Modelo normal linear multivariado: Parte 2

Prof. Caio Azevedo

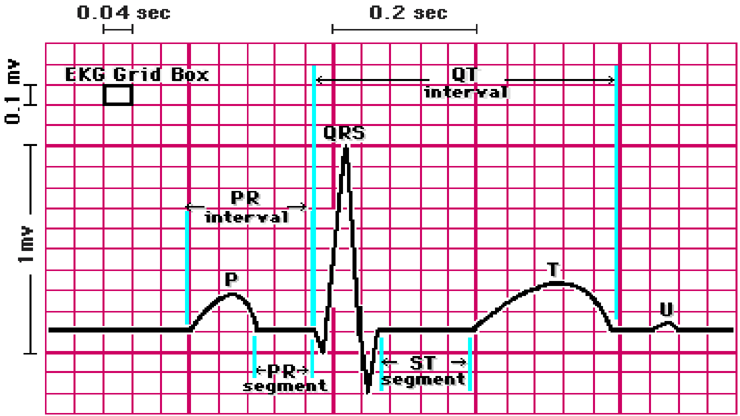
Exemplo 3: Amitriptilina

- Amitriptilina é prescrita por alguns médicos como antidepressivo.
- Entretanto existem alguns efeitos colaterais que podem estar associados ao seu uso como: batimento cardíaco irregular, pressão sanguínea anormal e ondas irregulares no eletrocardiograma.
- Os dados consistem na medição de algumas características de interesse de 17 pacientes que deram entrada em um hospital depois de uma overdose de amitriptilina.

Indivíduo	Tot	Ami	Gen	Amt	Pr	Diap	Qrs
1	3389	3149	1	7500	220	0	140
2	1101	653	1	1975	200	0	100
3	1131	810	0	3600	205	60	111
4	596	448	1	675	160	60	120
5	896	844	1	750	185	70	83
6	1767	1450	1	2500	180	60	80
7	807	493	1	350	154	80	98
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
15	781	501	0	4500	180	0	100
16	1070	405	0	1500	170	90	120
17	1754	1520	1	3000	180	0	129

Descrição das variáveis

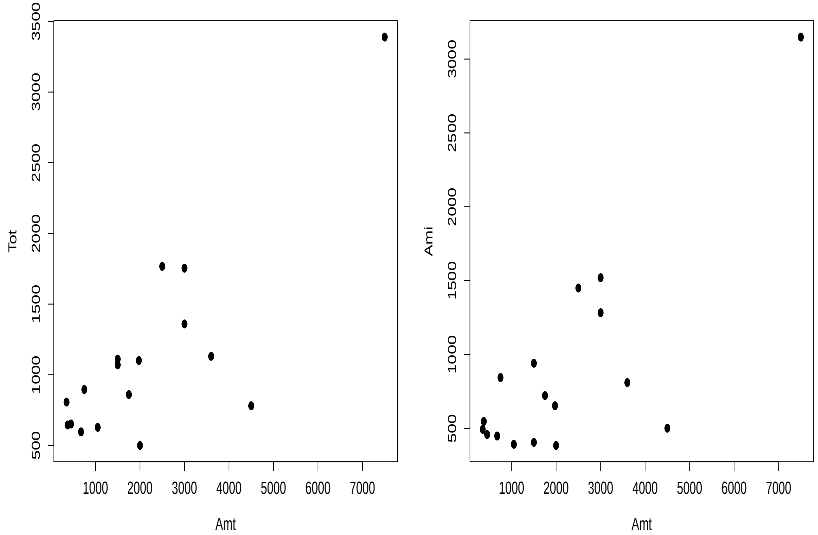
- Tot: nível total no plasma TCAD (ou tricíclicos anti-depressivos - classe de fármacos usados no tratamento sintomático da depressão e outras síndromes depressivas.).
- Ami: quantidade presente de amitriptilina no nível TCAD no plasma.
- Gen: Gênero, 1 - (feminino), 0 - (masculino).
- Amt: quantidade de antidepressivos tomados no momento da overdose.
- Pr: medida da onda PR (eletrocardiograma).
- Diap: Pressão diastólica.
- QRS: medida da onda QRS (eletrocardiograma).



Fonte: http://cal.vet.upenn.edu/projects/lgcardiac/ecg_tutorial/printerval.htm

Modelagem

- Objetivo: modelar o Tot e o Ami em função do Amt.
- Correlação entre Tot e Ami (variáveis - resposta): 0,976.
- Pode-se considerar outras covariáveis.
- Na presente modelagem, torna-se um pouco mais complicado a seleção de covariáveis.


Tot
3500
3000
2500
2000
1500
1000
500
1000
2000
3000
4000
5000
6000
7000
Amt
Ami
3000
2500
2000
1500
1000
500
1000
2000
3000
4000
5000
6000
7000
Amt

Nosso exemplo

- $Y_{ij} = \beta_{0i} + \beta_{1i} x_j + \xi_{ij}, \boldsymbol{\xi}_j = (\xi_{1j}, \xi_{2j}) \sim N_2(\mathbf{0}, \boldsymbol{\Sigma}), i = 1, 2, j = 1, ..., 17.$
- $\boldsymbol{X} = \begin{bmatrix} \mathbf{1}_{(17\times1)} & \boldsymbol{x}_{(17\times1)} \end{bmatrix}$, em que $\boldsymbol{x} = (x_1, x_2, ..., x_{17})'$ (é a mesma para as duas variáveis.
- $\boldsymbol{B} = \begin{bmatrix} \beta_{01} & \beta_{02} \\ \beta_{11} & \beta_{12} \end{bmatrix}$
- β_{0i} : valor esperado de Tot (i=1) ou de Ami (i=2) para uma quantidade nula de Amt.
- β_{1i} : incremento (positivo ou negativo) no valor esperado de Tot (i =1) ou de Ami (i=2) para o aumento em uma unidade na quantidade de Amt ingerido.

Inferência

- Nesse caso a metodologia MANOVA é útil para testar se existe regressão, ou seja, para testar $H_0 : \beta_{11} = \beta_{12} = 0$ vs H_1 : há pelo menos uma diferença.

Estatística	**Valor**	**Aproxim. pela dist. F.**	**p-valor**
Wilks	0,35	13,12	0,0006
Pillai	0,65	13,12	0,0006
Hotelling-Lawley	1,87	13,12	0,0006
Roy	1,87	13,12	0,0006

Assim, rejeita-se H_0.

Testes de hipótese individuais e intervalos de confiança

- Com base nos resultados disponíveis em http://www.ime.unicamp.br/~cnaber/aula_MNLM_Ana_Multi_2S_2017.pdf e http://www.ime.unicamp.br/~cnaber/aula_Intro_MRLM_REG_2S_2016_parte_1.pdf, podemos construir intervalos de confiança de testar hipóteses individuias para os parâmetros de regressão, à semelhança do que é feito para o modelo univariado.

Testes de hipótese individuais e intervalos de confiança

- $\widehat{\beta}_{ki} \sim N(\beta_{ki}, \sigma_i^2 \psi_{ki})$, $\frac{(n-q)\widehat{\sigma}_i^2}{\sigma_i^2} \sim \chi^2_{(n-q)}$ e $\widehat{\beta}_{ki} \perp \frac{(n-q)\widehat{\sigma}_i^2}{\sigma_i^2}$ (k: parâmetro, i: variável), em que ψ_{ki} é um elemento apropriado da diagonal principal da matriz $(\boldsymbol{X}'\boldsymbol{X})^{-1}$.
- Logo, $\frac{\widehat{\beta}_{ki} - \beta_{ki}}{\sqrt{\widehat{\sigma}_i^2 \psi_{ki}}} \sim t_{(n-q)}$, portanto (considerando-se $P(X \leq t_{\frac{1+\gamma}{2}}) = \frac{1+\gamma}{2}, X \sim t_{(n-q)}$), temos que $IC(\beta_{ki}, \gamma) = \left[\widehat{\beta}_{ki} - t_{\frac{1+\gamma}{2}} \sqrt{\widehat{\sigma}_i^2 \psi_{ki}}; \widehat{\beta}_{ki} + t_{\frac{1+\gamma}{2}} \sqrt{\widehat{\sigma}_i^2 \psi_{ki}}\right]$

Testes de hipóteses

- Suponha que queremos testar $H_0 : \beta_{ki} = \beta_{ki0}$ vs $H_1 : \beta_{ki} \neq \beta_{ki0}$, para alguns k e i, em que β_{ki0} é um valor fixado.
- Estatística do teste $T_t = \frac{\widehat{\beta}_{ki} - \beta_{ki0}}{\sqrt{\widehat{\sigma}_i^2 \psi_{ki}}}$, em que $\widehat{\beta}_{ki}$ é o estimador de MQO de β_{ki} e $\widehat{\sigma}_i^2$ é o i-ésimo elemento da diagonal principal da matriz $\widehat{\boldsymbol{\Sigma}}$.

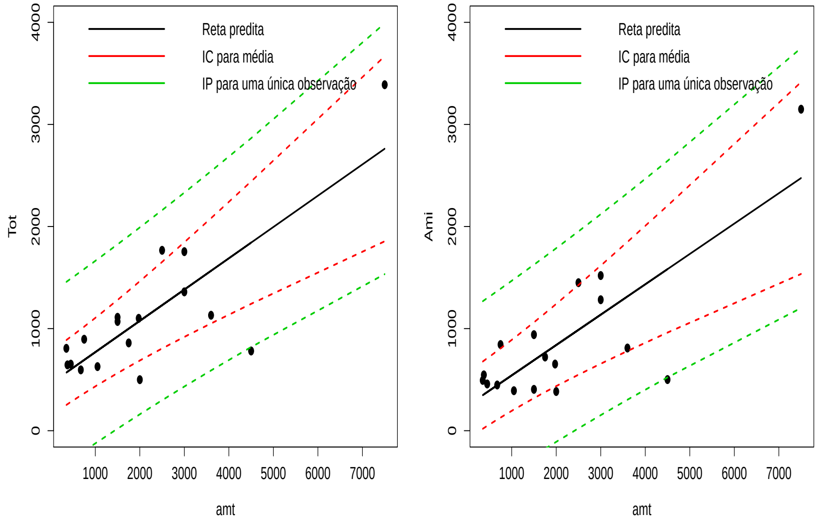
Cont.

Variável-resposta: Tot

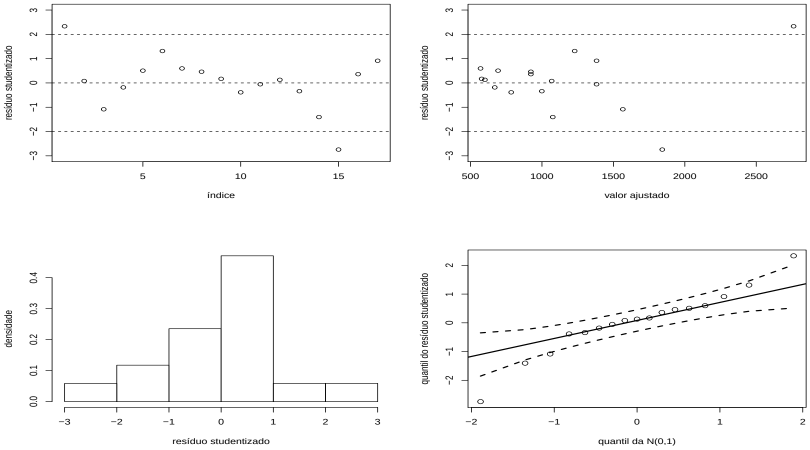
Parâmetro	Estimativa	EP	IC(95%)	Estat. t	p-valor
β_{01}	462,89	160,91	[119,92 ; 805,86]	2,88	0,0115
β_{11}	0,31	0,06	[0,18 ; 0,43]	5,30	0,0001

Variável-resposta: Ami

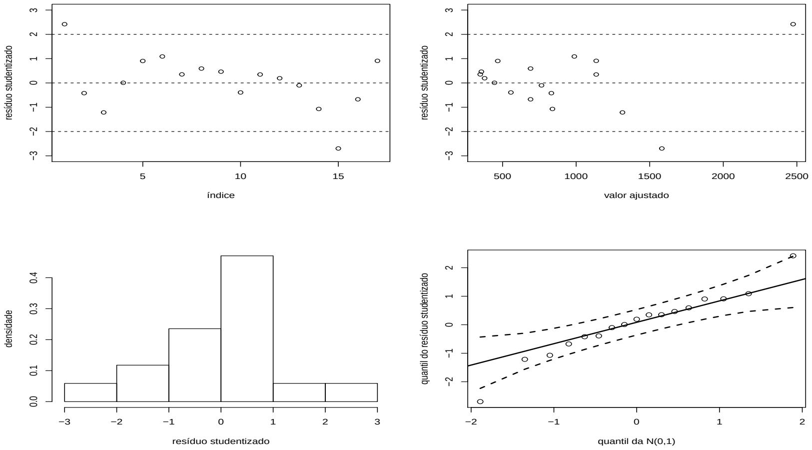
Parâmetro	Estimativa	EP	IC(95%)	Estat. t	p-valor
β_{02}	244,31	166,76	[-111,13 ; 599,76]	1,47	0,1636
β_{12}	0,30	0,06	[0,17 ; 0,43]	4,96	0,0002


Reta predita
IC para média
IP para uma única observação
Tot
amt
0
1000
2000
3000
4000
5000
6000
7000
Reta predita
IC para média
IP para uma única observação
Ami
amt

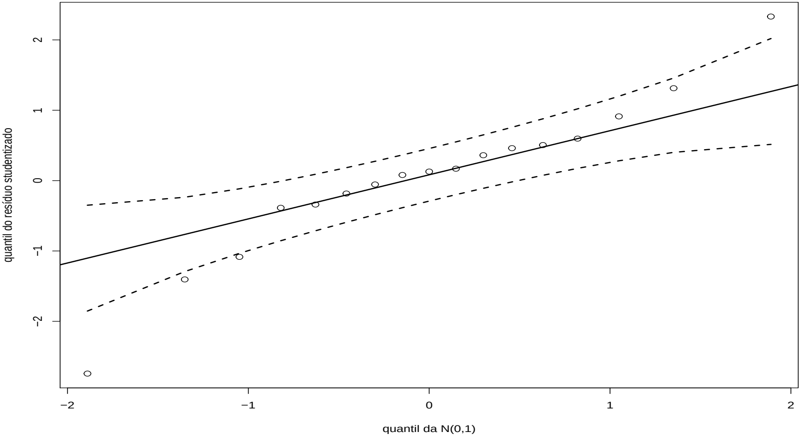
Análise de resíduos (RS) - Tot



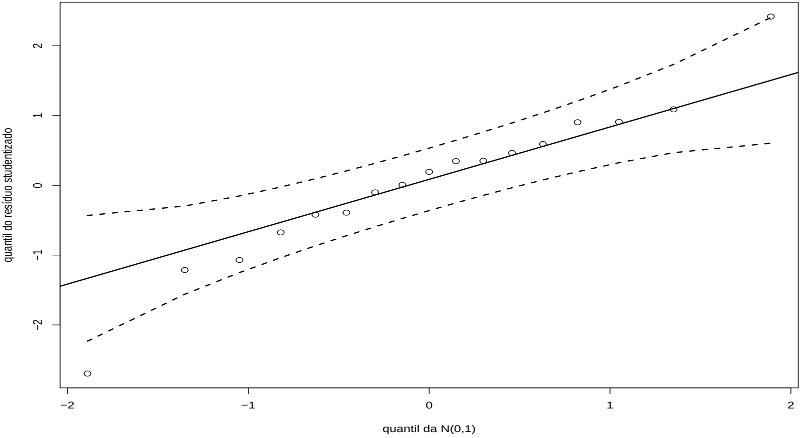
Análise de resíduos (RS) - Amt



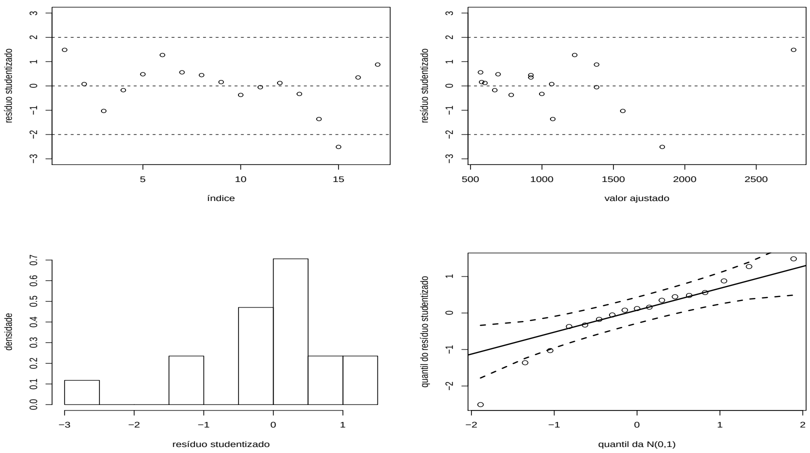
Envelopes (RS) - Tot



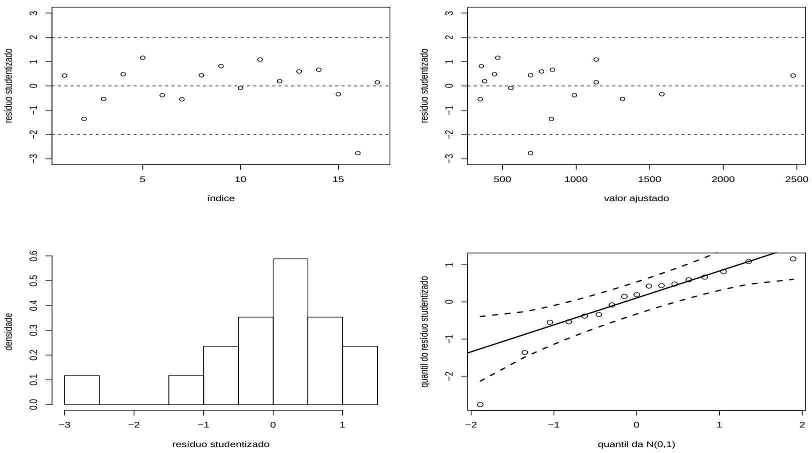
Envelopes (RS) - Amt



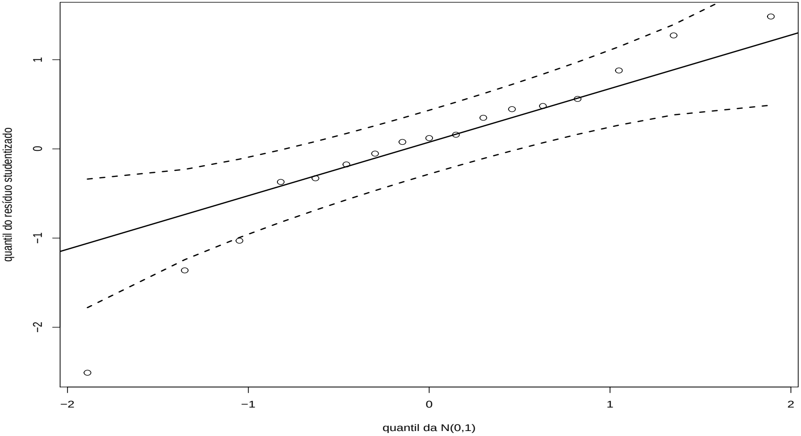
Análise de resíduos (RSM) - Tot



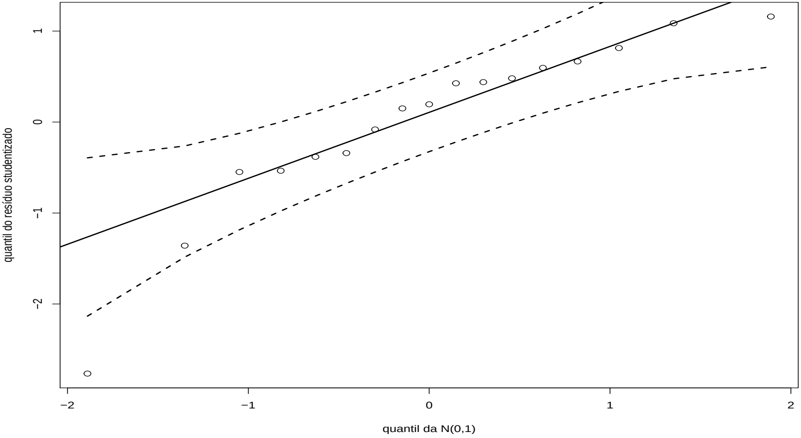
Análise de resíduos (RSM) - Amt



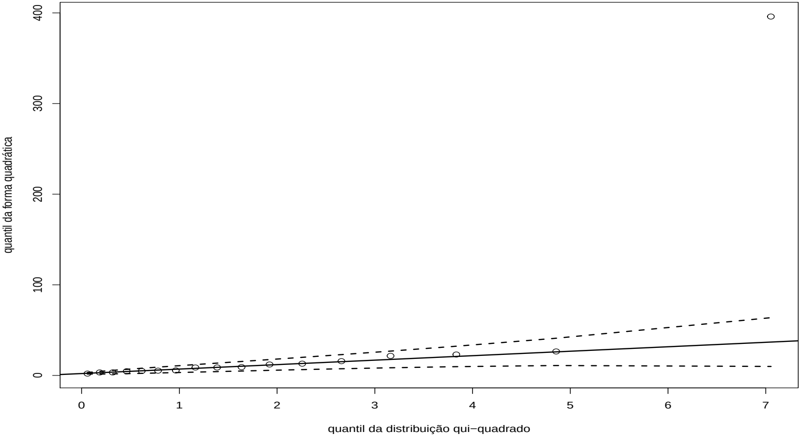
Envelopes (RSM) - Tot



Envelopes (RSM) - Amt



Envelopes (RSM) - distância de Mahalanobis



- Teste para igualdade dos coeficientes angulares: $H_0 : \beta_{11} = \beta_{12}$ vs $H_1 : \beta_{11} \neq \beta_{12}$.
- Teste $\boldsymbol{CBU} = \boldsymbol{M}$, $\boldsymbol{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}$, $\boldsymbol{U} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\boldsymbol{M} = 0$.

  Resultado: $q_c = 0,18$(0,6682). Não se rejeita a igualdade.
- Veja também a função "linearHypothesis" do pacote "car".
- Exercício: ajustar um modelo quadrático e compará-lo com o modelo linear.