




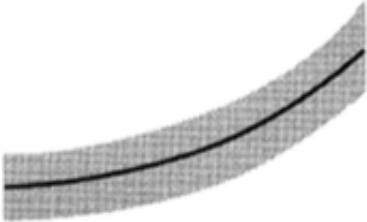
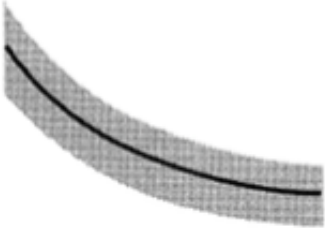
ME613 - Análise de Regressão

Parte 4

Benilton S Carvalho - 1S2020

Transformações

Transformações: relações não lineares

	Prototype Regression Pattern	Transformations of X
(a)		$X' = \log_{10} X$ $X' = \sqrt{X}$
(b)		$X' = X^2$ $X' = \exp(X)$
(c)		$X' = 1/X$ $X' = \exp(-X)$

Transformações: relações não lineares

- Linearizar uma relação linear, em casos que a suposição de normalidade dos erros com variância constante é adequada .
- Tentar encontrar transformação em X .
- Avaliar os gráficos de resíduo para decidir qual a melhor transformação.

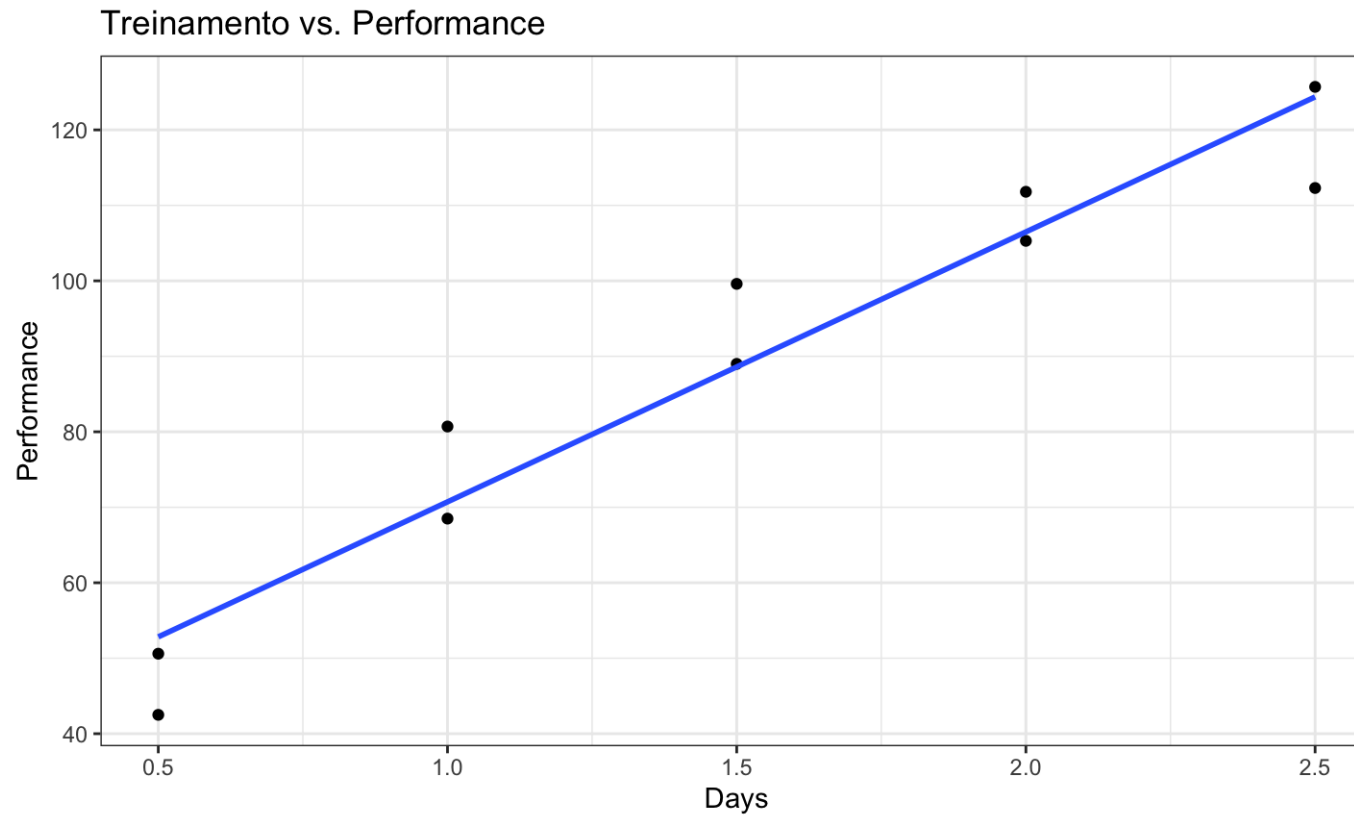
Exemplo: Treinamento e Performance

X : número de dias de treinamento recebido.

Y : performance nas vendas.

Parte 1		Parte 2	
Days	Performance	Days	Performance
0.5	42.5	1.5	99.6
0.5	50.6	2.0	105.3
1.0	68.5	2.0	111.8
1.0	80.7	2.5	112.3
1.5	89.0	2.5	125.7

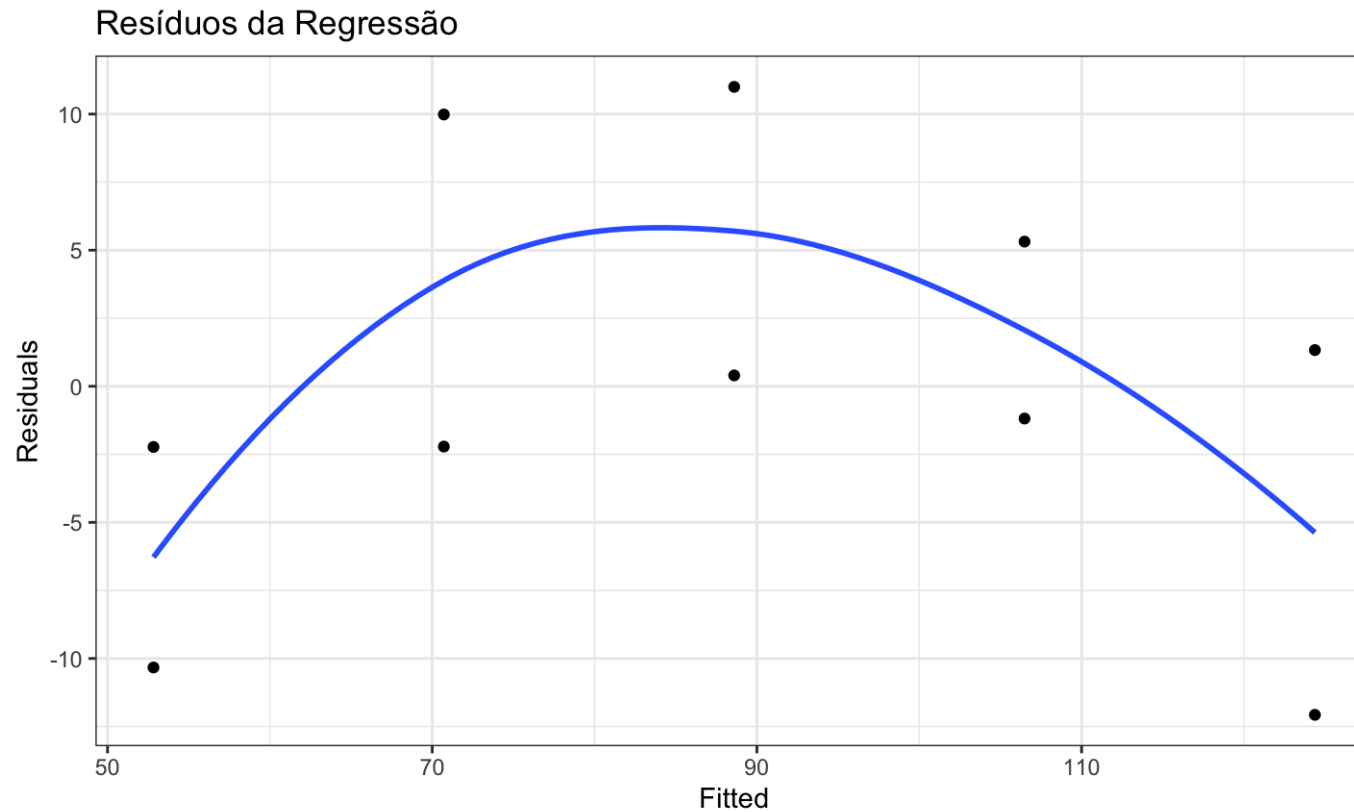
Exemplo: Treinamento e Performance



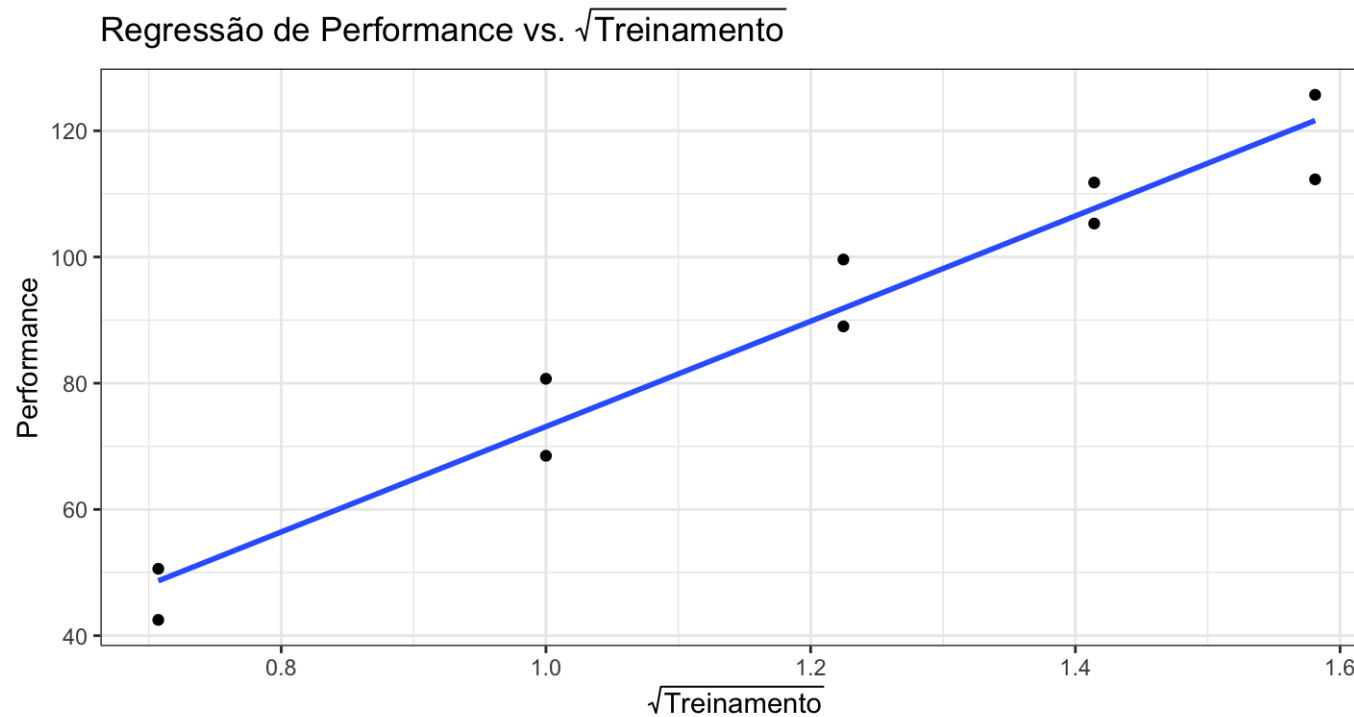
$$\hat{Y} = 34.945 + 35.77X$$

Exemplo: Treinamento e Performance

Gráfico de resíduos: e_i versus \hat{Y}_i .



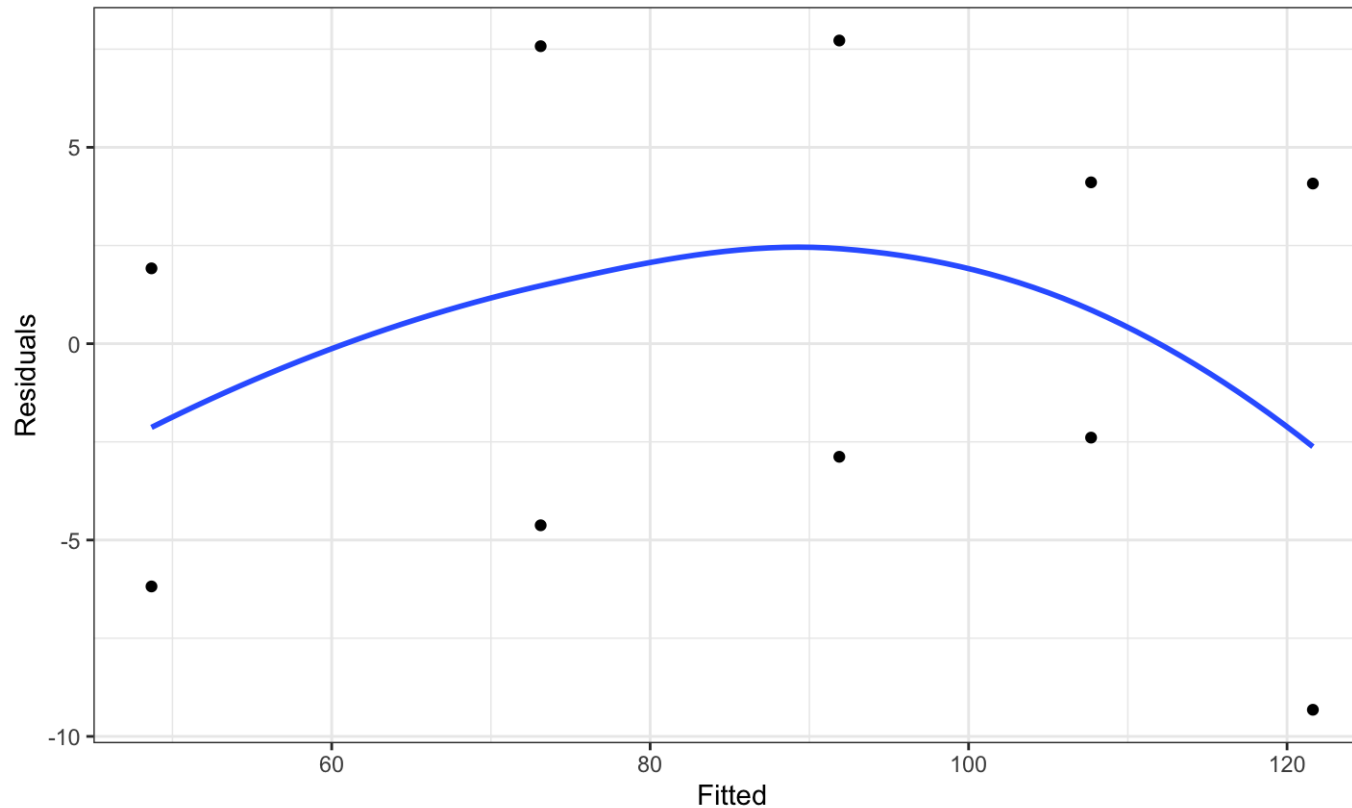
Exemplo: Treinamento e Performance



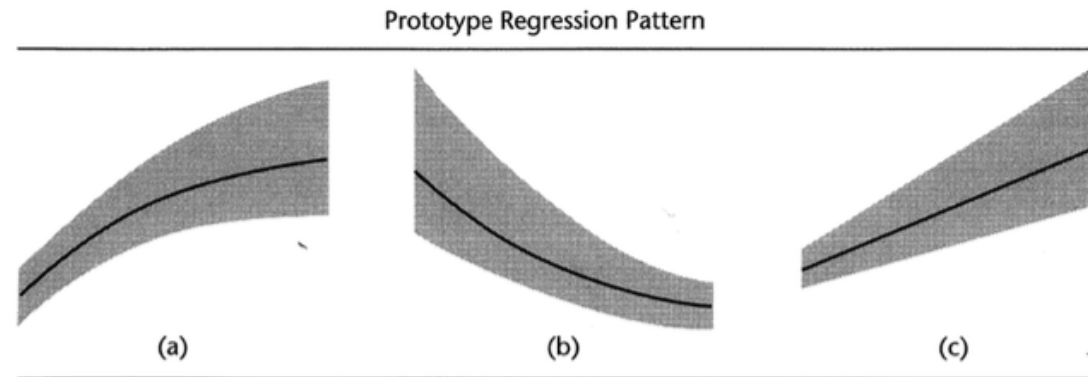
$$\hat{Y} = -10.33 + 83.45\sqrt{X}$$

Exemplo: Treinamento e Performance

Gráfico de resíduos: e_i versus \hat{Y}_i usando $\sqrt{\text{Treinamento}}$ no modelo.



Transformações: não normalidade e variância não constante



Transformações: não normalidade e variância não constante

- Tentar encontrar transformação em Y .
- Pode ser combinada com uma transformação também em X .

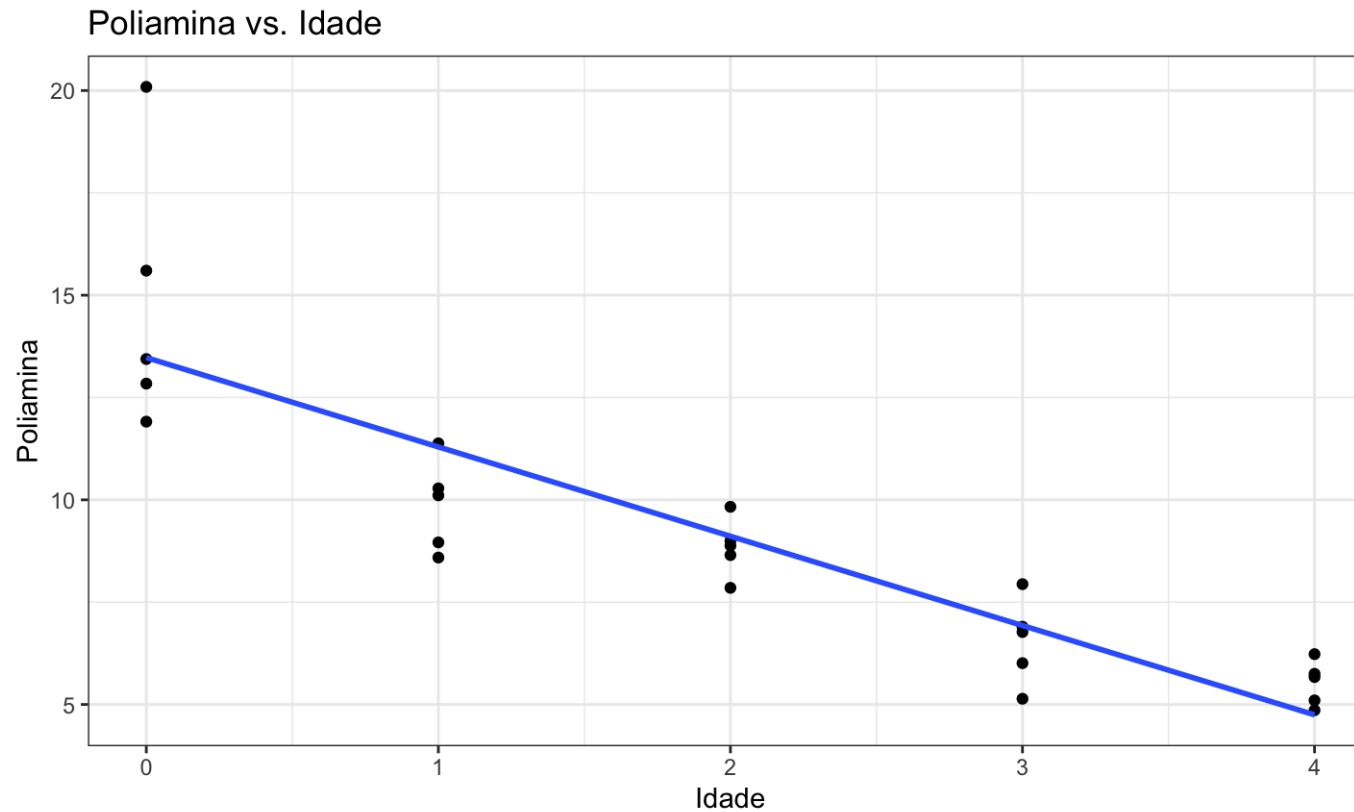
Exemplo: Idade e nível de poliamina

X : Idade

Y : nível de poliamina no plasma

Parte 1		Parte 2		Parte 3		Parte 4	
Idade	Poliamina	Idade	Poliamina	Idade	Poliamina	Idade	Poliamina
0	13.44	1	10.28	2	8.88	4	5.10
0	12.84	1	8.96	3	7.94	4	5.67
0	11.91	1	8.59	3	6.01	4	5.75
0	20.09	2	9.83	3	5.14	4	6.23
0	15.60	2	9.00	3	6.90	NA	NA
1	10.11	2	8.65	3	6.77	NA	NA
1	11.38	2	7.85	4	4.86	NA	NA

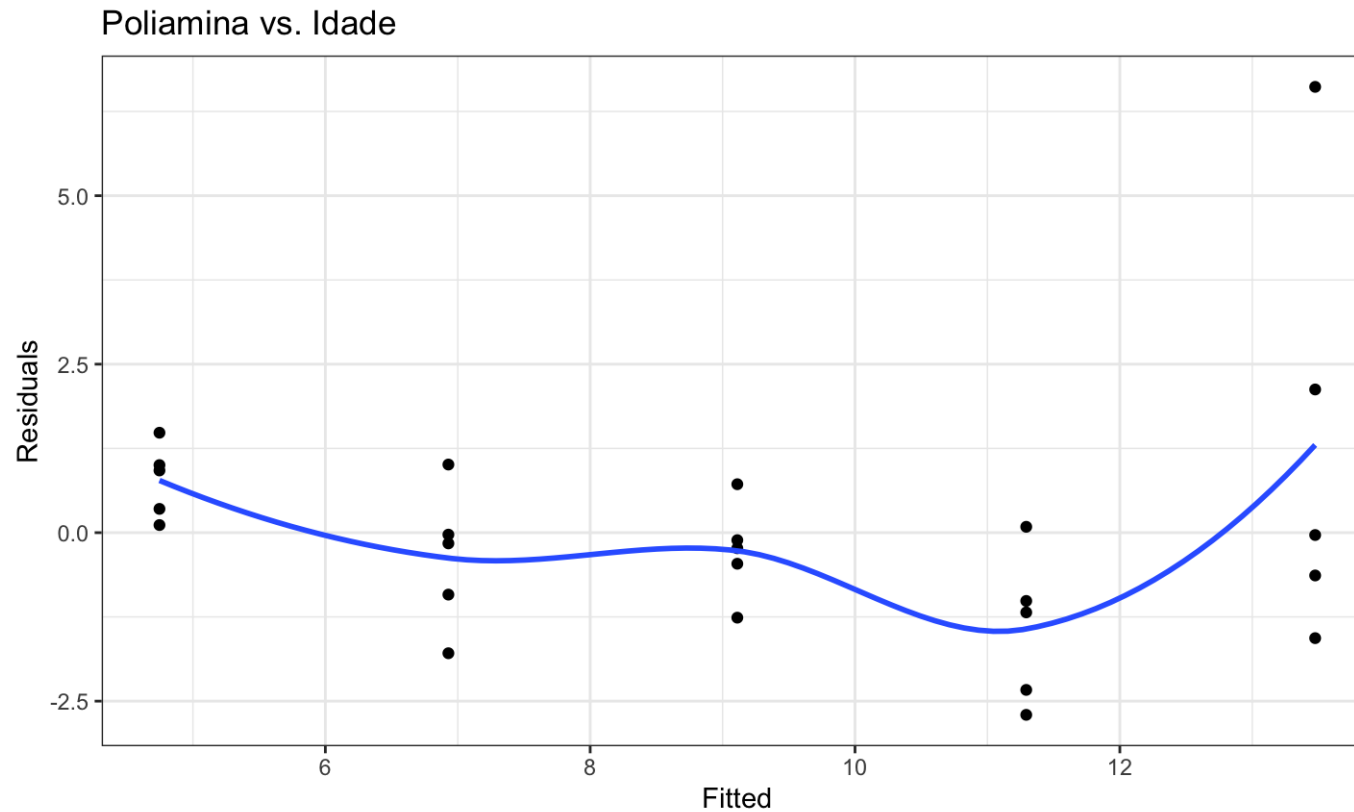
Exemplo: Idade e nível de poliamina



$$\hat{Y} = 13.4752 - 2.182X$$

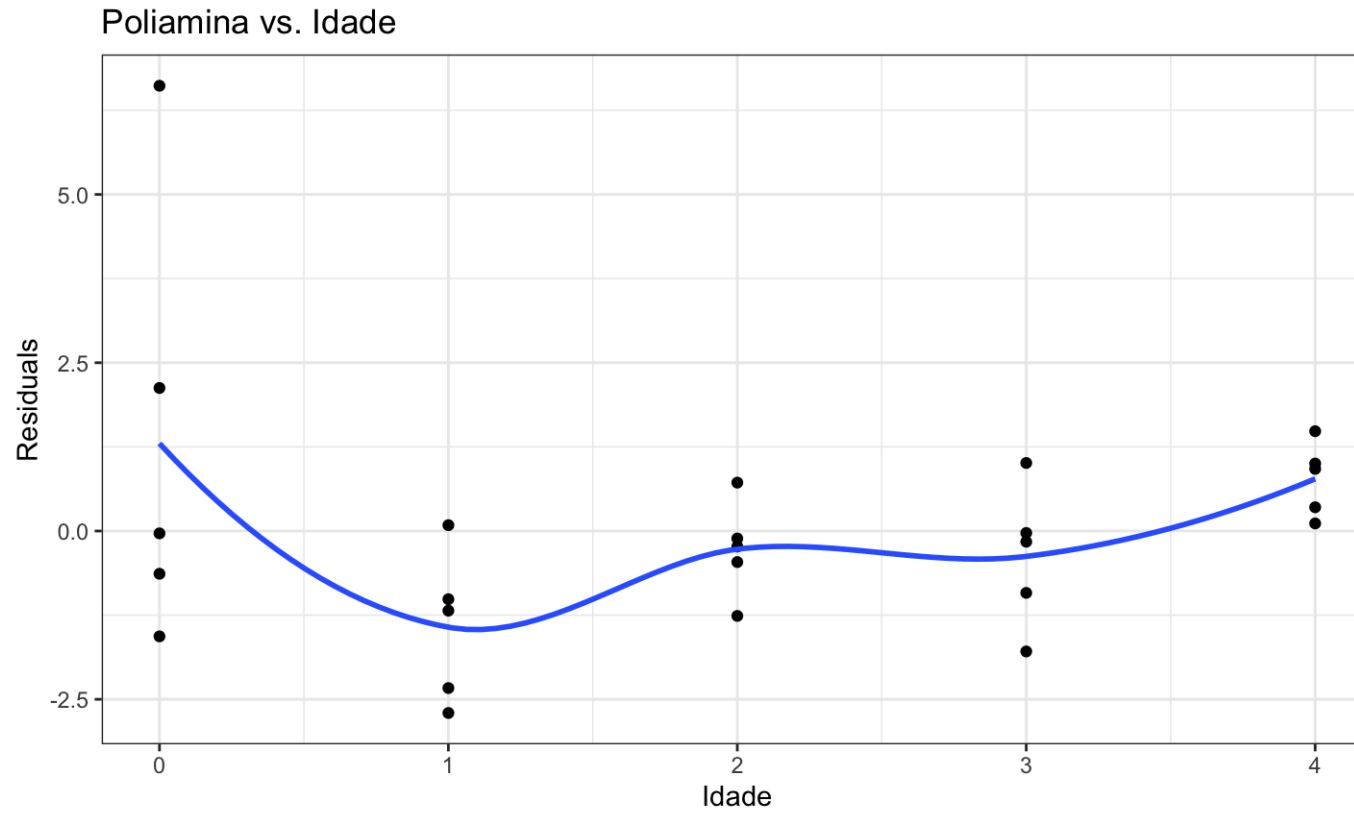
Exemplo: Idade e nível de poliamina

Gráfico de resíduos: e_i versus \hat{Y}_i .



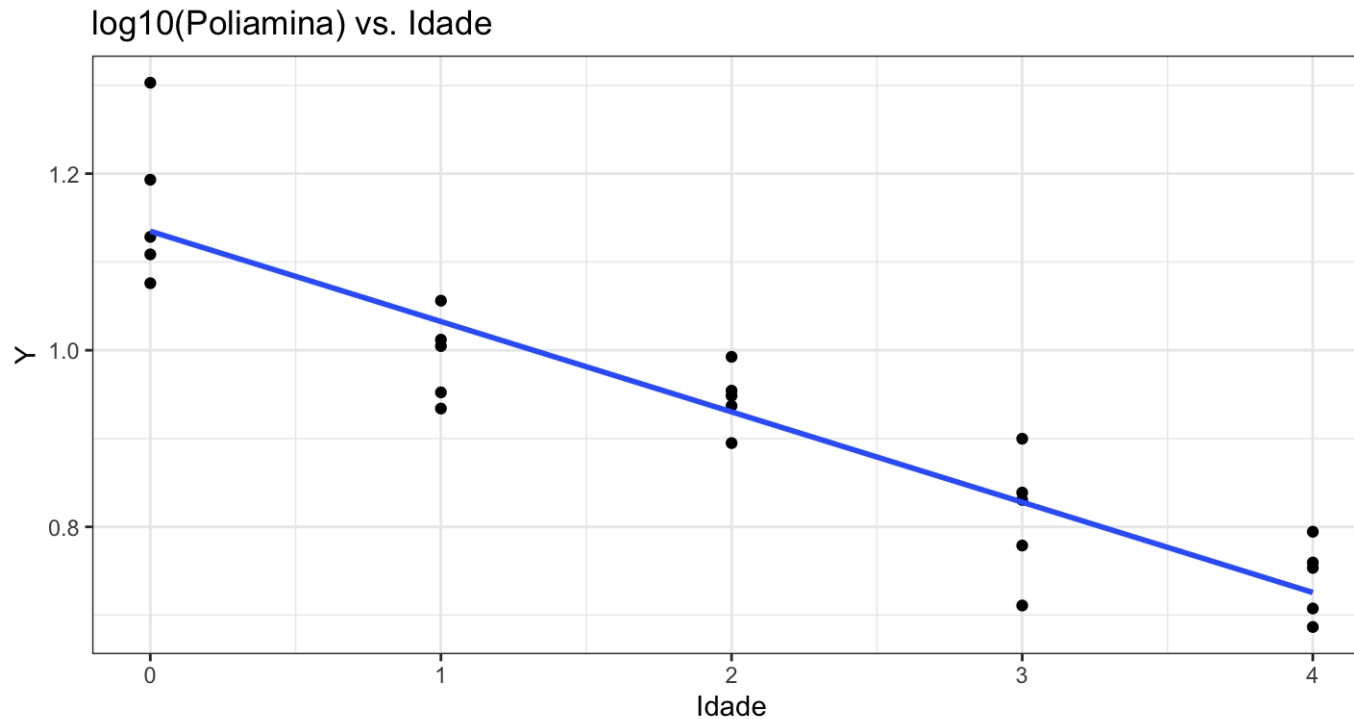
Exemplo: Idade e nível de poliamina

Gráfico de resíduos: e_i versus X_i .



Exemplo: Idade e nível de poliamina

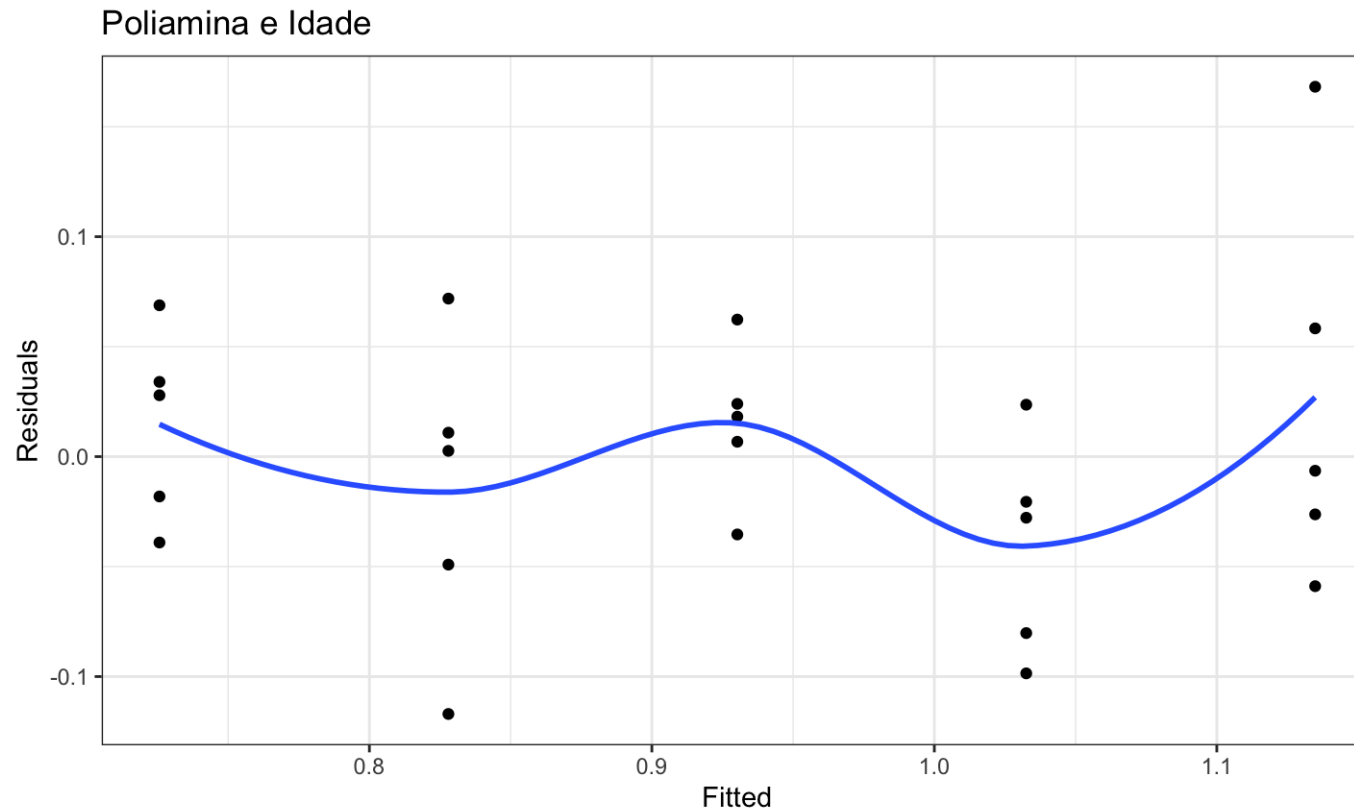
$$Y' = \log_{10} Y$$



$$\hat{Y}' = 1.13 - 0.1X$$

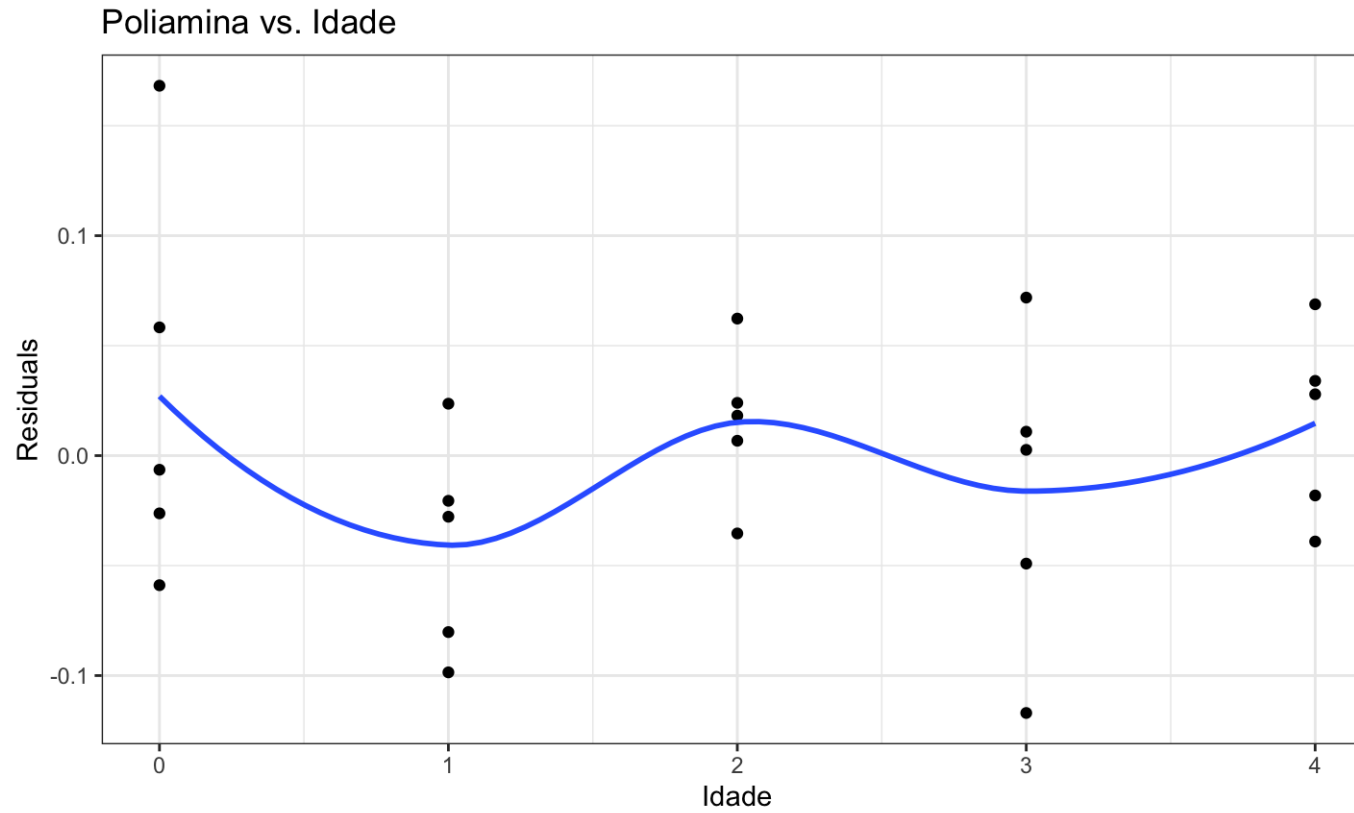
Exemplo: Idade e nível de poliamina

Gráfico de resíduos: e_i versus \hat{Y}'_i usando Y' no modelo.



Exemplo: Idade e nível de poliamina

Gráfico de resíduos: e_i versus X_i .



Algumas transformações em Y

- $\log_e(Y)$: para estabilizar a variância quando esta tende a crescer à medida que Y cresce.
- \sqrt{Y} : estabilizar a variância quando esta é proporcional à média dos Y 's.
- $\frac{1}{Y}$: estabilizar a variância, minimizando o efeito de valores muito altos de Y .
- Y^2 : estabilizar a variância quando esta tende a decrescer com a média de Y 's.
- $\arcsin \sqrt{Y}$: estabilizar a variância quando os dados são proporções.
- etc...

Transformações de Box-Cox

- Muitas vezes é difícil determinar, através de gráficos, qual a melhor transformação a ser feita.
- O procedimento de [Box-Cox](#) identifica automaticamente uma transformação:

$$Y' = Y^\lambda$$

em que λ é um parâmetro a ser determinado a partir dos dados.

- Modelo com dados transformados:

$$Y_i^\lambda = \beta_0 + \beta_1 X_i + \varepsilon_i$$

- O procedimento de Box-Cox utiliza o método de máxima verossimilhança para estimar λ .

Agradecimento

- Slides criados por Samara F Kiihl / IMECC / UNICAMP

Leitura

- Applied Linear Statistical Models: 3.8-3.11.
- Weisberg - [Applied Linear Regression](#): Capítulo 8.
- Faraway - [Linear Models with R](#): Capítulo 9.

