

# Questão 10

## Introdução

Em uma amostra de 71 coelhos europeus da espécie *Oryctolagus Cuniculuse*, se mediu o peso da lente ocular seca (em mg) e a idade do animal (em dias). Temos como objetivo explicar o peso da lente em função da idade. Para isso ajustaremos dois modelos, onde a variável resposta é o peso da lente e a explicativa a idade. Ajustaremos um modelo normal não linear homocedástico e iremos comparar seu resultado com o ajuste de um modelo linear normal homocedástico com transformação da variável explicativa. Para mais informação sobre o ajuste e diagnóstico desses modelos veja Azevedo (2019). Todas as análises serão feitas com auxílio computacional do R.

## Análise Descritiva

Pelo gráfico de dispersão, figura 1, podemos observar uma relação positiva entre o peso e a idade. Esta relação porém aparente não ser linear, sendo melhor ajustada por uma função logarítmica ou por uma razão.

## Análise Inferencial

O modelo linear foi ajustado segundo o método dos mínimos quadrados ordinários e o modelo não linear foi ajustado segundo o método iterativo de Gauss-Newton. Seja  $Y_i$  o peso da lente ocular do  $i$ -ésimo coelho temos dois modelos.

- Modelo linear:

$$Y_i = \beta_0 + \beta_1 \log x_i + \varepsilon_i$$

Onde  $\varepsilon_i$  é a parte aleatória,  $x_i$  é a idade do  $i$ -ésimo coelho,  $\beta_0$  é o peso esperado da lente ocular do coelho quando a idade é igual a 1 e  $\beta_1$  é o incremento esperado no peso quando  $\log x_i$  aumenta em uma unidade. Porém esse modelo é problemático quando temos  $x_i \leq 1$ , resultando em valores negativos do peso.

- Modelo linear:

$$Y_i = \frac{\phi_0 * x_i}{\phi_1 + x_i} + \varepsilon_i$$

Onde  $\varepsilon_i$  é a parte aleatória,  $x_i$  é a idade do  $i$ -ésimo coelho, porém por causa da forma da equação perdemos a interpretação dos parâmetros  $\phi_0$  e  $\phi_1$ .

Os resultados dos ajustes dos modelos podem ser vistos nas tabelas 1 e 2. Todos os parâmetros foram considerados significativos. Como notamos o modelo linear com transformação na variável apresenta problemas para valores pequeno de  $x_i$

Tabela 1: Estimativa para os parâmetros do modelo não linear

Termo	Estimativa	Erro Padrão	IC (95%)	Estatística t	p-valor
$\phi_0$	295,746	4,799	[ 286,341 ; 305,152 ]	61,629	< 0.001
$\phi_1$	168,382	7,019	[ 154,623 ; 182,14 ]	23,988	< 0.001

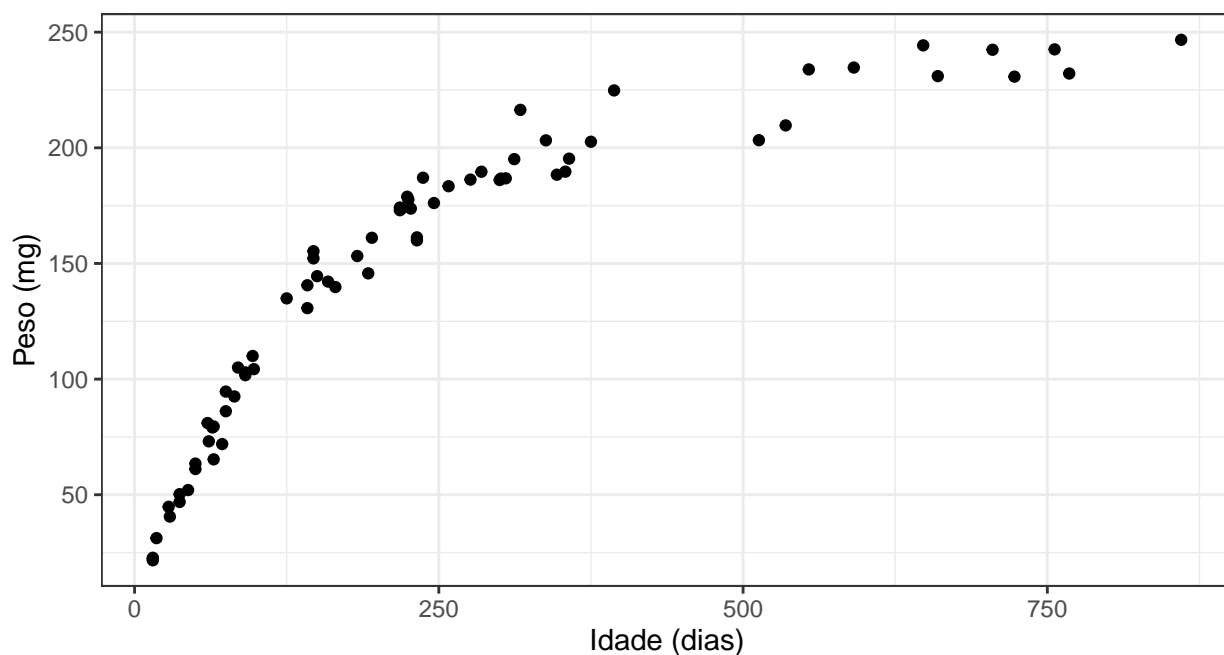


Figura 1: Gráfico de dispersão do peso pela idade.

Tabela 2: Estimativa para os parâmetros do modelo linear com transformação.

Termo	Estimativa	Erro Padrão	IC (95%)	Estatística t	p-valor
$\beta_0$	-169,137	6,851	[ -182,565 ; -155,71 ]	-24,689	< 0.001
$\beta_1$	62,167	1,327	[ 59,567 ; 64,768 ]	46,854	< 0.001

Como podemos observar pela figura 2 o modelo não linear pareceu se ajustar melhor aos dados. Pela análise de resíduos nas figuras 3 e 4 as suposições de normalidade dos dados e de homoscedasticidade não parecem corretas. No gráfico dos resíduos pelos valores ajustados vemos um aumento na variância, que indica heterocedasticidade. Os gráficos quantis-quantis e histogramas dos resíduos mostram uma assimetria, que indica não normalidade. Porém no quesito de normalidade o modelo não linear se aproxima muito mais de uma normal que o linear. Para as medidas de ajuste do modelo, vistas na tabela 3, o modelo não linear resultou em melhores valores, que indica um melhor ajuste.

## Conclusão

O modelo não linear quebrou algumas suposições mas se ajustou melhor aos dados que a alternativa linear. Uma sugestão seria ajustar um modelo não normal, com alguma distribuição que comportasse melhor essa variância. A partir desse modelo podemos concluir que existe uma relação positiva entre o peso da lente ocular do coelho e sua idade. Porém ela não é linear e conforme o coelho envelhece a variação na idade resulta em uma menor variação no peso.

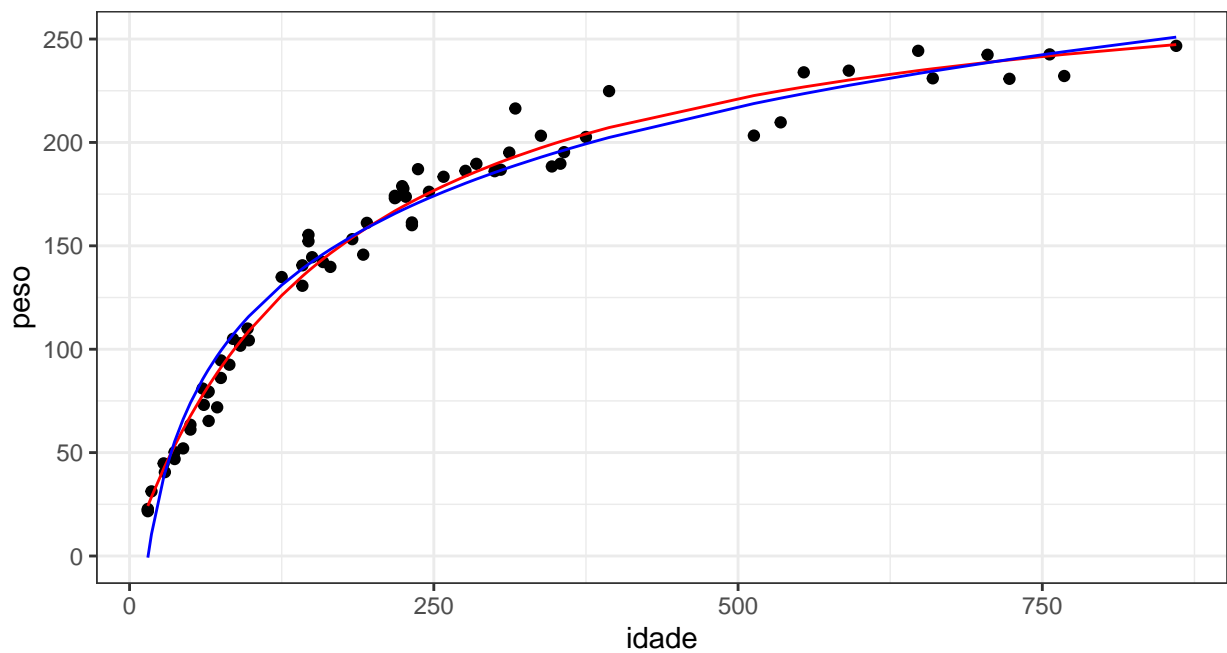


Figura 2: Gráfico de dispersão do peso pela idade com os valores preditos. Modelo linear em azul e não linear em vermelho

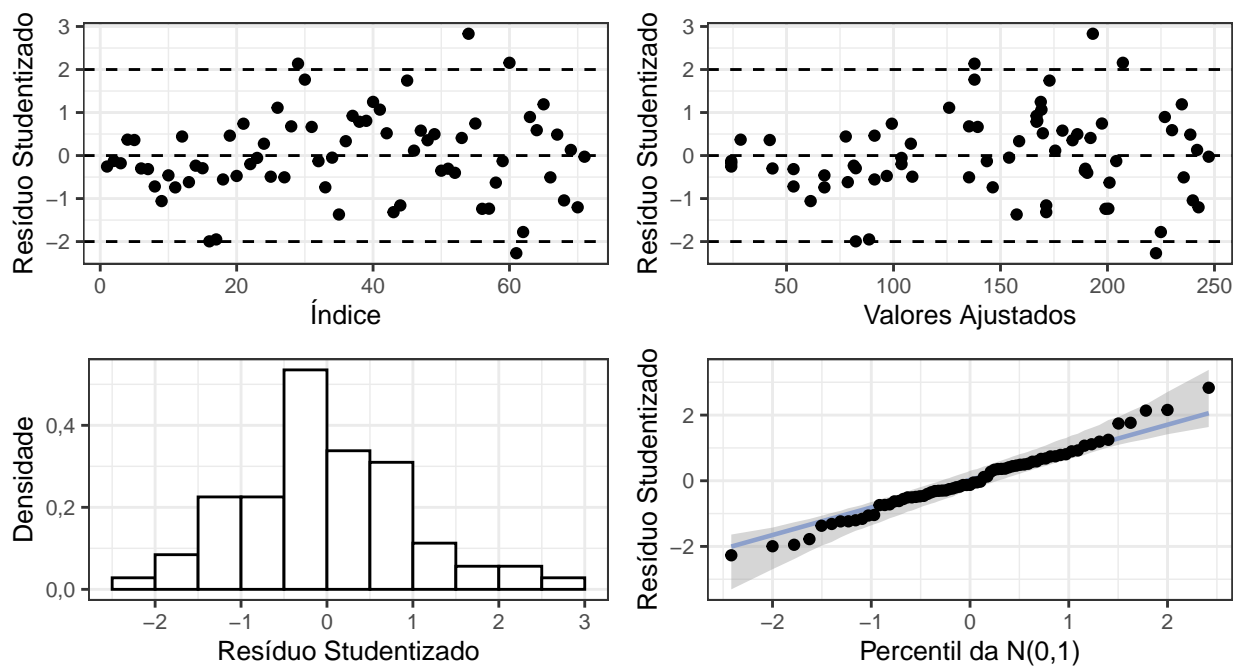


Figura 3: Gráfico de resíduos para o modelo não linear

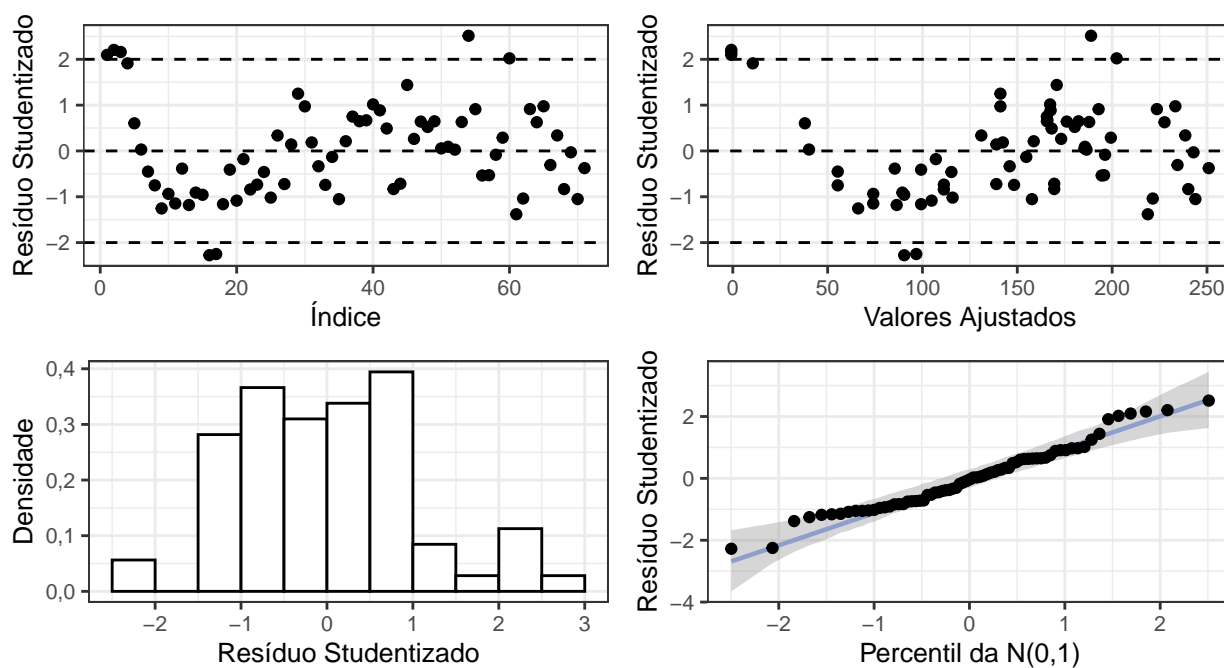


Figura 4: Gráfico de resíduos para o modelo linear

Tabela 3: Medidas de comparação do modelo.

	Linear	Não Linear
AIC	552,0755	506,8978
BIC	558,8635	513,6858
AICc	552,2520	507,0743
SABIC	548,3003	503,1226
HQCIC	551,8751	506,6974
-2log.lik	546,0755	500,8978

## Referências

- Azevedo, C. L. N (2019). Notas de aula sobre Análise de regressão, [http://www.ime.unicamp.br/~cnaber/Material\\_ME613\\_1S\\_2019.htm](http://www.ime.unicamp.br/~cnaber/Material_ME613_1S_2019.htm)
- Paula, G. A. (2013). Modelos de regressão com apoio computacional, versão pré-eliminar, [https://www.ime.usp.br/~giapaula/texto\\_2013.pdf](https://www.ime.usp.br/~giapaula/texto_2013.pdf)