Modelos de Regressão Múltipla - Parte VI

Erica Castilho Rodrigues

7 de Fevereiro de 2017

Transformação de Variáveis

- Transformação para estabilizar a variância

Transformação para corrigir não-linearidade

- Podemos fazer uma transformação na variável resposta Y e/ou na preditora X para:
  - solucionar problemas de variância não constante;
  - não normalidade dos erros;
  - não linearidade do modelo.
- Violações que geralmente ocorrem ao mesmo tempo:
  - variância constante e distribuição normal.
- Suposições básicas

$$Y_i \text{ tem distribuição normal} \quad Var(Y_i) = \sigma^2.$$

Transformação para estabilizar a variância

- Frequentemente a variância das observações varia com sua média.
- Situações comuns:

$$Y \text{ são contagens} \Rightarrow Var(Y) \propto E(Y)$$

$$Y \text{ são proporções} \Rightarrow Var(Y) \propto E(Y)(1 - E(Y))$$

- Uma solução: fazer transformações em Y.
- Vimos que uma outra possibilidade é usar Mínimos Quadrados Generalizdos.

- A tabela a seguir mostra:
  - tipos de relação mais comuns entre variância e esperança
  - tipo de transformação geralmente adequada.

Relação entre $Var(Y)$ e $E(Y)$	Transformação
$Var(Y) \propto E(Y)$	$Y^* = \sqrt{Y}$
$Var(Y) \propto E(Y)(1 - E(Y))$	$Y^* = \arcsin Y$
$Var(Y) \propto E(Y)^2$	$Y^* = \ln Y$
$Var(Y) \propto E(Y)^3$	$Y^* = 1/\sqrt{Y}$
$Var(Y) \propto E(Y)^4$	$Y^* = 1/Y$

- Vejamos com mais detalhes qual transformação é adequada para situações específicas.

$\sqrt{Y}$

- É usada se a variância cresce um pouco com o valor de Y.
- Mais especificamente se

$$Var(Y) \propto E(Y)\,.$$

- Observe que Y tem que ser positiva nesse caso.
- Essa transformação é adequada se a variável resposta é uma contagem.
- Muito usada quando Y segue uma distribuição de Poisson.

$\log(Y)$

- Usada para estabilizar a variância quando ela cresce mais rapidamente com Y.
- Mais espeficicamente se

$$Var(Y) \propto E(Y)^2 \quad \text{ou} \quad \sqrt{Var(Y)} \propto E(Y)$$

- Observe que Y precisa ser positiva.
- É usada quando intervalo de variação de Y é muito grande, de 1 a 10000 por exemplo.

$1/Y$

- Usada para estabilizar a variância quando ela cresce muito com Y
$$Var(Y) \propto E(Y)^4 .$$
- Pode ser usada, por exemplo, quando muitos valores de Y ficam próximos de zero e alguns são muito grandes.
- Os efeitos dos valores grandes são reduzidos nesse caso.
- Se Y_i é muito grande então $1/Y_i$ é muitro próximo de zero.
- Considere dois valores Y_i e Y_j muito distantes.
- Essa diferença se torna pequena quando pegamos $1/Y_i$ e $1/Y_j$.
- Essa transformação é geralmente adequada quando Y representa tempo de espera até um evento (modelo exponencial).

$\arcsin(Y)$

- É usada se Y uma proporção ou uma taxa.
- Nesse caso temos que

$$Var(Y) \propto E(Y)(1 - E(Y))\,.$$

- Como Y é uma proporção temos que

$$Y_i \in [0, 1] \quad \forall i\,.$$

Observações:

- Se Y atinge valor negativo, não podemos usar as transformações
$$\log(Y) \quad \sqrt{Y}\,.$$
- Uma solução é somar uma pequena constante c em Y
$$\log(Y+c) \quad \sqrt{Y+c}\,.$$
- Essa constante tem que ser pequena comparada com o intervalo dos dados.
- Se não for possível encontrar essa constante, deve ser feita outra transformação.

- As tranformações

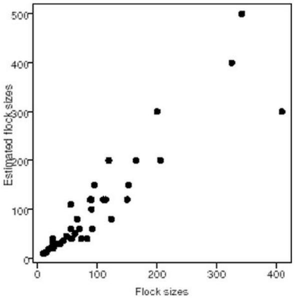
$$\log(Y) \quad \sqrt{Y} \quad 1/Y$$

  são adequadas quanco a variância cresce com Y.
- Mas cada uma delas é gradualmente mais forte que a anterior.
- A transformação raiz é a mais sutil.
- A transformação inversa é a mais forte de todas.
- A transformação mais usada é o logaritmo.

Exemplo:

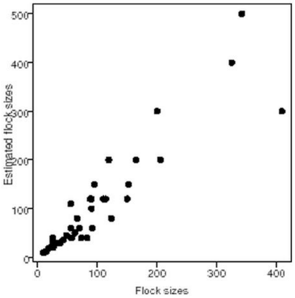
- Deseja-se estimar o número de gansos de neve em uma área do Canadá.
- Um avião sobrevoa a área e um observador estima o tamanho dos bandos.
- Queremos estimar a precisão desse método.
- Dois observadores resgistraram o número de gansos em 45 bandos.
- Além disso, uma foto foi tirada de cada um dos bandos.
- A imagem fornece o número exato de gansos.
- O objetivo é verificar qual a relação entre o valor real do número de gansos e o valor estimado.

Exemplo: (continuação)



- A figura ao lado mostra:
  - o gráfico de dispersão entre os valores reais e os valores estimados por um dos pesquisadores.
- Um modelo linear usual parece adequado nesse caso?

Exemplo: (continuação)

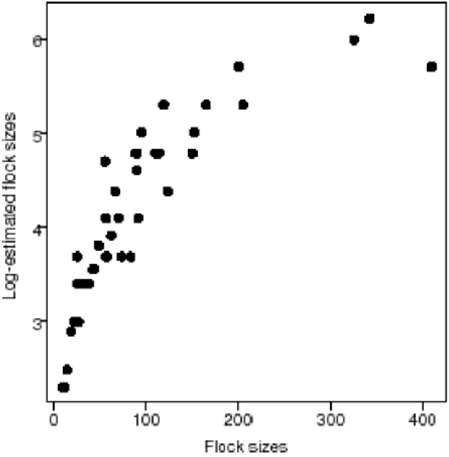


- A figura ao lado mostra:
  - o gráfico de dispersão entre os valores reais e os valores estimados por um dos pesquisadores.
- Um modelo linear usual parece adequado nesse caso? Não.
- A relação entre as variáveis parece linear.
- Nota-se, porém, que a variância aumenta com o valor de X.

Exemplo: (continuação)

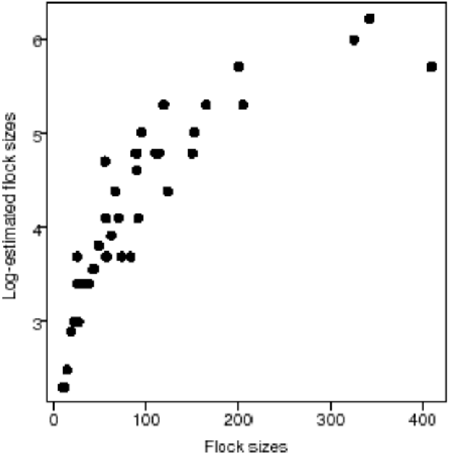
- A variância está aumentando muito com Y.
- Precisamos de uma transformação que reduza esse efeito.
- Vamos usar o logaritmo.
- A figura seguir mostra o gráfico de dispersão entre as variáveis após a transformação.

Exemplo: (continuação)



- O problema parece ter sido resolvido?

Exemplo: (continuação)



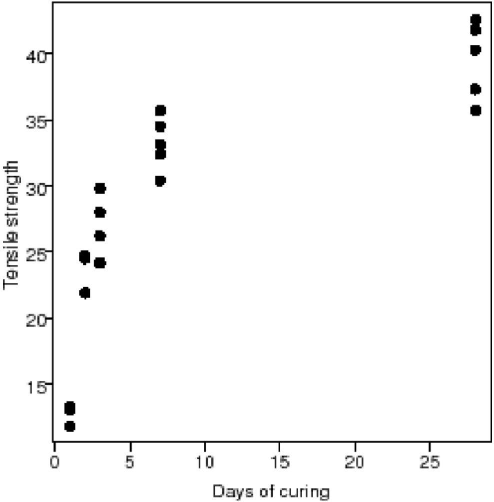
- O problema parece ter sido resolvido?
- Sim. A variância não parece estar mais aumentando com X.

Exemplo: (continuação)

- A relação se afastou um pouco da linearidade.
- Porém, no momento, estamos preocupados apenas em tornar a variância constante.
- Esse tipo de transformação é conhecida como:
  - transformação para estabilização da variância.
- Estudos anteriores podem indicar a transformação adequada.
- Mas em geral, essa escolha é feita com base nos:
  - gráficos de dispersão;
  - gráficos de resíduos.

Exemplo

- O tempo de cura do cimento afeta sua resistência.
- A resistência de lotes de cimento com tempo de curas distintos foram mensuradas.



Exemplo (continuação)

- A relação entre as variáveis é linear?

Exemplo (continuação)

- A relação entre as variáveis é linear?
- Não, se aproxima mais de uma curva.
- A variância é constante?

Exemplo (continuação)

- A relação entre as variáveis é linear?
- Não, se aproxima mais de uma curva.
- A variância é constante?
- Não, ela está aumentando com Y.
- Como a variância não está aumentando muito, não precisamos de uma transformação muito forte.
- Qual podemos pegar?

Exemplo (continuação)

- A relação entre as variáveis é linear?
- Não, se aproxima mais de uma curva.
- A variância é constante?
- Não, ela está aumentando com Y.
- Como a variância não está aumentando muito, não precisamos de uma transformação muito forte.
- Qual podemos pegar? Log.
- A figura a seguir mostra o gráfico de dispersão entre as variáveis quando pegamos o log da variável resposta.

Exemplo (continuação)

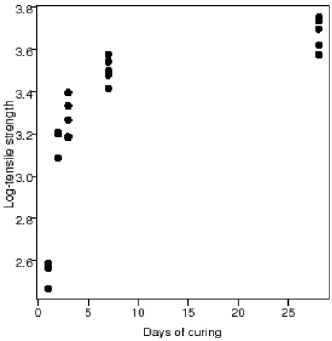


Figura: Gráfico do logaritmo da resistência do cimento em função do tempo de cura.

- A transformação alcançou objetivo pretendido?

Exemplo (continuação)

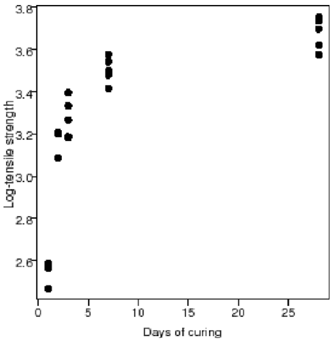


Figura: Gráfico do logaritmo da resistência do cimento em função do tempo de cura.

- A transformação alcançou objetivo pretendido?
- Sim, a variância fica aproximadamente constante.
- A relação agora é linear?

Exemplo (continuação)

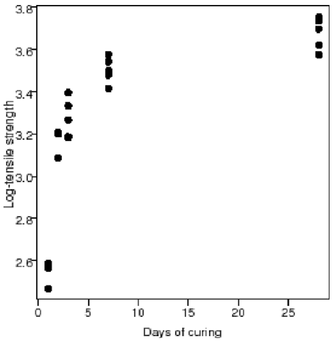


Figura: Gráfico do logaritmo da resistência do cimento em função do tempo de cura.

- A transformação alcançou objetivo pretendido?
- Sim, a variância fica aproximadamente constante.
- A relação agora é linear?
- Não, esse problema não foi corrigido.

Exemplo:

- Estamos analisando duas variáveis:
  - nível de poliamina no sangue;
  - idade.
- Qual é a varável respota?

Exemplo:

- Estamos analisando duas variáveis:
  - nível de poliamina no sangue;
  - idade.
- Qual é a varável respota? Nível de poliamina.
- A figura a seguir mostra o gráfico de dipersão entre as variáveis.

Exemplo (continuação)

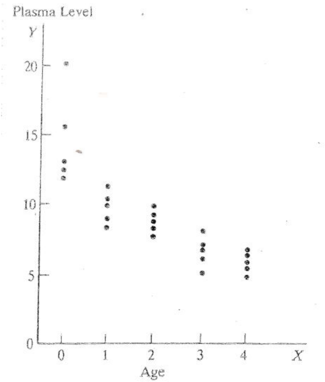


Figura: Gráfico do nível de poliamina em função da idade.

- A relação entre as variáveis é linear?

Exemplo (continuação)

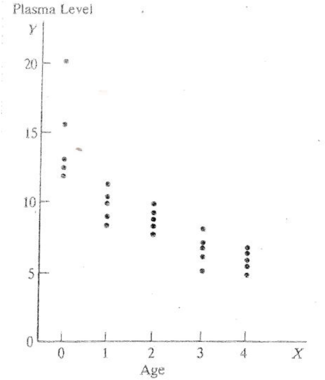


Figura: Gráfico do nível de poliamina em função da idade.

- A relação entre as variáveis é linear?
- Não, a relação se aproxima mais de uma curva.
- A variância é constante?

Exemplo (continuação)

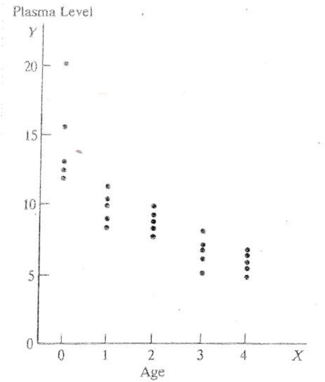


Figura: Gráfico do nível de poliamina em função da idade.

- A relação entre as variáveis é linear?
- Não, a relação se aproxima mais de uma curva.
- A variância é constante?
- Não, ela diminui com o aumento de X.

Exemplo (continuação)

- Vamos aplicar a transformação logaritmo.

$$Y^* = \log Y\,.$$

- O modelo ajustado é o seguinte

$$\hat{Y}^* = 1.135 - 0.1023X\,.$$

- Se quisermos voltar para a escala original, basta fazer

$$Y =$$

Exemplo (continuação)

- Vamos aplicar a transformação logaritmo.

$$Y^* = \log Y \, .$$

- O modelo ajustado é o seguinte

$$\hat{Y}^* = 1.135 - 0.1023X \, .$$

- Se quisermos voltar para a escala original, basta fazer

$$Y = \exp\{1.135 - 0.1023X\} \, .$$

- Vejamos a seguir o gráfico de dispersão para as variáveis na escala transformada.

Exemplo (continuação)

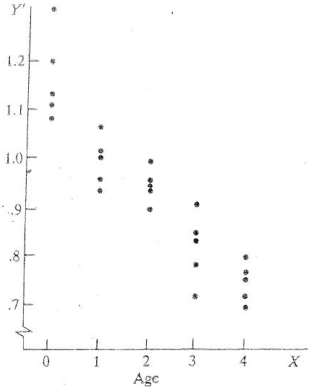


Figura: Gráfico do logaritmo dos níveis de poliamina em função da idade.

- A transformação alcançou objetivo pretendido?

Exemplo (continuação)

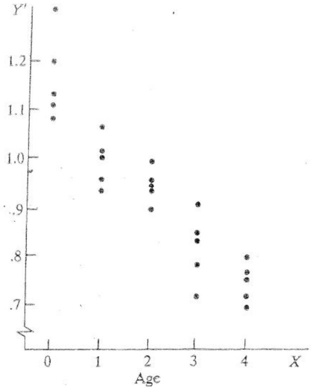


Figura: Gráfico do logaritmo dos níveis de poliamina em função da idade.

- A transformação alcançou objetivo pretendido?
- Sim, a variância fica aproximadamente constante.
- A relação agora é linear?

Exemplo (continuação)

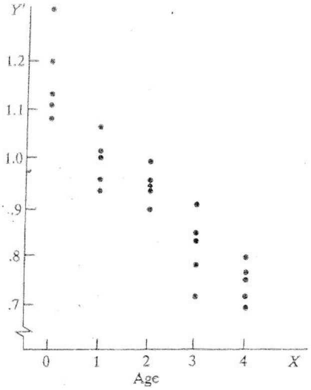
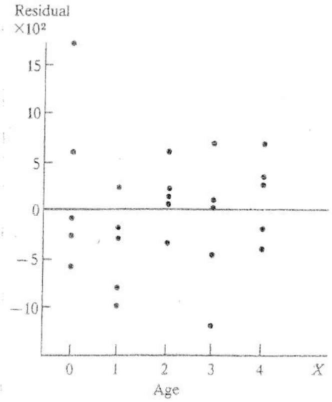
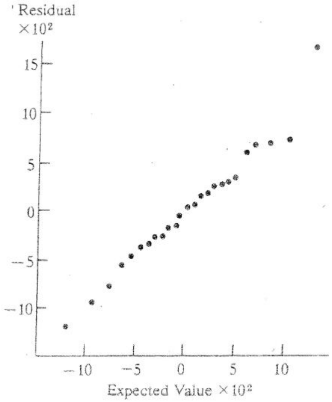


Figura: Gráfico do logaritmo dos níveis de poliamina em função da idade.

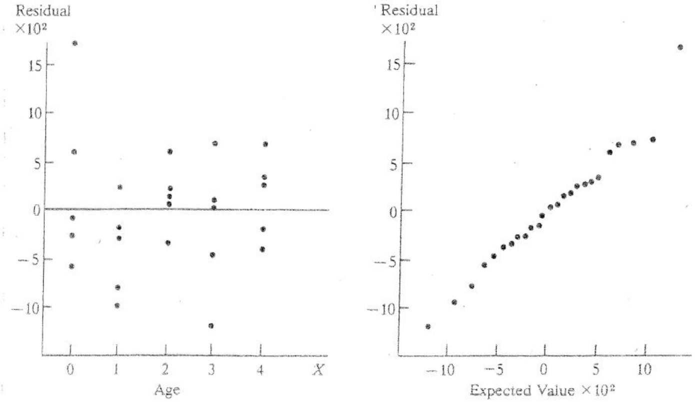
- A transformação alcançou objetivo pretendido?
- Sim, a variância fica aproximadamente constante.
- A relação agora é linear?
- Sim, a relação torna-se aproximadamente linear.

- A figura abaixo mostra os gráficos de resíduos para esse modelo.





- A figura abaixo mostra os gráficos de resíduos para esse modelo.



- As suposições do modelo não parecem ter sido violadas.

Métodos analíticos para seleção de transformação

- Em muitos casos é possível escolher a transformação empiricamente.
- O gráficos e o tipo de dado, nos mostram indícios de qual é a transformação mais adequada.
- Pode ser interessante ter um técnica mais objetiva e automática de escolha da transformação.
- Uma técnica usa para isso é:
  - **procedimento de Box-Cox.**

Procedimento de Box-Cox.

- É um método para escolher transformações de maneira automática.
- Essas transformações são escolhidas na família de transformações potência.
- Pode ser aplicada apenas se a variável resposta assumi só valores positivos.
- A transformação da variável Y é dada por

$$g_\lambda(y) = \begin{cases} \frac{(y^\lambda - 1)}{\lambda} & \lambda \neq 0 \\ \log(\lambda) & \lambda = 0\,. \end{cases}$$

- O parâmetro λ precisa ser estimado.
- Ele é escolhido de maneira a minimizar a função

$$L(\lambda) = -\frac{n}{2} \log \left(\frac{SQR_\lambda}{n} \right) + (\lambda - 1) \sum_{i=1}^{n} \log(y_i)$$

  onde SQR_λ é a soma de quadrados da regressão quando $g_\lambda(y)$ é a variável resposta.
- O valor estimado $\hat{\lambda}$ pode ser encontrado através de métodos numéricos de maximização de funções.
- O intervalo de 95% de confiança para λ é

$$\left\{ \lambda : L(\lambda) > L(\hat{\lambda}) - \frac{1}{2} \chi^2_{1;1-\alpha} \right\}$$

- Observe que para $\lambda = 1$ temos que

$$g_\lambda(y) = \begin{cases} \frac{(y^\lambda - 1)}{\lambda} & \lambda \neq 0 \\ \log(\lambda) & \lambda = 0\,. \end{cases}$$

fica

$$g_\lambda(y) =$$

- Observe que para $\lambda = 1$ temos que

$$g_\lambda(y) = \begin{cases} \frac{(y^\lambda - 1)}{\lambda} & \lambda \neq 0 \\ \log(\lambda) & \lambda = 0\,. \end{cases}$$

  fica

$$g_\lambda(y) = \frac{(y^1 - 1)}{1} = y\,.$$

- Construímos o IC para λ.
- Se o valor 1 pertencer ao intervalo, o que isso indica?

- Observe que para $\lambda = 1$ temos que

$$g_\lambda(y) = \begin{cases} \frac{(y^\lambda - 1)}{\lambda} & \lambda \neq 0 \\ \log(\lambda) & \lambda = 0 \end{cases}.$$

fica

$$g_\lambda(y) = \frac{(y^1 - 1)}{1} = y .$$

- Construímos o IC para λ.
- Se o valor 1 pertencer ao intervalo, o que isso indica?
- Que não é necessário fazer a transformação.

Exemplo:

- Está sendo analisado o número de espécies (Specie) de tartaruga em 30 ilhas de Gálapagos.
- As variáveis explicativas são as seguintes:
  - área da ilha em km^2 (Area);
  - altura da elevação mais alta em metros (Elevation);
  - distância da ilha mais próxima em km (Nearest);
  - distância da ilha de Santa Cruz em km (Scruz);
  - área da ilha adjacente em km^2 (Adjacent).

Exemplo: (continuação)

- Aplicando a transformação Box-Cox obteve-se

$$\hat{\lambda} = 0,33\,.$$

- O intervalo de confiança para o parâmetro é dado por

$$IC_{\lambda}(95\%) = [0,15; 0,55]\,.$$

- Devemos aplicar a transformação?

Exemplo: (continuação)

- Aplicando a transformação Box-Cox obteve-se

$$\hat{\lambda} = 0,33\,.$$

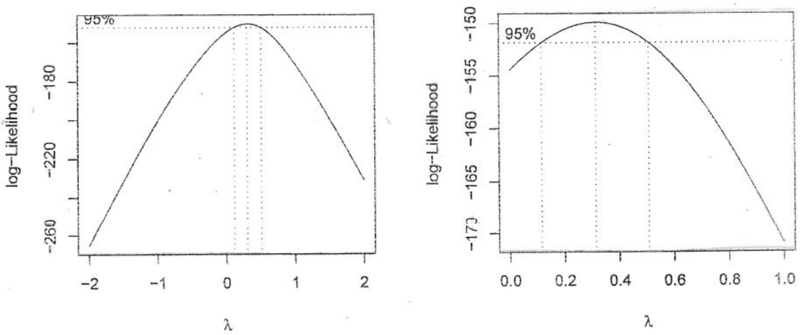
- O intervalo de confiança para o parâmetro é dado por

$$IC_{\lambda}(95\%) = [0,15; 0,55]\,.$$

- Devemos aplicar a transformação?
- Sim, pois o um não pertence ao intervalo.

Exemplo: (continuação)

- A figura mostra o gráfico de $L(\lambda)$ e o intervalo de 95% de confiança para λ são mostrados abaixo.



Exemplo:

- Estamos interessados em analisar a taxa de aplicação em poupança em 50 países.
- Essa taxa é dada por

$$\frac{\text{valor poupado}}{\text{renda}}.$$

- As variáveis explicativas são as seguintes:
  - porcentagem da população abaixo de 15 anos de idades;
  - porcentagem da poulação acima de 75 anos;
  - renda *per capita* em dólares;
  - taxa de crescimento da renda *per capita*.

Exemplo: (continuação)

- Aplicando a transformação Box-Cox obteve-se

$$\hat{\lambda} = 0,95\,.$$

- O intervalo de confiança para o parâmetro é dado por

$$IC_\lambda(95\%) = [0,06;\ 1,4]\,.$$

- Devemos aplicar a transformação?

Exemplo: (continuação)

- Aplicando a transformação Box-Cox obteve-se

$$\hat{\lambda} = 0,95\,.$$

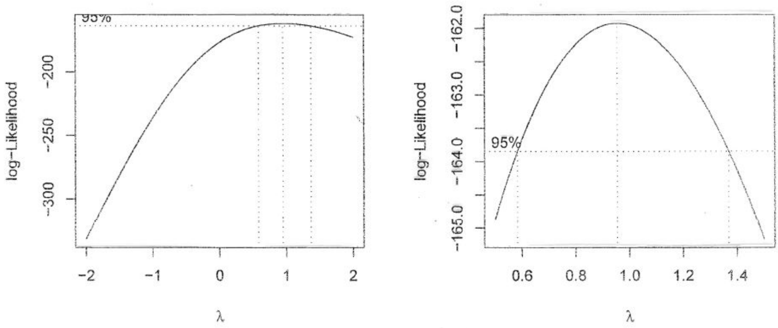
- O intervalo de confiança para o parâmetro é dado por

$$IC_\lambda(95\%) = [0,06;\ 1,4]\,.$$

- Devemos aplicar a transformação?
- Não, pois o 1 pertence ao intervalo.

Exemplo: **(continuação)**

- A figura mostra o gráfico de $L(\lambda)$ e o intervalo de 95% de confiança para λ são mostrados abaixo.



Transformação para corrigir não-linearidade

- Vimos até agora transformações para estabilizar a variância.
- Porém essas transformações também afetam as relações entre as variáveis.
- Resultando geralmente em relações curvas entre as variáveis.
- Maneira de voltar a termos uma relação linear é:
  - fazer transformações na variável explicativa.
- Variáveis explicativas são aleatórias?

- Vimos até agora transformações para estabilizar a variância.
- Porém essas transformações também afetam as relações entre as variáveis.
- Resultando geralemente em relações curvas entre as variáveis.
- Maneira de voltar a termos uma relação linear é:
  - fazer transformações na variável explicativa.
- Variáveis explicativas são aleatórias? Não.
- Ao fazermos transformações nelas vamos afetar a variância?

- Vimos até agora transformações para estabilizar a variância.
- Porém essas transformações também afetam as relações entre as variáveis.
- Resultando geralmente em relações curvas entre as variáveis.
- Maneira de voltar a termos uma relação linear é:
  - fazer transformações na variável explicativa.
- Variáveis explicativas são aleatórias? Não.
- Ao fazermos transformações nelas vamos afetar a variância? Não.

- Transformar a variável resposta afeta a variância e a linearidade.
- Transformar a variável explicativa só afeta linearidade.
- Então devemos primeiro trabalhar com a variância:
  - transformar a variável Y.
- Em seguida corrigir a linearidade:
  - transformar a variável explicativa.

Exemplo:

- Estamos querendo analisar a relação entre duas variáveis:
  - a velocidade do vento registrada em um gerador;
  - e a velocidade real do vento.

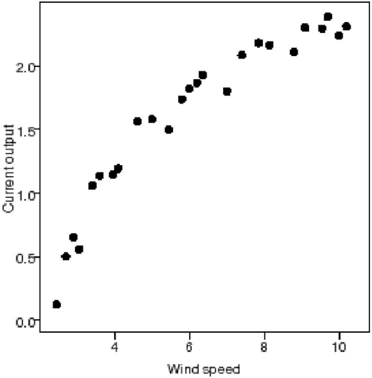


Figura: Gráfico de dispersao da velocidade registrada pelo gerador e a velocidade real do vento.

Exemplo: (continuação)

- A relação entre as variáveis é linear?

Exemplo: (continuação)

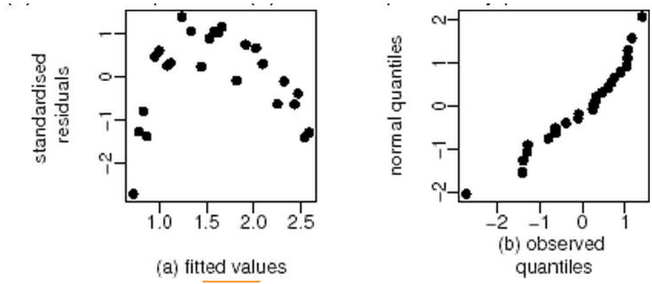
- A relação entre as variáveis é linear?
- Aproximadamente sim.
- O modelo de regressão linear simples foi ajustado

$$\hat{y} = 0,131 + 0,241x \,.$$

- Vamos verificar os gráficos de resíduos para ver se as suposições do modelo são válidas.

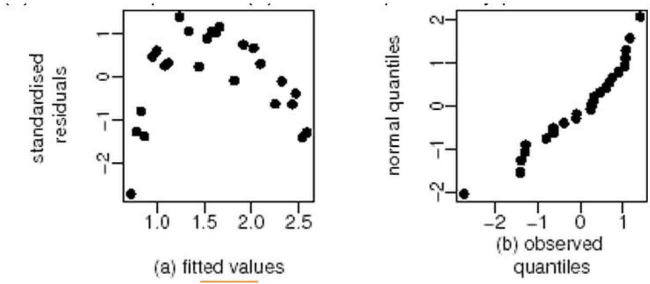
Exemplo: (continuação)

- A figura seguir mostra o gráfico de resíduos vs valores ajustados e o gráfico de probabilidade normal.



Exemplo: (continuação)

- A figura seguir mostra o gráfico de resíduos vs valores ajustados e o gráfico de probabilidade normal.



- Os resíduos não parecem ser normais e nem ter variância constante.

Exemplo: (continuação)

- O modelo linear simples não parece ser o mais adequado nesse caso.
- Vamos analisar o gráfico de dispersão para encontrar uma transformação adequada em X.
- Essa transformação deve ser monótona e crescente.
- Algumas possiblidades

$$\log(x) \quad \sqrt{x} \quad 1/x\,.$$

- Devemos testar cada uma delas e analisar os resíduos e gráfico de dispersão.
- A mais adequada nesse caso é a transformação inversa $1/x$.

Exemplo: (continuação)

- O modelo ajustado fica

$$\hat{Y} = 2.98 - 6.93/x \, .$$

- Vejamos como ficam os resíduos.

Exemplo: (continuação)

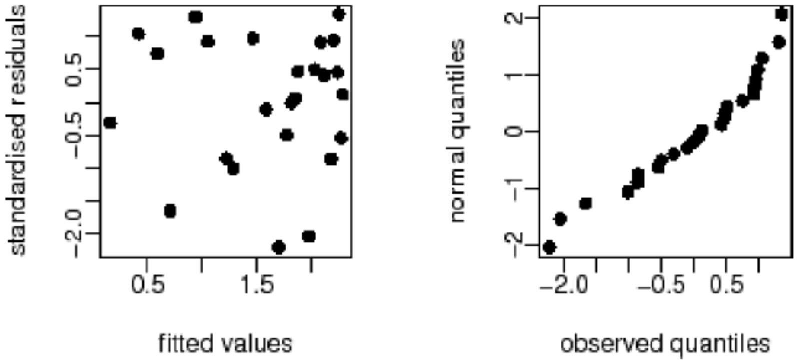


Figura: Gráfico de resíduos vs valores ajustados e gráfico de probabilidade normal.

Exemplo: (continuação)

- Conclusões:

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que a variância é constante e a relação entre as variáveis é linear;

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que a variância é constante e a relação entre as variáveis é linear;
  - no segundo gráfico os pontos caem em torno de uma reta;
  - isso é um indício de que

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que a variância é constante e a relação entre as variáveis é linear;
  - no segundo gráfico os pontos caem em torno de uma reta;
  - isso é um indício de que os erros tem distribuição normal.
- Percebemos então que a transformação apenas em X resolveu o problema de ajuste do modelo.

Exemplo:

- Vamos retomar o exemplo do cimento.

- Aplicamos a transformação logaritmo na variável resposta.
- A variância ficou constante.
- Mas a relação deixou de ser linear.
- Para corrigir esse problema, vamos aplicar uma transformação em X.

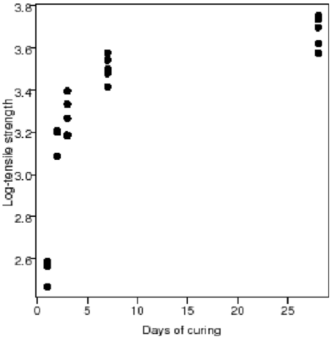
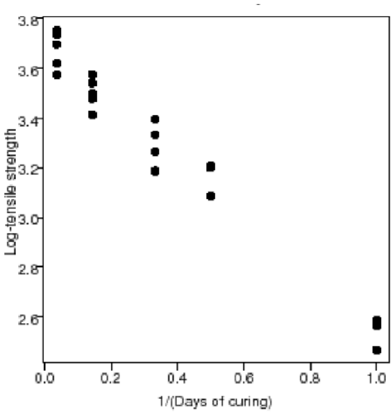


Figura: Gráfico do logaritmo da resistência do cimento em função do tempo de cura.

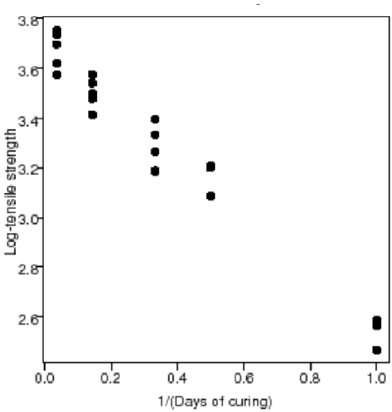
Exemplo: (continuação)

- Vamos aplicar a transformação $1/x$ e verificar como fica o gráfico de dispersão.



Exemplo: (continuação)

- Vamos aplicar a transformação $1/x$ e verificar como fica o gráfico de dispersão.



- A relação entre as variáveis parece linear.

Exemplo: (continuação)

- O modelo ajustado foi o seguinte

$$\log(\hat{Y}) = 3.69 - 1.15/x \,.$$

- Vejamos agora o gráfico de resíduos para verificar se as suposições do modelo são válidas.

Exemplo: (continuação)

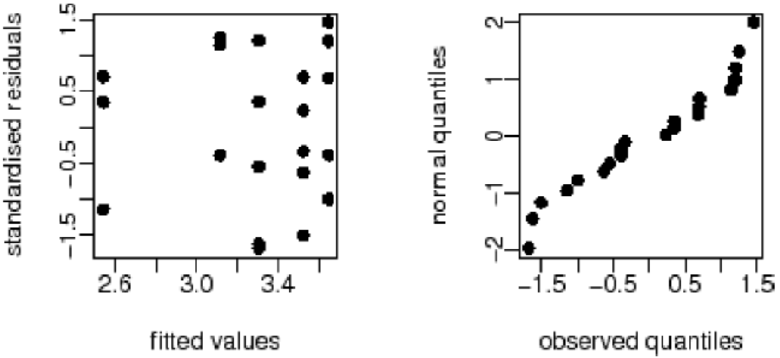


Figura: Gráfico de resíduos vs valores ajustados e gráfico de probabilidade normal.

Exemplo: (continuação)

- Conclusões:

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que a variância é constante e a relação entre as variáveis é linear;

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que a variância é constante e a relação entre as variáveis é linear;
  - no segundo gráfico os pontos caem em torno de uma reta;
  - isso é um indício de que

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que a variância é constante e a relação entre as variáveis é linear;
  - no segundo gráfico os pontos caem em torno de uma reta;
  - isso é um indício de que os erros tem distribuição normal.
- Percebemos então que a transformação apenas em X resolveu o problema de ajuste do modelo.

Transformação log-log

- Já vimos que podemos transformar a variável resposta e a explicativa.
- A transformação mais comum é a log-log.
- Tiramos o log da variável resposta e da explicativa.
- Essa transformação é últil quando a verdadeira relação entre X e Y é dada por

$$Y_i = \beta_0 x_i^{\beta_1} \times \epsilon_i \; .$$

- O erro aleatório é multiplicativo.

- Tomando o logaritmo ficamos com

- Tomando o logaritmo ficamos com

$$\log(Y_i) = \log(\beta_0) + \beta_1 \log x_i + \log(\epsilon_i) \, .$$

- Portanto o erro volta ser aditivo como no modelo original.
- Essa transformação é últil quando:
  - a relação entre X e Y parece linear, mas a variância está aumentando;
  - os intervalos de valores de X e Y são muito grandes;
  - ambos possuem muitos valores pequenos e alguns extremamente grandes.

Exemplo:

- Estamos analisando a relação entre as seguintes medidas de mamíferos:
  - peso do cérebro;
  - peso do corpo.

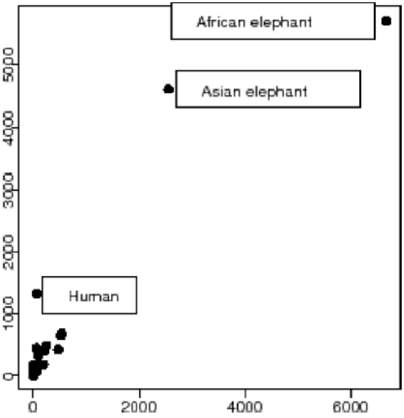


Figura: Peso do cérebro em função do peso do corpo.

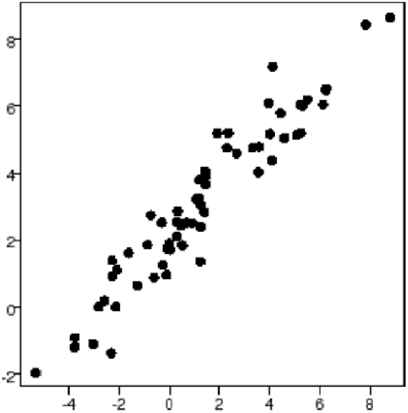
Exemplo: (continuação)

- Esse gráfico não é muito informativo.
- Quase todas observações estão concentradas no canto esquerdo inferior.
- Como as duas variáveis apresentam muitos valores pequenos e alguns muito altos, devemos aplicar a transformação log-log.
- O modelo ajustado é o seguinte

$$\log(\hat{Y}) = 2.092 + 0.763 \log(x) .$$

Exemplo: (continuação)

- A figura a seguir mostra o gráfico de dispersão para as duas variáveis na escala logaritma.



- Notamos que agora a relação se aproxima muito mais de uma reta.

Exemplo: (continuação)

- Vejamos agora os gráficos de resíduos.

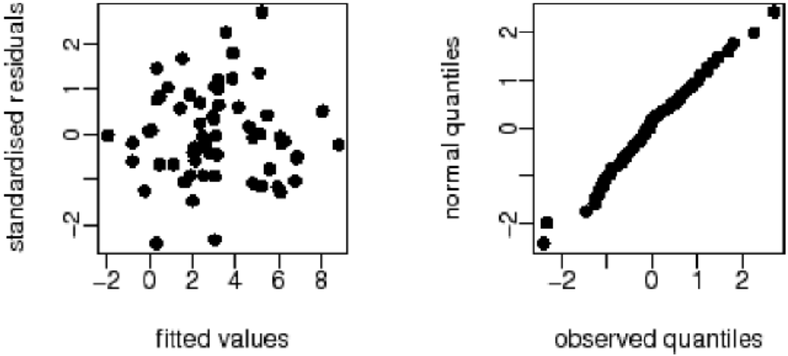


Figura: Gráfico de resíduos vs valores ajustados e gráfico de probabilidade normal.

Exemplo: (continuação)

- Conclusões:

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que a variância é constante e a relação entre as variáveis é linear;

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que a variância é constante e a relação entre as variáveis é linear;
  - no segundo gráfico os pontos caem em torno de uma reat;
  - isso é um indício de que

Exemplo: (continuação)

- Conclusões:
  - no primeiro gráfico os pontos estão distribuídos de maneira aleatória;
  - isso é um indício de que a variância é constante e a relação entre as variáveis é linear;
  - no segundo gráfico os pontos caem em torno de uma reat;
  - isso é um indício de que os erros tem distribuição normal.
- Percebemos então que a transformação apenas em X e em Y problema de ajuste do modelo.