DM DEPARTAMENTO DE MATEMÁTICA TÉCNICO LISBOA	**Probabilidade e Estatística** **LEC, LEFT, LEGI, LMAC**	1º Semestre – 2023/2024 29/01/2024 **10:30–12:30**

Duração: **120** minutos **Exame de época de recurso**

- 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 múltiplas têm de ser assinaladas na folha de instruções previamente distribuída.
- Para as restantes questões, terá de justificar convenientemente as suas respostas.

Pergunta 1 2 valores

Num certo ramo de negócios, 10% das empresas apresentam irregularidades na sua contabilidade fiscal que é, anualmente, sujeita a auditorias. O auditor A_1 assinala corretamente irregularidades com probabilidade 0.8 em empresas cuja contabilidade não está em ordem e assinala incorretamente irregularidades com probabilidade 0.1 em empresas cuja contabilidade está em ordem. As mesmas probabilidades são de 0.7 e 0.05, respetivamente, para o auditor A_2. Considere ainda que, dada a existência ou ausência de irregularidades na contabilidade de uma empresa, as conclusões dos dois auditores são independentes.

A contabilidade fiscal de uma empresa, selecionada ao acaso, vai ser analisada pelos dois auditores. Calcule a probabilidade de ambos assinalarem irregularidades nessa contabilidade.

- **Quadro de acontecimentos e probabilidades**

Acontecimento	Probabilidade
I = haver irregularidades na contabilidade de empresa (i.e., a contabilidade não está em ordem)	$P(I) = 0.1$
A_i = auditor A_i assinala irregularidades na contabilidade de empresa ($i = 1, 2$)	$P(A_i) = ?$
	$P(A_1 \mid I) = 0.8$
	$P(A_1 \mid \bar{I}) = 0.1$
	$P(A_2 \mid I) = 0.7$
	$P(A_2 \mid \bar{I}) = 0.05$

- **Prob. pedida**

  Ao invocarmos a lei da probabilidade total (LPT) e a independência condicional (IC), temos

$$
\begin{aligned}
P(A_1 \cap A_2) &\overset{LPT}{=} P(A_1 \cap A_2 \mid I) \times P(I) + P(A_1 \cap A_2 \mid \bar{I}) \times P(\bar{I}) \\
&\overset{IC}{=} P(A_1 \mid I) \times P(A_2 \mid I) \times P(I) + P(A_1 \mid \bar{I}) \times P(A_2 \mid \bar{I}) \times P(\bar{I}) \\
&= 0.8 \times 0.7 \times 0.1 + 0.1 \times 0.05 \times 0.9 \\
&= 0.0605.
\end{aligned}
$$

Pergunta 2 2 valores

Admita que, após a construção de um dique, o número de anos observados até à primeira ocorrência de cheia em determinada região ribeirinha é uma variável aleatória X que possui distribuição geométrica com valor esperado igual a 20 anos. Qual é a probabilidade de não ocorrer cheia nos primeiros 10 anos após a construção do dique? **A**: 0.5987 **B**: 0.6302 **C**: 0.0299 **D**: 0.4013

- **V.a. de interesse**

  X = no. de anos observados até à primeira ocorrência de cheia em determinada região ribeirinha...

- **Distribuição**

  $X \sim \text{geométrica}(p)$

  $p = P(\text{registo de cheia no ano observado})$

  $p \in (0,1) : E(X) = 20 \Leftrightarrow \frac{1}{p} = 20 \Leftrightarrow p = 0.05$

- **F.p. de X**

  $P(X = x) = (1-0.05)^{x-1} \times 0.05, \quad x = 1,2,3,\ldots$

- **Prob. pedida**

$$
\begin{aligned}
P(X > 10) &= 1 - P(X \le 10) \\
&= 1 - \sum_{x=1}^{10} (1-0.05)^{x-1} \times 0.05 \\
&= 1 - 0.05 \times \sum_{x=1}^{10} (1-0.05)^{x-1} \\
&= 1 - 0.05 \times (1-0.05)^{1-1} \, \frac{1-(1-0.05)^{10}}{1-(1-0.05)} \\
&= (1-0.05)^{10} \\
&\simeq 0.5987.
\end{aligned}
$$

Pergunta 3	2 valores

Considere a variável aleatória X com função de distribuição

$$
F_X(x) = \begin{cases} 0, & x < 0 \\ 1-(1-x)^2, & 0 \le x < 1 \\ 1, & x \ge 1. \end{cases}
$$

Calcule $P(X \le 2 \times q_{0.25} \mid X > q_{0.25})$, em que $q_{0.25}$ representa o quantil de probabilidade 0.25 de X.

- **V.a. e f.d.**

  X

$$
F_X(x) = \begin{cases} 0, & x < 0 \\ 1-(1-x)^2, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}
$$

- **Primeiro quartil de X**

$$
q_{0.25} \in [0,1] \quad : \quad F_X(q_{0.25}) = 0.25 \quad \Leftrightarrow \quad 1 - \left(1 - q_{0.25}\right)^2 = 0.25 \quad \Leftrightarrow \quad \left(1 - q_{0.25}\right)^2 = 0.75
$$

$$
q_{0.25} = 1 - \sqrt{0.75} \quad \Leftrightarrow \quad q_{0.25} = 1 - \frac{\sqrt{3}}{2} \quad [\Leftrightarrow \quad q_{0.25} \simeq 0.1340].
$$

- **Prob. pedida**

$$
\begin{aligned}
P\left(X \le 2q_{0.25} \mid X > q_{0.25}\right) &= \frac{P\left(q_{0.25} < X \le 2q_{0.25}\right)}{P\left(X > q_{0.25}\right)} = \frac{F_X\left(2q_{0.25}\right) - F_X\left(q_{0.25}\right)}{1 - F_X\left(q_{0.25}\right)} = \frac{F_X\left(2-\sqrt{3}\right) - 0.25}{0.75} \\
&= \frac{1 - \left(1 - 2 + \sqrt{3}\right)^2 - 0.25}{0.75} = \frac{1 - 1 + 2\sqrt{3} - 3 - 0.25}{3/4} = \frac{8\sqrt{3} - 13}{3} \simeq 0.2855.
\end{aligned}
$$

Pergunta 4	2 valores

De acordo com registos de uma grande empresa de construção civil, a função de probabilidade conjunta do número anual de acidentes envolvendo um trabalhador de um estaleiro (X) e do número de anos completos de experiência de trabalho nesse mesmo estaleiro (Y) é dada pela tabela à direita.

Determine o coeficiente de correlação entre X e Y e comente o valor obtido.

	Y		
X	0	1	2
0	0.05	0.1	0.15
1	0.2	0.1	0.05
2	0.15	0.1	0.1

- **Par aleatório e f.p. marginais**

  (X, Y)

  X = número anual de acidentes envolvendo um trabalhador de um estaleiro

  Y = número de anos completos de experiência de trabalho nesse mesmo estaleiro

- **Correlação pedida**

$$corr(X,Y) = \frac{cov(X,Y)}{\sqrt{V(X)\,V(Y)}} = \frac{E(XY) - E(X)\times E(Y)}{\sqrt{V(X)\,V(Y)}}.$$

  Logo 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 $P(Y = y) = \sum_x P(X = x, Y = y)$.

	Y			
X	0	1	2	$P(X = x)$
0	0.05	0.1	0.15	0.3
1	0.2	0.1	0.05	0.35
2	0.15	0.1	0.1	0.35
$P(Y = y)$	0.4	0.3	0.3	1

- **Valor esperado e variância de X**

$$\begin{aligned} E(X) &= \sum_{x=0}^{2} x \times P(X = x) = 1 \times 0.35 + 2 \times 0.35 = 1.05 \\ V(X) &= E(X^2) - E^2(X) \\ &= \sum_{x=0}^{2} x^2\, P(X = x) - E^2(X) = \left(1^2 \times 0.35 + 2^2 \times 0.35\right) - 1.05^2 = 1.75 - 1.1025 = 0.6475 \end{aligned}$$

- **Valor esperado e variância de Y**

$$\begin{aligned} E(Y) &= \sum_{y=0}^{2} y \times P(Y = y) = 1 \times 0.3 + 2 \times 0.3 = 0.9 \\ V(Y) &= E(Y^2) - E^2(Y) = \sum_{y=0}^{2} y^2\, P(Y = y) - E^2(Y) = \left(1^2 \times 0.3 + 2^2 \times 0.3\right) - 0.9^2 = 1.5 - 0.81 = 0.69 \end{aligned}$$

- **Valor esperado de XY**

$$E(XY) = \sum_{x=0}^{2}\sum_{y=0}^{2} xy \times P(X = x, Y = y) = 1 \times 1 \times 0.1 + 1 \times 2 \times 0.05 + 2 \times 1 \times 0.1 + 2 \times 2 \times 0.1 = 0.8$$

- **Covariância entre X e Y**

$$cov(X,Y) = E(XY) - E(X) \times E(Y) = 0.8 - 1.05 \times 0.9 = -0.145$$

- **Correlação pedida** (cont.)

$$\begin{aligned} corr(X,Y) &= \frac{cov(X,Y)}{\sqrt{V(X)\,V(Y)}} \\ &= \frac{-0.145}{\sqrt{0.6475 \times 0.69}} \\ &\simeq -0.216932 \end{aligned}$$

- **Comentários**
  - [É sabido que: caso X e Y sejam v.a. independentes, então $corr(X,Y)=0$.]

    Uma vez que $corr(X,Y) \neq 0$, concluímos que X e Y são v.a. dependentes.
  - Dado que $corr(X,Y) < 0$, podemos adiantar que X e Y tenderão a variar em sentidos opostos relativamente aos respectivos valores esperados.

    [Este resultado é expectável uma vez que quanto maior o número de anos completos de experiência de trabalho num estaleiro menor deverá o número anual de acidentes envolvendo um trabalhador de um estaleiro.]
  - Como $|corr(X,Y)| \simeq 0.21$ dista significativamente de 1, podemos afirmar que as v.a. estão fracamente correlacionadas.

    [Importa notar que o coeficiente de correlação quantifica a associação linear entre as v.a. X e Y, logo não captura uma eventual relação não linear entre estas duas v.a.]

Pergunta 5 — 2 valores

Um engenheiro civil admite que os números de buracos que requerem reparação urgente em dois troços de auto-estrada são duas variáveis aleatórias independentes com distribuição de Poisson cujas variâncias são iguais a 2 e 4.

Obtenha o valor exato para a probabilidade de o número total de buracos que requerem reparação urgente nos dois troços exceder 7.

- **V.a.**

  X_i = número de buracos que requerem reparação urgente no troço i, $i=1,2$
- **Distribuição de X_i**

  $X_i \overset{indep}{\sim} \text{Poisson}(\lambda_i), \quad i=1,2$

  $\lambda_1 = E(X_1) = V(X_1) = 2$

  $\lambda_2 = E(X_2) = V(X_2) = 4$
- **V.a. de interesse**

  $Y = X_1 + X_2$ = número total de buracos que requerem reparação urgente nos dois troços
- **Distribuição de Y**

  Tratando-se Y da soma de duas v.a. independentes com distribuição de Poisson, temos

  $$Y \sim \text{Poisson}(E(Y)),$$

  onde $E(Y) = E(X_1) + E(X_2) = 2 + 4 = 6$.
- **Prob. pedida**

  $$P(Y > 7) \quad = \quad 1 - P(Y \leq 7) = 1 - F_{Poisson(6)}(7) \overset{tabela/calc.}{=} 1 - 0.7440 = 0.2560.$$

Pergunta 6 — 2 valores

Admita que o tempo até falha por fadiga de determinado material de construção é representado pela variável aleatória X (em anos) com quantil de probabilidade p $(0 < p < 1)$ dado por

$$F_X^{-1}(p) = \frac{1}{4}\left\{\gamma \times \Phi^{-1}(p) + \sqrt{4 + \left[\gamma \times \Phi^{-1}(p)\right]^2}\right\}^2,$$

onde γ é uma constante positiva desconhecida e $\Phi^{-1}(p)$ denota o quantil de probabilidade p da normal padrão. A concretização de uma amostra aleatória de dimensão $n = 25$ proveniente de X conduziu à seguinte estimativa de máxima verosimilhança de γ: $\hat{\gamma} = 3.0098$.

Qual é o valor da estimativa de máxima verosimilhança do terceiro quartil de X?

A: 1.1214 **B**: 5.0305 **C**: 1.2200 **D**: 5.9534

- **V.a. de interesse**

  X = tempo até falha por fadiga de determinado material de construção

  $F_X^{-1}(p) = \frac{1}{4}\left[\gamma \times \Phi^{-1}(p) + \sqrt{4 + \left(\gamma \times \Phi^{-1}(p)\right)^2}\right]^2, \quad 0 < p < 1$

- **Parâmetro desconhecido**

  $\gamma \quad (\gamma > 0)$

- **Amostra**

  $\underline{x} = (x_1, \ldots, x_n)$, amostra de dimensão $n = 25$ proveniente da população X e tal que $\hat{\gamma} = 3.0098$.

- **Outro parâmetro desconhecido**

$$\begin{aligned} h(\gamma) &= F_X^{-1}(3/4) \\ &= \frac{1}{4}\left\{\gamma \times \Phi^{-1}(p) + \sqrt{4 + \left[\gamma \times \Phi^{-1}(p)\right]^2}\right\}^2, \quad 0 < p < 1 \end{aligned}$$

- **Estimativa de MV de $h(\gamma)$**

  Ao invocar a propriedade de invariância dos estimadores de MV, concluímos que a estimativa de MV de $h(\gamma)$ é

$$\begin{aligned} \widehat{h(\gamma)} &= h(\hat{\gamma}) \\ &= \frac{1}{4}\left\{\hat{\gamma} \times \Phi^{-1}(3/4) + \sqrt{4 + \left[\hat{\gamma} \times \Phi^{-1}(3/4)\right]^2}\right\}^2 \\ &\overset{tabela/calc.,\ \hat{\gamma}=3.0098}{=} \frac{1}{4}\left[3.0098 \times 0.6745 + \sqrt{4 + (3.0098 \times 0.6745)^2}\right]^2 \\ &\simeq 5.9534. \end{aligned}$$

Pergunta 7	2 valores

Para avaliar o tempo de vida de certas baterias a usar num painel de uma aeronave foi efetuado um teste a 60 destas componentes, escolhidas ao acaso de um lote de grande dimensão. Da amostra obteve-se uma média igual a $\bar{x} = 126$ horas e um desvio padrão de $s = 6.4$ horas.

Determine um intervalo de confiança aproximado a 96% para o valor esperado do tempo de vida das baterias.

- **V.a. de interesse**

  X = tempo de vida da bateria

- **Situação**

  X com distribuição arbitrária

  $\mu = E(X)$ DESCONHECIDO

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

  $n = 60 \geq 30$

- **Obtenção de IC aproximado para μ**

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

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

  Passo 2 — Obtenção dos quantis de probabilidade

  Como $(1-\alpha)\times 100\% = 96\% \Leftrightarrow \alpha = 0.04$, lidaremos com os quantis seguintes:

  $$\begin{cases} a_\alpha = \Phi^{-1}(\alpha/2) = -\Phi^{-1}(1-\alpha/2) = -\Phi^{-1}(0.98) \overset{tabela/calc.}{=} -2.0537 \\ b_\alpha = \Phi^{-1}(1-\alpha/2) = \Phi^{-1}(0.98) = 2.0537. \end{cases}$$

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

  $$P(a_\alpha \leq Z \leq b_\alpha) \simeq 1-\alpha$$
  $$P\left(a_\alpha \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq b_\alpha\right) \simeq 1-\alpha$$
  $$P\left[\bar{X} - b_\alpha \times S/\sqrt{n} \leq p \leq \bar{X} - a_\alpha \times S/\sqrt{n}\,\right] \simeq 1-\alpha$$

  Passo 4 — Concretização

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

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

  Atendendo a que $n = 60$, $\alpha = 0.04$, $\Phi^{-1}(1-\alpha/2) = 2.0537$, $\bar{x} = 126$ e $s = 6.4$, segue-se

  $$\begin{aligned} IC_{96\%}(\mu) &\simeq \left[126 - 2.0537 \times 6.4/\sqrt{60},\, 126 - 2.0537 \times 6.4/\sqrt{60}\right] \\ &= [124.303,\, 127.697]. \end{aligned}$$

Pergunta 8 — 2 valores

Uma engenheira civil admite que o tempo (em horas) de transporte de resíduos por camião é uma variável aleatória X com distribuição normal com valor esperado μ e desvio padrão σ desconhecidos.

Teste a hipótese $H_0 : \sigma = 0.15$ contra $H_1 : \sigma > 0.15$, num dia em que foram recolhidos casualmente 16 tempos de transporte de resíduos por camiões e o desvio padrão amostral foi de $s = 0.19$. Obtenha o valor-p (ou um intervalo para o valor-p) do teste. Para que níveis de significância não se deve rejeitar H_0?

- **V.a. de interesse**

  X = tempo (em horas) de transporte de resíduos por camião

- **Situação**

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

  μ desconhecido

  σ DESCONHECIDO

- **Hipóteses**

  $H_0 : \sigma = \sigma_0 = 0.15$ vs. $H_1 : \sigma > \sigma_0$

- **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 unilateral superior ($H_1 : \sigma^2 > \sigma_0^2$), logo a região de rejeição de H_0 é do tipo $W = (c, +\infty)$.

- **Valor-p e comentário**

  Atendendo a que o valor observado da estatística de teste e o valor-p são iguais a

$$\begin{aligned} t &= \frac{(n-1)\, s^2}{\sigma_0^2} = \frac{(16-1) \times 0.19^2}{0.15^2} \simeq 24.07 \\ valor-p &= P(T > t \mid H_0) = 1 - F_{\chi^2_{(n-1)}}(t) = 1 - F_{\chi^2_{(15)}}(24.07) \stackrel{calc.}{=} 0.0639, \end{aligned}$$

  é suposto:

  - não rejeitar H_0 a qualquer n.s. $\alpha_0 \leq valor-p = 6.39\%$, designadamente aos níveis de significâncias de 1% e 5%[;
  - rejeitar H_0 a qualquer n.s. $\alpha_0 > valor-p = 6.39\%$, nomeadamente ao nível de significância de 10%.]

  [Alternativamente, poderíamos recorrer às tabelas de quantis da distribuição do qui-quadrado e adiantar um intervalo para o valor-p:

$$\begin{aligned} F^{-1}_{\chi^2_{(15)}}(0.925) = 23.45 &< t = 24.07 < 25.00 = F^{-1}_{\chi^2_{(15)}}(0.95) \\ 0.925 &< F_{\chi^2_{(15)}}(24.07) < 0.95 \\ 1 - 0.95 &< 1 - F_{\chi^2_{(15)}}(24.07) < 1 - 0.925 \\ 0.05 &< valor-p < 0.075. \end{aligned}$$

  Assim, é suposto:

  - não rejeitar H_0 a qualquer n.s. $\alpha_0 \leq valor-p = 5\%$, designadamente aos níveis de significâncias de 1% e 5%;
  - rejeitar H_0 a qualquer n.s. $\alpha_0 \geq valor-p = 7.5\%$, nomeadamente ao nível de significância de 10%.]

Pergunta 9 — 2 valores

Numa sondagem, foi pedido a 150 habitantes de uma cidade para classificarem o estado atual das suas finanças pessoais relativamente ao ano anterior em uma de cinco categorias. As frequências registadas foram:

Muito pior	Pior	Igual	Melhor	Muito melhor
35	40	30	28	17

Usando os dados recolhidos, teste a hipótese H_0 de as cinco categorias serem equiprováveis ao nível de significância de 7.5%.

- **V.a. de interesse; f.p. conjecturada**

  X = categoria em que foi classificado do estado atual das finanças pessoais...

  Muito pior (1), Pior (2), Igual (3), Melhor (4), Muito melhor (5).

$$P_0(X = x) = \begin{cases} 1/5, & x = 1, 2, 3, 4, 5 \text{ (cinco categorias equiprováveis)} \\ 0, & \text{caso contrário} \end{cases}$$

- **Hipóteses**

  $H_0 : X \sim$ uniforme discreta($\{1, 2, 3, 4, 5\}$) vs. $H_1 : X \not\sim$ uniforme discreta($\{1, 2, 3, 4, 5\}$)

- **N.s.**

  $\alpha_0 = 7.5\%$

- **Estatística de teste**

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

  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.

- **Cálculo das frequências absolutas esperadas sob H_0**

  As frequências absolutas esperadas sob H_0 são iguais a $E_i = n \times p_i^0 = n \times P(X = i \mid H_0) = n \times P_0(X = i) = 150 \times 1/5 = 30$, para $i = 1,2,3,4,5$.

- **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 $W = (c, +\infty)$, onde

  $$c = F^{-1}_{\chi^2_{(k-1)}}(1-\alpha_0) = F^{-1}_{\chi^2_{(5-1)}}(1-0.075) = F^{-1}_{\chi^2_{(4)}}(0.925) \overset{tabela/calc.}{=} 8.496.$$

- **Decisão**

  No cálculo do valor observado da estatística de teste convém recorrer à seguinte tabela auxiliar:

i	Categoria	Freq. abs. obs. o_i	Freq. abs. esper. sob H_0 $E_i = n \times p_i^0$	Parcelas valor obs. estat. teste $\frac{(o_i-E_i)^2}{E_i}$
1	Muito pior	35	30	$\frac{(35-30)^2}{30} \simeq 0.833$
2	Pior	40	30	3.333
3	Igual	30	30	0
4	Melhor	28	30	0.133
5	Muito Melhor	17	30	5.633
		$\sum_{i=1}^k o_i = n = 150$	$\sum_{i=1}^k E_i = n = 150$	$t = \sum_{i=1}^k \frac{(o_i-E_i)^2}{E_i} \simeq 9.932$

  Como $t = 9.932 \in W = (8.496, +\infty)$ devemos rejeitar H_0 ao n.s. $\alpha_0 = 7.5\%$ [bem como a qualquer n.s. superior a 7.5%].

Pergunta 10	2 valores

Para analisar a possível associação entre o número anual de quedas de satélites em órbita terrestre baixa (Y) e o número de manchas solares (x), foram recolhidas observações referentes ao período 1969–2004. Admitindo a validade do modelo de regressão linear simples, $Y = \beta_0 + \beta_1 x + \epsilon$, foram calculadas as seguintes quantidades:

$$\textstyle\sum_{i=1}^{36} x_i = 2671, \quad \sum_{i=1}^{36} x_i^2 = 280527, \quad \hat\beta_1 \simeq 0.163507, \quad \hat\sigma^2 \simeq 5.943501^2.$$

Tendo presente estas quantidades e as hipóteses de trabalho convenientes, teste a hipótese $H_0 : \beta_1 = 0$ contra $H_1 : \beta_1 \neq 0$, ao nível de significância de 5%. Interprete o resultado do teste.

Nota: $F^{-1}_{t_{(34)}}(0.975) \simeq 2.032$.

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

  Y = número anual de quedas de satélites em órbita terrestre baixa (v.a. resposta)

  x = número anual de manchas solares (variável explicativa)

  $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \ldots, n$

$\epsilon_i \overset{i.i.d.}{\sim} \text{normal}(0, \sigma^2), \quad i = 1, \ldots, n]$

- **Hipóteses**

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

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

- **Nível de significância**

  $\alpha_0 = 5\%$

- **Estatística de teste**

$$T = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\,\bar{x}^2}}} \sim_{H_0} t_{(n-2)}$$

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

  Estamos a lidar com um teste bilateral ($H_1 : \beta_1 \neq \beta_{1,0}$), logo a região de rejeição de H_0 é do tipo $W = (-\infty, -c) \cup (c, +\infty)$, onde $c : P(\text{Rejeitar } H_0 \mid H_0) = \alpha_0$, i.e.,

$$c \;=\; F^{-1}_{t_{(n-2)}}(1 - \alpha_0/2) = F^{-1}_{t_{(36-2)}}(1 - 0.05/2) = F^{-1}_{t_{(34)}}(0.975) \overset{tabela/calc.}{\simeq} 2.032.$$

- **Decisão**

  Atendendo ao enunciado, $\sum_{i=1}^{36} x_i = 2\,671$, $\sum_{i=1}^{36} x_i^2 = 280\,527$, $\hat{\beta}_1 \simeq 0.163507$ e $\hat{\sigma}^2 \simeq 5.943501^2$. Logo, o valor observado da estatística de teste é igual a

$$\begin{aligned} t &= \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\,\bar{x}^2}}} \\ &\simeq \frac{0.163507 - 0}{\sqrt{\frac{5.943501^2}{280\,527 - 36 \times (2\,671/36)^2}}} \\ &\simeq 7.894696. \end{aligned}$$

  Como $t \simeq 7.894696 \in W = (-\infty, -2.032) \cup (2.032, +\infty)$ devemos rejeitar H_0 ao n.s. de $\alpha_0 = 5\%$[, assim como a qualquer n.s. superior a 5%.]

- **Interpretação pedida**

  Há forte evidência a favor de H_1, i.e., da hipótese de o número anual de manchas solares influenciar o número anual de quedas de satélites em órbita terrestre baixa.