Modelos de regressão para dados correlacionados

Cibele Russo

cibele@icmc.usp.br

ICMC USP

Mini-curso oferecido no
Workshop on Probabilistic and Statistical Methods

28 a 30 de janeiro de 2013

Conteúdo da aula

- Modelo linear com efeitos mistos
- Modelo marginal $\times$ modelo hierárquico
- Aplicações

Referências:

- Pinheiro and Bates (2009) 'Mixed-effects Models in S and S-PLUS', Springer.
- Searle, S. R., Casella, G. McCulloch, C. E. (2006), Variance Components. Wiley Series in Probability and Statistics.
- Verbeke G. and Molenberghs G. (2000) 'Linear mixed models for longitudinal data,' Springer Series in Statistics, Springer-Verlag, New-York.

Medidas repetidas e dados longitudinais

Medidas repetidas (repeated measurements) são obtidas quando uma resposta é observada repetidamente em um conjunto de unidades experimentais

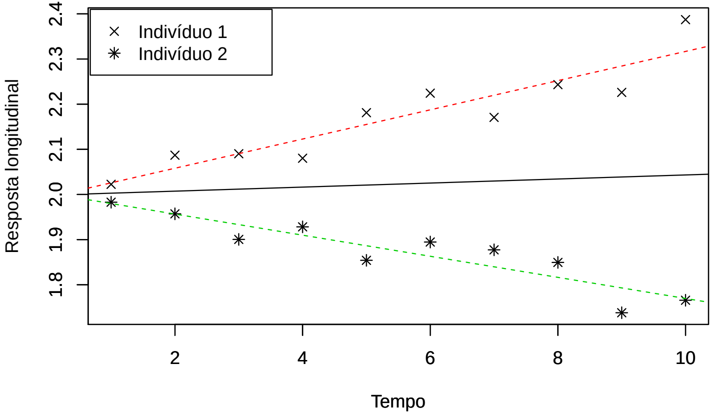
- Unidades experimentais: Sujeitos, pacientes, participantes
- Animais, plantas
- Grupos (clusters): famílias, cidades, filiais de uma companhia

Dados longitudinais são um caso especial de medidas repetidas

Medidas repetidas e dados longitudinais

Quando as medidas são repetidas no mesmo indivíduo
(ou em uma mesma unidade experimental),
as características individuais induzem uma correlação nos dados
observados naquele indivíduo.

Modelos para dados correlacionados



Medidas longitudinais

(adaptado de Rizopoulos, D. 'An Introduction to Joint Models for Longitudinal & Survival Data, with Applications in R', short course 27th IWSM, Prague.)

Modelos para dados correlacionados

Como os indivíduos são amostrados aleatoriamente de uma população, pode-se supor que coeficientes de regressão sujeito-específicos também amostrados de uma população de coeficientes de regressão:

$$\boldsymbol{\beta}_i \sim N(\boldsymbol{\beta}, D), \ 1, \ldots, n.$$

Modelos para dados correlacionados

Poderíamos então ajustar um modelo diferente para as medidas de cada indivíduo?

Sim, mas existe o interesse em estimar o comportamento geral da população, e ao mesmo tempo identificar um padrão específico de cada indivíduo.

Modelo com efeitos mistos

Um modelo com efeitos mistos para Y_{ij}, a j-ésima resposta observada para o i-ésimo indivíduo, utilizando o valor da covariável x, x_{ij}, é dado por

$$Y_{ij} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})x_{ij} + \epsilon_i,\ i = 1, \ldots, n;\ j = 1, \ldots, m_i$$

Modelo com efeitos mistos

$$Y_{ij} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})x_{ij} + \epsilon_i, \; i = 1, \ldots, n; \; j = 1, \ldots, m_i$$

Efeitos fixos

- β_0 é o efeito fixo referente ao intercepto da reta
- β_1 é o efeito fixo referente ao coeficiente angular da reta

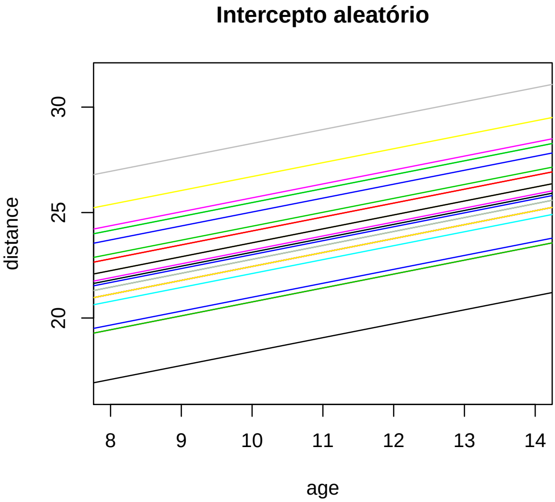
Efeitos aleatórios

- b_{i0} é um efeito aleatório adicionado ao intercepto
- b_{i1} é um efeito aleatório adicionado ao coeficiente angular

Modelos com efeitos mistos

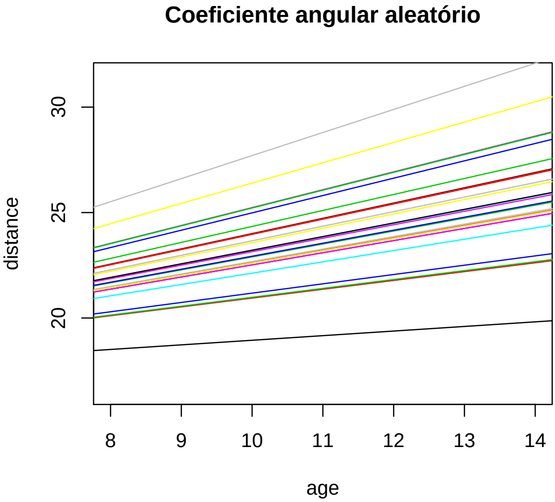
Qual o resultado da inclusão dos efeitos aleatórios no intercepto e no coeficiente angular?

Interpretação dos efeitos adicionados ao modelo



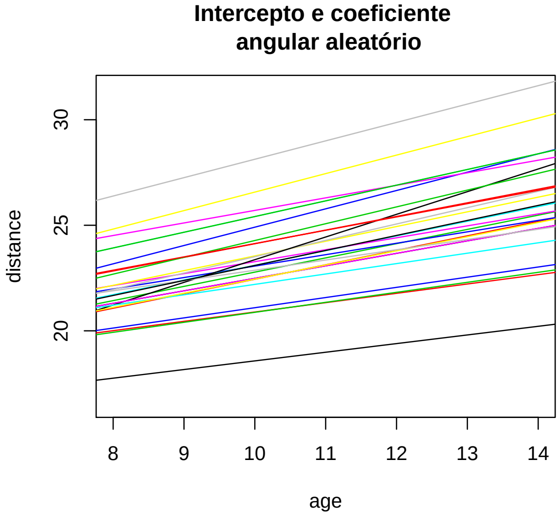
Exemplo de retas resultantes da inclusão de intercepto aleatório

Interpretação dos efeitos adicionados ao modelo



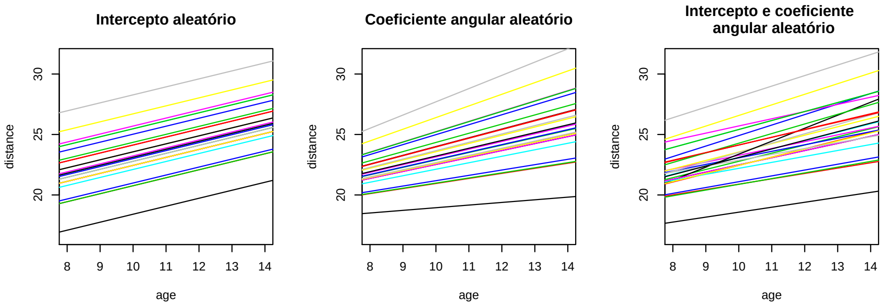
Exemplo de retas resultantes da inclusão de coeficiente angular aleatório

Interpretação dos efeitos adicionados ao modelo



Exemplo de retas resultantes da inclusão de intercepto e coeficiente angular aleatório

Interpretação dos efeitos adicionados ao modelo



Retas resultantes da inclusão dos efeitos aleatórios

Modelo com efeitos mistos

Reescrevendo o modelo

$$Y_{ij} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})x_{ij} + \epsilon_{ij},\ \ i = 1, \ldots, n;\ \ j = 1, \ldots, m_i$$

temos

$$Y_{ij} = (\beta_0 + \beta_1 x_{ij}) + (b_{i0} + b_{i1} x_{ij}) + \epsilon_{ij},\ \ i = 1, \ldots, n;\ \ j = 1, \ldots, m_i$$

em que m_i observações pertencem ao indivíduo i. Matricialmente,

$$\mathbf{Y}_i = X_i \boldsymbol{\beta} + Z_i \mathbf{b}_i + \boldsymbol{\epsilon}_i,\ \ i = 1, \ldots, n;$$

Modelo com efeitos mistos

Modelo linear com efeitos mistos

(1 covariável, intercepto e coeficiente angular aleatórios)

$$\mathbf{Y}_i = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \ \ i = 1, \ldots, n$$

$$\text{com } \mathbf{Y}_i = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} \ X_i = Z_i = \begin{bmatrix} 1 & x_{i1} \\ 1 & x_{i2} \\ \vdots & \vdots \\ 1 & x_{in} \end{bmatrix},$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \boldsymbol{\epsilon}_i = \begin{bmatrix} \epsilon_{i1} \\ \vdots \\ \epsilon_{in} \end{bmatrix}, \mathbf{b}_i = \begin{bmatrix} b_{i0} \\ b_{i1} \end{bmatrix}.$$

Modelo com efeitos mistos

Modelo linear com efeitos mistos

(1 covariável, intercepto e coeficiente angular aleatórios)

$$\mathbf{Y}_i = X_i\beta + Z_i\mathbf{b}_i + \epsilon_i, \ \ i = 1, \ldots, n$$

- $\mathbf{Y}_{i(m_i \times 1)}$: vetor de respostas do indivíduo i (aleatório)
- $X_{i(m_i \times 2)}$: matriz do modelo referente aos efeitos fixos do indivíduo i (fixa e conhecida)
- $\beta_{(2 \times 1)}$: vetor de parâmetros (efeitos fixos)
- $Z_{i(m_i \times 2)}$: matriz do modelo referente aos efeitos aleatórios para o indivíduo i (fixa e conhecida)
- $\mathbf{b}_{i(2 \times 1)}$: vetor de efeitos aleatórios para o indivíduo i (aleatório)
- $\epsilon_{i(m_i \times 1)}$: vetor de erros aleatórios para o indivíduo i (aleatório)

Modelo com efeitos mistos

Modelo linear com efeitos mistos

(1 covariável, intercepto e coeficiente angular aleatórios)

$$\mathbf{Y}_i = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \ \ i = 1, \ldots, n$$

Suposições comuns:

- $\mathbf{b}_i \overset{ind}{\sim} N_2(\mathbf{0}, D)$
- $\boldsymbol{\epsilon}_i \overset{ind}{\sim} N_{m_i}(\mathbf{0}, \sigma^2 I)$
- $\mathbf{b}_i$ independente de $\boldsymbol{\epsilon}_i$.

Parâmetros a serem estimados: $\boldsymbol{\beta}$, σ^2, elementos da matriz D.

Em alguns trabalhos considera-se D diagonal.

Modelo com efeitos mistos

Modelo linear com efeitos mistos

(forma geral, p covariáveis, q efeitos aleatórios)

$$\mathbf{Y}_i = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \ \ i = 1, \ldots, n$$

- $\mathbf{Y}_{i(m_i \times 1)}$: vetor de respostas do indivíduo i (aleatório)
- $X_{i(m_i \times (p+1))}$: matriz do modelo referente aos efeitos fixos do indivíduo i (fixa e conhecida)
- $\boldsymbol{\beta}_{((p+1) \times 1)}$: vetor de parâmetros (efeitos fixos)
- $Z_{i(m_i \times q)}$: matriz do modelo referente aos efeitos aleatórios para o indivíduo i (fixa e conhecida)
- $\mathbf{b}_{i(q \times 1)}$: vetor de efeitos aleatórios para o indivíduo i (aleatório)
- $\boldsymbol{\epsilon}_{i(m_i \times 1)}$: vetor de erros aleatórios para o indivíduo i (aleatório)

Modelo com efeitos mistos

Modelo linear com efeitos mistos

(forma geral, p covariáveis, q efeitos aleatórios)

$$\mathbf{Y}_i = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \;\; i = 1, \ldots, n$$

Suposições comuns:

- $\mathbf{b}_i \sim N_q(\mathbf{0}, D)$
- $\boldsymbol{\epsilon}_i \sim N_{m_i}(\mathbf{0}, \sigma^2 I)$

Parâmetros a serem estimados: $\boldsymbol{\beta}$, σ^2, elementos da matriz D.

Em alguns trabalhos considera-se D diagonal.

Modelo com efeitos mistos - Interpretação

Efeitos fixos

- β's têm as mesmas propriedades encontradas nos modelos com efeitos fixos

Efeitos aleatórios

- b_{ij}'s são interpretados como o parâmetro i do j-ésimo sujeito se desvia do parâmetro populacional.

Características interessantes

- β's descreve comportamentos populacionais
- $(\beta + b_i)$'s descreve comportamentos individuais

Modelo com efeitos mistos

Modelo linear com efeitos mistos

(forma geral, p covariáveis, q efeitos aleatórios)

$$\mathbf{Y}_i = X_i\beta + Z_i\mathbf{b}_i + \epsilon_i, \;\; i = 1, \ldots, n$$

Estimação dos parâmetros

Existem várias formas de estimar os parâmetros, as mais utilizadas são

- Formulação de dois estágios
- Enfoque marginal no modelo linear geral misto
- Enfoque hierárquico no modelo linear geral misto

Modelos lineares com efeitos mistos

Formulação de dois estágios

Estágio 1: Um modelo de regressão linear é ajustado para cada indivíduo separadamente. Esses modelos descrevem a variabilidade intraindivíduos

$$\mathbf{Y}_i = X_i\boldsymbol{\beta}_i + Z_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \;\; i = 1, \ldots, n$$

Estágio 2: Um modelo de regressão multivariado é ajustado utilizando as estimativas $\boldsymbol{\beta}_i$ obtidas no Estágio 1 como covariáveis.

$$\boldsymbol{\beta}_i = K_i\boldsymbol{\beta} + \mathbf{b}_i$$

(Verbeke G. and Molenberghs G. (2000) 'Linear mixed models for longitudinal data,' Springer Series in Statistics, Springer-Verlag, New-York.)

Modelos lineares com efeitos mistos

Formulação de dois estágios

$$\begin{cases} \mathbf{Y}_i = X_i\boldsymbol{\beta}_i + Z_i\mathbf{b}_i + \boldsymbol{\epsilon}_i \\ \boldsymbol{\beta}_i = K_i\boldsymbol{\beta} + \mathbf{b}_i, \;\; i = 1, \ldots, n \end{cases}$$

com $\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ e $\mathbf{b}_i \sim N(\mathbf{0}, D)$.

Modelo linear geral com efeitos mistos

Formulação geral e suposições

$$\begin{cases} \mathbf{Y}_i = X_i\beta + Z_i\mathbf{b}_i + \epsilon_i, \;\; i = 1, \ldots, n \\ \mathbf{b}_i \sim N_q(\mathbf{0}, D) \\ \epsilon_i \sim N_{m_i}(\mathbf{0}, \sigma^2 I) \\ \mathbf{b}_1, \ldots, \mathbf{b}_n, \epsilon_1, \ldots, \epsilon_n \text{ são independentes} \end{cases}$$

Modelo linear geral com efeitos mistos

O modelo linear geral com efeitos mistos com as suposições usuais pode ser reescrito como

$$\begin{bmatrix} \mathbf{Y}_i \\ \mathbf{b}_i \\ \boldsymbol{\epsilon}_i \end{bmatrix} \sim N_{mi+q+mi}\left(\begin{bmatrix} X_i\boldsymbol{\beta} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} Z_iDZ_i' + \sigma^2 I & Z_iD & \sigma^2 I \\ DZ_i' & D & \mathbf{0} \\ \sigma^2 I & \mathbf{0} & \sigma^2 I \end{bmatrix} \right)$$

Modelo hierárquico

Outro enfoque seria o **Modelo hierárquico usual**

$$\begin{cases} \mathbf{Y}_i | \mathbf{b}_i \sim N(X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i, \sigma^2 I), \ \ i = 1, \ldots, n \\ \mathbf{b}_i \sim N(\mathbf{0}, D) \end{cases}$$

Modelo hierárquico

- Uma possibilidade para estimação: Enfoque Bayesiano

$$\begin{cases} \mathbf{Y}_i | \boldsymbol{\beta}, \mathbf{b}_i \sim N(X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i, \sigma^2 I) \\ \mathbf{b}_i \sim N(\mathbf{0}, D) \\ \boldsymbol{\beta} \sim N(\boldsymbol{\beta}_0, B) \end{cases}$$

Modelo hierárquico

Para estimar $\boldsymbol{\beta}$, usamos a distribuição a posteriori de $\boldsymbol{\beta}|\mathbf{y}, B, D, \sigma^2$, que pode ser obtida de

$$f_{\boldsymbol{\beta}}(\boldsymbol{\beta}|\mathbf{y}) = \frac{\int f(\mathbf{y}|\boldsymbol{\beta}, b) f_{\boldsymbol{\beta}}(\boldsymbol{\beta}) f_b(b) db}{\int \int f(\mathbf{y}|\boldsymbol{\beta}, b) f_{\boldsymbol{\beta}}(\boldsymbol{\beta}) f_b(b) db d\boldsymbol{\beta}}.$$

Para isso, usa-se o fato de que

$$E(\boldsymbol{\beta}|\mathbf{y}_i, B, D, \sigma^2) = (X_i V_i^{-1} X_i)^{-1} X_i' V_i^{-1} \mathbf{y}$$

com $V_i = Z_i D Z_i' + \sigma^2 I$.

Modelo marginal

Por propriedades da distribuição normal, partindo do modelo em sua formulação geral podemos obter a distribuição marginal de $\mathbf{Y}_i$

Formulação geral

$$\begin{cases} \mathbf{Y}_i = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \ \ i = 1, \ldots, n \\ \mathbf{b}_i \sim N_q(\mathbf{0}, D) \\ \boldsymbol{\epsilon}_i \sim N_{m_i}(\mathbf{0}, \sigma^2 I) \\ \mathbf{b}_1, \ldots, \mathbf{b}_n, \boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n \text{ são independentes} \end{cases}$$

Esse modelo pode ser reescrito como um modelo marginal

$$\mathbf{Y}_i \sim N(X_i\boldsymbol{\beta}, Z_iDZ_i' + \sigma^2 I)$$

Modelo marginal

$$\mathbf{Y}_i \sim N(X_i\boldsymbol{\beta}, Z_iDZ_i' + \sigma^2 I)$$

Podemos então estimar os parâmetros $\boldsymbol{\beta}$ pelo método da máxima verossimilhança, utilizando a função de verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\tau})'$ com $\boldsymbol{\tau}$ o vetor de elementos de $V_i = Z_iDZ_i' + \sigma^2 I$.

Modelo marginal

A função de verossimilhança é dada por

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left\{ (2\pi)^{-n_i/2} |V_i(\boldsymbol{\tau})|^{-1/2} \exp\left\{ -\frac{1}{2}(y_i - X\beta)' V_i^{-1}(\boldsymbol{\tau})(y_i - X\beta) \right\} \right\}$$

Se $\boldsymbol{\tau}$ é conhecido,

$$\widehat{\boldsymbol{\beta}} = (\sum_{i}^{n} X_i V_i^{-1}(\boldsymbol{\tau}) X_i)^{-1} \sum_{n}^{n} X_i V_i^{-1}(\boldsymbol{\tau}) \mathbf{y}_i.$$

Se $\boldsymbol{\tau}$ não é conhecido, pode-se considerar um processo iterativo.

Modelo marginal

Se $\boldsymbol{\tau}$ é conhecido,

$$\widehat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sum_{i=1}^{n} X_i V_i^{-1} X_i)$$

Na prática, substitui-se $\boldsymbol{\tau}$ por um estimador com boas propriedades e obtém-se inferências aproximadas.

Propriedades

- Vários modelos hierárquicos diferentes podem levar ao mesmo modelo marginal,
- Um modelo marginal bem ajustado pode não ser uma evidência de um bom modelo hierárquico,
- Um tratamento satisfatório do modelo hierárquico só é possível em um contexto Bayesiano,
- Componentes de variância negativos (de D) podem ser obtidos sob o enfoque marginal, o que não acontece no modelo hierárquico.

Predição dos efeitos aleatórios

Método de Bayes empírico

Distribuição a posteriori de $\mathbf{b}_i|\mathbf{y}_i$:

$$\mathbf{b}_i|\mathbf{y}_i \sim N(DZ_i'V_i^{-1}(\mathbf{y}_i - X_i\boldsymbol{\beta}), \Lambda_i)$$

com Λ_i uma matriz positiva definida.

Podemos então utilizar a média a posteriori de $\mathbf{b}_i|\mathbf{y}_i$ como preditor de $\mathbf{b}_i$

$$\widehat{\mathbf{b}_i} = DZ_i'V_i^{-1}(\mathbf{y}_i - X_i\boldsymbol{\beta}).$$

Valores preditos para a variável resposta

Podemos obter um preditor para $\mathbf{Y}_i$ fazendo

$$\begin{aligned}
\widehat{\mathbf{Y}}_i &= X_i\widehat{\boldsymbol{\beta}} + \mathbf{Z}_i\widehat{\mathbf{b}}_i \\
&= X_i\widehat{\boldsymbol{\beta}} + \mathbf{Z}_i\widehat{\mathbf{D}}\mathbf{Z}_i^T\widehat{V}_i^{-1}(\mathbf{y}_i - X_i\widehat{\boldsymbol{\beta}}) \\
&= (\mathbf{I}_{m_i} - \mathbf{Z}_i\widehat{\mathbf{D}}\mathbf{Z}_i^T\widehat{V}_i^{-1})X_i\widehat{\boldsymbol{\beta}} + \mathbf{Z}_i\widehat{\mathbf{D}}\mathbf{Z}_i^T\widehat{V}_i^{-1}\mathbf{y}_i \\
&= \widehat{\sigma}^2\widehat{V}_i^{-1}X_i\widehat{\boldsymbol{\beta}} + (\mathbf{I}_{m_i} - \widehat{\sigma}^2\widehat{V}_i^{-1})\mathbf{y}_i,
\end{aligned}$$

que pode ser interpretado como uma média ponderada entre o perfil populacional $X_i\widehat{\boldsymbol{\beta}}$ e os dados observados $\mathbf{y}_i$, com pesos $\widehat{\sigma}^2\widehat{V}_i^{-1}$ e $(\mathbf{I}_{m_i} - \widehat{\sigma}^2\widehat{V}_i^{-1})$, respectivamente.

Modelos com efeitos mistos em R

Pacote **nlme**

- Possibilita o ajuste de modelos lineares e não lineares com efeitos mistos
- Várias possibilidades de estruturas de variância e covariância

Pacote **lme4**

- Possibilita o ajuste de modelos lineares, não lineares e lineares generalizados com efeitos mistos
- Utiliza apenas efeitos aleatórios
- Permite outras estruturas para os efeitos aleatórios

Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

(1 covariável, intercepto aleatório)

$$\begin{cases} \mathbf{Y}_i = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i + \boldsymbol{\epsilon} \\ \mathbf{b}_i \sim N(\mathbf{0}, D), \quad \boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \sigma^2 I) \end{cases}$$

Nesse caso,

$$X_i = \begin{bmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \end{bmatrix} \text{ para } i = 1, \ldots 11, \quad Z_i = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

(1 covariável, intercepto aleatório)

Comandos em R:

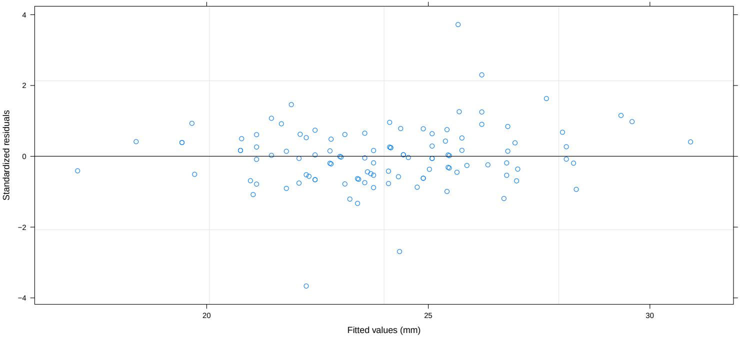
```
> library(nlme)
> attach(Orthodont)
> fit.IA<-lme(distance~ age,data=Orthodont,random=~ 1|Subject)
> ranef(fit.IA)
> coef(fit.IA)
> plot(fit.IA)
> plot(augPred(fit.IA),aspect="xy",grid=T)
> qqnorm(fit.IA,~ resid(.) |Sex)
```

⚠ Certifique-se que os comandos foram passados corretamente para o R.

Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

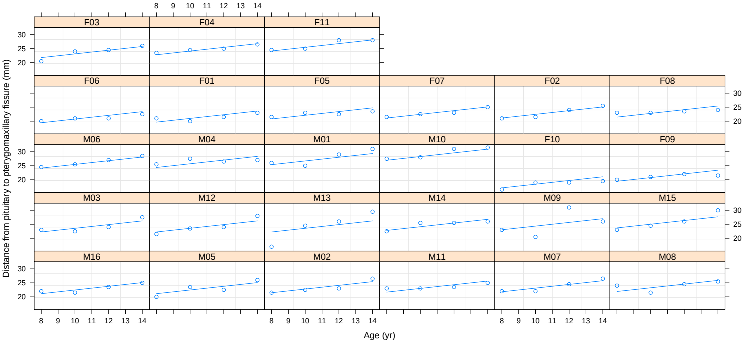
(1 covariável, intercepto aleatório)



Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

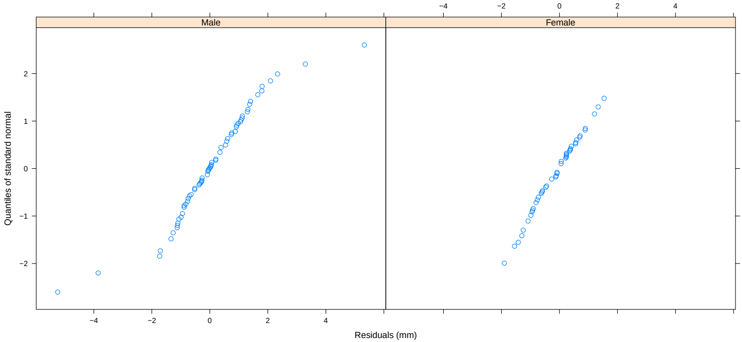
(1 covariável, intercepto aleatório)



Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

(1 covariável, intercepto aleatório)



Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

(1 covariável, coeficiente angular aleatório)

$$\begin{cases} \mathbf{Y}_i = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i + \boldsymbol{\epsilon} \\ \mathbf{b}_i \sim N(\mathbf{0}, D), \quad \boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \sigma^2 I) \end{cases}$$

Nesse caso,

$$X_i = \begin{bmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \end{bmatrix} \text{ para } i = 1, \ldots 11, \quad Z_i = \begin{bmatrix} 8 \\ 10 \\ 12 \\ 14 \end{bmatrix}$$

Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

(1 covariável, coeficiente angular aleatório)

Comandos em R:

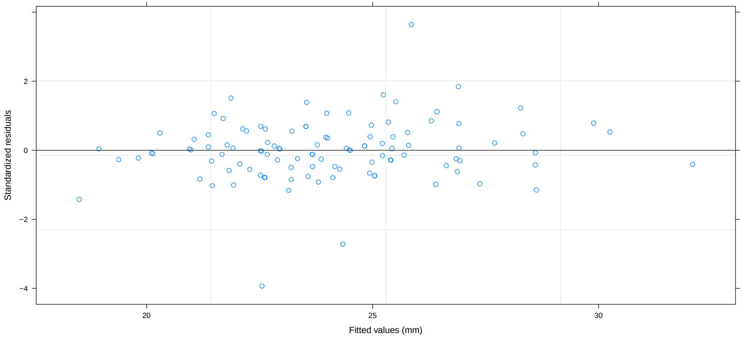
```
> library(nlme)
> attach(Orthodont)
> fit.CA<-lme(distance~ age,data=Orthodont,random=~ -1+age|Subject)
> ranef(fit.CA)
> coef(fit.CA)
> plot(fit.CA)
> plot(augPred(fit.CA),aspect="xy",grid=T)
> qqnorm(fit.CA,~ resid(.) |Sex)
```

⚠ Certifique-se que os comandos foram passados corretamente para o R.

Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

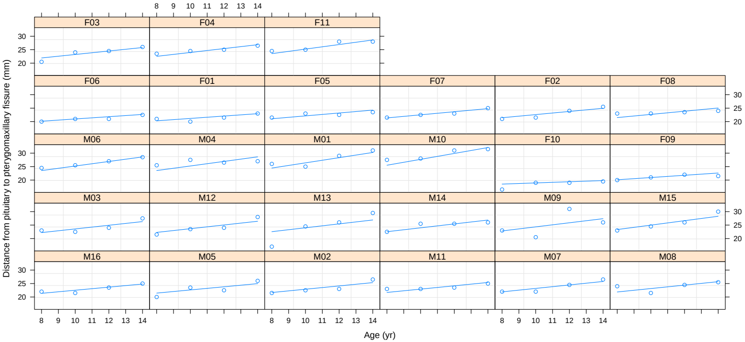
(1 covariável, coeficiente angular aleatório)



Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

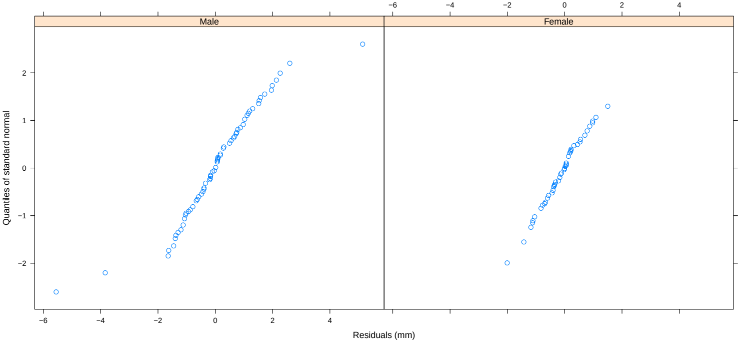
(1 covariável, coeficiente angular aleatório)



Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

(1 covariável, coeficiente angular aleatório)



Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

(1 covariável, intercepto e coeficiente angular aleatórios)

$$\begin{cases} \mathbf{Y}_i = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i + \boldsymbol{\epsilon} \\ \mathbf{b}_i \sim N(\mathbf{0}, D), \quad \boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \sigma^2 I) \end{cases}$$

Nesse caso,

$$X_i = \begin{bmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \end{bmatrix} \text{ para } i = 1, \ldots 11, \quad Z_i = \begin{bmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \end{bmatrix}$$

Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

(1 covariável, intercepto e coeficiente angular aleatórios)

Comandos em R:

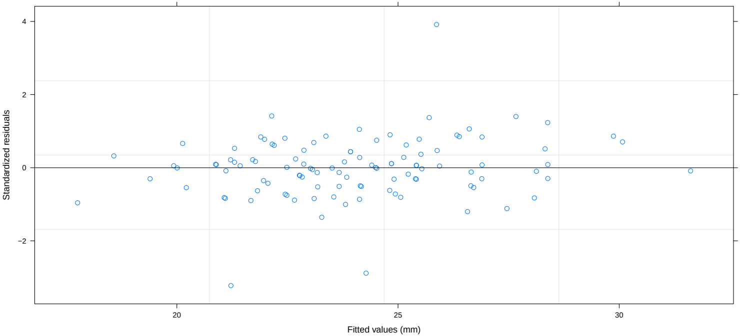
```
> library(nlme)
> attach(Orthodont)
> fit.doisefeitos<-lme(distance~ age,data=Orthodont,random=~ age|Subject)
> summary(fit.doisefeitos)
> ranef(fit.doisefeitos)
> coef(fit.doisefeitos)
> plot(fit.doisefeitos)
> plot(augPred(fit.doisefeitos),aspect="xy",grid=T)
> qqnorm(fit.doisefeitos,~ resid(.)|Sex)
```

⚠ Certifique-se que os comandos foram passados corretamente para o R.

Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

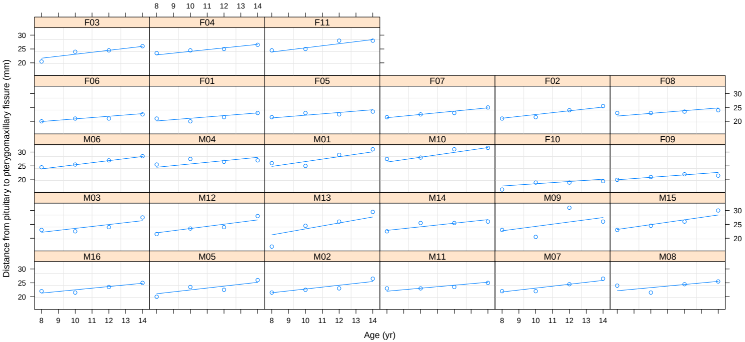
(1 covariável, intercepto e coeficiente angular aleatório)



Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

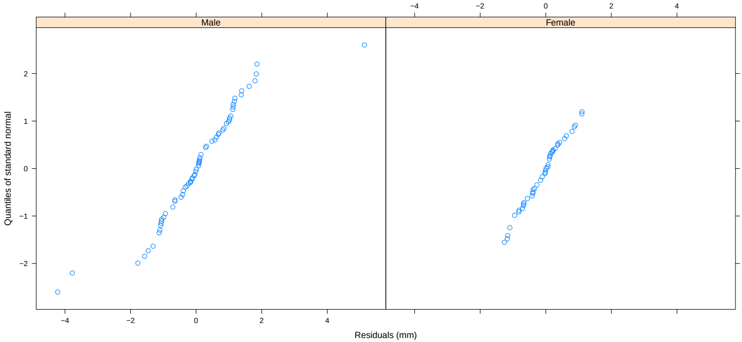
(1 covariável, intercepto e coeficiente angular aleatório)



Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

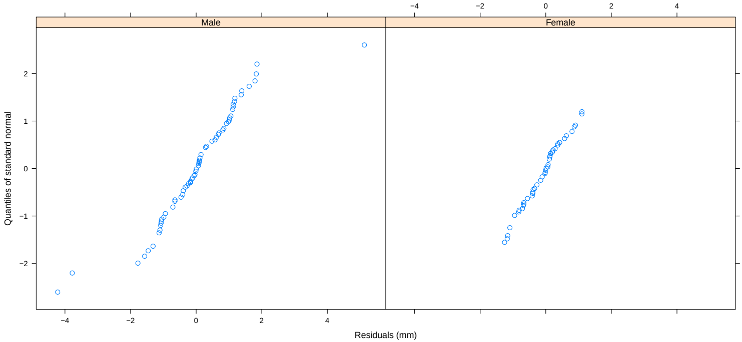
(1 covariável, intercepto e coeficiente angular aleatório)



Exemplo: Dados ortodônticos

Modelo linear com efeitos mistos

(1 covariável, intercepto e coeficiente angular aleatório)



Exemplo: Dados de produtividade industrial

Consideramos um modelo

$$Y_{ijk} = \beta_j + b_i + \epsilon ijk, \;\; i = 1, \ldots, 6; \; j = 1, 2, 3; \; k = 1, 2, 3,$$

com

$$b_i \sim N(0, \sigma_b^2), \;\; \epsilon_{ijk} \sim N(0, \sigma^2).$$

Existe um efeito fixo para cada tipo de máquina e um efeito aleatório para cada trabalhador.

Exemplo: Dados de produtividade industrial

Comandos em R:

```
> library(nlme)
> attach(Machines)
> fit.Machine1<-lme(score~ Machine,random=~ 1|Worker)
> fit.Machine1
> summary(fit.Machine1)
```

⚠ Certifique-se que os comandos foram passados corretamente para o R.

Exercício: Dados de produtividade industrial

Outro possível modelo seria

$$Y_{ijk} = \beta_j + b_i + b_{ij} + \epsilon_{ijk}, \;\; i = 1, \ldots, 6; \;\; j = 1, 2, 3; \;\; k = 1, 2, 3,$$

com

$$b_i \sim N(0, \sigma_1^2), \;\; b_{ij} \sim N(0, \sigma_2^2), \;\; \epsilon_{ijk} \sim N(0, \sigma^2).$$

Esse modelo tem efeitos aleatórios em dois níveis: os efeitos b_i para o trabalhador e b_{ij} para o tipo de máquina dentro de cada trabalhador. Devemos ler isso como "trabalhador" e "máquina dentro de trabalhador".

Exercício: Dados de produtividade industrial

Comandos em R:

Podemos simplesmente atualizar o modelo anterior com a função update

```
> fit.Machine2<-update(fit.Machine1,random=~ 1|Worker/Machine)
> fit.Machine2
> summary(fit.Machine2)
```

⚠ Certifique-se que os comandos foram passados corretamente para o R.

Pergunta:

Qual é o melhor modelo?