

1. Questão 1

a)  $\beta_1$  : é o incremento no valor esperado da resposta para o aumento em uma unidade da covariável,  $i=1,2,\dots,m$ ;  $\beta_2$  : é o incremento no valor esperado da resposta para o aumento em uma unidade da covariável,  $i=m+1,\dots,n$ .

b) Temos que:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{bmatrix}; \mathbf{X} = \begin{bmatrix} x_1 & 0 \\ x_2 & 0 \\ x_3 & 0 \\ 0 & x_4 \\ 0 & x_5 \end{bmatrix}; \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}; \boldsymbol{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{bmatrix}$$

c) Temos que  $Q = \sum_{i=1}^n (Y_i - \beta_1 x_i I_{\{1,2,\dots,m\}}(i) - \beta_2 x_i I_{\{m+1,m+2,\dots,n\}}(i))^2$ , (seja  $\eta_i = \beta_1 x_i I_{\{1,2,\dots,m\}}(i) + \beta_2 x_i I_{\{m+1,m+2,\dots,n\}}(i)$ ):

$$\begin{aligned} \frac{\partial Q}{\partial \beta_1} &= - \sum_{i=1}^n (Y_i - \eta_i) x_i I_{\{1,2,\dots,m\}}(i) \\ &= \sum_{i=1}^m Y_i x_i - \beta_1 \sum_{i=1}^m x_i^2 - \beta_2 \sum_{i=1}^n x_i^2 I_{\{1,2,\dots,m\}}(i) I_{\{m+1,m+2,\dots,n\}}(i) \\ &= \sum_{i=1}^m Y_i x_i - \beta_1 \sum_{i=1}^m x_i^2 \end{aligned} \quad (1)$$

e

$$\frac{\partial Q}{\partial \beta_2} = - \sum_{i=1}^n (Y_i - \eta_i) x_i I_{\{m+1,m+2,\dots,n\}}(i) = \sum_{i=m+1}^n Y_i x_i - \beta_2 \sum_{i=m+1}^n x_i^2$$

Portanto,  $\hat{\beta}_1 = \frac{\sum_{i=1}^m Y_i x_i}{\sum_{i=1}^m x_i^2}$  e  $\hat{\beta}_2 = \frac{\sum_{i=m+1}^n Y_i x_i}{\sum_{i=m+1}^n x_i^2}$  (assumindo que este vetor é ponto

de mínimo). Temos que (combinação linear de normais)  $\hat{\beta}_1 \sim N(\mu_1, \psi_1)$  e  $\hat{\beta}_2 \sim N(\mu_2, \psi_2)$ , em que:

$$\mu_1 = \beta_1, \mu_2 = \beta_2; \psi_1 = \frac{\sigma^2}{\sum_{i=1}^m x_i^2}; \psi_2 = \frac{\sigma^2}{\sum_{i=m+1}^n x_i^2}$$

Temos que,  $\hat{\beta}_1 = f(Y_1, \dots, Y_m)$  e  $\hat{\beta}_2 = g(Y_{m+1}, \dots, Y_n)$ , logo  $\hat{\beta}_1 \perp \hat{\beta}_2$ . Portanto,

$$f(\boldsymbol{\beta}) = g(\beta_1)h(\beta_2) = \frac{1}{\sqrt{2\pi\psi_1}} \exp \left\{ - \left( \frac{(\hat{\beta}_1 - \mu_1)^2}{2\psi_1} \right) \right\} \frac{1}{\sqrt{2\pi\psi_2}} \exp \left\{ - \left( \frac{(\hat{\beta}_2 - \mu_2)^2}{2\psi_2} \right) \right\}$$

Normal bivariada com  $\rho = 0$ .

- d) Temos que  $Z_t = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\psi_1 + \psi_2}} \sim N(0, 1)$ , por resultados anteriores, sob  $H_0$ . Assim, rejeitamos  $H_0$  se  $|z_t| > z_c$ , em que  $z_c = \frac{\tilde{\beta}_1 - \tilde{\beta}_2}{\sqrt{\psi_1 + \psi_2}}$  e  $P(Z \leq z_c) = 1 - \frac{\alpha}{2}$ ,  $Z \sim N(0, 1)$  ou, de modo equivalente, se  $P(Z \geq |z_t| | H_0) = \alpha$ .

2. Temos que:

- a) Temos que  $\hat{\boldsymbol{\beta}}_G = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{Y}$ . Logo,  $\hat{\boldsymbol{\beta}}_G \sim N_p(\boldsymbol{\beta}, (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1})$  e  $\mathbf{C}\hat{\boldsymbol{\beta}}_G \sim N_q(\mathbf{C}\boldsymbol{\beta}, \mathbf{C}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}')$ . Seja  $Q_t = (\mathbf{C}\hat{\boldsymbol{\beta}}_G)' (\mathbf{C}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}')^{-1} (\mathbf{C}\hat{\boldsymbol{\beta}}_G)$  e defina  $\mathbf{A} = (\mathbf{C}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}')^{-1}$  e  $\boldsymbol{\Sigma}_\beta = \mathbf{C}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}'$ . Note que  $\mathbf{A}$  é não aleatório e que

$$\mathbf{A}' = (\mathbf{C}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}')^{-1'} = (\mathbf{C}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}')^{-1} = \mathbf{A}$$

Além disso, note que  $\mathbf{A}\boldsymbol{\Sigma}_\beta = \mathbf{I}$ , a qual é idempotente. Assim, sob  $H_0$ ,  $Q_t \sim \chi_q^2$ ,  $q = r(\mathbf{C})$ . Portanto, rejeita-se  $H_0$  se  $q_t > q_c$ , em que  $q_t = (\mathbf{C}\tilde{\boldsymbol{\beta}}_G)' (\mathbf{C}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}')^{-1} (\mathbf{C}\tilde{\boldsymbol{\beta}}_G)$  e  $P(Q > q_c | H_0) = \alpha$ ,  $Q \sim \chi_q^2$ . Analogamente, rejeita-se  $H_0$  se  $p \geq \alpha$ , em que  $p = P(Q \geq q_t | H_0)$ .

b) Sejam

$$\begin{aligned}\widehat{\sigma}_W^2 &= \frac{1}{n-p} (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_W)' \mathbf{W} (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}_W) \\ &= \frac{1}{n-p} \mathbf{Y}' (\mathbf{W} - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}) \mathbf{Y} = \frac{1}{n-p} \mathbf{Y}' \mathbf{A} \mathbf{Y}\end{aligned}$$

e  $\widehat{\boldsymbol{\beta}}_W = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \mathbf{X}'\mathbf{W}\mathbf{Y} = \mathbf{B}\mathbf{Y}$ . Note que  $\widehat{\boldsymbol{\beta}}_W \sim N_p(\boldsymbol{\beta}, \sigma^2 (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1})$  e

$$\begin{aligned}\frac{\mathbf{A}}{\sigma^2} \boldsymbol{\Sigma} \frac{\mathbf{A}}{\sigma^2} \boldsymbol{\Sigma} &= (\mathbf{W} - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}) \mathbf{W}^{-1} (\mathbf{W} - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}) \mathbf{W}^{-1} \\ &= (\mathbf{I} - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}') (\mathbf{I} - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}') \\ &= \mathbf{I} - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}' - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}' + \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}' \\ &= \mathbf{I} - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}' = \mathbf{A}\mathbf{B}\end{aligned}$$

Portanto  $\frac{\mathbf{A}}{\sigma^2} \boldsymbol{\Sigma}$  é idempotente. Além disso,  $\frac{\mathbf{A}}{\sigma^2} = \frac{1}{\sigma^2} (\mathbf{W} - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W})$  é simétrica, não aleatória e (assumindo que)  $r(\mathbf{A}) = n - p$ . Dessa forma,

$$\frac{(n-p)\widehat{\sigma}_W^2}{\sigma^2} \sim \chi_{n-p}^2 \quad (2)$$

. Além disso:

$$\begin{aligned}\mathbf{B}\boldsymbol{\Sigma}\mathbf{A} &= (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \mathbf{X}'\mathbf{W}\mathbf{W}^{-1} (\mathbf{W} - \mathbf{W}'\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}) \\ &= (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \mathbf{X}'\mathbf{W} - (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \mathbf{X}'\mathbf{W} = \mathbf{0}\end{aligned}$$

Portanto,  $\widehat{\boldsymbol{\beta}}_W \perp \widehat{\sigma}_W^2$ . Defina agora:

$$Q = \frac{1}{\sigma^2} (\mathbf{C}\widehat{\boldsymbol{\beta}}_W)' (\mathbf{C}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{C}')^{-1} (\mathbf{C}\widehat{\boldsymbol{\beta}}_W) \quad (3)$$

Pelo item a), temos que  $Q \sim \chi_q^2$ . Por outro lado, seja:

$$Q_t = \frac{Q/q}{(n-p)\widehat{\sigma}_W^2/[\sigma^2(n-p)]} = \frac{1}{q\widehat{\sigma}_W^2} (\mathbf{C}\widehat{\beta}_W)' \left( \mathbf{C}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{C}' \right)^{-1} (\mathbf{C}\widehat{\beta}_W)$$

Como  $\widehat{\beta}_W \perp \widehat{\sigma}_W^2$ , por (2) e (3), temos que, sob  $H_0$ ,  $Q_t \sim F(q, n-p)$ . Portanto, rejeita-se  $H_0$  se  $q_t > q_c$ , em que  $q_t = \frac{1}{q\widehat{\sigma}_W^2} (\mathbf{C}\widehat{\beta}_W)' \left( \mathbf{C}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{C}' \right)^{-1} (\mathbf{C}\widehat{\beta}_W)$  e  $P(Q > q_c | H_0) = \alpha$ ,  $Q \sim F(q, n-p)$ . Analogamente, rejeita-se  $H_0$  se  $p \geq \alpha$ , em que  $p = P(Q \geq q_t | H_0)$ .

- c) Se  $\Sigma$  for conhecida, pelo item a), já temos um teste de hipótese adequado. Em sendo  $\widehat{\Sigma}$  um estimador consistente para  $\Sigma$ , então (omitindo alguns detalhes técnicos), temos que

$$Q_t = (\mathbf{C}\widehat{\beta}_G)' \left( \mathbf{C}(\mathbf{X}'\widehat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}' \right)^{-1} (\mathbf{C}\widehat{\beta}_G) \xrightarrow[n \rightarrow \infty]{D} \chi_q^2$$

Assim, rejeita-se  $H_0$  se  $q_t \leq q_c$ , em que  $q_t = (\mathbf{C}\widehat{\beta}_G)' \left( \mathbf{C}(\mathbf{X}'\widehat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}' \right)^{-1} (\mathbf{C}\widehat{\beta}_G)$  e  $P(Q \geq q_c | H_0) \approx \alpha$ ,  $Q \sim \chi_q^2$  ou, de forma equivalente, se  $p \geq \alpha$ , em que  $p \approx P(Q \geq q_t | H_0)$

### 3. Análise dos dados

Análise Descritiva: Parecem que as variâncias são diferentes, entre os anos, o que poderia indicar uma heterocedasticidade entre os anos. A média de 1962 é menor do que a de 1961, o que indicaria uma diminuição no número médio de acidentes. Essa aparente diminuição é ratificada pelas Figuras 1 e 2, as quais também sugerem assimetria positiva das distribuições do número de acidentes em cada um dos dois anos. Com base nos testes de Bartlett e Levene identifica-se uma heterocedasticidade marginal entre os anos.

Ano	Média	DP	Var.	CV(%)	Mínimo	Med.	Máximo
1961	26,05	9,09	82,66	34,91	8,00	25,00	47,00
1962	18,05	6,69	44,71	37,05	7,00	17,00	41,00

Testes: Levene - 4,34 (0,0401); Bartlett - 3,86 (0,0495).

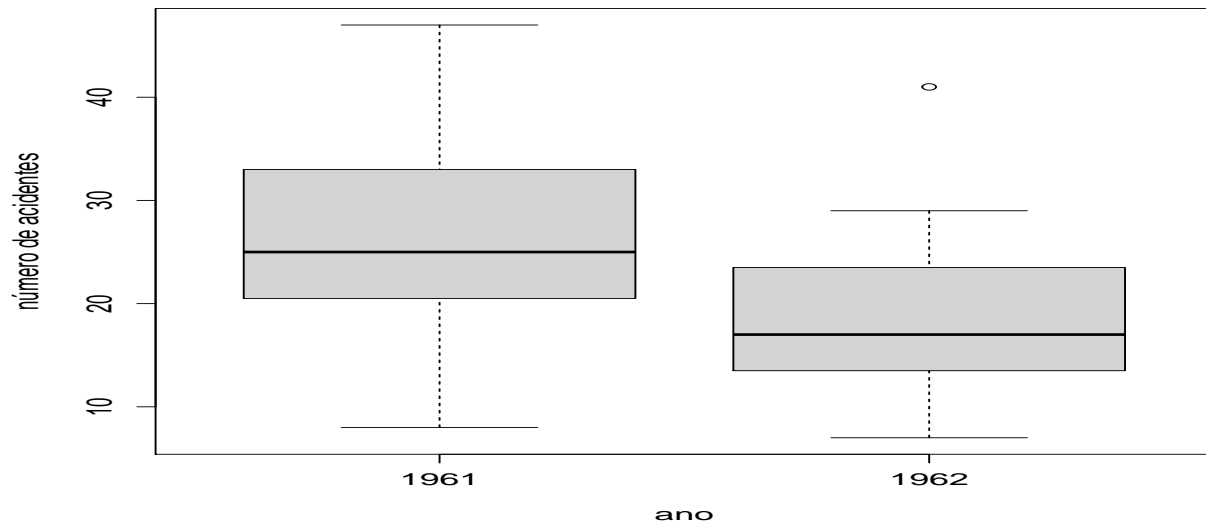


Figura 1: Box-plots do número de acidentes em cada ano

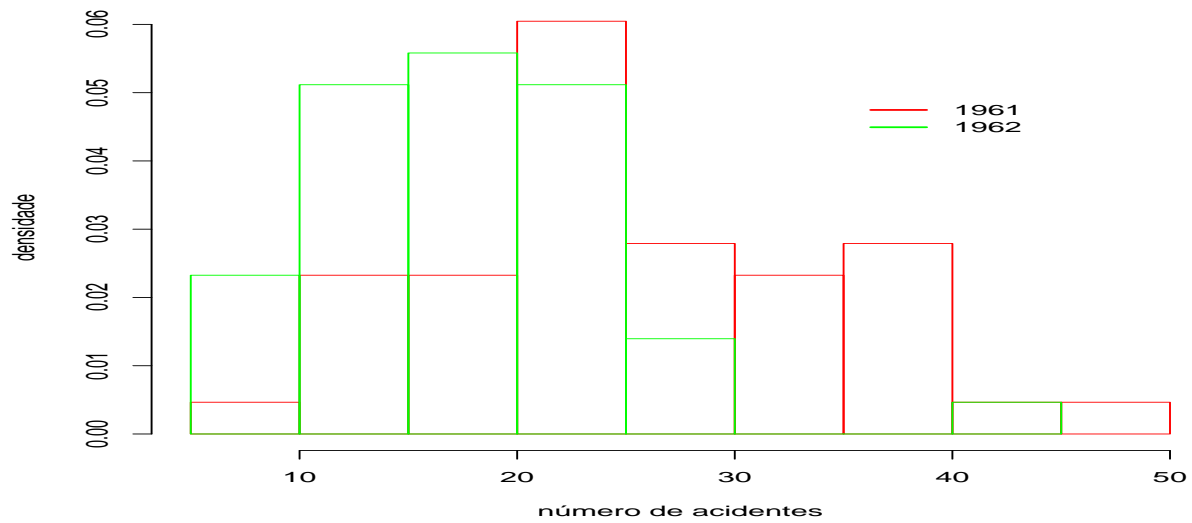


Figura 2: Histogramas do número de acidentes em cada ano

Análise Inferencial: O modelo parece não estar bem ajustado pelas Figuras 3 e 4, as quais indicam assimetria positiva, heterocedasticidade (leve) e não normalidade dos resíduos. Apesar do mau ajuste do modelo, seguiremos com ele, dadas as restrições. Pela Tabela 2 vemos que a média do número de acidentes de 1962 é significativamente menor do que a de 1961 ( $\alpha_2 \neq 0$ ,  $p < 0,0001$ ). Podemos concluir que a imposição de limites de velocidade parece ter levado (ainda que conjuntamente com outros fatores) à redução do número (médio) de acidentes. As Figuras 5 e 6 indicam que, aparentemente, não há pontos influentes. Finalmente, a Tabela 2 apresenta as médias preditas dos números de acidentes o que confirma que a média de 1962 aparenta ser significativamente menor do que a 1961.

Modelo

$$Y_{ij} = \mu + \alpha_2 + \xi_{ij}, i (\text{ano}) = 1 (1961), 2 (1962); j (\text{dia}) = 1, 2, \dots, 4342$$

$$\xi_{ij} \stackrel{iid}{\sim} N(0, \sigma^2), \alpha_1 = 0$$

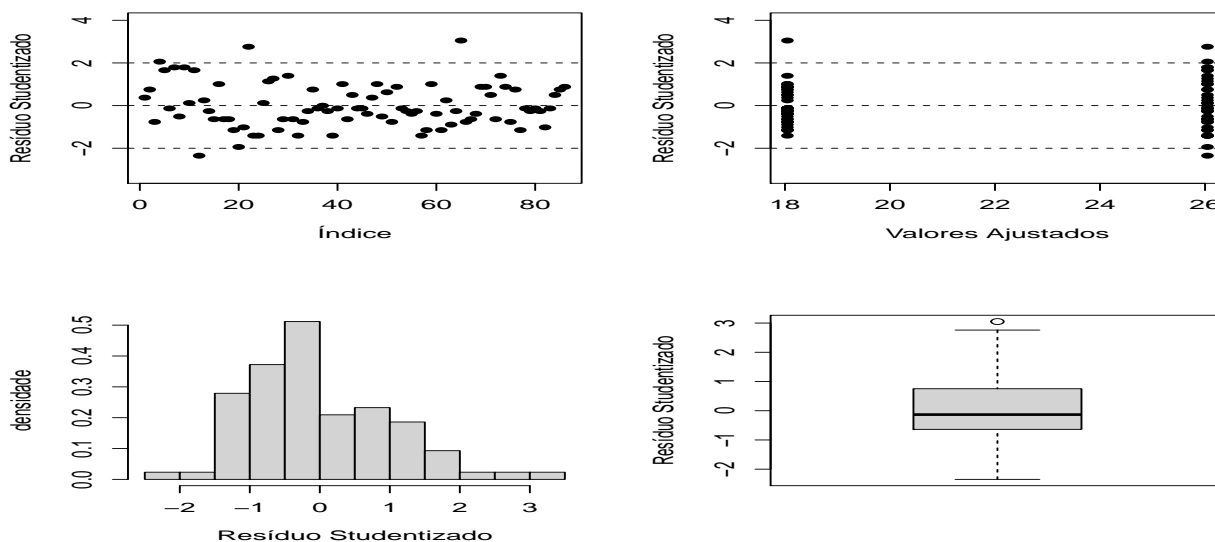


Figura 3: Gráficos de diagnóstico para o resíduo studentizado

$$R^2 = 0,205 \text{ e } \bar{R}^2 = 0,196.$$

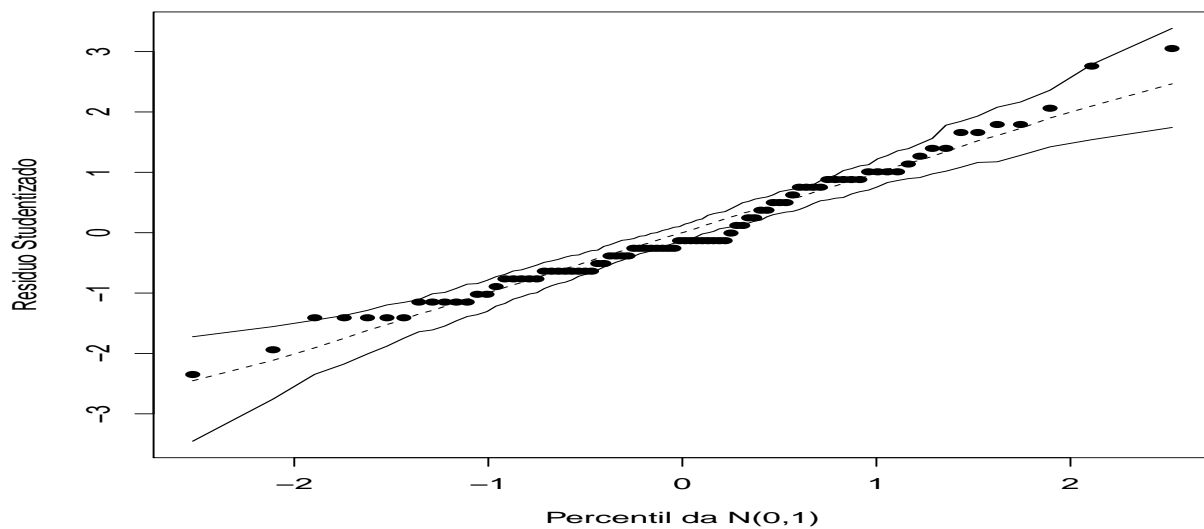


Figura 4: Gráficos de quantil quantil com envelopes para o resíduo studentizado

Tabela 2: Estimativas pontuais e intervalares e testes de hipótese de nulidade para os parâmetros de regressão

Par.	Est.	EP	IC(95%)	Estat. t	p-valor
$\mu$	26,05	1,22	[23,63;28,47]	21,40	< 0,0001
$\alpha_2$	-8,00	1,72	[-11,42;-4,58]	-4,65	< 0,0001

Tabela 3: Estimativas pontuais e intervalares das médias dos números de acidentes

Ano	Estimativa	IC(95%)
1961	26,05	[23,63;28,47]
1962	18,05	[15,63;20,47]

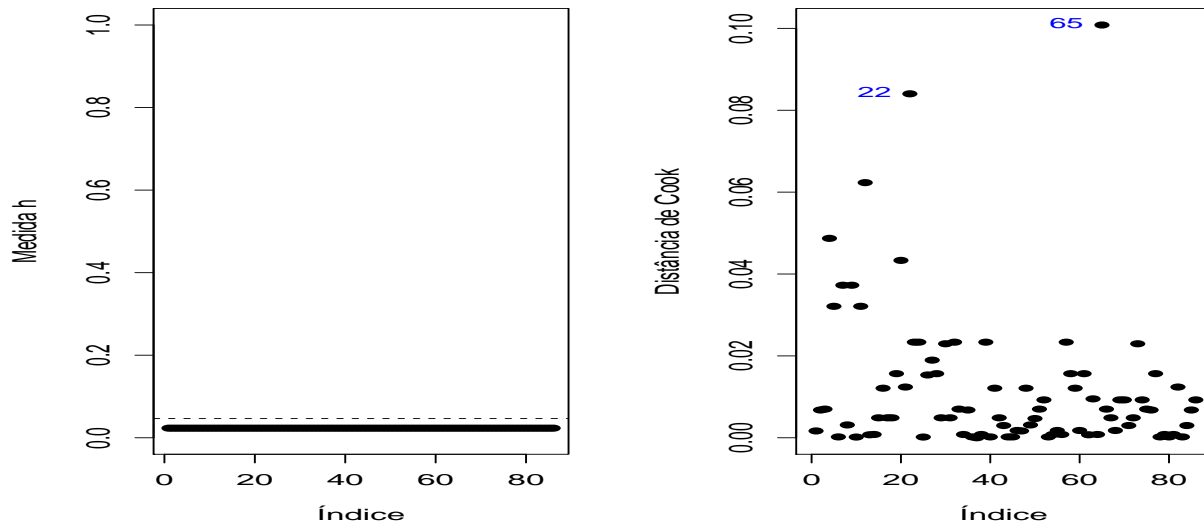


Figura 5: Análise de influência para o modelo

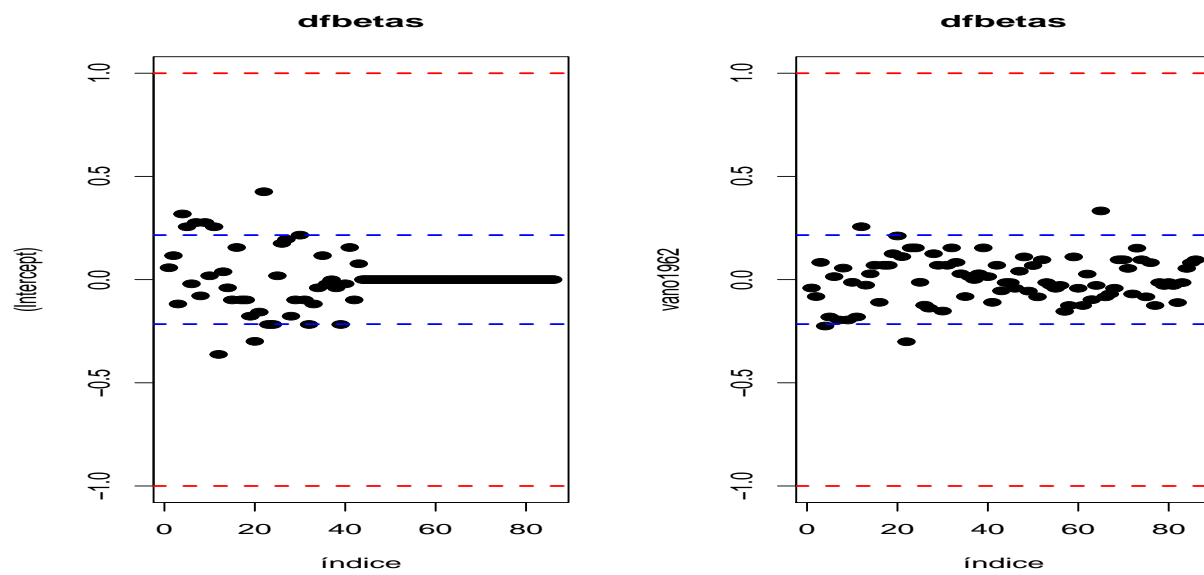


Figura 6: Gráficos DFBETAS para o modelo