Diagnostic and treatment for linear mixed models

Julio M Singer

in collaboration with
Francisco M.M. Rocha and **Juvêncio S Nobre**

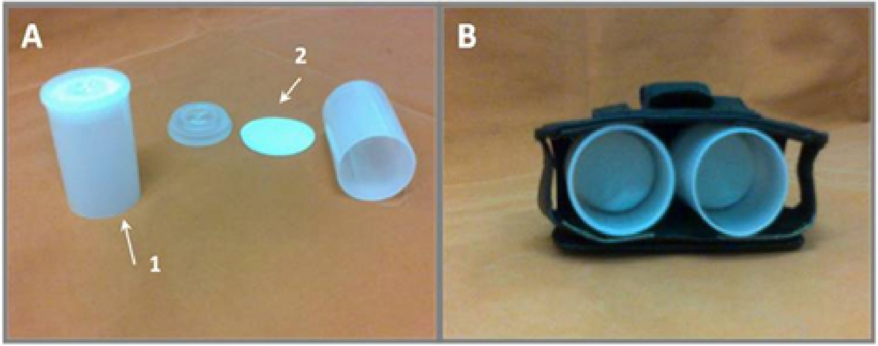
Departamento de Estatística
Universidade de São Paulo, Brazil
www.ime.usp.br/∼jmsinger

Outline

- Two examples
- Gaussian linear mixed models (LMM)
- Diagnostic tools
  - Residual analysis
  - Global influence analysis
  - Local influence analysis
- Treatment
  - Fine tuning of the model
  - LMM with elliptically-symmetric random effects
  - LMM with skew-symmetric random effects
  - Generalized linear mixed models (GLMM)
  - Generalized estimating equations based models
- Practical issues
- Where do we go from here

Ozone example

- Ozone concentration: measured with expensive instruments
- Alternative: reflectance in passive filters / calibration curve



Ozone example

- Experiment LPAE/FMUSP: predict period expected reflectance (latent value) accounting for possible outliers

Period	Reflectance	Period	Reflectance
1	27.0	6	47.9
1	34.0	6	60.4
1	17.4	6	47.3
2	24.8	7	50.4
2	29.9	7	50.7
2	32.1	7	55.9
3	35.4	8	54.9
3	63.2	8	43.2
3	27.4	8	52.1
4	51.2	9	38.8
4	54.5	9	59.9
4	52.2	9	61.1
5	77.7		
5	53.9		
5	48.2		

Ozone example

- Linear mixed model:

$$y_{ij} = \mu + a_i + e_{ij}, \quad i = 1, \ldots, 9, \; j = 1, 2, 3$$

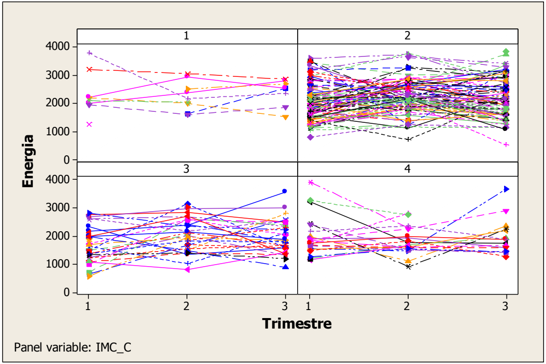
  - $a_i \sim N(0, \sigma_a^2)$ independent
  - $e_{ij} \sim N(0, \sigma^2)$ independent
  - a_i and e_{ij} independent

- Consequently
  - $\mathbb{V}(y_{ij}) = \sigma_a^2 + \sigma^2$
  - $\mathbb{C}\mathbb{o}\mathbb{v}(y_{ij}, y_{ik}) = \sigma_a^2$
  - $\mathbb{C}\mathbb{o}\mathbb{v}(y_{lj}, y_{ik}) = 0$
  - Reliability of the mean: $\rho_m = \sigma_a^2/(\sigma_a^2 + \sigma^2/3)$

Kcal intake example

- Study conducted at FMUSP
- Compare average daily kcal intake during pregnancy

Após a exclusão das gestantes com consumo maior que 4000 Kcal:



Valores de estatísticas descritivas para Energia

Trimestre	N	Média	Desvio padrão	Mínimo	Mediana	Máximo
1	160	1973,2	689,5	584,4	1903,7	3905,0
2	146	2148,8	598,1	699,4	2107,0	3754,1
3	135	2124,8	649,4	509,2	2038,5	3820,2

Kcal intake example

- Linear mixed model

$$y_{ijk} = \mu_j + \alpha_i + a_{ijk} + e_{ijk},\ \alpha_1 = 0$$

  - i indexes BMI ($1 = BMI < 24.9kg/m^2$, $2 = BMI \geq 24.9kg/m^2$)
  - j indexes period ($1 = 1st$ trimester and $2 = 2nd$ or $3rd$ trimesters)
  - k indexes women ($k = 1, \ldots, n_i$)

- $\mathbf{b}_{ik} = (a_{i1k}, a_{i2k}, a_{i3k})^\top \sim N_3(\mathbf{0}, \mathbf{G})$ independent

$$\mathbf{G} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}$$

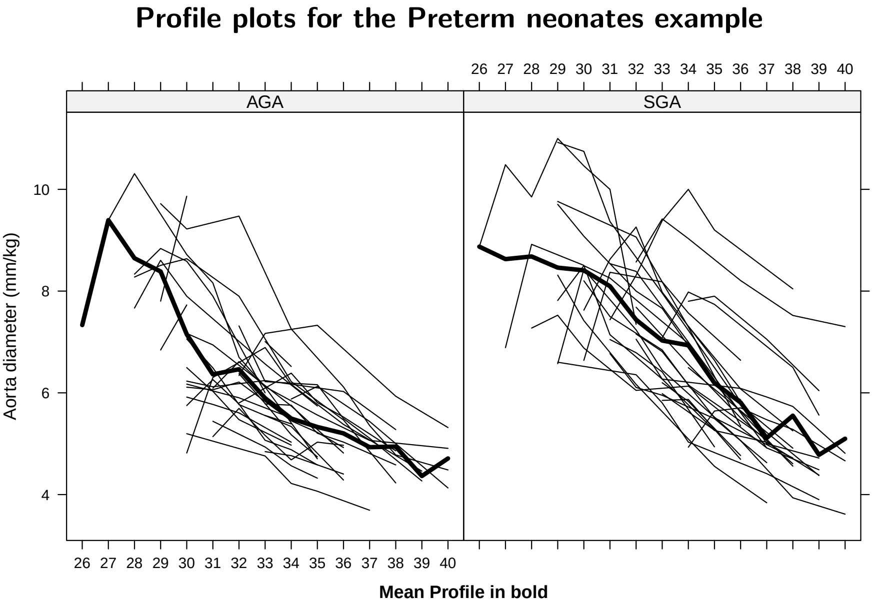
- $\mathbf{e}_{ij} = (e_{i1k}, e_{i2k}, e_{i3k})^\top \sim N_3[\mathbf{0}, \sigma^2 \mathbf{I}_3]$ independent
- $\mathbf{b}_{ik}$ and $\mathbf{e}_{ik}$ independent

Preterm neonates example

Aorta diameter per unit weight (mm/kg)

	Weeks post conception													
Weight	27	28	29	30	31	32	33	34	35	36	37	38	39	40
AGA	9.4	10.3		8.7	8.2	6.7	6.1		5.6		5.1			4.9
AGA				6.1	6.1	6.2		5.4	5.2	4.9				
AGA				5.8	6.3	5.8	5.1	4.9	4.6					
AGA			9.7	9.2		9.5		7.3		6.1	5.4	4.8		4.5
AGA						6.4	5.8	5.2	4.7					
AGA					5.4		4.9	4.6	4.3					
AGA		8.3	8.5	8.6		7.9		6.2		5.5		4.2		
AGA		7.7	8.6	7.9			6.6		5.7					
AGA								5.9	6.1		5.4			4.1
AGA							7.0	6.5						
AGA				5.2			4.8	4.2	4.1		3.7			
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
AGA				6.2	6.1		6.2			6.0		5.3		
SGA						7.2	6.8		5.5			4.7		
SGA							7.1	8.0	7.7			6.5	5.6	
SGA					7.4	8.3	9.4	10.0	9.2			8.0		
SGA						7.7		6.6		5.5		4.6		
SGA								6.5					4.4	
SGA				7.6	8.6	9.3	8.0		6.6		5.0		4.7	
SGA				6.6	8.4		8.2	7.6		6.6				
SGA						7.1	6.3			6.1	5.9	5.7		4.8
SGA					8.5	8.4					4.9			
SGA			8.3	7.4		6.2			4.6		3.8			
SGA			9.8			9.1		7.3		5.3				
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SGA	8.5													
SGA			10.9	10.7	9.4		8.0			5.8		4.9		

Preterm neonates example



Preterm neonates example

- Objective: evaluate evolution of aorta diameter of preterm neonates from birth to 40-th week post conception
- Linear mixed model suggested via exploratory analysis (Rocha & Singer, 2013, under revision)

$$y_{ijk} = \alpha_i + \beta_1(t_{1jk} - 26) + \gamma_2(t_{2jk} - 26)^2 + a_{ij} + b_{ij}(t_{ijk} - 26) + e_{ijk}$$

  - i indexes group (1=AGA and 2=SGA)
  - j indexes neonates ($j = 1, \dots, n_i$)
  - k indexes week ($k = 1, \dots, m_{ij}$)
- $\mathbf{b}_{ij} = (a_{ij}, b_{ij})^\top \sim N_2(\mathbf{0}, \mathbf{G}_i)$ independent

$$\mathbf{G}_i = \begin{bmatrix} \sigma^2_{a_i} & \sigma_{ab_i} \\ \sigma_{ab_i} & \sigma^2_{b_i} \end{bmatrix}$$

- $\mathbf{e}_{ij} = (e_{ij1}, \dots, e_{ijm_{ij}})^\top \sim N_{m_{ij}}[\mathbf{0}, \sigma^2 \mathbf{I}_{m_{ij}}]$ independent
- $\mathbf{b}_{ij}$ and $\mathbf{e}_{ij}$ independent

Gaussian LMM

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \mathbf{e}_i, \quad i = 1, \ldots, n$$

$\mathbf{y}_i$: $(m_i \times 1)$ response profile for i-th unit

$\boldsymbol{\beta}$: $(p \times 1)$ (fixed effects)

$\mathbf{X}_i$: $(m_i \times p)$ fixed effects specification matrix

$\mathbf{Z}_i$: $(m_i \times q)$ random effects specification matrix

$\mathbf{b}_i$: $(q \times 1)$ random effects, $\mathbf{b}_i \sim N_q(\mathbf{0}, \mathbf{G})$ independent

$\mathbf{e}_i$: $(m_i \times 1)$ random errors, $\mathbf{e}_i \sim N_{m_i}(\mathbf{0}, \mathbf{R}_i)$ independent

- $\mathbf{b}_i$ and $\mathbf{e}_i$ independent
- $\mathbf{G} = \mathbf{G}(\boldsymbol{\theta})$ and $\mathbf{R}_i = \mathbf{R}_i(\boldsymbol{\theta})$, $\boldsymbol{\theta}$: covariance parameters
- Marginal variance: $\mathbb{V}(\mathbf{y}_i) = \mathbf{V}_i = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^\top + \mathbf{R}_i$
- Usually $\mathbf{R}_i = \sigma^2\mathbf{I}_{m_i}$: homoskedastic conditional independence model

Gaussian LMM

Compactly

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \mathbf{e}$$

with

$$\mathbf{y} = (\mathbf{y}_1^\top, \cdots, \mathbf{y}_n^\top)^\top \quad (N \times 1,\ N = \textstyle\sum_{i=1}^n m_i)$$
$$\mathbf{X} = (\mathbf{X}_1^\top, \cdots, \mathbf{X}_n^\top)^\top \quad (N \times p)$$
$$\mathbf{Z} = \oplus_{i=1}^n \mathbf{Z}_i \quad (N \times nq)$$
$$\mathbf{b} = (\mathbf{b}_1^\top, \cdots, \mathbf{b}_n^\top)^\top \quad (nq \times 1)$$
$$\mathbf{e} = (\mathbf{e}_1^\top, \cdots, \mathbf{e}_n^\top)^\top \quad (N \times 1)$$
$$\boldsymbol{\Gamma} = \mathbf{I}_n \otimes \mathbf{G}(\boldsymbol{\theta}) \quad (nq \times nq)$$
$$\boldsymbol{\Sigma} = \oplus_{i=1}^n \mathbf{R}_i(\boldsymbol{\theta}) \quad (N \times N)$$

Consequently

$$\mathbb{V}(\mathbf{y}) = \mathbf{V} = \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}^\top + \boldsymbol{\Sigma}$$

Covariance matrix structure

1. **Uniform**

$$\mathbf{R}_i(\boldsymbol{\theta}), \mathbf{G}(\boldsymbol{\theta}) = \begin{bmatrix} \sigma^2 & \tau & \tau \\ \tau & \sigma^2 & \tau \\ \tau & \tau & \sigma^2 \end{bmatrix}$$

2. **Unstructured**

$$\mathbf{R}_i(\boldsymbol{\theta}), \mathbf{G}(\boldsymbol{\theta}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}$$

3. **AR(1)**

$$\mathbf{R}_i(\boldsymbol{\theta}), \mathbf{G}_i(\boldsymbol{\theta}) = \sigma^2 \begin{bmatrix} 1 & \phi & \phi^2 \\ \phi & 1 & \phi \\ \phi^2 & \phi & 1 \end{bmatrix}$$

Inference for Gaussian LMM

- Given $\boldsymbol{\theta}$ $[\boldsymbol{\Gamma}(\boldsymbol{\theta})$, $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ and $\mathbf{V}(\boldsymbol{\theta})]$
- BLUE of $\boldsymbol{\beta}$: $\widehat{\boldsymbol{\beta}} = \left(\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{y}$
- BLUP of $\mathbf{b}$: $\widehat{\mathbf{b}} = \boldsymbol{\Gamma} \mathbf{Z}^\top \mathbf{V}^{-1} [\mathbf{I} - \mathbf{X} \left(\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{V}^{-1}] \mathbf{y}$
    - Ozone example BLUP

$$\widehat{y}_{ij} = \overline{y} + k(\overline{y}_i - \overline{y})$$

    - Shrinkage constant

$$k = \frac{\sigma_a^2}{\sigma_a^2 + \sigma^2/3}$$

- Substituting $\widehat{\boldsymbol{\Gamma}}$ and $\widehat{\boldsymbol{\Sigma}}$ in the expressions for $\widehat{\boldsymbol{\beta}}$ and $\widehat{\mathbf{b}}$ we obtain empirical BLUE and BLUP

Inference for Gaussian LMM

Restricted maximum likelihood (REML) for $\boldsymbol{\theta}$ ($\boldsymbol{\Gamma}$, $\boldsymbol{\Sigma}$, $\mathbf{V}$):

$$-\frac{1}{2}\sum_{i=1}^{N} tr\{\mathbf{V}_i^{-1}(\widehat{\boldsymbol{\theta}})\dot{\mathbf{V}}_i(\widehat{\boldsymbol{\theta}})\} - \frac{1}{2}\sum_{i=1}^{N}[\partial Q_i(\boldsymbol{\theta})/\partial\boldsymbol{\theta}_j|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}]$$

$$-\frac{1}{2}\sum_{i=1}^{N} tr\{\mathbf{V}_i^{-1}(\widehat{\boldsymbol{\theta}})\mathbf{X}_i^\top\mathbf{V}_i^{-1}(\widehat{\boldsymbol{\theta}})\dot{\mathbf{V}}_i(\widehat{\boldsymbol{\theta}})\mathbf{V}_i^{-1}(\widehat{\boldsymbol{\theta}})\mathbf{X}_i\} = 0,$$

- $\dot{\mathbf{V}}_i(\widehat{\boldsymbol{\theta}}) = [\partial\mathbf{V}_i(\boldsymbol{\theta})/\partial\boldsymbol{\theta}_j]^\top_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$
- $Q_i(\boldsymbol{\theta}) = [\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta})]^\top\mathbf{V}_i^{-1}(\boldsymbol{\theta})[\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta})]$

- Newton-Raphson algorithm:

$$\boldsymbol{\theta}^{(l)} = \boldsymbol{\theta}^{(l-1)} - \mathbf{H}^{-1}[\boldsymbol{\theta}^{(l-1)}]\mathbf{u}[\boldsymbol{\theta}^{(l-1)}], l = 1, 2, \ldots$$

  - Score function: $\mathbf{u}(\boldsymbol{\theta}) = \partial l[\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \boldsymbol{\theta}]/\partial\boldsymbol{\theta}$
  - Hessian matrix: $\mathbf{H}(\boldsymbol{\theta}) = \partial^2 l[\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \boldsymbol{\theta}]/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top$
  - Stopping rule: $\|\boldsymbol{\theta}^{(l)} - \boldsymbol{\theta}^{(l-1)}\| < \varepsilon$, $\varepsilon > 0$
- Fisher Scoring algorithm: $\mathbf{H}(\boldsymbol{\theta})$ replaced by its expectation
- EM algorithm

Statistical properties of estimators

- $\widehat{\boldsymbol{\beta}}(\boldsymbol{\theta}) \sim N\{\boldsymbol{\beta}, \mathbf{V}_{\widehat{\boldsymbol{\beta}}}(\boldsymbol{\theta})\}$ with $\mathbf{V}_{\widehat{\boldsymbol{\beta}}}(\boldsymbol{\theta}) = [\sum_{i=1}^{N} \mathbf{X}_i^\top \mathbf{V}_i^{-1}(\boldsymbol{\theta})\mathbf{X}_i]^{-1}$
- $\widehat{\boldsymbol{\theta}} \approx N\{\boldsymbol{\theta}, \mathbf{V}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\}$ with $\mathbf{V}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})$ denoting an $(m \times m)$ matrix for which the element $(r, s), r, s = 1, \ldots, m$, is $[\mathbf{V}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})]_{rs} = \frac{1}{2} \sum_{i=1}^{N} tr\{\mathbf{V}_i^{-1}(\boldsymbol{\theta})\dot{\mathbf{V}}_{ir}(\boldsymbol{\theta})\mathbf{V}_i^{-1}(\boldsymbol{\theta})\dot{\mathbf{V}}_{is}(\boldsymbol{\theta})\}$
- $\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\theta}}) \approx N\{\boldsymbol{\beta}, \mathbf{V}_{\widehat{\boldsymbol{\beta}}}(\boldsymbol{\theta})\}$
- $\mathbf{V}_{\widehat{\boldsymbol{\beta}}}(\boldsymbol{\theta})$ and $\mathbf{V}_{\widehat{\boldsymbol{\theta}}}(\boldsymbol{\theta})$ may be estimated by the inverse of Fisher's information matrix
- Asymptotic results hold even without normality provided n is sufficiently large

Types of residuals in Gaussian LMM

- Three types of residuals that accommodate the extra source of variability present in linear mixed models, namely:

  i) Marginal residuals, $\widehat{\boldsymbol{\xi}} = \mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}$ predictors of marginal errors, $\boldsymbol{\xi} = \mathbf{y} - \mathbb{E}[\mathbf{y}] = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} = \mathbf{Z}\mathbf{b} + \mathbf{e}$

  ii) Conditional residuals, $\widehat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}} - \mathbf{Z}\widehat{\mathbf{b}}$ predictors of conditional errors $\mathbf{e} = \mathbf{y} - \mathbb{E}[\mathbf{y}|\mathbf{b}] = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{b}$

  iii) BLUP, $\mathbf{Z}\widehat{\mathbf{b}}$, predictors of random effects, $\mathbf{Z}\mathbf{b} = \mathbb{E}[\mathbf{y}|\mathbf{b}] - \mathbb{E}[\mathbf{y}] = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{b}) - (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$

Confounded Residuals

- Hilden-Minton (1995, PhD thesis, UCLA): residual is pure for an error if it depends only on fixed components and on error it is supposed to predict
- Otherwise: confounded residuals
- Given that

$$\begin{aligned}\widehat{\boldsymbol{\xi}} &= [\mathbf{I} - \mathbf{X}(\mathbf{X}^\top \widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}^\top \widehat{\mathbf{V}}^{-1}]\boldsymbol{\xi}, \\ \widehat{\mathbf{e}} &= \widehat{\boldsymbol{\Sigma}}\widehat{\mathbf{Q}}\mathbf{e} + \widehat{\boldsymbol{\Sigma}}\widehat{\mathbf{Q}}\mathbf{Z}\mathbf{b}, \\ \mathbf{Z}\widehat{\mathbf{b}} &= \mathbf{Z}\widehat{\boldsymbol{\Gamma}}\mathbf{Z}^\top \widehat{\mathbf{Q}}\mathbf{Z}\mathbf{b} + \mathbf{Z}\widehat{\boldsymbol{\Gamma}}\mathbf{Z}^\top \widehat{\mathbf{Q}}\mathbf{e},\end{aligned}$$

where $\mathbf{Q} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^\top\mathbf{V}^{-1}$, we have

- $\widehat{\mathbf{e}}$ is confounded with $\widehat{\mathbf{b}}$
- $\mathbf{Z}\widehat{\mathbf{b}}$ is confounded with $\widehat{\mathbf{e}}$
- Exception: columns of $\mathbf{Z}$ belong to the space generated by the columns of $\mathbf{X}$

Marginal Residuals

- Since $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\xi}$, plots of marginal residuals ($\widehat{\boldsymbol{\xi}}_{ij}$) versus explanatory variables or fitted values ($\widehat{\mathbf{y}}_{ij} = \mathbf{x}_{ij}^\top \widehat{\boldsymbol{\beta}}$) to check for linearity
- Index plots of $\widehat{\boldsymbol{\xi}}_{ij}$ to detect outlying observations
- Lesaffre and Verbeke (1998, Biometrics): (unit) index plots of $\mathcal{R}_i = \widehat{\mathbf{V}}_i^{-1/2}\widehat{\boldsymbol{\xi}}_i$ useful to check appropriateness of the within-unit covariance matrix
  - When $\mathcal{V}_i = ||\mathbf{I}_{m_i} - \mathcal{R}_i\mathcal{R}_i^\top||^2$ is small, within-units covariance matrix is acceptable for unit i [$\mathbb{E}(\widehat{\boldsymbol{\xi}}_i\widehat{\boldsymbol{\xi}}_i^\top) = \widehat{\mathbf{V}}_i$??]
  - *In lieu* of $\widehat{\mathbf{V}}_i$, we suggest using $\widehat{\mathbb{V}}_i(\boldsymbol{\xi})$, the i-th diagonal block of

$$\widehat{\mathbb{V}}(\widehat{\boldsymbol{\xi}}) = [\widehat{\mathbf{V}} - \mathbf{X}(\mathbf{X}\widehat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}^\top]$$

  - Also, we suggest using (unit) index plots of $\mathcal{V}_i^* = \mathcal{V}_i / \sum_{i=1}^n \mathcal{V}_i$ to allow comparison among different models

Conditional Residuals

- Conditional studentized residuals (Nobre and Singer, 2007, Biometrical Journal): $\widehat{e}^*_{ij} = \widehat{e}_{ij}/\sqrt{\widehat{p}_{ij}}$
  - p_{ij}: ij-th element of the main diagonal of $\widehat{\mathbb{V}}(\widehat{\mathbf{e}}) = \widehat{\boldsymbol{\Sigma}}\widehat{\mathbf{Q}}\widehat{\boldsymbol{\Sigma}}$
  - We suggest $\widehat{\mathbf{e}}^*_i = [\widehat{\mathbb{V}}_i(\widehat{\mathbf{e}})]^{-1/2}\widehat{\mathbf{e}_i}$ where $\widehat{\mathbb{V}}_i(\widehat{\mathbf{e}})$ is the i-th block of $\widehat{\mathbb{V}}(\widehat{\mathbf{e}})$
- Index plot of $\widehat{e}^*_{ij}$ to detect outlying observations
- Plot of $\widehat{e}^*_{ij}$ versus predicted values $(\widehat{\mathbf{y}}^*_{ij} = \mathbf{x}_{ij}^\top\widehat{\boldsymbol{\beta}} + \mathbf{z}_{ij}^\top\widehat{\mathbf{b}}_i)$ to check for homoskedasticity of conditional errors: $(\mathbf{R}_i = \sigma^2\mathbf{I}_{m_i})$
- Check for normality of conditional errors
  - Must take confounding into consideration
  - $\widehat{\mathbf{e}}$ may not be adequate to check for normality of $\mathbf{e}$
  - When $\mathbf{b}$ is non-normal, $\widehat{\mathbf{e}}$ may not be normal even when $\mathbf{e}$ is

Conditional Residuals

- Hilden-Minton (1995, PhD thesis, UCLA): ability to check for normality of $\mathbf{e}$, using $\widehat{\mathbf{e}}$, decreases as $\mathbb{V}[\boldsymbol{\Sigma}\mathbf{Q}\mathbf{Z}^\top\mathbf{b}] = \boldsymbol{\Sigma}\mathbf{Q}\mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}^\top\mathbf{Q}\boldsymbol{\Sigma}$ increases in relation to $\mathbb{V}[\boldsymbol{\Sigma}\mathbf{Q}\mathbf{e}] = \boldsymbol{\Sigma}\mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}\boldsymbol{\Sigma}$
- Fraction of confounding for the ij-th conditional residual $\widehat{e}_{ij}$

$$0 \leq F_{ij} = \frac{\mathbf{u}_{ij}^\top\boldsymbol{\Sigma}\mathbf{Q}\mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}^\top\mathbf{Q}\boldsymbol{\Sigma}\mathbf{u}_{ij}}{\mathbf{u}_{ij}^\top\boldsymbol{\Sigma}\mathbf{Q}\boldsymbol{\Sigma}\mathbf{u}_{ij}} = 1 - \frac{\mathbf{u}_{ij}^\top\boldsymbol{\Sigma}\mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}\boldsymbol{\Sigma}\mathbf{u}_{ij}}{\mathbf{u}_{ij}^\top\boldsymbol{\Sigma}\mathbf{Q}\boldsymbol{\Sigma}\mathbf{u}_{ij}} \leq 1$$

  where $\mathbf{u}_{ij}$ is ij-th column of $\mathbf{I}_N$
- Least confounded residuals: linear transformation $\mathbf{C}\widehat{\mathbf{e}}$ that maximizes

$$\lambda_{ij} = \frac{\mathbf{c}_{ij}^\top\boldsymbol{\Sigma}\mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}\boldsymbol{\Sigma}\mathbf{c}_{ij}}{\mathbf{c}_{ij}^\top\boldsymbol{\Sigma}\mathbf{Q}\boldsymbol{\Sigma}\mathbf{c}_{ij}}, \quad i = 1, \ldots, n, \; j = 1, \ldots, m_i$$

- Least confounded residuals: homoskedastic, uncorrelated, variance σ^2
- QQ plots and histograms to check for normality of conditional residuals

EBLUP

- EBLUP: reflects the difference between the predicted expected response (latent value) for i-th unit and population average
- Unit index plots of $\mathcal{M}_i = \widehat{\mathbf{b}}_i^\top \{\widehat{\mathbb{V}}[\widehat{\mathbf{b}}_i]\}^{-1}\widehat{\mathbf{b}}_i$ to detect outlying units
- We suggest (unit) index plots of $\mathcal{M}_i^* = \mathcal{M}_i / \sum_{i=1}^n \mathcal{M}_i$ to allow comparison among different models
- To assess normality of random effects:
  - No confounding: χ_q^2 QQ plot of $\mathcal{M}_i$
  - With confounding: ?

Residual diagnostics for Gaussian LMM

Diagnostic for	Residual	Plot
Linearity of effects fixed ($\mathbb{E}[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta}$)	Marginal	$\widehat{\boldsymbol{\xi}}_{ij}^*$ *vs* fitted values or explanatory variables
Presence of outlying observations	Marginal	$\widehat{\boldsymbol{\xi}}_{ij}^*$ *vs* observation indices
Within-subjects covariance matrix ($\mathbf{V}_i$)	Marginal	$\mathcal{V}_i^*$ *vs* unit indices
Presence of outlying observations	Conditional	$\widehat{e}_{ij}^*$ *vs* observation indices
Homoskedasticity of conditional errors ($\mathbf{e}_i$)	Conditional	$\widehat{e}_{ij}^*$ *vs* predicted values
Normality of conditional errors ($\mathbf{e}_i$)	Conditional	Gaussian QQ plot for $\widehat{\mathbf{e}}_{ij}^*$ or $\mathbf{c}_{ij}^\top \widehat{\mathbf{e}}^*$
Presence of outlying units	EBLUP	$\mathcal{M}_i^*$ *vs* unit indices
Normality of the random effects ($\mathbf{b}_i$)	EBLUP	χ_q^2 QQ plot for $\mathcal{M}_i$

Global influence analysis (leverage)

- Generalized leverage matrix (for fixed effects), $\widehat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}$

$$\mathbf{L}_1 = \frac{\partial \widehat{\mathbf{y}}}{\partial \mathbf{y}^\top} = \mathbf{X}\left(\mathbf{X}^\top \mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^\top \mathbf{V}^{-1}, \quad tr(\mathbf{L}_1) = p$$

- High leverage unit: $tr(\mathbf{L}_{1i})/m_i > 2p/n$ where $\mathbf{L}_{1i}$ is i-th block of $\mathbf{L}_1$
- High leverage observation: $\mathbf{L}_{1ijj} > 2p/N$ where $\mathbf{L}_{1ijj}$ is j-th diagonal element of $\mathbf{L}_{1i}$
- Since $\widehat{\mathbf{y}}^* = \mathbf{X}\widehat{\boldsymbol{\beta}} + \mathbf{Z}\widehat{\mathbf{b}}$: Generalized joint leverage matrix

$$\mathbf{L} = \frac{\partial \widehat{\mathbf{y}}^*}{\partial \mathbf{y}^\top} = \frac{\partial \widehat{\mathbf{y}}}{\partial \mathbf{y}^\top} + \frac{\partial \mathbf{Z}\widehat{\mathbf{b}}}{\partial \mathbf{y}^\top} = \mathbf{L}_1 + \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}^\top\mathbf{Q}$$

- $\mathbf{L}_2 = \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}^\top$: variability explained by random effects
- Demidenko and Stukel (2005, SIM) suggest using $\mathbf{H}_2 = \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}^\top\mathbf{Q}$ as generalized leverage matrix for random effects
- Since $\mathbf{H}_2 = \mathbf{L}_2\mathbf{V}^{-1}[\mathbf{I}_n - \mathbf{L}_1]$, Nobre and Singer (2011, JAS) argue for $\mathbf{L}_2 = \mathbf{Z}\boldsymbol{\Gamma}\mathbf{Z}^\top$

Global influence analysis (case deletion)

- Impact of a unit on some characteristic (*e.g.*, parameter estimate)
- $\widehat{\boldsymbol\beta} - \widehat{\boldsymbol\beta}_{(I)} = (\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{U}_I(\mathbf{U}_I^\top\mathbf{Q}\mathbf{U}_I)^{-1}\mathbf{U}_I^\top\mathbf{Q}\mathbf{y}$
- Cook´s distance

$$D_I = \frac{(\widehat{\boldsymbol\beta} - \widehat{\boldsymbol\beta}_{(I)})^\top(\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X})(\widehat{\boldsymbol\beta} - \widehat{\boldsymbol\beta}_{(I)})}{p} = \frac{(\widehat{\mathbf{y}} - \widehat{\mathbf{y}}_{(I)})^\top\mathbf{V}^{-1}(\widehat{\mathbf{y}} - \widehat{\mathbf{y}}_{(I)})}{p}$$

- Conditional Cook distance (for $\mathbf{R}_i = \sigma^2\mathbf{I}_{m_i}$)

$$D_{i(j)}^{cond} = \sum_{i=1}^{n}\frac{(\widehat{\mathbf{y}}_i^* - \widehat{\mathbf{y}}_{i(j)}^*)^\top(\widehat{\mathbf{y}}_i^* - \widehat{\mathbf{y}}_{i(j)}^*)}{\sigma^2(n+p)}$$

where $\widehat{\mathbf{y}}_i^* = \mathbf{X}_i\widehat{\boldsymbol\beta} + \mathbf{Z}_i\widehat{\mathbf{b}}$, $\widehat{\mathbf{y}}_{i(j)}^* = \mathbf{X}_i\widehat{\boldsymbol\beta}_{(i(j))} + \mathbf{Z}_i\widehat{\mathbf{b}}_{(i(j))}$

- Ratio of variance ellipsoids

$$\rho_{(I)} = \frac{\left|\widehat{\mathbb{V}}(\widehat{\boldsymbol\beta}_{(I)})\right|}{|\widehat{\mathbb{V}}(\widehat{\boldsymbol\beta})|} = \left|\mathbf{I}_N + \mathbf{X}^\top\mathbf{V}^{-1}\mathbf{U}_I\left(\mathbf{U}_I^\top\mathbf{Q}\mathbf{U}_I\right)^{-1}\mathbf{U}_I^\top\mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\right|$$

Global influence analysis

Diagnostic for effect on	Global influence measure	Index plot of
Fixed portion of fitted value ($\mathbf{X}\widehat{\boldsymbol{\beta}}$)	Generalized marginal leverage matrix $\mathbf{L}_1$	$\mathbf{L}_{1i(jj)}$ $[tr(\mathbf{L}_{1i})/m_i]$ *vs* observations (units)
Random portion of fitted value ($\mathbf{Z}\widehat{\mathbf{b}}$)	Generalized random component marginal leverage matrix $\mathbf{L}_2$	$\mathbf{L}_{2i(jj)}$ $[tr(\mathbf{L}_{2i})/m_i]$ *vs* observations (units)
Regression coefficients ($\widehat{\boldsymbol{\beta}}$)	Cook's distance D_I	$D_{i(j)}$ $[D_i]$ *vs* observations (units)
Covariance matrix of regression coefficients $[\widehat{\mathbb{V}}(\widehat{\boldsymbol{\beta}})]$	Ratio of variance ellipsoids $\rho_{(I)}$	$\rho_{i(j)}$ $[\rho_i]$ *vs* observations (units)

Diagnostic for effect on	Index plot of
Regression coefficients ($\widehat{\boldsymbol{\beta}}$)	$D_{1i(j)}^{cond}$ *vs* observations
Random effects ($\widehat{\mathbf{b}}$)	$D_{2i(j)}^{cond}$ *vs* observations
Changes in covariance between $\widehat{\boldsymbol{\beta}}$ and $\widehat{\mathbf{b}}$	$D_{3i(j)}^{cond}$ *vs* observations

Local influence analysis

- Changes in the analysis resulting from small perturbations on the data or on some element of the model
- Behaviour of likelihood displacement: $LD(\boldsymbol{\omega}) = 2\left\{L(\widehat{\boldsymbol{\psi}}) - L(\widehat{\boldsymbol{\psi}}_{\boldsymbol{\omega}}|\boldsymbol{\omega})\right\}$
    - L: likelihood for proposed model
    - $\boldsymbol{\psi}$: parameter vector
    - $\boldsymbol{\omega}$: vector of perturbations
    - $\widehat{\boldsymbol{\psi}}$ and $\widehat{\boldsymbol{\psi}}_{\boldsymbol{\omega}}$: MLE of $\boldsymbol{\psi}$ based on $L(\boldsymbol{\psi})$ and $L(\boldsymbol{\psi}|\boldsymbol{\omega})$
- Usual perturbation schemes
    - Response variables: $\mathbf{y}_i(\boldsymbol{\omega}_i) = \mathbf{y}_i + \boldsymbol{\omega}_i$
    - Explanatory variables: $\mathbf{X}_i(\mathbf{W}_i) = \mathbf{X}_i + \mathbf{W}_i$
    - Random effects covariance matrix: $\mathbf{G}(\omega_i) = \omega_i \mathbf{G}$
    - Error covariance matrix: $\mathbf{R}_i(\omega_i) = \omega_i \mathbf{R}_i$
- Index plots of normalized eigenvectors corresponding to the largest eigenvalue of $-\mathbf{H}^\top \ddot{\mathbf{F}}^{-1}\mathbf{H}$ where $\ddot{\mathbf{F}} = [\partial^2 L(\boldsymbol{\psi})/\partial\boldsymbol{\psi}^\top \partial\boldsymbol{\psi}]_{\boldsymbol{\psi}=\widehat{\boldsymbol{\psi}}}$ and $\mathbf{H} = [\partial^2 L(\boldsymbol{\psi}|\boldsymbol{\omega})/\partial\boldsymbol{\psi}^\top \partial\boldsymbol{\omega}]_{\boldsymbol{\omega}=\boldsymbol{\omega}_0;\boldsymbol{\psi}=\widehat{\boldsymbol{\psi}}}$ suggest units or observations with greater impact

Treatment: fine tuning of model

- Diagnostic tools depend on correct specification of covariance structure
- Examination of linear models fitted to individual profiles (or to rows of within-unit covariance matrix when available) [Rocha and Singer (2012, submitted)]
  - Use of simple t-tests
  - Coefficients significantly different from zero are candidates for random effects
- Plots of covariances and correlations *versus* lags (Grady and Helms (1995, SIM): useful to identify auto-regressive structures
- Modelling covariance structure as function of explanatory variables [Singer and Cúri (2006, Environ Ecol Stat)]

Treatment: heavy-tailed distributions

- Elliptically-symmetric distributions
  - Useful to accommodate outliers
  - Density function: $f(\mathbf{y}) = |\boldsymbol{\Sigma}|^{-1/2} g[(\mathbf{y}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})]$
    - g non-negative valued function
    - $\mathbb{E}(\mathbf{y}) = \boldsymbol{\mu}$, $\mathbb{V}(\mathbf{y}) = \boldsymbol{\Omega} = \alpha\boldsymbol{\Sigma}$
    - α convenient constant ($= \nu/(\nu-2), \nu > 2$ for multivariate-t with ν df)
  - Includes multivariate-t, slash, contaminated normal etc
  - LMM may be defined hierarchically
    - $\mathbf{y}_i|\mathbf{b}_i \sim ES_{m_i}[\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i, \mathbf{R}_i(\boldsymbol{\theta}), \alpha_i, \gamma_i]$
    - $\mathbf{b}_i \sim ES_q[\mathbf{0}, \mathbf{G}(\boldsymbol{\theta}), \alpha_i]$
    - $\mathbf{e}_i \sim ES_{m_i}[\mathbf{0}, \mathbf{R}_i(\boldsymbol{\theta}), \gamma_i]$
  - Joint distribution of $(\mathbf{y}_i^\top, \mathbf{b}_i^\top)^\top$ may not belong to the same class
  - Maximum likelihood estimation similar to Gaussian case (but more problematic)
- Asymptotic properties of estimators still deserves study
- Local influence considered by Osorio et al. (2007, CSDA)
- Residual analysis and global influence analysis ?

Treatment: asymmetric distributions

- Skew-elliptical distributions
- Fitting of general case is difficult in practice [Jara et al. (2008, CSDA)]
- In general: Bayesian methods
- Alternative: skew-normal hierarchical LMM
  - $\mathbf{y}_i|[\boldsymbol{\beta}, \mathbf{b}_i, \mathbf{R}_i(\boldsymbol{\theta}), \boldsymbol{\Lambda}_{e_i}] \sim SN_{m_i}[\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i, \mathbf{R}_i(\boldsymbol{\theta}), \boldsymbol{\Lambda}_{e_i}]$
  - $\mathbf{b}_i|[\mathbf{G}(\boldsymbol{\theta}), \boldsymbol{\Lambda}_b] \sim SN_q[\mathbf{0}, \mathbf{G}(\boldsymbol{\theta}), \boldsymbol{\Lambda}_b]$
  - $\boldsymbol{\Lambda}_{e_i}$ and $\boldsymbol{\Lambda}_b$ are asymmetry parameters
- Asymptotic properties of estimators still deserves study
- Local influence based on EM algorithm considered by Bolfarine et al. (2007, Sankhya)
- Residual analysis and global influence analysis ?

Treatment: GLMM models

- GLMM allow non-normal distributions and non-linear models
- Models may be defined in a two-stage approach
  - $f(y_{ij}|\mathbf{b}_i) = \exp\{\phi[y_{ij}\vartheta_{ij} - a(\vartheta_{ij})] + c(y_{ij}, \phi)\}$
  - a and c: known functions and ϕ: scale parameter
  - $\mathbb{E}(y_{ij}|\mathbf{b}_i) = \mu_{ij} = da(\vartheta_{ij})/d\vartheta_{ij}$
  - $\mathbb{V}(y_{ij}|\mathbf{b}_i) = \phi^{-1}d^2(\vartheta_{ij})/d\vartheta_{ij}^2$
  - $g(\mu_{ij}) = \eta_{ij} = \mathbf{x}_{ij}^\top\boldsymbol{\beta} + \mathbf{z}_{ij}^\top\mathbf{b}_i$
  - Second stage, usually: $\mathbf{b}_i \sim N[\mathbf{0}, \mathbf{G}(\boldsymbol{\theta})]$
- Fitting is complicated (usually based on EM algorithm)
- Interpretation of parameters is not straightforward
- Residual analysis?
- Case deletion analysis [Xu et al. (2006, CSDA)]

Treatment: GEE based models

- GEE-based models focus on marginal distributions: not mixed model
- No need to specify form of underlying distribution
  - $\mathbb{E}(y_{ij}) = \mu_{ij}$ and $\mathbb{V}(y_{ij}) = \phi^{-1}\nu(\mu_{ij})$
  - $\nu(\mu_{ij})$: known function of the mean and ϕ: scale parameter.
  - Relation between response mean and explanatory variables: $g(\mu_{ij}) = \eta_{ij} = \mathbf{x}_{ij}^\top \boldsymbol{\beta}$
  - Working correlation matrix: $\mathbf{V}_{Wi}(\boldsymbol{\theta}) = \phi \mathbf{A}_i^{1/2} \mathbf{R}_W(\boldsymbol{\theta}) \mathbf{A}_i^{1/2}$
  - $\mathbf{A}_i = \oplus_{j=1}^{m_i} \nu(\mu_{ij})$
  - $\mathbf{R}_W(\boldsymbol{\theta})$: known positive-definite matrix
- Generalized estimating equations

$$\sum_{i=1}^{n} \mathbf{X}_i^\top \boldsymbol{\Delta}_i [\mathbf{V}_{Wi}(\boldsymbol{\theta})]^{-1} [\mathbf{y}_i - \boldsymbol{\mu}_i(\mathbf{X}_i^\top \boldsymbol{\beta})] = \mathbf{0}$$

- $\widehat{\boldsymbol{\beta}}$ asymptotically normal even if working covariance matrix misspecified
- Residual analysis: Venezuela et al. (2007, JSCSimulation)
- Local influence: Venezuela et al. (2011, CSDA)

Computation

Library	Function	Fits	Random effects distribution	$\mathbf{G}$ or $\mathbf{R}_W$ matrix	Error distribution	$\mathbf{R}_i$ matrix
lme4	lmer	LMM	gaussian	unstructured $\mathbf{G}$	gaussian	$\sigma^2 \mathbf{I}_{m_i}$
	nlmer	NLMM	gaussian	unstructured $\mathbf{G}$	gaussian	structured
	glmer	GLMM	gaussian	unstructured $\mathbf{G}$	exponential family	NA
nlme	lme	LMM	gaussian	structured $\mathbf{G}$	gaussian	structured
	nlme	NLMM	gaussian	structured $\mathbf{G}$	gaussian	structured
	gls	LM	NA	NA	gaussian	structured
gee	gee	GEE-based model	NS	structured $\mathbf{R}_W$	exponential family or NS	NA
geepack	geeglm	GEE-based model	NS	structured $\mathbf{R}_W$	exponential family or NS	NA
heavy	heavyLme	ES-LMM	elliptically symmetric	unstructured $\mathbf{G}$	elliptically symmetric	NA NA

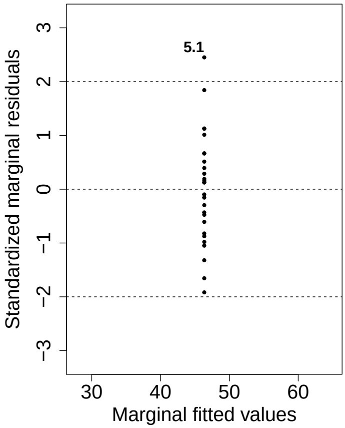
NA: not applicable
NS: not specified

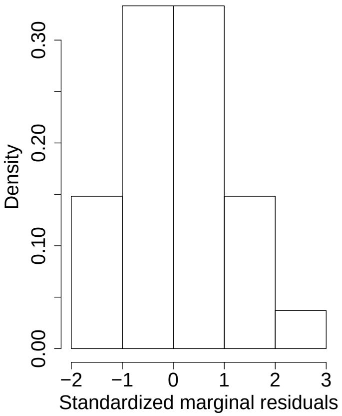
- Functions for diagnostic available only from authors
- Difficult to use in more complicated problems
- First version of functions for residual diagnostic based on lme4 and nlme being developed

Ozone example - standard model (A)

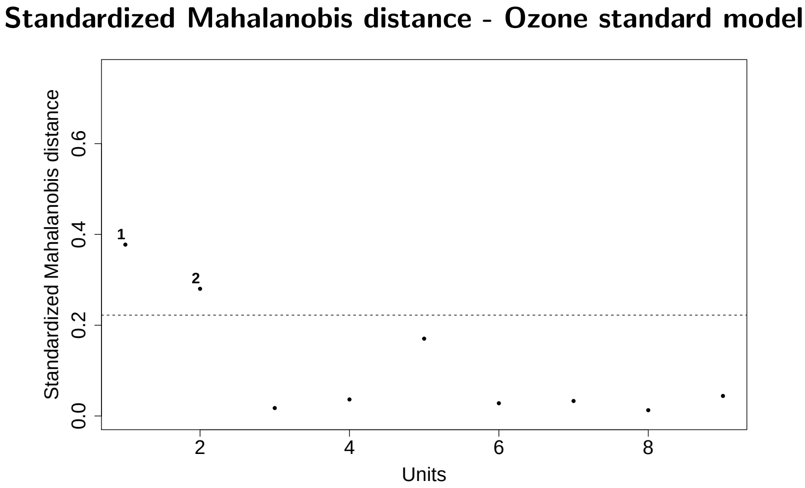
Results (standard model): $\widehat{\mu} = 46.1$, $\widehat{\sigma}_a^2 = 100.4$, $\widehat{\sigma}^2 = 104.8$, $k = 0.75$

Standardized marginal residuals - Ozone standard model

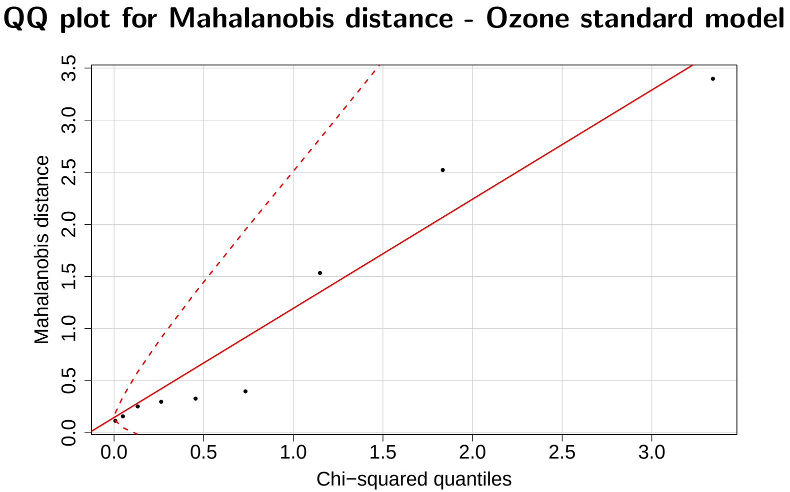




Ozone example - standard model (B)

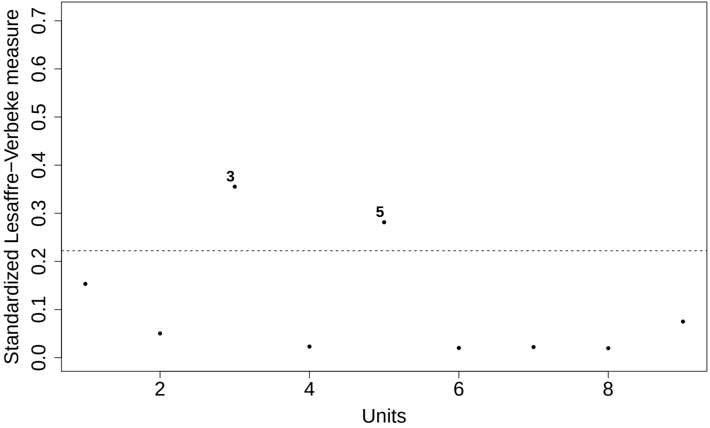


Ozone example standard model (C)

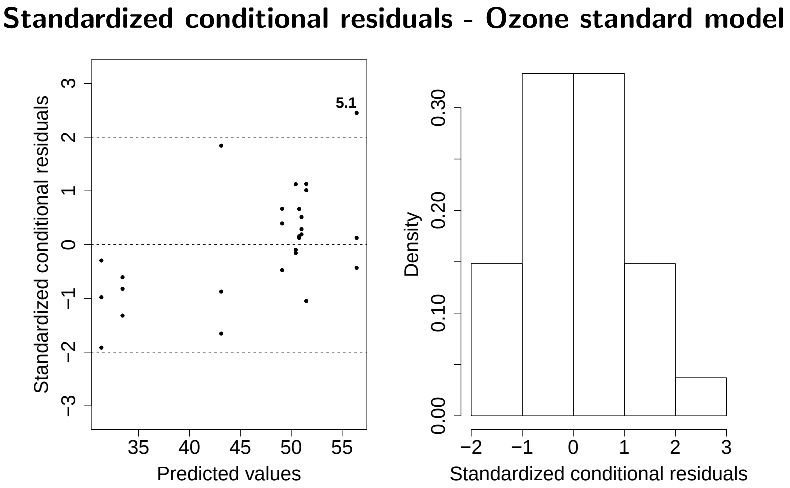


Ozone example standard model (D)

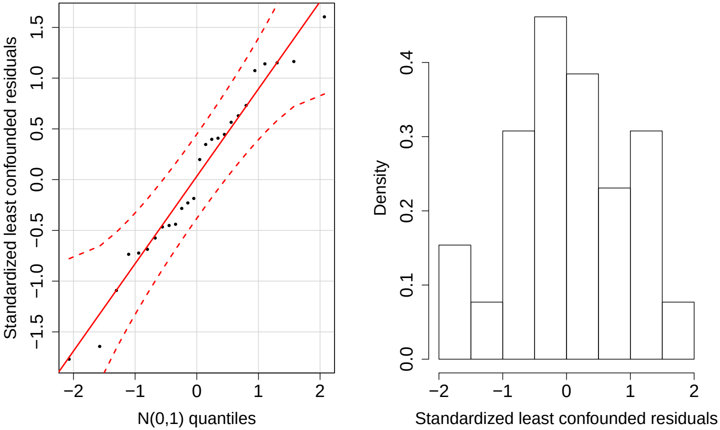
Standardized LeSaffre-Verbeke measure - Ozone standard model



Ozone example standard model (E)



Standardized least confounded conditional residuals - Ozone standard model



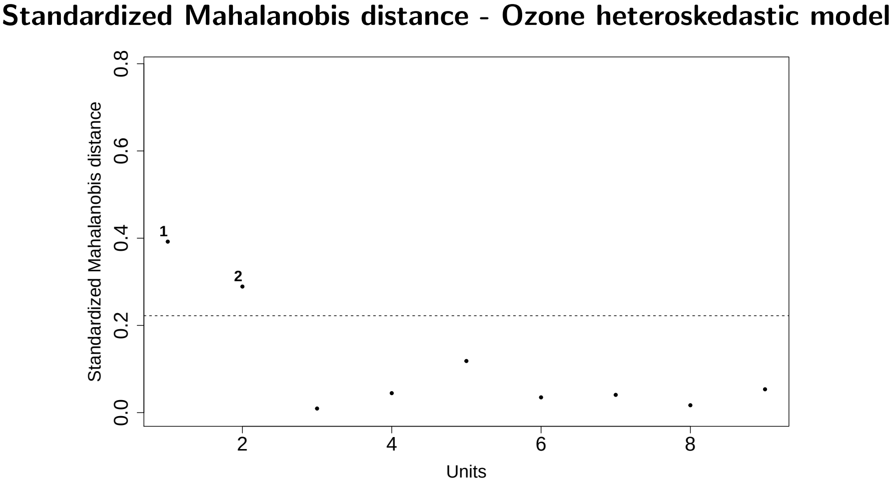
Ozone example heteroskedastic model (A)

- Suggested (heteroskedastic) model

$$y_{ij} = \mu + a_i + e_{ij} \text{ with } e_{ij} \sim N(0, \sigma_i^2)$$

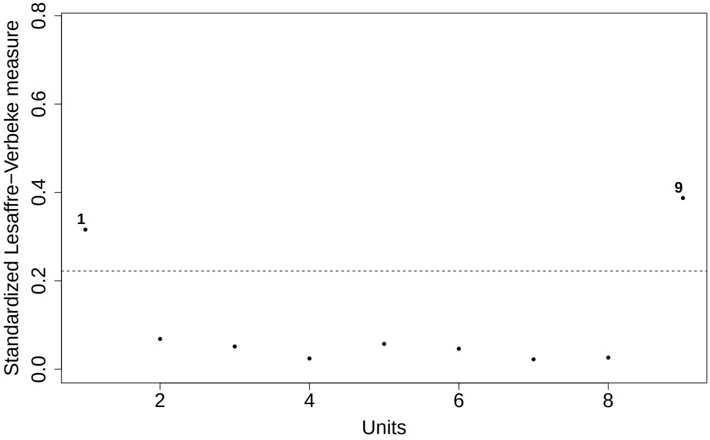
- For parsimony: $\sigma_i^2 = \tau^2, i = 3, 5$, $\sigma_i^2 = \sigma^2$, otherwise
- Shrinkage constant: $k_i = \sigma_a^2/(\sigma_a^2 + \sigma_i^2/3)$
- Results heteroskedastic model: $\widehat{\mu} = 46.4$, $\widehat{\sigma}_a^2 = 114.3$, $\widehat{\sigma}^2 = 49.6$, $\widehat{\tau}^2 = 274.0$, $k_{i\neq 3,5} = 0.87$, $k_{i=3,5} = 0.56$
- Results homoskedastic model): $\widehat{\mu} = 46.1$, $\widehat{\sigma}_a^2 = 100.4$, $\widehat{\sigma}^2 = 104.8$, $k = 0.75$
- Predicted latent values
  - Heteroskedastic model: $\widehat{y}_{3*} = 43.8$, $\widehat{y}_{5*} = 53.9$
  - Standard model: $\widehat{y}_{3*} = 42.8$, $\widehat{y}_{5*} = 56.4$

Ozone example heteroskedastic model (B)



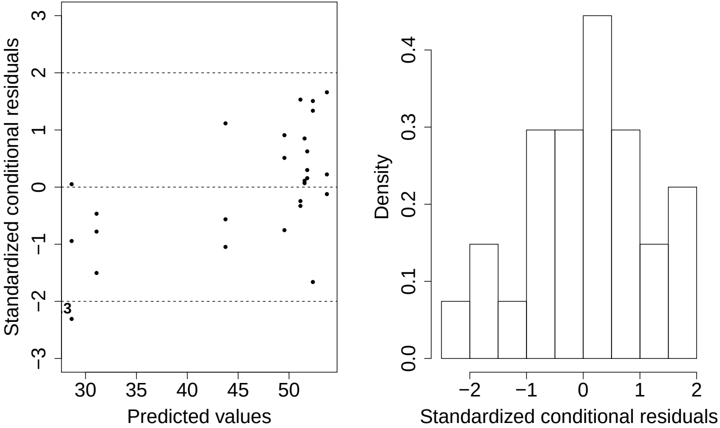
Ozone example heteroskedastic model (C)

Standardized Lesaffre-Verbeke measure - Ozone heteroskedastic model



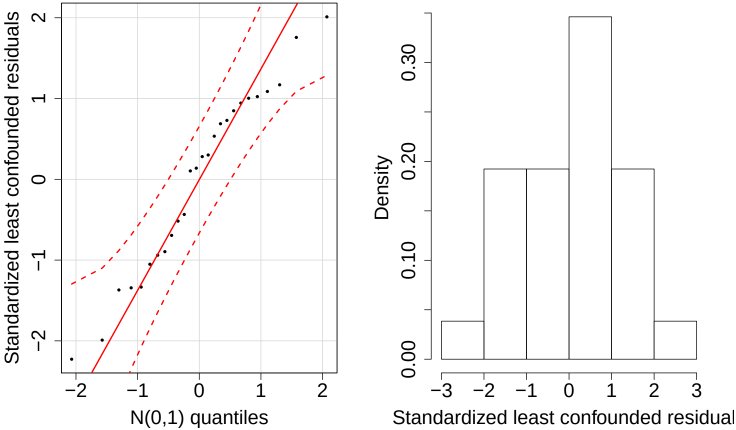
Ozone example heteroskedastic model (D)

Standardized conditional residuals - Ozone heteroskedastic model



Ozone example heteroskedastic model (E)

Standardized least confounded conditional residuals - Ozone heteroskedastic model



Ozone example heteroskedastic model (F)

Period	Reflectance	Period	Reflectance
1	27.0	6	47.9
1	34.0	6	60.4
1	17.4	6	47.3
2	24.8	7	50.4
2	29.9	7	50.7
2	32.1	7	55.9
3	35.4	8	54.9
3	63.2	8	43.2
3	27.4	8	52.1
4	51.2	9	38.8
4	54.5	9	59.9
4	52.2	9	61.1
5	77.7		
5	53.9		
5	48.2		

Kcal intake example (A)

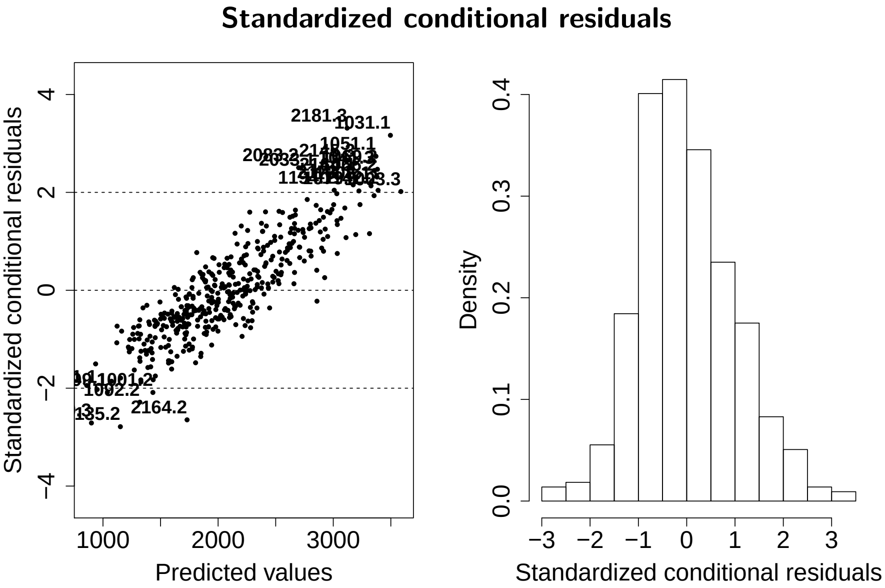
- Model: $y_{ijk} = \mu_j + \alpha_i + a_{ijk} + e_{ijk}, \ \alpha_1 = 0$

$$\mathbf{G} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}, \quad \mathbf{R} = \sigma^2 \mathbf{I}_3$$

- Results:

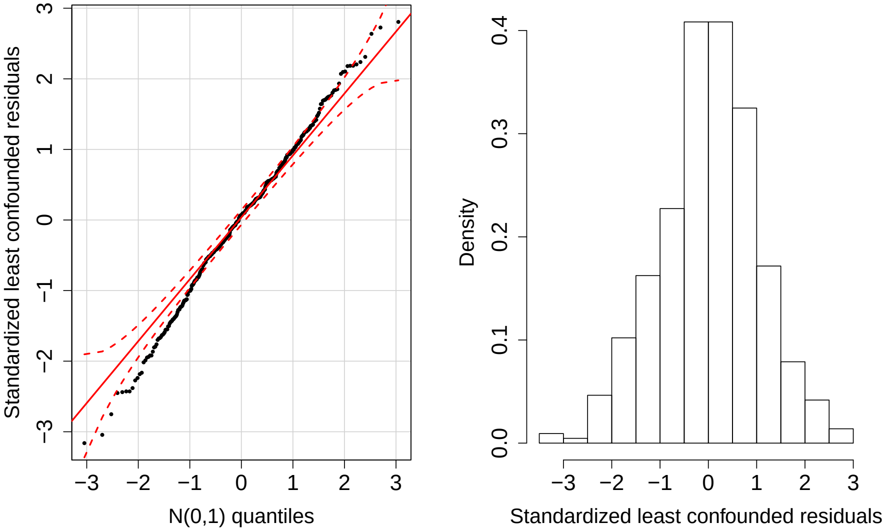
$$\widehat{\mu}_1 = 2085 \pm 62$$
$$\widehat{\mu}_2 = 2246 \pm 52$$
$$\widehat{\alpha}_2 = -290 \pm 84$$

Kcal intake example (B)



Kcal intake example (C)

Standardized least confounded conditional residuals



Kcal intake example (D)

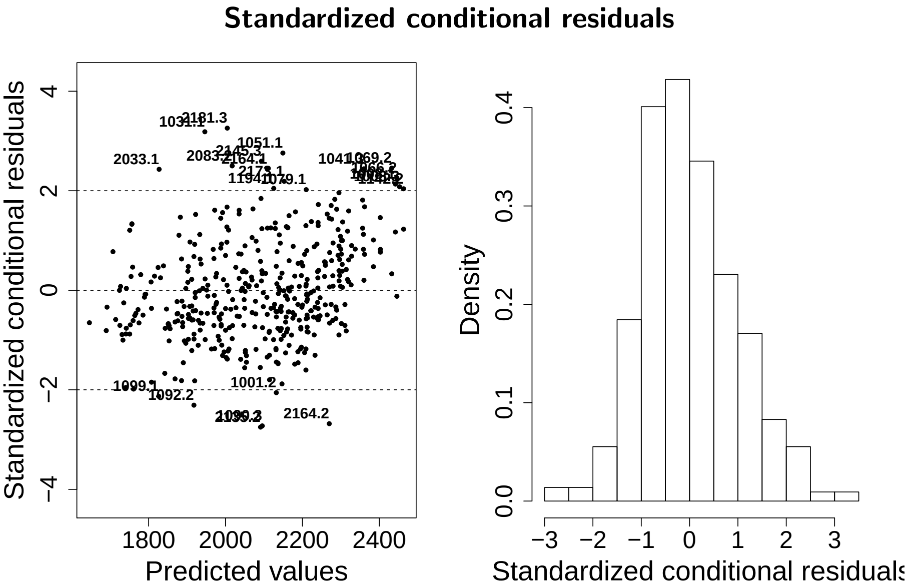
- Residual analysis suggests inadequacy of correlation structure for conditional errors
- Alternative model:

$$\mathbf{G} = \tau^2 \mathbf{I}_3, \quad \mathbf{R} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 & \rho\sigma_2\sigma_3 \\ \rho\sigma_1\sigma_3 & \rho\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}$$

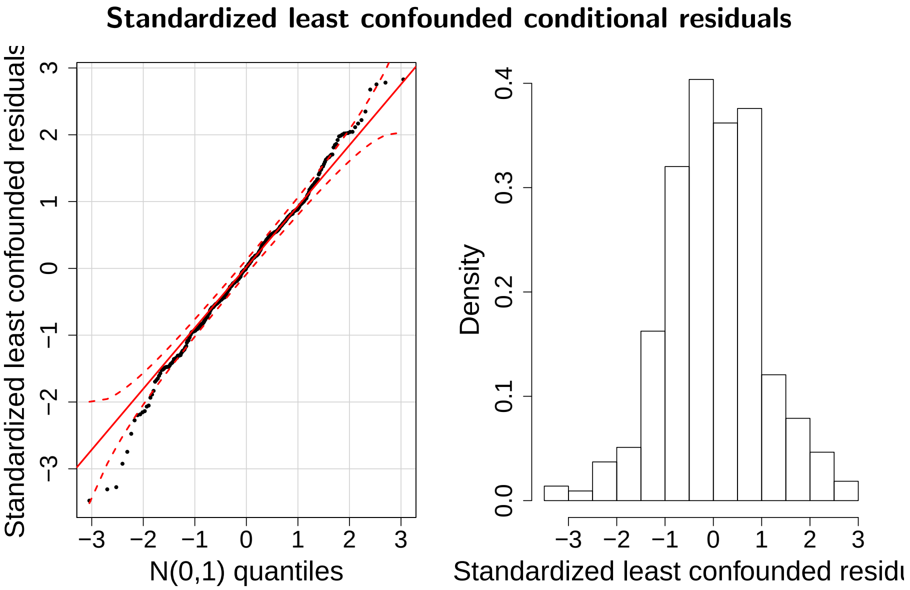
- Results:

$$\widehat{\mu}_1 = 2083 \pm 62$$
$$\widehat{\mu}_2 = 2244 \pm 52$$
$$\widehat{\alpha}_2 = -287 \pm 84$$
$$\widehat{\rho} = 0.36$$

Kcal intake example (E)



Kcal intake example (F)



Preterm neonates example

- Model:

$$y_{ijk} = \alpha_i + \beta_1(t_{1jk} - 26) + \gamma_2(t_{2jk} - 26)^2 + a_{ij} + b_{ij}(t_{ijk} - 26) + e_{ijk}$$

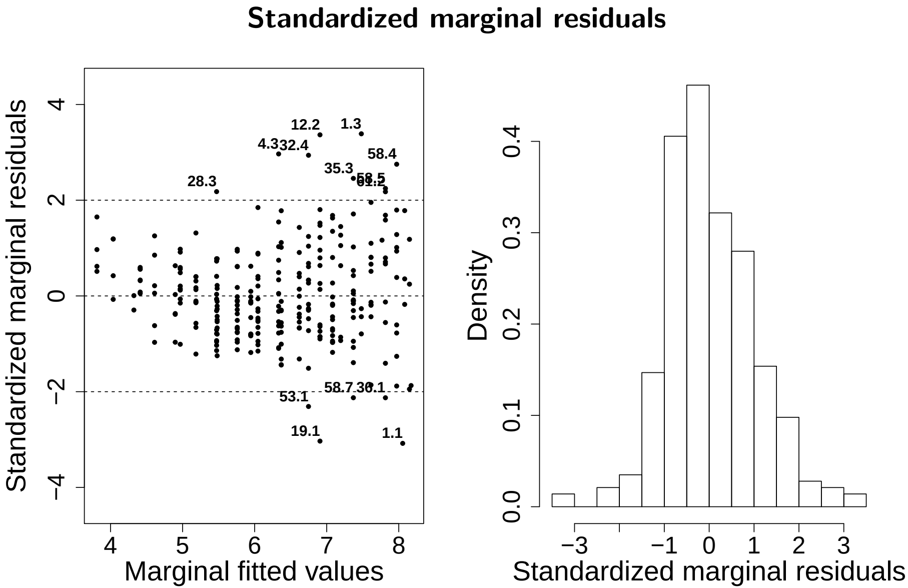
  - i indexes group (1=AGA and 2=SGA)
  - j indexes neonates ($j = 1, \ldots, n_i$)
  - k indexes week ($k = 1, \ldots, m_{ij}$)

- $\mathbf{b}_{ij} = (a_{ij}, b_{ij})^\top \sim N_2(\mathbf{0}, \mathbf{G}_i)$ independent

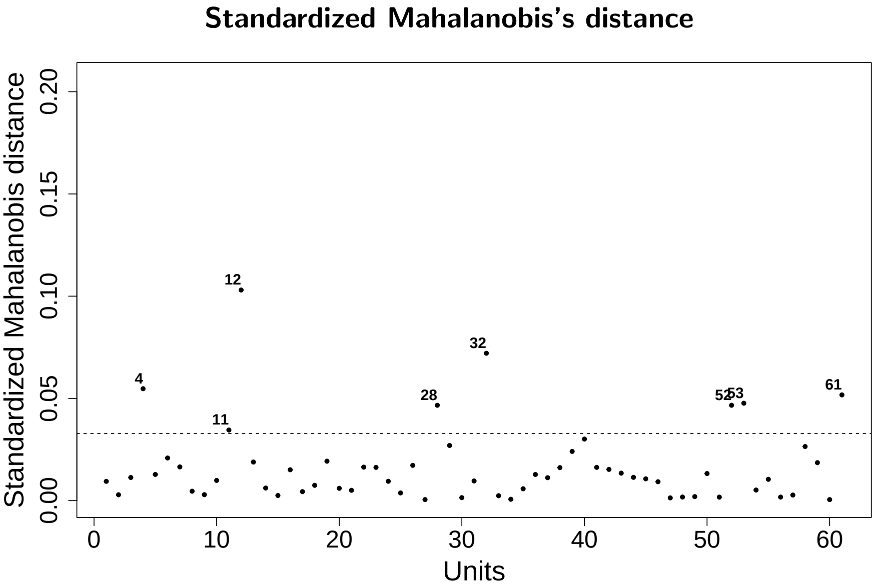
$$\mathbf{G}_i = \begin{bmatrix} \sigma^2_{a_i} & \sigma_{ab_i} \\ \sigma_{ab_i} & \sigma^2_{b_i} \end{bmatrix}$$

- $\mathbf{e}_{ij} = (e_{ij1}, \ldots, e_{ijm_{ij}})^\top \sim N_{m_{ij}}[\mathbf{0}, \sigma^2\mathbf{I}_{m_{ij}}]$ independent
- $\mathbf{b}_{ij}$ and $\mathbf{e}_{ij}$ independent

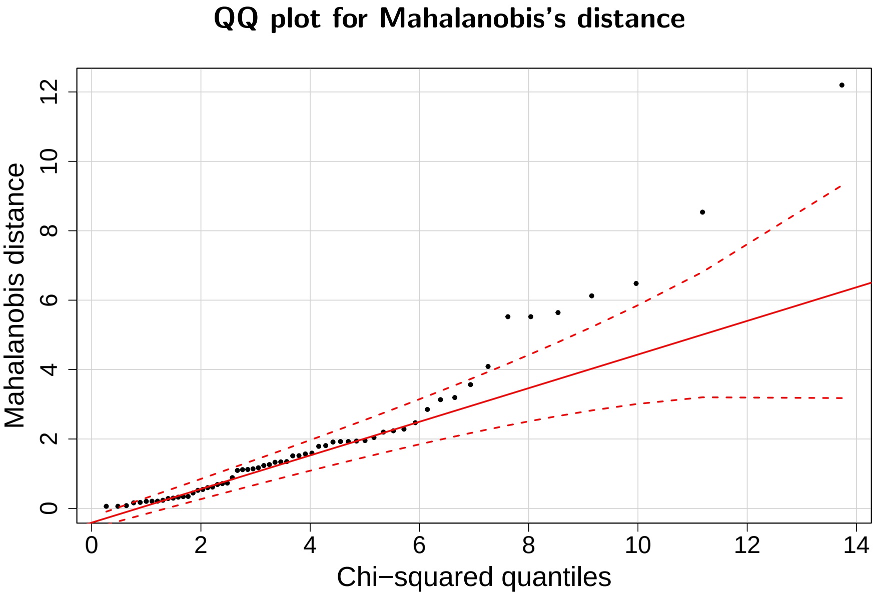
Preterm neonates example (A)



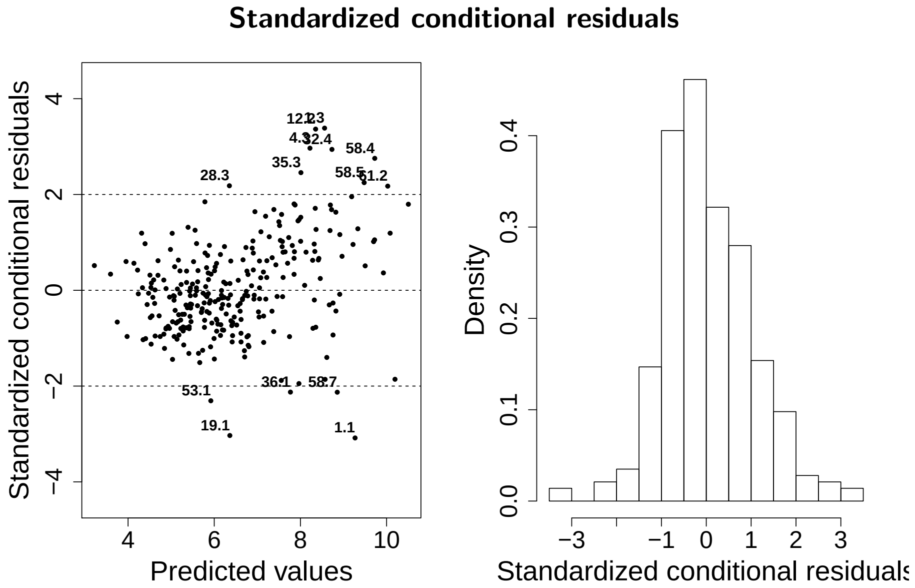
Preterm neonates example (B)



Preterm neonates example (C)

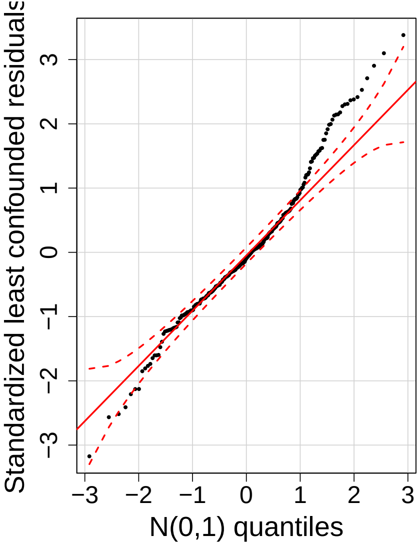


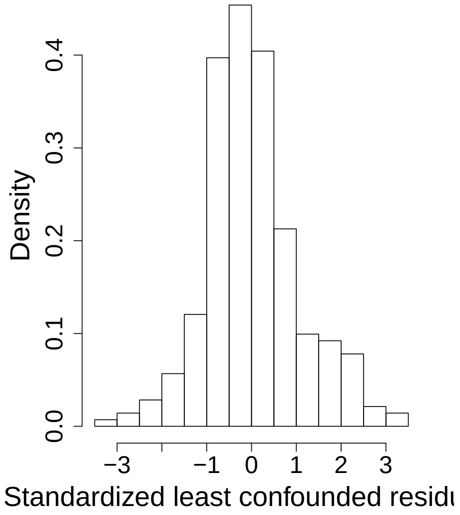
Preterm neonates example (D)



Preterm neonates example (E)

QQ plot and histogram for conditional least confounded residuals





Preterm neonates example (F)

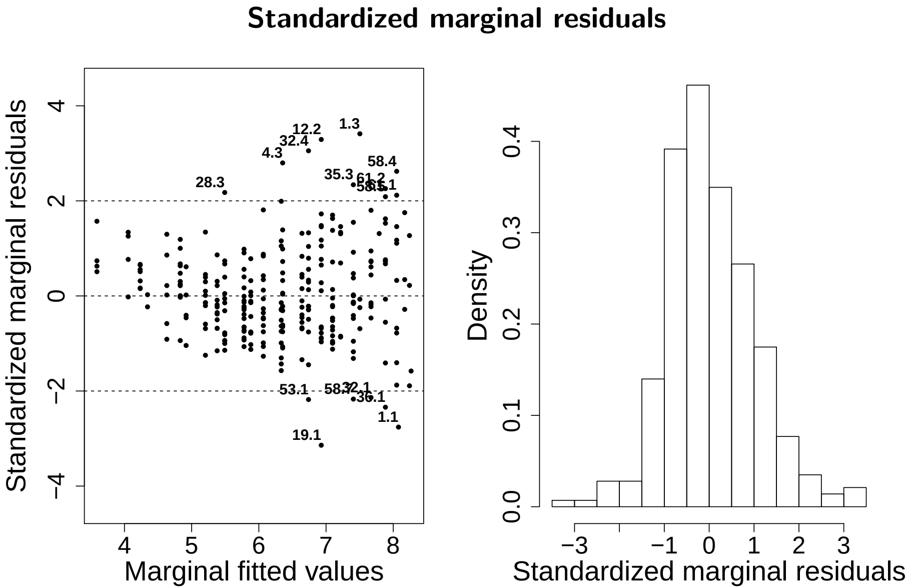
- Recall: $\mathbb{V}(\mathbf{y}_{ij}) = \mathbf{Z}_{ij}\mathbf{G}_i\mathbf{Z}_{ij}^\top$
- Alternative model: examine balance between random effects and errors

$$y_{ijk} = \alpha_i + \beta_1(t_{1jk} - 26) + \gamma_2(t_{2jk} - 26)^2 + a_{ij} + c_{ij}(t_{ijk} - 26)^2 + e_{ijk}$$

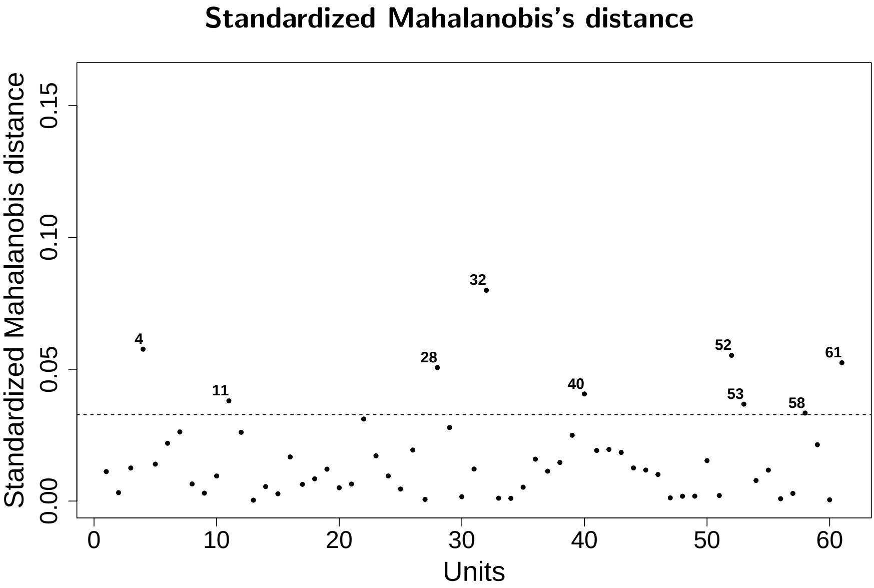
- $\mathbf{b}_{ij} = (a_{ij}, c_{ij})^\top \sim N_2(\mathbf{0}, \mathbf{G}_i)$ independent

$$\mathbf{G}_i = \begin{bmatrix} \sigma_{a_i}^2 & \sigma_{ac_i} \\ \sigma_{ac_i} & \sigma_{c_i}^2 \end{bmatrix}$$

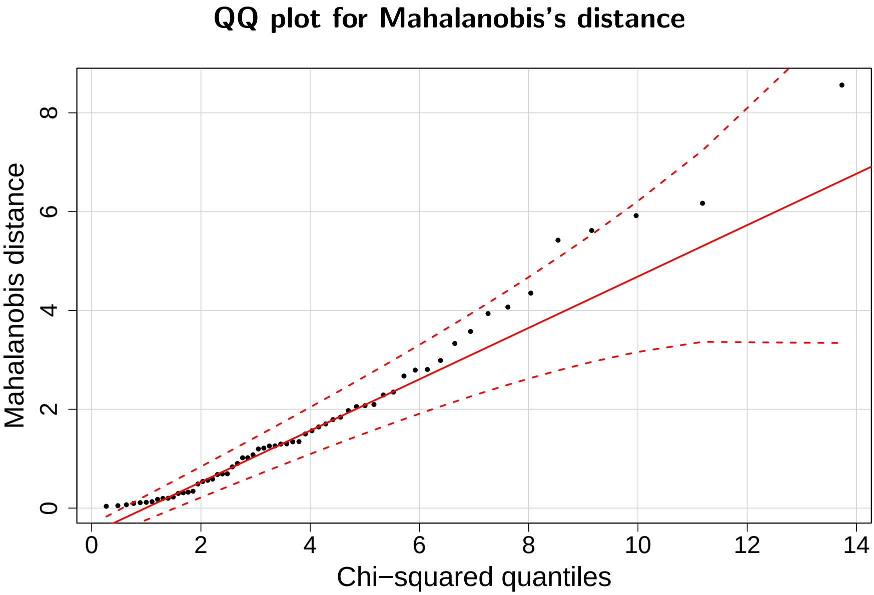
Preterm neonates example (A)



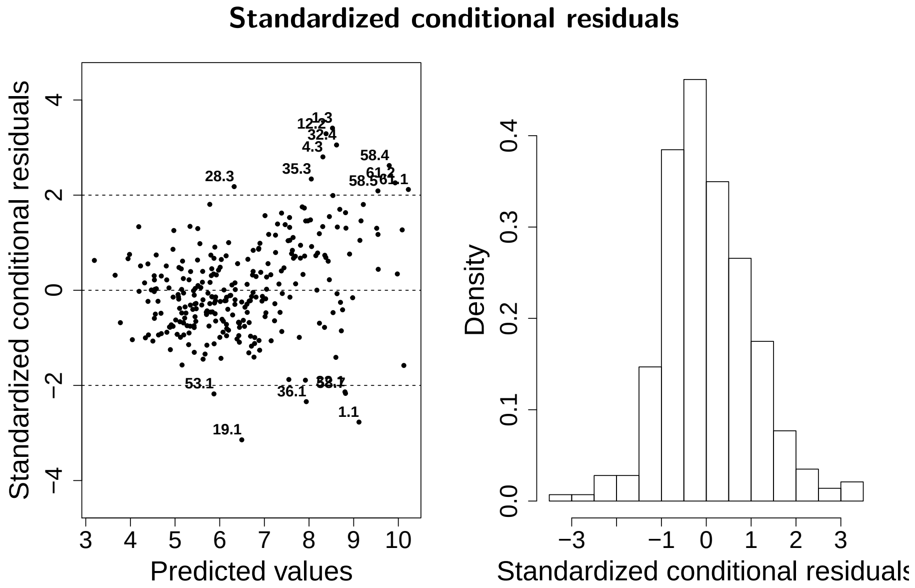
Preterm neonates example (B)



Preterm neonates example (C)

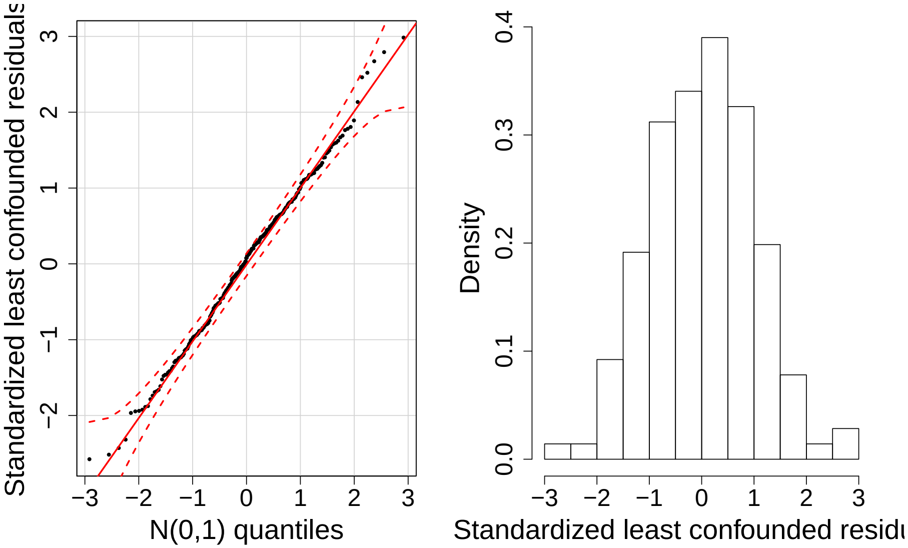


Preterm neonates example (D)



Preterm neonates example (E)

QQ plot and histogram for conditional least confounded residuals



Preterm neonates example (F)

- To accommodate possible "outliers": t (df=4) distribution for random effects and errors
- 15 units eliminated because of limitations in "heavy" function
- No available residual diagnostic tools

	Gaussian		t (df=4)		Change in	
Parameter	Estimate	SE	Estimate	SE	Estimate	SE
α_1	8.08	0.29	8.48	0.29	5%	0%
α_2	8.27	0.22	8.15	0.24	0%	0%
β_1	-0.28	0.03	-0.28	0.03	-1.5%	9.1%
γ_2	-0.02	0.002	-0.2	0.002	0%	0%