

# Transformação de variáveis, outros modelos e outros métodos de estimação

Prof. Caio Azevedo

# Introdução

- Vamos admitir que as observações (os erros) da variável resposta sejam não correlacionadas.
- Quando as suposições de normalidade e homocedasticidade não são satisfeitas, pode-se aplicar alguma transformação (em geral na variável resposta, embora também possa ser feita na variável explicativa) a fim de obter-se ambas as propriedades.
- Outras vezes a transformação é aplicada para se “linearizar” um modelo (multiplicativo/não-linear). Ou seja, um modelo não-linear é mais apropriado mas, deseja-se utilizar a estrutura dos modelos lineares.

# Introdução

- Uma outra forma de se tratar a heterocedasticidade, dentro da nossa classe de modelos, é estimar os parâmetros (inclusive relativos à variância) através da metodologia de mínimos quadrados ponderados, como veremos adiante.
- Pode-se utilizar outros modelos para incorporar a não normalidade e/ou a heterocedasticidade (por exemplo, modelos lineares generalizados, veja [http://www.ime.unicamp.br/~cnaber/Material\\_MLG\\_1S\\_2016.htm](http://www.ime.unicamp.br/~cnaber/Material_MLG_1S_2016.htm)).

# Transformações

- Transformações mais comuns para obter-se normalidade dos dados:
  - Dados positivos: logaritmo e raiz quadrada.
  - Dados restritos ao intervalo (a,b): coloca-se-os no intervalo (0,1) e usa-se o logito ( $\ln(Y/(1 - Y))$ ).
- Transformações mais comuns para obter-se homocedasticidade:
  - Dados positivos: as mesmas que no caso anterior, função inversa ( $1/Y$ ) e funções trigonométricas ( $\arcsin(\sqrt{y})$ ).
  - Dados restritos ao intervalo (a,b): colocá-los no intervalo (0,1) e usar o logito ( $\ln(Y/(1 - Y))$ ).

# Transformação de Box -Cox

- Uma família de transformações geral (se  $y$  for positivo) é dada por

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda}, \lambda \neq 0 \\ \ln y, \lambda = 0 \end{cases}$$

- O parâmetro  $\lambda$  pode ser estimado através dos dados via máxima verossimilhança ou mínimos quadrados, utilizando -se  $y^{(\lambda)}$  com resposta.

# Problemas com as transformações

- Mudam a natureza/escala dos dados (a não ser que faça sentido em termos do problema, isso é ruim).
- As interpretações dos parâmetros, na escala original dos dados, é perdida.
- As estimativas, em geral e independentemente do método de estimação adotado, tornam-se menos precisas (vício, variância e erro quadrático médio).
- Muitas vezes a transformação não é eficaz para se conseguir normalidade/homocedasticidade.

# Problemas com as transformações

- Se aplicarmos transformações na variável resposta (e/ou explicativa), e considerarmos o modelo de regressão normal linear usual, teremos

$$\mathcal{E}(g(Y_i)) = \beta_0 + \sum_{j=1}^{p-1} \beta_j h(x_{ij})$$

em que  $g(\cdot)$  e  $h(\cdot)$  são as transformações adotadas. Note que

$$\mathcal{E}(g(Y_i)) \neq g(\mathcal{E}(Y_i)).$$

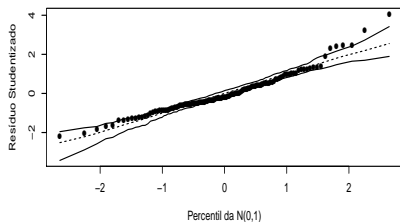
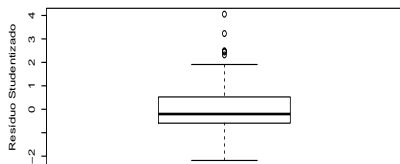
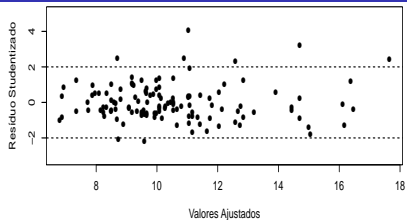
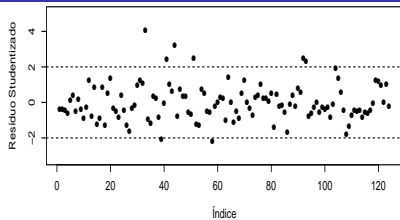
## Exemplo 1: considerando as etiologias cardíacas

$$Y_{ij} = \beta_{0i} + \beta_{1i}x_{ij} + \xi_{ij}, i = 1, \dots, ; j = 1, \dots, n_i$$

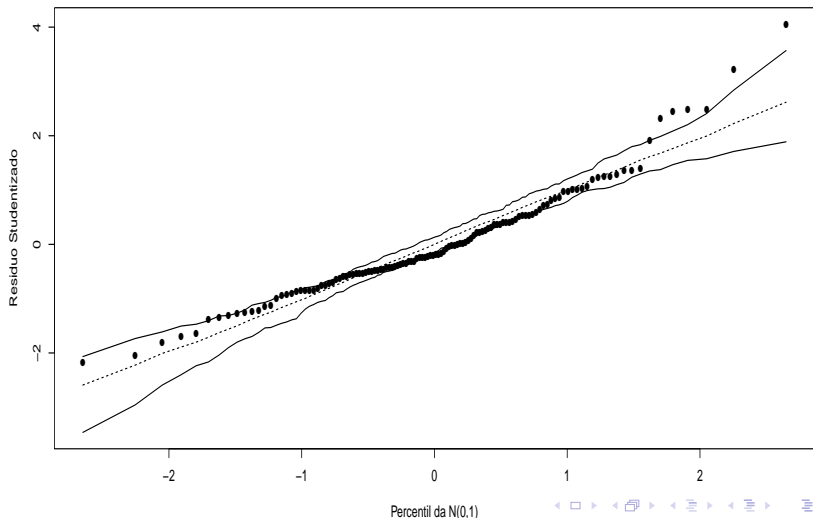
- Etiologias : CH ( $i = 1$ ), ID ( $i = 2$ ), IS ( $i = 3$ ), C: ( $i = 4$ ).
- $\xi_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$ .
- $x_{ij}$ : carga à que o paciente  $j$  que apresenta a etiologia cardíaca  $i$  foi submetido (conhecido e não aleatório).
- $\beta_{0i}$  : consumo esperado para pacientes da  $i$ -ésima etiologia submetidos à uma carga igual a 0.
- $\beta_{1i}$  : incremento (positivo ou negativo) no consumo esperado, de pacientes da  $i$ -ésima etiologia, para o aumento em uma unidade da carga.



# Gráficos de resíduos



# Gráfico de envelopes para os resíduos

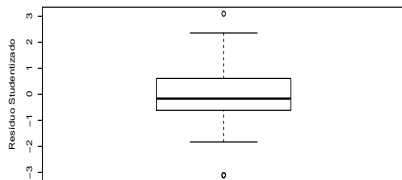
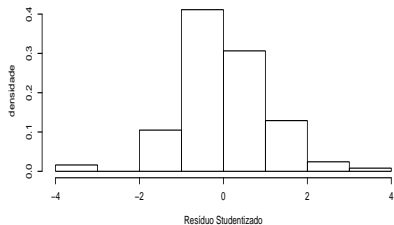
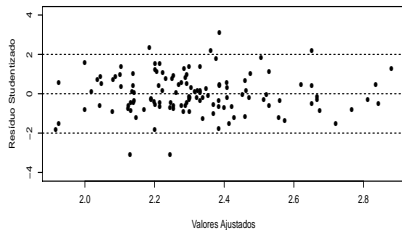
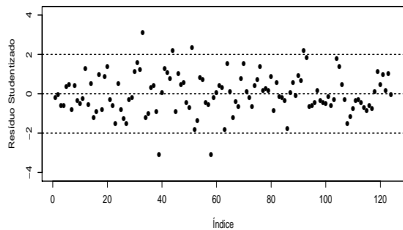


## Exemplo 1: (transformação logarítmica)

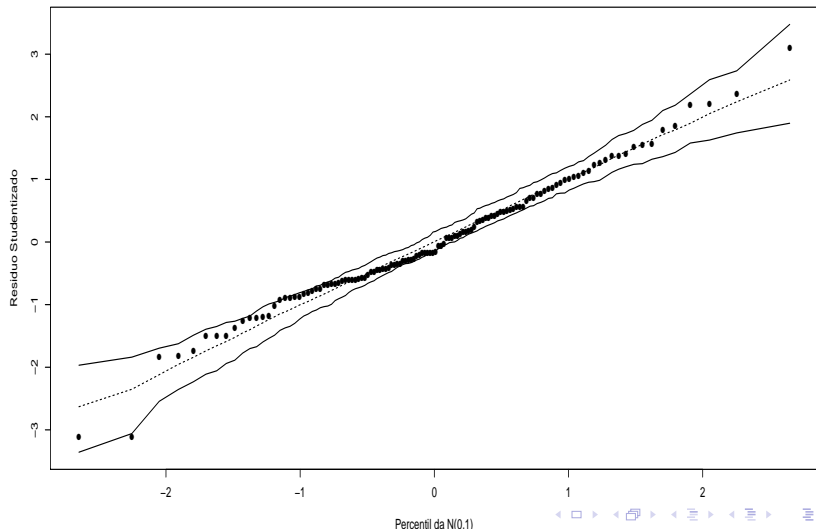
$$\ln Y_{ij} = \beta_{0i} + \beta_{1i}x_{ij} + \xi_{ij}, i = 1, \dots, ; j = 1, \dots, n_i$$

- Etiologias : CH ( $i = 1$ ), ID ( $i = 2$ ), IS ( $i = 3$ ), C: ( $i = 4$ ).
- $\xi_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$ .
- $x_{ij}$ : carga à que o paciente  $j$  que apresenta a etiologia cardíaca  $i$  foi submetido (conhecido e não aleatório).
- $\beta_{0i}$ : valor esperado do logaritmo natural do consumo para pacientes da  $i$ -ésima etiologia submetidos à uma carga igual a 0.
- $\beta_{1i}$ : incremento (positivo ou negativo) no valor esperado do logaritmo natural do consumo, de pacientes da  $i$ -ésima etiologia, para o aumento em uma unidade da carga.

# Gráficos de resíduos



# Gráfico de envelopes para os resíduos

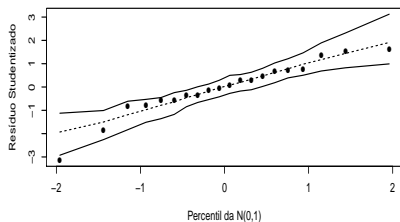
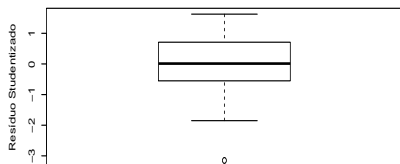
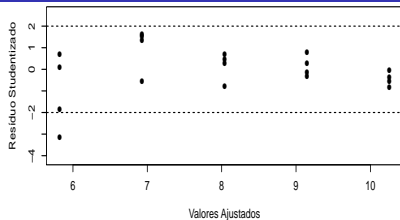
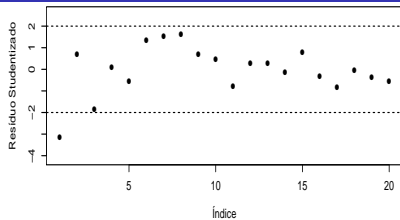


## Exemplo 3, Modelo linear 1: reta

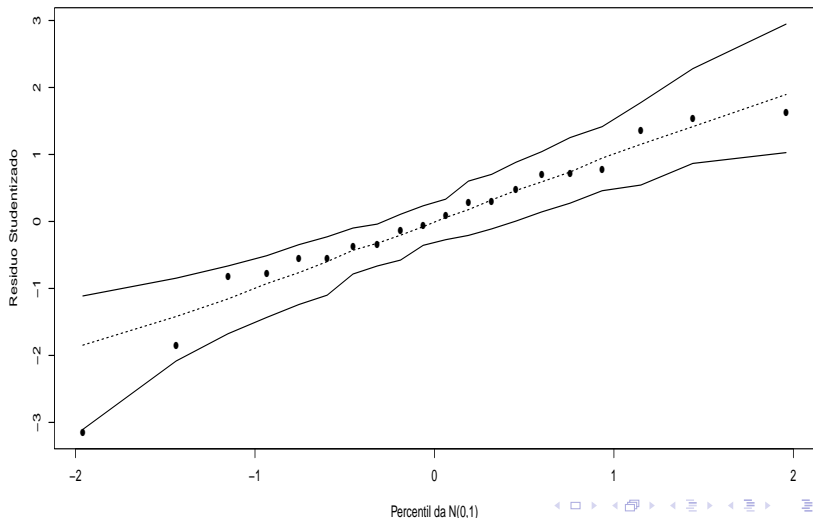
$$Y_i = \beta_0 + \beta_1 x_i + \xi_i, i = 1, 2, \dots, 20$$

- $x_i$  : quantidade de fósforo ministrada a i-ésima parcela.
- $\beta_0$  : valor esperado (média) da produção de milho quando a quantidade de fósforo aplicada é igual a 0.
- $\beta_1$  : incremento no valor esperado da produção de milho quando a quantidade de fósforo aplicada aumenta em uma unidade.
- $\xi_i \stackrel{iid}{\sim} N(0, \sigma^2)$ .

# Gráficos de resíduos



# Gráfico de envelopes para os resíduos



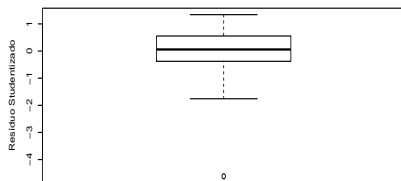
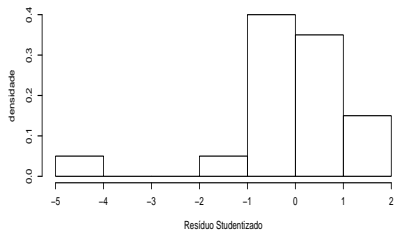
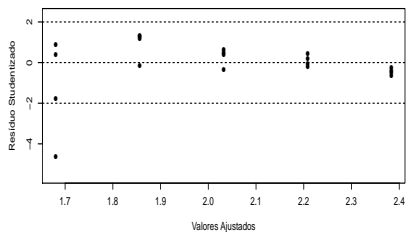
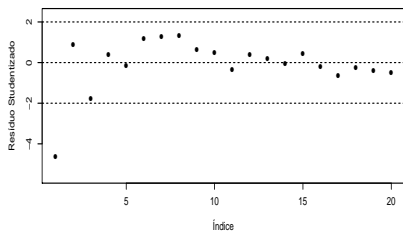


## Exemplo 3: (transformação logarítmica)

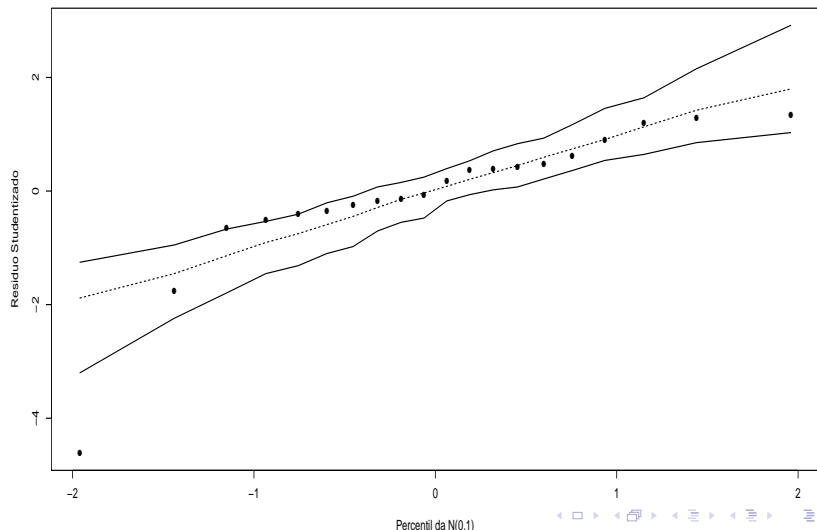
$$\ln Y_i = \beta_0 + \beta_1 x_i + \xi_i, i = 1, 2, \dots, 20$$

- $x_i$  : quantidade de fósforo ministrada a  $i$ -ésima parcela.
- $\beta_0$  : valor esperado (média) do logaritmo natural da produção de milho quando a quantidade de fósforo aplicada é igual a 0.
- $\beta_1$  : incremento no valor esperado do logaritmo natural da produção de milho quando a quantidade de fósforo aplicada aumenta em uma unidade.
- $\xi_i \stackrel{iid}{\sim} N(0, \sigma^2)$ .

# Gráficos de resíduos



# Gráfico de envelopes para os resíduos

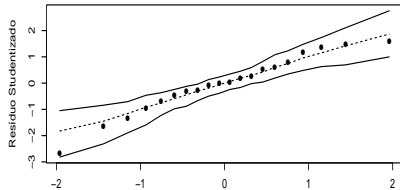
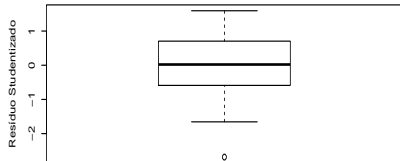
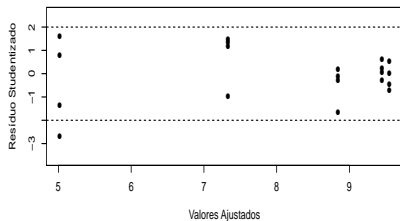
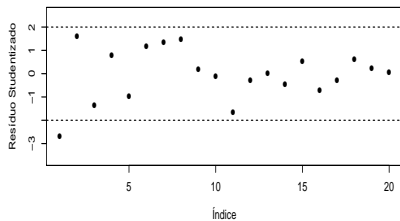


## Exemplo 3, Modelo linear 2: parábola

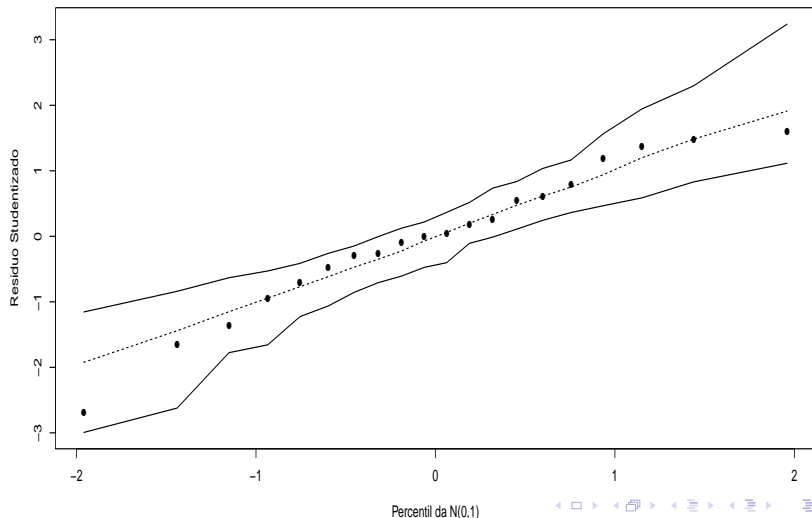
$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \xi_i, i = 1, 2, \dots, 20$$

- $x_i$  : quantidade de fósforo ministrada a  $i$ -ésima parcela.
- $\beta_0$  : valor esperado (média) da produção de milho quando a quantidade de fósforo aplicada é igual a 0.
- A interpretação isolada dos parâmetros  $\beta_1$  e  $\beta_2$  é complicada mas, podemos dizer que  $\frac{-\beta_1}{2\beta_2}$  é a quantidade de fósforo que leva à produção máxima de milho.
- $\xi_i \stackrel{iid}{\sim} N(0, \sigma^2)$ .

# Gráficos de resíduos



# Gráfico de envelopes para os resíduos

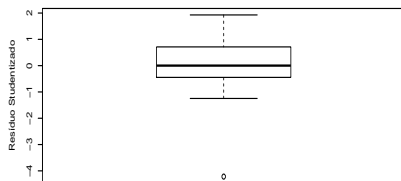
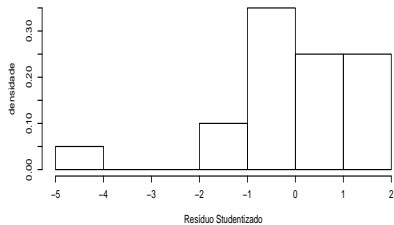
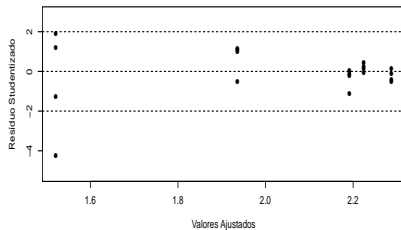
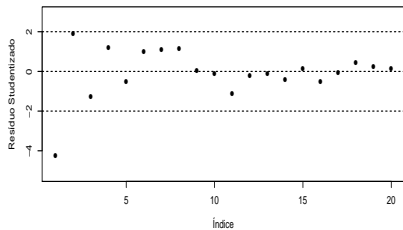


## Exemplo 3: (transformação logarítmica)

$$\ln Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \xi_i, i = 1, 2, \dots, 20$$

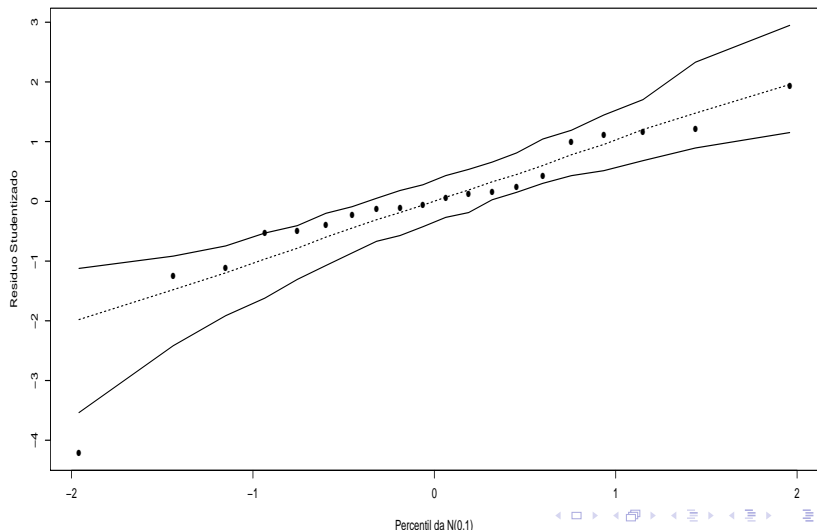
- $x_i$  : quantidade de fósforo ministrada a i-ésima parcela.
- $\beta_0$  : valor esperado (média) do logaritmo natural da produção de milho quando a quantidade de fósforo aplicada é igual a 0.
- A interpretação isolada dos parâmetros  $\beta_1$  e  $\beta_2$  é complicada mas, podemos dizer que  $\frac{-\beta_1}{2\beta_2}$  é a quantidade de fósforo que leva ao máximo do log da produção de milho.
- $\xi_i \stackrel{iid}{\sim} N(0, \sigma^2)$ .

# Gráficos de resíduos





# Gráfico de envelopes para os resíduos



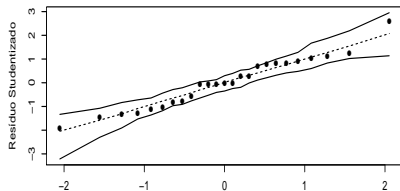
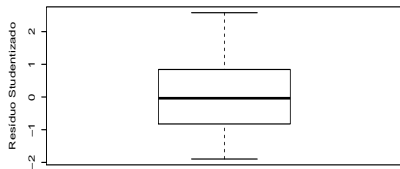
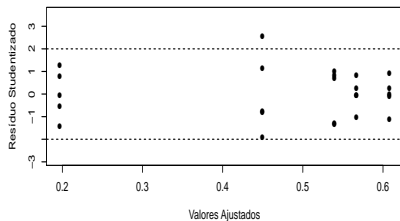
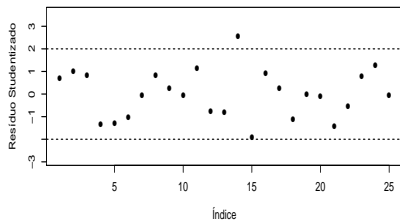
## Exemplo 5: Modelo (casela de referência)

$$Y_{ij} = \mu + \alpha_i + \xi_{ij},$$

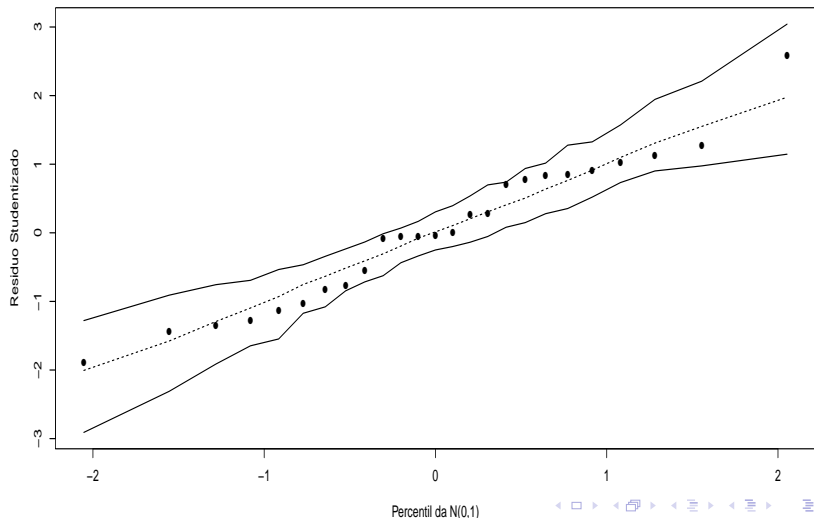
$i = 1, 2, \dots, 5$  (grupos);  $j = 1, \dots, 5$  (unidades experimentais)

- $\xi_{ij} \stackrel{ind.}{\sim} N(0, \sigma^2)$ . Parte sistemática:  $\mu_i = \mu + \alpha_i$ , é a média populacional relacionada ao  $i$ -ésimo fator,  $\alpha_1 = 0$  (restrição de identificabilidade) .
- $\mu$  : é a média populacional do grupo de referência,  $\mu_1 = \mu$ .
- $\alpha_i = \mu_i - \mu_1, i = 2, \dots, 5$ , é o incremento (positivo ou negativo) entre a média do grupo  $i$  e a média do grupo de referência.
- Grupos : grupo 1(E50), grupo 2(E70), grupo 3(EAW), grupo 4(M1M), grupo 5(MAW).

# Gráficos de resíduos



# Gráfico de envelopes para os resíduos



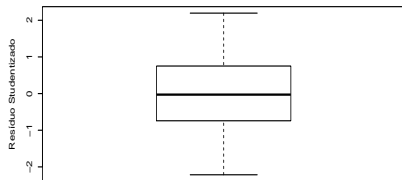
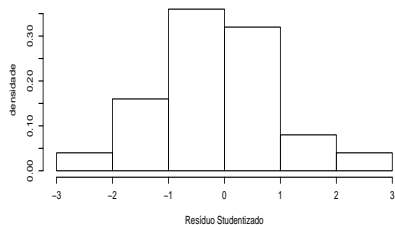
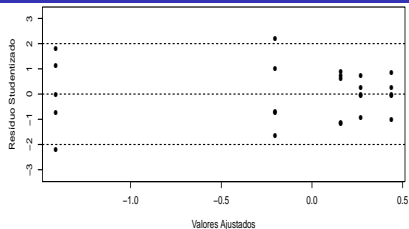
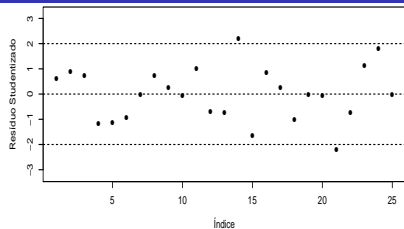
## Exemplo 5: (transformação logito)

$$\ln(Y_{ij}/(1 - Y_{ij})) = \mu + \alpha_i + \xi_{ij},$$

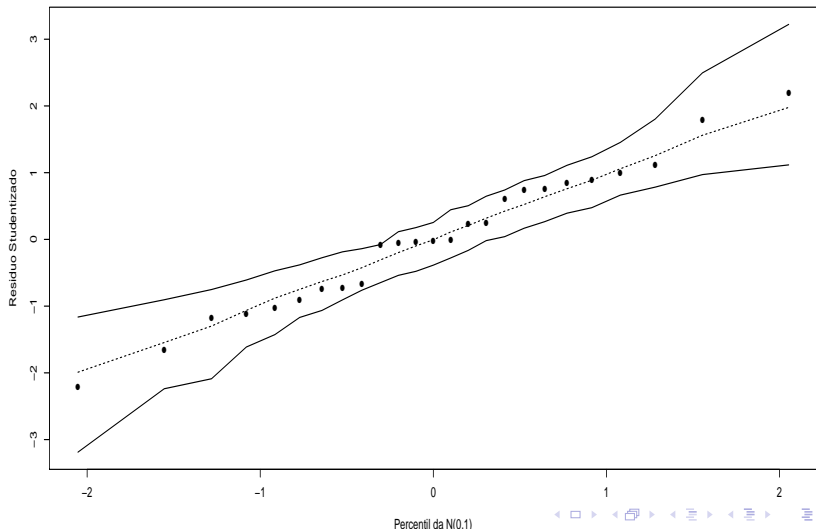
$i = 1, 2, \dots, 5$  (grupos);  $j = 1, \dots, 5$  (unidades experimentais)

- $\xi_{ij} \stackrel{ind.}{\sim} N(0, \sigma^2)$ . Parte sistemática:  $\mu_i = \mu + \alpha_i$ , é a média populacional relacionada ao  $i$ -ésimo fator,  $\alpha_1 = 0$  (restrição de identificabilidade) .
- $\mu$  : é a média populacional do logito da absorvância do grupo de referência,  $\mu_1 = \mu$ .
- $\alpha_i = \mu_i - \mu_1, i = 2, \dots, 5$ , é o incremento (positivo ou negativo) entre a média do logito da absorvância do grupo  $i$  e a média do logito da absorvância do grupo de referência.

# Gráficos de resíduos



# Gráfico de envelopes para os resíduos



## Exemplo 2: desconsiderando o sexo (modelo log-normal)

$$Y_{ij} = \beta_{0i} x_{ij}^{\beta_{1i}} \xi_{ij}, i = 1(\text{convencional}), 2(\text{hugger})(\text{tipo de escova}),$$
$$j = 1, \dots, 26(\text{criança}).$$

- $\xi_{ij} \stackrel{i.i.d.}{\sim} LN(0, 1)$ , em que  $LN(\mu, \sigma^2)$  representa uma distribuição log-normal tal que  $\mathcal{E}(\ln \xi_{ij}) = \mu$  e  $\mathcal{V}(\ln \xi_{ij}) = \sigma^2$ .
- $x_{ij}$ : IPB pré-escovação da criança  $j$  utilizando a escova do tipo  $i$ .
- $Y_{ij}$ : IPB pós-escovação da criança  $j$  utilizando a escova do tipo  $i$ .



## Exemplo 2: desconsiderando o sexo (modelo log-normal) - cont.

- $\beta_{0i}$  : proporcional ao IPB pós-escovação esperado quando se utiliza a escova do tipo  $i$  para um IPB pré-escovação igual a 1.
- $\beta_{1i}$  : regula a taxa de crescimento ( $> 0$ ) ou decréscimo ( $< 0$ ) do IPB pós-escovação para o aumento em uma unidade no IPB pré-escovação.

## Exemplo 2: (transformado)

$$\ln Y_{ij} = \ln(\beta_{0i}) + \beta_{1i} \ln(x_{ij}) + \ln \xi_{ij}, \quad i = 1(\text{convencional}), 2(\text{hugger})(\text{tipo de escova})$$
$$j = 1, \dots, 26(\text{criança}).$$

- $\ln \xi_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$ .
- $x_{ij}$ : IPB pré-escovação da criança  $j$  utilizando a escova do tipo  $i$ .
- $Y_{ij}$ : IPB pós-escovação da criança  $j$  utilizando a escova do tipo  $i$ .

## Exemplo 2: (transformado)

- $\ln \beta_{0i}$  : valor esperado do logaritmo natural do IPB pós-escovação quando se utiliza a escova do tipo  $i$  para um IPB pré-escovação igual a 1.
- $\beta_{1i}$  : incremento (positivo ou negativo) no valor esperado do logaritmo natural do IPB pós-escovação quando se utiliza a escova do tipo  $i$ , para o aumento em uma unidade no logaritmo natural do IPB pré-escovação.

# Estudo de simulação

$$Y_i = \beta_0 x_i^{\beta_1} \xi_i, i = 1, \dots, 100, \xi_{ij} \stackrel{i.i.d.}{\sim} LN(0, 1)$$

- $\beta_0 = 0,5$  e  $\beta_1 = 1,05$ .
- Foram geradas 1000 réplicas de tamanho 100.
- Em cada réplica, os parâmetros  $\beta_0$  e  $\beta_1$  foram estimados por máxima verossimilhança, no modelo acima, e também por mínimos quadrados ordinários a partir do modelo  
$$\ln Y_i = \ln \beta_0 + \beta_1 \ln x_i + \ln \xi_i, \ln \xi_i \stackrel{i.i.d.}{\sim} N(0, 1).$$
- Para cada método de estimação e parâmetros, obtivemos a média e variância das estimativas das réplicas, bem como o vício e a raiz quadrática do erro quadrático médio (rqeqm).

# Estimação por MV

## ■ Verossimilhança

$$L(\beta_0, \beta_1) \propto \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (\ln y_i - \beta_0 x_i^{\beta_1})^2 \right\}$$

## ■ Logverossimilhança

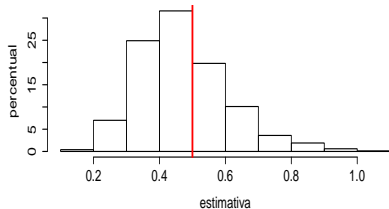
$$l(\beta_0, \beta_1) = \left\{ -\frac{1}{2} \sum_{i=1}^n (\ln y_i - \beta_0 x_i^{\beta_1})^2 \right\} + \text{constante} \quad (1)$$

# Estimação por MV

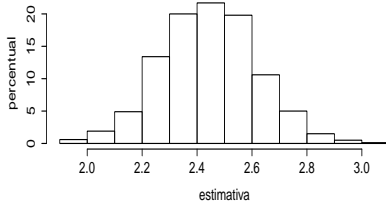
- Devemos derivar a função (1) com relação à  $(\beta_0, \beta_1)'$ , igualá-la a  $(0, 0)'$  e resolver o sistema resultante. Nesse caso, a solução passa pela utilização de métodos numéricos.
- Usou-se a função `optim` e o algoritmo de otimização de Nelder-Mead.
- Exercício: encontrar o vetor escore, a matriz Hessiana e a Informação de Fisher. Indique como obter as estimativas via algoritmo Escore de Fisher.

# Histogramas das estimativas

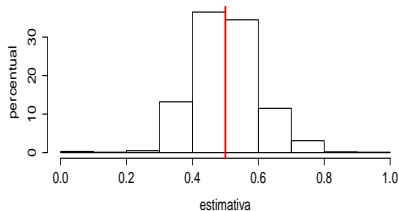
Estimador de MQ  $\beta_0$



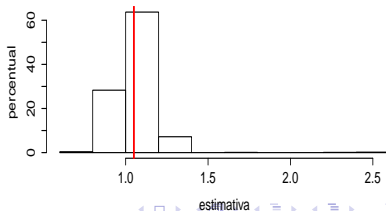
Estimador de MQ  $\beta_1$



Estimador de MV  $\beta_0$



Estimador de MV  $\beta_1$

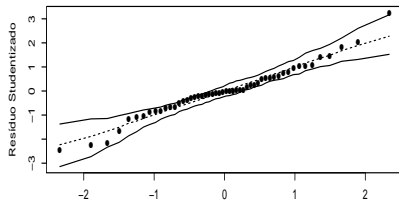
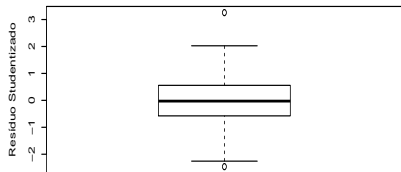
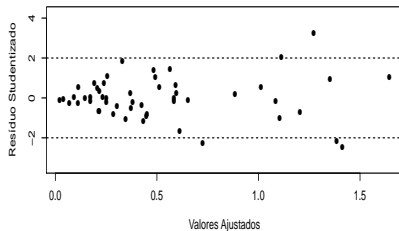
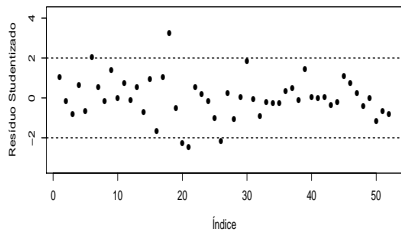


Mod/Mét. de est.	Par.	Valor verd.	Média	Vício	Var.	RQEQM
Normal/MQ	$\beta_0$	0,5	0,471	-0,029	0,018	0,139
	$\beta_1$	1,05	2,442	1,392	0,030	1,403
Log-normal/MV	$\beta_0$	0,5	0,503	0,002	0,010	0,100
	$\beta_1$	1,05	1,058	0,008	0,015	0,124

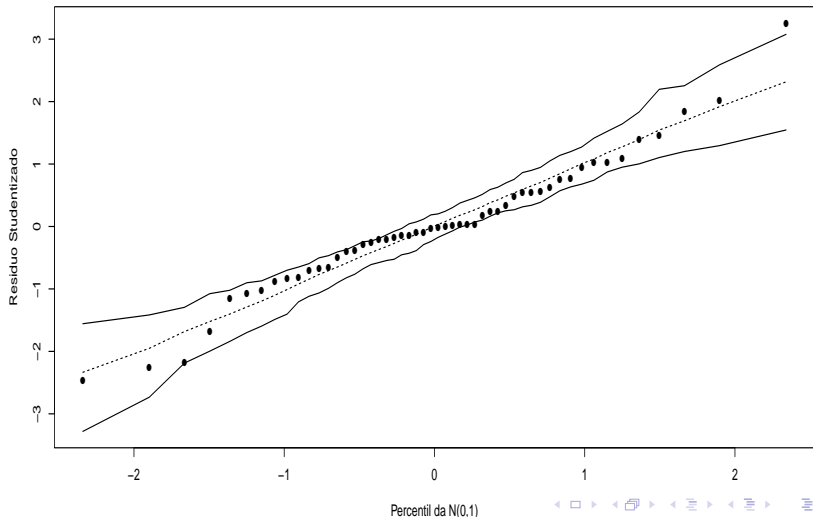
Percebe-se que estimar os parâmetros utilizando-se o modelo original via MV fornece melhores resultados.



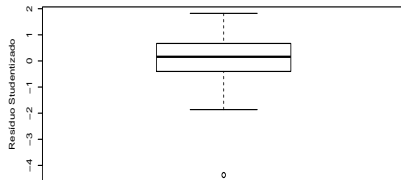
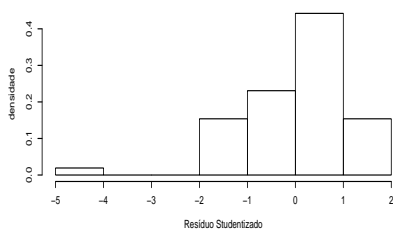
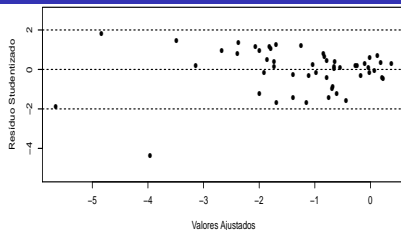
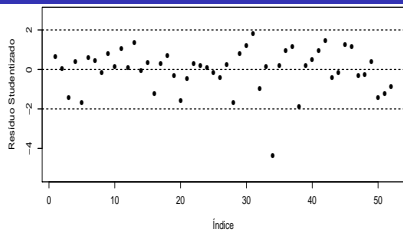
# Gráficos de resíduos (sem transformação)



# Gráfico de envelopes para os resíduos (sem transformação)



# Gráficos de resíduos (com transformação)



# Gráfico de envelopes para os resíduos (com transformação)

