Heavy-Tailed Longitudinal Linear Mixed Models for Multiple Censored Responses Data

Larissa Avila Matos

Co-Authors: Victor H. Lachos, Luis M. Castro and Ming-Hui Chen

October/2016

Summary

Summary

Motivation

- In AIDS studies it is quite common to observe viral load measurements that are collected irregularly over time. Moreover, these measurements can be subjected to some upper and/or lower detection limits depending on the quantification assays.
- A complication arises when these continuous repeated measures have a heavy-tailed behavior.
- For such data structures, we propose a robust censored linear mixed model for multiple responses based on the class of multivariate scale mixtures of normal distributions.

Motivating data

A5055 clinical trial

- **A5055 study**

  $\rightarrow$ 44 infected patients with the human immunodeficiency virus type 1 (HIV-1) treated with one of the two potent antiretroviral (ARV) therapies;

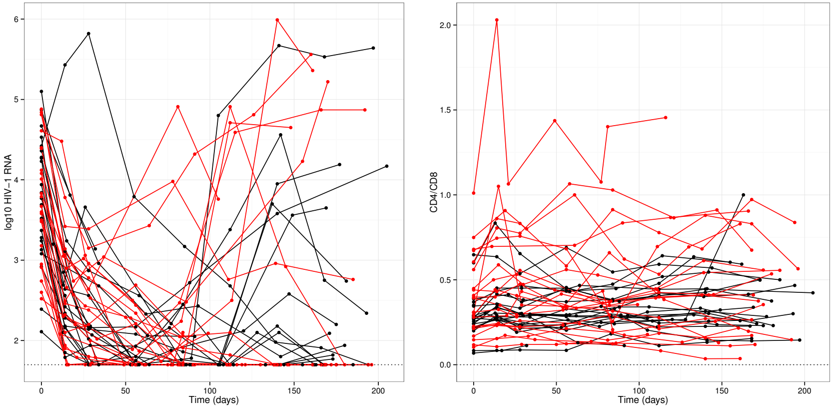
  $\rightarrow$ 2 outcomes: $\log_{10}$(RNA) and CD4/CD8, where CD4 and CD8 are two immunologic markers frequently used for monitoring disease progression in AIDS studies;

  $\rightarrow$ 33% (106 out of 316) of measurements lying below the limits of assay quantification (left-censored).

Motivating data

A5055 clinical trial

A5055



Recent works

Longitudinal Models

Censored longitudinal models with heavy-tailed distribution

- Lachos *et al.* (2011) [*Biometrics*]
- **Garay et al. (2014)** [*Statistical Methods in Medical Research*]
- **Matos et al. (2013b)** [*Statistica Sinica*]
- **Wang et al. (2015)** [*Statistical Methods in Medical Research*]

Summary

Scale mixture of normal distributions (SMN)

Andrews & Mallows (1974); Lange & Sinsheimer (1993)

Stochastic representation

$$\mathbf{y} = \boldsymbol{\mu} + \kappa(\mathrm{U})^{1/2}\mathbf{Z}, \qquad (1)$$

where,

- $\mu \in \mathbb{R}$ is a location vector;
- $Z \sim N(0, \boldsymbol{\Sigma})$;
- U is a positive random variable with cumulative distribution function (cdf) $\mathrm{H}(u|\boldsymbol{\nu})$ and probability density function (pdf) $h(u|\boldsymbol{\nu})$ independent of Z;
- $\kappa(\mathrm{U})$ is the weight function;
- Notation: $\mathbf{y} \sim \mathrm{SMN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathrm{H})$.

$$\begin{aligned} \mathbf{y}|\mathrm{U} = u &\sim N(\boldsymbol{\mu}, \kappa(u)\boldsymbol{\Sigma}), \\ \mathrm{U} = u &\sim h(u|\boldsymbol{\nu}). \end{aligned} \qquad (2)$$

Scale mixture of normal distributions (SMN)

Special cases: $\mathbf{y} \in \mathbb{R}^p$ and $\kappa(u) = 1/u$;

- **The multivariate normal distribution**
  - Distribution function: $\mathrm{N}(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \phi_p(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- **The multivariate Student's-t distribution**
  - Distribution function:

$$\mathrm{T}(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \frac{\Gamma\left(\frac{p+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\pi^{p/2}} \nu^{-p/2} |\boldsymbol{\Sigma}|^{-1/2} \left(1 + \frac{d}{\nu}\right)^{-(p+\nu)/2}.$$

- **The multivariate slash distribution**
  - Distribution function:

$$\mathrm{SL}(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \nu \int_0^1 u^{\nu-1} \phi_p(\mathbf{y}; \boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma}) du, \quad u \in (0, 1), \quad \nu > 0.$$

- **The multivariate contaminated normal distribution**
  - Distribution function:

$$\mathrm{CN}(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\nu}) = \nu \phi_p(\mathbf{y}; \boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Sigma}) + (1 - \nu)\phi_p(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

SAEM Algorithm

SAEM Algorithm - Delyon *et al.* (1999)

Let $\boldsymbol{\theta}$ be the parameter vector and $\boldsymbol{y}_c = (\mathbf{y}^\top, \boldsymbol{q}^\top)$ be the vector of complete data, i.e., the observed data $\mathbf{y}^\top$ and the missing/censored data (or the latent variables, depending on the situation) $\boldsymbol{q}^\top$. The SAEM algorithm consists in:

- **E-Step:**

  1. Simulation-step: Draw $\mathbf{q}^{(k,l)}$ $(l = 1, \ldots, m)$ from the conditional distribution $f(\mathbf{q}|\mathbf{y}, \widehat{\boldsymbol{\theta}}^{(k-1)})$;

  2. Stochastic-approximation-step: Update $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)})$ according to

$$Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) = Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k-1)}) + \delta_k \left[\frac{1}{m} \sum_{l=1}^{m} \ell_c(\boldsymbol{\theta}|\mathbf{q}^{(k,l)}, \mathbf{y}) - Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k-1)}) \right],$$

  where $\ell_c(\boldsymbol{\theta} \mid \mathbf{y}_c) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta} \mid \mathbf{y}_c)$ is the complete log-likelihood function and δ_k is a smoothness parameter, *i.e.*, a decreasing sequence of positive numbers such that $\sum_{k=1}^{\infty} \delta_k = \infty$ and $\sum_{k=1}^{\infty} \delta_k^2 < \infty$.

SAEM Algorithm

SAEM Algorithm - Delyon *et al.* (1999)

- **M-Step:** Update $\boldsymbol{\theta}^{(k)}$ according to

$$\widehat{\boldsymbol{\theta}}^{(k+1)} = \underset{\theta}{\text{argmax}} Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}).$$

- As proposed by Galarza *et al.* (2015), we will consider the following smoothing parameter

$$\delta_k = \begin{cases} 1, & \text{if} \quad 1 \leq k \leq cW; \\ \frac{1}{k-cW}, & \text{if} \quad cW + 1 \leq k \leq W, \end{cases} \tag{3}$$

where,

  - W is the maximum number of iterations; e
  - c is a cut point $(0 \leq c \leq 1)$ which determines the percentage of the initial iterations.

Summary

SMN-MLMEC model

Let $\mathbf{Y}_i = [\mathbf{y}_{i1} : \ldots : \mathbf{y}_{ir}]$, then

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \ldots, n; \tag{4}$$

where:

- $\mathbf{y}_i = \text{vec}(\mathbf{Y}_i) = (\mathbf{y}_{i1}^\top, \ldots, \mathbf{y}_{ir}^\top)^\top$, where $\mathbf{y}_{ij} = (y_{ij1}, \ldots, y_{ijn_i})^\top$ is a $n_i \times 1$ vector of the jth outcome for the ith subject;
- $\mathbf{X}_i = \text{Bdiag}\{\mathbf{X}_{i1}, \ldots, \mathbf{X}_{ir}\}$, where $\mathbf{X}_{ij}$ is an $n_i \times p_j$ design matrix for fixed effects corresponding to the jth outcome of the ith subject;
- $\mathbf{Z}_i = \text{Bdiag}\{\mathbf{Z}_{i1}, \ldots, \mathbf{Z}_{ir}\}$, onde $\mathbf{Z}_{ij}$ is an $n_i \times q_j$ design matrix for random effects corresponding to the jth outcome of the ith subject, which is generally a subset of $\mathbf{X}_{ij}$;
- $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \ldots, \boldsymbol{\beta}_r^\top)^\top$ is the $p \times 1$ vector of fixed effects associated with the design matrix $\mathbf{X}_i$, $p = \sum_{j=1}^r p_j$;
- $\mathbf{b}_i = (\mathbf{b}_{i1}^\top, \ldots, \mathbf{b}_{ir}^\top)^\top$ is the $q \times 1$ vector of random effects associated with the design matrix $\mathbf{Z}_i$, $q = \sum_{j=1}^r q_j$;
- $\boldsymbol{\epsilon}_i = \text{vec}(\mathrm{E}_i) = (\boldsymbol{\epsilon}_{i1}^\top, \ldots, \boldsymbol{\epsilon}_{ir}^\top)^\top$ is the vector of random errors of dimension $(s_i \times 1)$ $(s_i = n_i \times r)$, where $\mathrm{E}_i = [\boldsymbol{\epsilon}_{i1} : \ldots : \boldsymbol{\epsilon}_{ir}]$ and $\boldsymbol{\epsilon}_{ij}$ corresponds to the error for the jth outcome for the ith subject.

SMN-MLMEC model

- Considering the multivariate SMN distributions for the random terms, the model can be expressed as

$$\begin{aligned} \mathbf{y}_i \mid \mathbf{b}_i &\overset{\text{ind.}}{\sim} \mathrm{SMN}_{s_i}(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i, \mathbf{R}_i; \mathrm{H}_1), \\ \mathbf{b}_i &\overset{\text{ind.}}{\sim} \mathrm{SMN}_q(\mathbf{0}, \mathbf{D}; \mathrm{H}_2), \quad i = 1, \ldots, n. \end{aligned} \tag{5}$$

- Using the stochastic representation (1), the hierarchical representation of the model defined in(4) is given by

$$\begin{aligned} \mathbf{y}_i \mid \mathbf{b}_i, \kappa_i &\overset{\text{ind.}}{\sim} \mathrm{N}_{s_i}(\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i, \kappa_i^{-1}\mathbf{R}_i), \\ \mathbf{b}_i \mid \tau_i &\overset{\text{ind.}}{\sim} \mathrm{N}_q(\mathbf{0}, \tau_i^{-1}\mathbf{D}), \\ \kappa_i &\overset{\text{ind.}}{\sim} \mathrm{H}_1(\nu), \\ \tau_i &\overset{\text{ind.}}{\sim} \mathrm{H}_2(\eta); \end{aligned} \tag{6}$$

where $\mathbf{R}_i = \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega}_i$.

SMN-MLMEC model

Correlation structures - Muñoz *et al.* (1992)

Damped exponential correlation (DEC):

$$\mathbf{\Omega}_i = \mathbf{\Omega}_i(\phi, \mathbf{t}_i) = \left[\phi_1^{|t_{ij}-t_{ik}|^{\phi_2}}\right], \; i = 1, \ldots, n, \;\; j,k = 1, \ldots, n_i, \quad (7)$$

For the DEC structure, we have that:

(a) if $\phi_2 = 0$, then $\mathbf{\Omega}_i$ generates the compound symmetry correlation structure;

(b) when $0 < \phi_2 < 1$, then $\mathbf{\Omega}_i$ presents a decay rate between the compound symmetry structure and the first-order AR (AR (1)) model;

(c) if $\phi_2 = 1$, then $\mathbf{\Omega}_i$ generates an AR(1) structure;

(d) when $\phi_2 > 1$, $\mathbf{\Omega}_i$ presents a decay rate faster than the AR(1) structure; and

(e) if $\phi_2 \to \infty$, then $\mathbf{\Omega}_i$ represents the first-order moving average model, MA(1).

SMN-MLMEC model

- Recall that we are interested in the case where left-censored observations can occur. That is, we assume that the observations are of the form

$$\begin{aligned} y_{ijk} &\leq V_{ijk} \quad \text{se} \quad C_{ijk} = 1, \\ y_{ijk} &= V_{ijk} \quad \text{se} \quad C_{ijk} = 0, \end{aligned}$$

  with $i = 1, \dots, n$, $j = 1, \dots, n_i$ and $k = 1, \dots, r$;
- The observed data for the i-th subject is represented by $(\mathbf{V}_i, \mathbf{C}_i)$, where $\mathbf{V}_i = [V_{i1} : \dots : V_{ir}]$ is an $n_i \times r$ matrix and $\mathbf{C}_i = [C_{i1} : \dots : C_{ir}]$ is an $n_i \times r$ matrix;
- The extensions to arbitrary censoring are immediate.

SMN-MLMEC model

Maximum likelihood estimation - SAEM

- The complete data log-likelihood function:
  - $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\phi}, \nu, \eta)$;
  - Augmenting data: $\mathbf{y}_c = (\mathbf{V}^\top, \mathbf{C}^\top, \mathbf{y}^\top, \boldsymbol{b}^\top, \boldsymbol{\kappa}^\top, \boldsymbol{\tau}^\top)^\top$;
  - $[\ \mathbf{V}\ \ \mathbf{C}\ \ \mathbf{y}\ ] \Rightarrow [\ \mathbf{y}\ ]$.

$$
\begin{aligned}
\ell_c(\boldsymbol{\theta}|\mathbf{y}_c) &= \sum_{i=1}^{n} [\log f(\mathbf{y}_i|\mathbf{b}_i, \kappa_i) + \log f(\mathbf{b}_i|\tau_i) + \log h_1(\kappa_i|\nu) + \log h_2(\tau_i|\eta)] \\
&= -\frac{1}{2}\sum_{i=1}^{n} \log|\mathbf{R}_i| - \frac{1}{2}\sum_{i=1}^{n} \kappa_i(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i)^\top \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i) \\
&- \frac{1}{2}\sum_{i=1}^{n} \log|\mathbf{D}| - \frac{1}{2}\sum_{i=1}^{n} \tau_i \mathbf{b}_i^\top \mathbf{D}^{-1}\mathbf{b}_i + \sum_{i=1}^{n} \log h_1(\kappa_i|\nu) + \sum_{i=1}^{n} \log h_2(\tau_i|\eta) + K,
\end{aligned}
$$

with K being a constant that does not depend on the parameter vector $\boldsymbol{\theta}$.

SMN-MLMEC model

Maximum likelihood estimation - SAEM

- Q-function: For the i-th subject,

$$
\begin{aligned}
Q_i\left(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}\right) \quad = \quad & \widehat{\ell h_{1i}}^{(k)} + \widehat{\ell h_{2i}}^{(k)} - \frac{1}{2}\log|\widehat{\mathbf{D}}^{(k)}| - \frac{1}{2}\text{tr}\left(\widehat{\tau \mathbf{b}_i^2}^{(k)}\widehat{\mathbf{D}}_i^{-1(k)}\right) - \frac{1}{2}\sum_{i=1}^{n}\log|\widehat{\mathbf{R}}_i^{(k)}| \\
- \quad & \frac{1}{2}\Bigg[\text{tr}\left(\widehat{\kappa \mathbf{y}_i^2}^{(k)}\widehat{\mathbf{R}}_i^{-1(k)}\right) - 2\widehat{\boldsymbol{\beta}}^{(k)\top}\mathbf{X}_i^\top\widehat{\mathbf{R}}_i^{-1(k)}\widehat{\kappa \mathbf{y}_i}^{(k)} + 2\widehat{\boldsymbol{\beta}}^{(k)\top}\mathbf{X}_i^\top\widehat{\mathbf{R}}_i^{-1(k)}\mathbf{Z}_i\widehat{\kappa \mathbf{b}_i}^{(k)} \\
- \quad & 2\text{tr}\left(\mathbf{Z}_i^\top\widehat{\mathbf{R}}_i^{-1(k)}\widehat{\kappa \mathbf{y}\mathbf{b}_i}^{(k)}\right) + \text{tr}\left(\mathbf{Z}_i^\top\widehat{\mathbf{R}}_i^{-1(k)}\mathbf{Z}_i\widehat{\kappa \mathbf{b}_i^2}^{(k)}\right) + \widehat{\kappa_i}^{(k)}\widehat{\boldsymbol{\beta}}^{(k)\top}\mathbf{X}_i^\top\widehat{\mathbf{R}}_i^{-1(k)}\mathbf{X}_i\widehat{\boldsymbol{\beta}}^{(k)}\Bigg],
\end{aligned}
$$

with

$$
\begin{aligned}
&\widehat{\ell h_{1i}}^{(k)} = E\left[\log h_1(\kappa_i|\nu)|\mathbf{V}_i, \mathbf{C}_i, \widehat{\boldsymbol{\theta}}^{(k)}\right], && \widehat{\ell h_{2i}}^{(k)} = E\left[\log h_2(\tau_i|\eta)|\mathbf{V}_i, \mathbf{C}_i, \widehat{\boldsymbol{\theta}}^{(k)}\right] \\
&\widehat{\kappa \mathbf{y}_i^2}^{(k)} = E\left[\kappa_i\mathbf{y}_i\mathbf{y}_i^\top|\mathbf{V}_i, \mathbf{C}_i, \widehat{\boldsymbol{\theta}}^{(k)}\right], && \widehat{\kappa \mathbf{y}_i}^{(k)} = E\left[\kappa_i\mathbf{y}_i|\mathbf{V}_i, \mathbf{C}_i, \widehat{\boldsymbol{\theta}}^{(k)}\right], \\
&\widehat{\kappa \mathbf{b}_i^2}^{(k)} = E\left[\kappa_i\mathbf{b}_i\mathbf{b}_i^\top|\mathbf{V}_i, \mathbf{C}_i, \widehat{\boldsymbol{\theta}}^{(k)}\right], && \widehat{\kappa \mathbf{b}_i}^{(k)} = E\left[\kappa_i\mathbf{b}_i|\mathbf{V}_i, \mathbf{C}_i, \widehat{\boldsymbol{\theta}}^{(k)}\right], \\
&\widehat{\tau \mathbf{b}_i^2}^{(k)} = E\left[\tau_i\mathbf{b}_i\mathbf{b}_i^\top|\mathbf{V}_i, \mathbf{C}_i, \widehat{\boldsymbol{\theta}}^{(k)}\right], && \widehat{\kappa \mathbf{y}\mathbf{b}_i}^{(k)} = E\left[\kappa_i\mathbf{y}_i\mathbf{b}_i^\top|\mathbf{V}_i, \mathbf{C}_i, \widehat{\boldsymbol{\theta}}^{(k)}\right], \\
&\widehat{\kappa_i}^{(k)} = E\left[\kappa_i|\mathbf{V}_i, \mathbf{C}_i, \widehat{\boldsymbol{\theta}}^{(k)}\right].
\end{aligned}
$$

SMN-MLMEC model

SAEM - E-step

- **Simulation-step:** Gibbs Sampler

Step 1. Sample $\mathbf{y}_i^c$ de $f(\mathbf{y}_i^c|\mathbf{V}_i^c, \mathbf{y}_i^o, \mathbf{b}_i^{(k,l-1)}, \kappa_i^{(k,l-1)}, \tau_i^{(k,l-1)}, \widehat{\boldsymbol{\theta}}^{(k)})$, where

$$\mathbf{y}_i^c|\mathbf{V}_i^c, \mathbf{y}_i^o, \mathbf{b}_i, \kappa_i, \tau_i, \boldsymbol{\theta} \sim \mathrm{TN}_{s_i^c}(\boldsymbol{\mu}_i, \kappa_i^{-1}\mathbf{S}_i; \mathbb{A}_i),$$

with $\mathbb{A}_i = \{\mathbf{y}_i^c = (y_{i1}^c, \ldots, y_{is_i^c}^c)^\top | y_{i1}^c \leq V_{i1}^c, \ldots, y_{is_i^c}^c \leq V_{is_i^c}^c\}$,

$$\begin{aligned} \boldsymbol{\mu}_i &= (\mathbf{X}_i^c\boldsymbol{\beta} + \mathbf{Z}_i^c\mathbf{b}_i) + \mathbf{R}_i^{co}(\mathbf{R}_i^{oo})^{-1}(\mathbf{y}_i^o - \mathbf{X}_i^o\boldsymbol{\beta} - \mathbf{Z}_i^o\mathbf{b}_i) \quad \text{e} \\ \mathbf{S}_i &= \mathbf{R}_i^{cc} - \mathbf{R}_i^{co}(\mathbf{R}_i^{oo})^{-1}\mathbf{R}_i^{oc}. \end{aligned}$$

Then $\mathbf{y}_i^{(k,l)} = (y_{i1}, \ldots, y_{is_i^o}, y_{is_i^o+1}^{c(k,l)}, \ldots, y_{is_i}^{c(k,l)})$ is the generated sample.

Step 2. Sample $\mathbf{b}_i^{(k,l)}$ from $f(\mathbf{b}_i|\mathbf{y}_i^{(k,l)}, \kappa_i^{(k,l-1)}, \tau_i^{(k,l-1)}, \widehat{\boldsymbol{\theta}}^{(k)})$, where

$$\mathbf{b}_i|\mathbf{y}_i, \kappa_i, \tau_i \sim \mathrm{N}_q(\boldsymbol{\Psi}_i\mathbf{Z}_i^\top\mathbf{R}_i^{-1}\kappa_i(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}), \boldsymbol{\Psi}_i),$$

with $\boldsymbol{\Psi} = (\kappa_i\mathbf{Z}_i^\top\mathbf{R}_i^{-1}\mathbf{Z}_i + \tau_i\mathbf{D}^{-1})^{-1}$ (Arellano-Valle *et al.*, 2005, Lemma 2).

SMN-MLMEC model

SAEM - E-step

Step 3. Sample $\kappa_i^{(k,l)}$ from $f(\kappa_i|\mathbf{y}_i^{(k,l)}, \mathbf{b}_i^{(k,l)}, \tau_i^{(k,l-1)}, \widehat{\boldsymbol{\theta}}^{(k)})$.

Step 4. Sample $\tau_i^{(k,l)}$ from $f(\tau_i|\mathbf{y}_i^{(k,l)}, \mathbf{b}_i^{(k,l)}, \kappa_i^{(k,l)}, \widehat{\boldsymbol{\theta}}^{(k)})$.

Observation: Note that since $\mathbf{y}_i \mid \mathbf{b}_i$ is independent of τ_i; $\mathbf{b}_i$ independent of κ_i; and κ_i and τ_i are mutually independent, then we have

$$f(\kappa_i|\mathbf{y}_i, \mathbf{b}_i, \tau_i) \propto f(\mathbf{y}_i|\mathbf{b}_i, \kappa_i) f(\kappa_i)$$

and

$$f(\tau_i|\mathbf{y}_i, \mathbf{b}_i, \kappa_i) \propto f(\mathbf{b}_i|\tau_i) f(\tau_i).$$

SMN-MLMEC model

SAEM - E-step

Distribution of $\boldsymbol{\epsilon}_i$	Distribution of κ_i	Distribution of $\kappa_i \mid \mathbf{y}_i, \mathbf{b}_i, \tau_i$
$\mathrm{T}_{si}(\mathbf{0}, \mathbf{R}_i, \nu)$	$\text{Gamma}(\nu/2, \nu/2)$	$\text{Gamma}\left((\nu + s_i)/2, (D^2_{e_i} + \nu)/2\right)$
$\mathrm{SL}_{si}(\mathbf{0}, \mathbf{R}_i, \nu)$	$\text{Beta}(\nu, 1)$	$\text{TGamma}\left(\nu + s_i/2, D^2_{e_i}/2, 1\right)$
$\mathrm{CN}_{si}(\mathbf{0}, \mathbf{R}_i, \nu_1, \nu_2)$	$\nu_1 \mathbb{I}_{\{\nu_2\}}(\kappa_i) + (1 - \nu_1)\mathbb{I}_{\{1\}}(\kappa_i)$	$P(\kappa_i = \nu_2) = 1 - P(\kappa_i = 1) = p_1/p_1 + p_2$ $p_1 = \nu_1 \nu_2^{s_i/2} exp\{-\frac{1}{2} D^2_{e_i} \nu_2\}$ $p_2 = (1 - \nu_1) exp\{-\frac{1}{2} D^2_{e_i}\}$

$D^2_{e_i} = (\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i)^\top \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i)$

Distribution of $\mathbf{b}_i$	Distribution of τ_i	Distribution of $\tau_i \mid \mathbf{y}_i, \mathbf{b}_i, \kappa_i$
$\mathrm{T}_q(\mathbf{0}, \mathbf{D}, \eta)$	$\text{Gamma}(\eta/2, \eta/2)$	$\text{Gamma}\left((\eta + q)/2, (D^2_{\mathbf{b}_i} + \eta)/2\right)$
$\mathrm{SL}_q(\mathbf{0}, \mathbf{D}, \eta)$	$\text{Beta}(\eta, 1)$	$\text{TGamma}\left(\eta + q/2, D^2_{\mathbf{b}_i}/2, 1\right)$
$\mathrm{CN}_q(\mathbf{0}, \mathbf{D}, \eta_1, \eta_2)$	$\eta_1 \mathbb{I}_{\{\eta_2\}}(\tau_i) + (1 - \eta_1)\mathbb{I}_{\{1\}}(\tau_i)$	$P(\tau_i = \eta_2) = 1 - P(\tau_i = 1) = q_1/q_1 + q_2$ $q_1 = \eta_1 \eta_2^{q/2} exp\{-\frac{1}{2} D^2_{\mathbf{b}_i} \eta_2\}$ $q_2 = (1 - \eta_1) exp\{-\frac{1}{2} D^2_{\mathbf{b}_i}\}$

$D^2_{\mathbf{b}_i} = \mathbf{b}_i^\top \mathbf{D}^{-1} \mathbf{b}_i$

SMN-MLMEC model

SAEM - E-step

- **Stochastic-approximation-step:** $(\mathbf{y}_i^{(k,l)}, \mathbf{b}_i^{(k,l)}, \kappa_i^{(k,l)}, \tau_i^{(k,l)})$, $l = 1, \ldots, m$:

$$
\begin{aligned}
\widehat{\kappa\mathbf{y}_i^2}^{(k)} &= \widehat{\kappa\mathbf{y}_i^2}^{(k-1)} + \delta_k \left(\frac{1}{m} \sum_{l=1}^{m} \kappa_i^{(k,l)} \mathbf{y}_i^{(k,l)} \mathbf{y}_i^{(k,l)\top} - \widehat{\kappa\mathbf{y}_i^2}^{(k-1)} \right), \\
\widehat{\kappa\mathbf{y}_i}^{(k)} &= \widehat{\kappa\mathbf{y}_i}^{(k-1)} + \delta_k \left(\frac{1}{m} \sum_{l=1}^{m} \kappa_i^{(k,l)} \mathbf{y}_i^{(k,l)} - \widehat{\kappa\mathbf{y}_i}^{(k-1)} \right), \\
\widehat{\kappa\mathbf{b}_i^2}^{(k)} &= \widehat{\kappa\mathbf{b}_i^2}^{(k-1)} + \delta_k \left(\frac{1}{m} \sum_{l=1}^{m} \kappa_i^{(k,l)} \mathbf{b}_i^{(k,l)} \mathbf{b}_i^{(k,l)\top} - \widehat{\kappa\mathbf{b}_i^2}^{(k-1)} \right), \\
\widehat{\kappa\mathbf{b}_i}^{(k)} &= \widehat{\kappa\mathbf{b}_i}^{(k-1)} + \delta_k \left(\frac{1}{m} \sum_{l=1}^{m} \kappa_i^{(k,l)} \mathbf{b}_i^{(k,l)} - \widehat{\kappa\mathbf{b}_i}^{(k-1)} \right), \\
\widehat{\kappa\mathbf{yb}_i}^{(k)} &= \widehat{\kappa\mathbf{yb}_i}^{(k-1)} + \delta_k \left(\frac{1}{m} \sum_{l=1}^{m} \kappa_i^{(k,l)} \mathbf{y}_i^{(k,l)} \mathbf{b}_i^{(k,l)\top} - \widehat{\kappa\mathbf{yb}_i}^{(k-1)} \right), \\
\widehat{\tau\mathbf{b}_i^2}^{(k)} &= \widehat{\tau\mathbf{b}_i^2}^{(k-1)} + \delta_k \left(\frac{1}{m} \sum_{l=1}^{m} \tau_i^{(k,l)} \mathbf{b}_i^{(k,l)} \mathbf{b}_i^{(k,l)\top} - \widehat{\tau\mathbf{y}_i^2}^{(k-1)} \right), \\
\widehat{\kappa_i}^{(k)} &= \widehat{\kappa_i}^{(k-1)} + \delta_k \left(\frac{1}{m} \sum_{l=1}^{m} \kappa_i^{(k,l)} - \widehat{\kappa_i}^{(k-1)} \right), \\
\widehat{\ell h_{1i}}^{(k)} &= \widehat{\ell h_{1i}}^{(k-1)} + \delta_k \left(\frac{1}{m} \sum_{l=1}^{m} \log h_1(\kappa_i^{(k,l)} | \widehat{\nu}^{(k-1)}) - \widehat{\ell h_{1i}}^{(k-1)} \right), \\
\widehat{\ell h_{2i}}^{(k)} &= \widehat{\ell h_{2i}}^{(k-1)} + \delta_k \left(\frac{1}{m} \sum_{l=1}^{m} \log h_2(\tau_i^{(k,l)} | \widehat{\eta}^{(k-1)}) - \widehat{\ell h_{2i}}^{(k-1)} \right).
\end{aligned}
$$

SMN-MLMEC model

SAEM - CM-step

Update $\widehat{\boldsymbol{\theta}}^{(k)}$ by the maximization of $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)})$, which leads to the following expressions:

$$\widehat{\boldsymbol{\beta}}^{(k+1)} = \left(\sum_{i=1}^{n} \widehat{\kappa_i}^{(k)} \mathbf{X}_i^\top \widehat{\mathbf{R}}_i^{-1(k)} \mathbf{X}_i\right)^{-1} \sum_{i=1}^{n} \mathbf{X}_i^\top \widehat{\mathbf{R}}_i^{-1(k)} \left(\widehat{\kappa \mathbf{y}_i}^{(k)} - Z_i \widehat{\kappa \mathbf{b}_i}^{(k)}\right),$$

$$\widehat{\sigma^2_{jl}}^{(k+1)} = \begin{cases} \left(\sum_{i=1}^n n_i\right)^{-1} \sum_{i=1}^n \mathrm{tr}\left(\widehat{\boldsymbol{\Omega}}_i^{-1(k)} \widehat{\kappa \boldsymbol{\epsilon}}_{ijl}^{(k)}\right) & \text{for } j = l, \\ \left(2\sum_{i=1}^n n_i\right)^{-1} \sum_{i=1}^n \mathrm{tr}\left[\widehat{\boldsymbol{\Omega}}_i^{-1(k)} \left(\widehat{\kappa \boldsymbol{\epsilon}}_{ijl}^{(k)} + \widehat{\kappa \boldsymbol{\epsilon}}_{ilj}^{(k)}\right)\right] & \text{for } j \neq l, \end{cases}$$

$$\widehat{\boldsymbol{\phi}}^{(k+1)} = \underset{\boldsymbol{\phi} \in (0,1) \times \mathbb{R}^+}{\mathrm{argmax}} \left\{ -\frac{r}{2} \sum_{i=1}^{n} \log |\boldsymbol{\Omega}_i(\boldsymbol{\phi}, \mathbf{t}_i)| - \frac{1}{2} \sum_{i=1}^{n} \mathrm{tr}\left[\left(\widehat{\boldsymbol{\Sigma}}^{(k)} \otimes \boldsymbol{\Omega}_i(\boldsymbol{\phi}, \mathbf{t}_i)\right)^{-1} \widehat{\kappa \mathrm{E}_i}\right] \right\},$$

$$\widehat{\mathbf{D}}^{(k+1)} = \frac{1}{n} \sum_{i=1}^{n} \widehat{\tau \mathbf{b}_i^2}^{(k)},$$

$$\widehat{\nu}^{(k+1)} = \underset{\nu}{\mathrm{argmax}} \sum_{i=1}^{n} \widehat{\ell h_{1i}}^{(k)}(\nu),$$

$$\widehat{\eta}^{(k+1)} = \underset{\eta}{\mathrm{argmax}} \sum_{i=1}^{n} \widehat{\ell h_{2i}}^{(k)}(\eta).$$

SMN-MLMEC model

The likelihood function

$$L_o(\boldsymbol{\theta};\mathbf{y}^{obs}) = \prod_{i=1}^{n}\int\left[\int_0^{\infty} f(\mathbf{y}_i|\mathbf{b}_i,\kappa_i;\boldsymbol{\theta})h_1(\kappa_i)d\kappa_i\right] f(\mathbf{b}_i|\boldsymbol{\theta})d\mathbf{b}_i.$$

Partitioning $\mathbf{y}_i$, we have

$$\begin{aligned} L_o(\boldsymbol{\theta};\mathbf{y}^{obs}) &= \prod_{i=1}^{n}\int\left[\int_0^{\infty}\phi_{s_i^o}(\mathbf{y}_i^o;\mathbf{X}_i^c\boldsymbol{\beta}-\mathbf{Z}_i^c\mathbf{b}_i,\kappa_i^{-1}\mathbf{R}_i^{oo})\Phi_{s_i^c}(\mathbf{V}_i^c;\boldsymbol{\mu}_i,\kappa_i^{-1}\mathbf{S}_i)h_1(\kappa_i)d\kappa_i\right] \\ &\times \quad f(\mathbf{b}_i|\boldsymbol{\theta})d\mathbf{b}_i = \prod_{i=1}^{n}\int g(\mathbf{y}_i|\mathbf{b}_i,\kappa_i;\boldsymbol{\theta})f(\mathbf{b}_i|\boldsymbol{\theta})d\mathbf{b}_i \qquad (8)\end{aligned}$$

where $g(\mathbf{y}_i|\mathbf{b}_i,\kappa_i;\boldsymbol{\theta}) = \int_0^{\infty}\phi_{s_i^o}(\mathbf{y}_i^o;\mathbf{X}_i^c\boldsymbol{\beta}-\mathbf{Z}_i^c\mathbf{b}_i,\kappa_i^{-1}\mathbf{R}_i^{oo})\Phi_{s_i^c}(\mathbf{V}_i^c;\boldsymbol{\mu}_i,\kappa_i^{-1}\mathbf{S}_i)h_1(\kappa_i|\nu)d\kappa_i$. The integral involved in (8) can be compute using an importance sampling strategy. In fact, we have

$$L_o(\boldsymbol{\theta};\mathbf{y}^{obs}) = \prod_{i=1}^{n}\int g(\mathbf{y}_i|\mathbf{b}_i,\kappa_i;\boldsymbol{\theta})\frac{f(\mathbf{b}_i|\boldsymbol{\theta})}{f^{\star}(\mathbf{b}_i|\boldsymbol{\theta})}d\mathbf{b}_i,$$

where $f^{\star}$ is the importance distribution. Consequently, $L_o(\boldsymbol{\theta};\mathbf{y}_i^{obs})$ is estimated through the following approximation

$$L_o(\boldsymbol{\theta};\mathbf{y}^{obs}) = \prod_{i=1}^{n}\left[\frac{1}{M}\sum_{m=1}^{M} g(\mathbf{y}_i|\mathbf{b}_{im},\kappa_i;\boldsymbol{\theta})\frac{f(\mathbf{b}_{im}|\boldsymbol{\theta})}{f^{\star}(\mathbf{b}_{im}|\boldsymbol{\theta})}\right],$$

with $\mathbf{b}_{i1},\ldots,\mathbf{b}_{im}$ being drawn from $f^{\star}(\mathbf{b}_i|\boldsymbol{\theta})$.

SMN-MLMEC model

Model selection criteria

- **AIC and BIC**

$$\text{AIC} = 2\,m - 2\,\ell_{max} \;\text{ and }\; \text{BIC} = m\,log\,N - 2\,\ell_{max}.$$

- **AIC and BIC decomposition (Zhang et al., 2014)**

  Let $\mathbf{y}_{i1}^\star = (\mathbf{y}_{i1}^\top, \ldots, \mathbf{y}_{ir^\star}^\top)^\top$ and $\mathbf{y}_{i2}^\star = (\mathbf{y}_{ir^\star+1}^\top, \ldots, \mathbf{y}_{ir}^\top)^\top$, where $\mathbf{y}_i = (\mathbf{y}_{i1}^{\star\top}, \mathbf{y}_{i2}^{\star\top})^\top$ and $r^\star \in \{1, \ldots, r\}$, then the AIC and BIC have the following decompositions:

$$\text{AIC} = \text{AIC}_{\mathbf{y}_1^\star} + \text{AIC}_{\mathbf{y}_2^\star|\mathbf{y}_1^\star} \;\text{ and }\; \text{BIC} = \text{BIC}_{\mathbf{y}_1^\star} + \text{BIC}_{\mathbf{y}_2^\star|\mathbf{y}_1^\star}.$$

- **Assessment criteria**

$$\Delta\text{AIC} = \text{AIC}_{\mathbf{y}_{2,0}^\star} - \text{AIC}_{\mathbf{y}_2^\star|\mathbf{y}_1^\star} \;\text{ and }\; \Delta\text{BIC} = \text{BIC}_{\mathbf{y}_{2,0}^\star} - \text{BIC}_{\mathbf{y}_2^\star|\mathbf{y}_1^\star}.$$

SMN-MLMEC model

A5055 clinical trial

◂ A5055

$$
\begin{aligned}
y_{i1k} &= \beta_{10} + \beta_{11} t_{ik} + \beta_{12}\text{treat}_i + \beta_{13} t_{ik}^{0.5} + \beta_{14}\text{treat}_i \times t_{ik} + b_{i10} + b_{i11} t_{ik} + e_{i1k},\\
y_{i2k} &= \beta_{20} + \beta_{21} t_{ik} + \beta_{22}\text{treat}_i + \beta_{23}\text{treat}_i \times t_{ik} + b_{i20} + b_{i21} t_{ik} + e_{i2k},
\end{aligned}
$$

i=1,. . . ,44,

- y_{i1k} is the $\log_{10}$ (RNA) outcome for subject i measured roughly at day_{ik};
- y_{i2k} is the log(CD4/CD8) outcome for subject i measured roughly at day_{ik};
- 316 observations;
- 33% of all viral load measurements are below the detection limit;
- $t_{ik} = \text{day}_{ik}/7$ (week), for $k = 1, \ldots, s_i$, where the weeks are: 0, 7, 14, 28, 56, 84, 112, 140 e 168;
- treat_i is the treatment indicator (= 0 FOR treatment 1; = 1 for treatment 2);
- b_{ij0} and b_{ij1} are the random intercept and random slope, respectively, for y_{ijk}, $j = 1, 2$.

- This dataset was previously analyzed by Wang *et al.* (2015).

SMN-MLMEC model

A5055 clinical trial

Information criteria for the *SMN-MLMEC* models under DEC structure:

	Distribution ϵ / Distribution $\mathbf{b}$								
	N/N	SL/N	T/N	N/SL	N/T	SL/SL	SL/T	T/SL	T/T
AIC	789.85	742.18	739.59	791.98	792.29	744.47	744.54	741.85	741.51
BIC	896.62	853.41	850.81	903.20	903.51	860.14	860.21	857.52	857.19

SMN-MLMEC model

A5055 clinical trial

ML estimates with standard errors for the *SMN-LMMC* model under the T/N distribution:

Structure	Parameters	Estimate (SE)	Parameters	Estimate (SE)	Parameters	Estimate (SE)
DEC	β_{10}	3.743 (0.134)	d_{11}	0.1446 (0.0829)	σ_{11}	0.409 (0.076)
	β_{11}	0.130 (0.026)	d_{21}	0.0011 (0.0133)	σ_{21}	-0.039 (0.020)
	β_{12}	-0.005 (0.067)	d_{22}	-0.0884 (0.1182)	σ_{22}	0.050 (0.011)
	β_{13}	-0.957 (0.098)	d_{31}	-0.0011 (0.0033)	ϕ_1	0.704 (0.065)
	β_{14}	-0.007 (0.025)	d_{32}	0.0034 (0.0027)	ϕ_2	0.632 (0.131)
	β_{20}	-1.284 (0.077)	d_{33}	-0.0122 (0.0116)	ν	4.737 (0.003)
	β_{21}	0.005 (0.005)	d_{41}	-0.0004 (0.0004)		
	β_{22}	0.252 (0.084)	d_{42}	0.2727 (0.0861)		
	β_{23}	-0.003 (0.007)	d_{43}	0.0008 (0.0015)		
			d_{44}	0.0001 (0.0001)		
	loglik	-344.79	AIC	739.59	BIC	850.81
UNC	β_{10}	3.718 (0.135)	d_{11}	0.4089 (0.1463)	σ_{11}	0.263 (0.053)
	β_{11}	0.129 (0.026)	d_{21}	-0.0112 (0.0153)	σ_{21}	-0.024 (0.012)
	β_{12}	0.003 (0.091)	d_{22}	-0.0964 (0.1251)	σ_{22}	0.028 (0.005)
	β_{13}	-0.955 (0.075)	d_{31}	0.0002 (0.0030)	ν	4.340 (0.004)
	β_{14}	-0.008 (0.027)	d_{32}	0.0054 (0.0029)		
	β_{20}	-1.278 (0.076)	d_{33}	-0.0132 (0.0116)		
	β_{21}	0.005 (0.004)	d_{41}	-0.0006 (0.0004)		
	β_{22}	0.286 (0.081)	d_{42}	0.2953 (0.0785)		
	β_{23}	-0.006 (0.006)	d_{43}	0.0002 (0.0015)		
			d_{44}	0.0001 (0.0001)		
	loglik	-357.97	AIC	761.94	BIC	864.26

SMN-MLMEC model

A5055 clinical trial

Decomposition of AIC and BIC for the best *SMN-MLMEC* model:

AIC	739.59	BIC	850.81
$\text{AIC}_{\mathbf{y}_2^\star \mid \mathbf{y}_1^\star}$	92.65	$\text{BIC}_{\mathbf{y}_2^\star \mid \mathbf{y}_1^\star}$	158.80
$\text{AIC}_{\mathbf{y}_{2,0}^\star}$	125.26	$\text{BIC}_{\mathbf{y}_{2,0}^\star}$	166.58
ΔAIC	32.61	ΔBIC	7.77

Summary

Introduction

Preliminaries

SMN-MLMEC model

References

References I

Andrews, D. F. & Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B*, pages 99–102.

Arellano-Valle, R. B., Bolfarine, H. & Lachos, V. (2005). Skew-normal linear mixed models. *Journal of Data Science*, **3**, 415–438.

Delyon, B., Lavielle, M. & Moulines, E. (1999). Convergence of a stochastic approximation version of the em algorithm. *Annals of Statistics*, pages 94–128.

Dempster, A., Laird, N. & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B,*, **39**, 1–38.

Garay, A. M., Castro, L. M., Leskow, J. & Lachos, V. H. (2014). Censored linear regression models for irregularly observed longitudinal data using the multivariate-t distribution. *Statistical Methods in Medical Research*, page DOI: 10.1177/0962280214551191.

Jank, W. (2006). Implementing and diagnosing the stochastic approximation EM algorithm. *Journal of Computational and Graphical Statistics*, **15**(4), 803–829.

Kuhn, E. & Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis*, **49**(4), 1020–1038.

Lachos, V. H., Bandyopadhyay, D. & Dey, D. K. (2011). Linear and nonlinear mixed-effects models for censored hiv viral loads using normal/independent distributions. *Biometrics*, **67**, 1594–1604.

Lange, K. L. & Sinsheimer, J. S. (1993). Normal/independent distributions and their applications in robust regression. *Journal of Computational and Graphical Statistics*, **2**, 175–198.

Matos, L., Lachos, V., Balakrishnan, N. & Labra, F. (2013a). Influence diagnostics in linear and nonlinear mixed-effects models with censored data. *Computational Statistical & Data Analysis*, **57**(1), 450–464.

Matos, L., Prates, M., Chen, M.-H. & Lachos, V. (2013b). Likelihood based inference for linear and nonlinear mixed-effects models with censored response using the multivariate-t distribution. *Statistica Sinica*, **23**, 1323–1345.

Muñoz, A., Carey, V., Schouten, J. P., Segal, M. & Rosner, B. (1992). A parametric family of correlation structures for the analysis of longitudinal data. *Biometrics*, **48**, 733–742.

Samson, A., Lavielle, M. & Mentré, F. (2006). Extension of the SAEM algorithm to left-censored data in nonlinear mixed-effects model: application to HIV dynamics model. *Computational Statistics & Data Analysis*, **51**(3), 1562–1574.

References II

Vaida, F. & Liu, L. (2009). Fast implementation for normal mixed effects models with censored response. *Journal of Computational and Graphical Statistics*, **18**(4), 797–817.

Vaida, F., Fitzgerald, A. & DeGruttola, V. (2007). Efficient hybrid EM for linear and nonlinear mixed effects models with censored response. *Computational Statistics & Data Analysis*, **51**(12), 5718–5730.

Wang, W.-L., Lin, T.-I. & Lachos, V. H. (2015). Extending multivariate-t linear mixed models for multiple longitudinal data with censored responses and heavy tails. *Statistical Methods in Medical Research*, page doi: 0962280215620229.

Wei, G. C. & Tanner, M. A. (1990). A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. *Journal of the American Statistical Association*, **85**(411), 699–704.

Zhang, D., Chen, M.-H., Ibrahim, J. G., Boye, M. E., Wang, P. & Shen, W. (2014). Assessing model fit in joint models of longitudinal and survival data with applications to cancer clinical trials. *Statistics in Medicine*, **33**(27), 4715–4733.

Thank you!