sum_{i=1}^k o_i = n = 200$	$\sum_{i=1}^k e_i = n = 200$	$t = \sum_{i=1}^k \frac{(o_i-e_i)^2}{e_i} \simeq 5.322$

Uma vez que $t \simeq 5.322 \not\in W = (7.815, +\infty)$, não devemos rejeitar H_0 ao n.s. de $\alpha_0 = 5\%$ [nem a qualquer outro n.s. inferior a α_0].

2. Um conjunto de seis medições independentes conduziu aos seguintes resultados referentes ao desgaste de um rolamento (Y) e à viscosidade do óleo usado no mesmo rolamento (x):

$$\sum_{i=1}^{6} x_i = 155.1\,, \quad \sum_{i=1}^{6} x_i^2 = 5\,264.81\,, \quad \sum_{i=1}^{6} y_i = 931\,, \quad \sum_{i=1}^{6} y_i^2 = 158\,759\,, \quad \sum_{i=1}^{6} x_i y_i = 20\,172\,,$$

onde $[\min_{i=1,\dots,6} x_i, \max_{i=1,\dots,5} x_i] = [1.6,\ 43]$.

(a) Admitindo a validade das hipóteses de trabalho habituais para o modelo de regressão linear simples de Y em x, obtenha as estimativas de máxima verosimilhança dos coeficientes β_0 e β_1. (1.5)

- **[Hipóteses de trabalho**

  $\epsilon_i \stackrel{i.i.d.}{\sim} \text{Normal}(0, \sigma^2),\ i = 1, \dots, n]$

- **Estimativas de MV de β_0 e β_1**

  Dado que

  $n = 6$

  $\sum_{i=1}^n x_i = 155.1$

  $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = \frac{155.1}{6} = 25.85$

  $\sum_{i=1}^n x_i^2 = 5\,264.81$

  $\sum_{i=1}^n x_i^2 - n\,(\bar{x})^2 = 5\,264.81 - 6 \times 25.85^2 = 1\,255.4750$

  $\sum_{i=1}^n y_i = 931$

  $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = \frac{931}{6} = 155.1(6)$

  $\sum_{i=1}^n y_i^2 = 158\,759$

  $\sum_{i=1}^n y_i^2 - n\,(\bar{y})^2 = 158\,759 - 6 \times 155.1(6)^2 = 14\,298.8(3)$

  $\sum_{i=1}^n x_i y_i = 20\,172$

  $\sum_{i=1}^n x_i y_i - n\,\bar{x}\,\bar{y} = 20\,172 - 6 \times 25.85 \times 155.1(6) = -3\,894.35,$

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

$$\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_{i=1}^n x_i y_i - n\,\bar{x}\,\bar{y}}{\sum_{i=1}^n x_i^2 - n\,(\bar{x})^2} \\
&= \frac{-3\,894.35}{1\,255.4750} \\
&\simeq -3.101894
\end{aligned}$$

$$\begin{aligned}
\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \times \bar{x} \\
&\simeq 155.1(6) - (-3.101894) \times 25.85 \\
&\simeq 235.350619.
\end{aligned}$$

(b) Obtenha um intervalo de confiança a 95% para o valor esperado do desgaste de um rolamento usando óleo com viscosidade igual a 35. (3.5)

Confronte $H_0 : \beta_0 + 35\beta_1 = 100$ e $H_1 : \beta_0 + 35\beta_1 \neq 100$, ao nível de significância de 5%, tirando partido do intervalo que obteve.

- **Obtenção do IC para** $E(Y \mid x = x_0) = \beta_0 + \beta_1 x_0$, com $x_0 = 35$

  Passo 1 — V.a. fulcral para $E(Y \mid x = x_0) = \beta_0 + \beta_1 x_0$

$$Z = \frac{(\hat{\beta}_0 + \hat{\beta}_1 x_0) - (\beta_0 + \beta_1 x_0)}{\sqrt{\hat{\sigma}^2 \times \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right]}} \sim t_{(n-2)}$$

  Passo 2 — Quantis de probabilidade

  Já que $(1-\alpha) \times 100\% = 95\%$, temos $\alpha = 0.05$ e lidaremos com os quantis

$$\begin{cases} a_\alpha = F^{-1}_{t_{(n-2)}}(\alpha/2) = -F^{-1}_{t_{(6-2)}}(1 - 0.05/2) = -F^{-1}_{t_{(4)}}(0.975) \overset{tabela/calc.}{=} -2.776 \\ b_\alpha = F^{-1}_{t_{(6-2)}}(1 - 0.05/2) = F^{-1}_{t_{(4)}}(0.975) \overset{tabela/calc.}{=} 2.776. \end{cases}$$

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

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

$$P\left[a_\alpha \leq \frac{(\hat{\beta}_0 + \hat{\beta}_1 x_0) - (\beta_0 + \beta_1 x_0)}{\sqrt{\hat{\sigma}^2 \times \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right]}} \leq b_\alpha\right] = 1 - \alpha$$

$$P\left[(\hat{\beta}_0 + \hat{\beta}_1 x_0) - b_\alpha \times \sqrt{\hat{\sigma}^2 \times \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right]} \leq \beta_0 + \beta_1 x_0 \right.$$
$$\left. \leq (\hat{\beta}_0 + \hat{\beta}_1 x_0) - a_\alpha \times \sqrt{\hat{\sigma}^2 \times \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right]}\right] = 1 - \alpha$$

- **Passo 4 — Concretização**

  Dado que a estimativa de σ^2 é igual a

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n-2}\left[\left(\sum_{i=1}^n y_i^2 - n\bar{y}^2\right) - (\hat{\beta}_1)^2 \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2\right)\right] \\ &\simeq \frac{1}{6-2}\left(14298.8(3) - (-3.101894)^2 \times 1\,255.4750\right) \\ &\simeq 554.742814 \end{aligned}$$

  e a expressão geral do IC pretendido é

$$IC_{(1-\alpha)\times 100\%}(\beta_0 + \beta_1 x_0) = \left[(\hat{\beta}_0 + \hat{\beta}_1 x_0) \pm F^{-1}_{t_{(n-2)}}(1 - \alpha/2) \times \sqrt{\hat{\sigma}^2 \times \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right]}\right],$$

  temos

$$\begin{aligned} & IC_{95\%}(\beta_0 + \beta_1 \times 35) \\ & \simeq \left[[235.350619 + (-3.101894) \times 35] \pm 2.776 \times \sqrt{554.742814 \times \left[\frac{1}{6} + \frac{(35 - 25.85)^2}{1\,255.4750}\right]}\right] \\ & \simeq [126.784339 \pm 2.776 \times 11.377645] \\ & \simeq [95.199997, 158.368681]. \end{aligned}$$

- **Hipóteses**

  $H_0 : \beta_0 + 35\beta_1 = 100$

  $H_1 : \beta_0 + 35\beta_1 \neq 100$

- **N.s.**

  $\alpha_0 = 0.05$

- **Decisão**
  Invocando a relação entre intervalos de confiança e testes de hipóteses (bilaterais), não devemos rejeitar a hipótese $H_0 : \beta_0 + 35\beta_1 = 100$ (a favor da hipótese $H_1 : \beta_0 + 35\beta_1 \neq 100$), ao n.s. de $\alpha \times 100\% = 100\% - 95\% = 5\%$ [ou a qualquer n.s. inferior a 5%] já que

  $$100 \in IC_{95\%}\beta_0 + 35\beta_1) = [95.199997,\ 158.368681].$$

(c) Obtenha e interprete o valor do coeficiente de determinação do modelo ajustado. (1.0)

- **Cálculo do coeficiente de determinação**

  $$\begin{aligned} r^2 &= \frac{\left(\sum_{i=1}^n x_i y_i - n\,\bar{x}\,\bar{y}\right)^2}{\left(\sum_{i=1}^n x_i^2 - n\,\bar{x}^2\right) \times \left(\sum_{i=1}^n y_i^2 - n\,\bar{y}^2\right)} \\ &\stackrel{(a)}{=} \frac{(-3894.35)^2}{1255.4750 \times 14298.8(3)} \\ &\simeq 0.844814. \end{aligned}$$

- **Interpretação coeficiente de determinação**
  Cerca de 84.5% da variação total da variável resposta Y é explicada pela variável x, através do modelo de regressão linear simples ajustado. Logo podemos afirmar que a recta estimada parece ajustar-se bem ao conjunto de dados.