University of Zurich[UZH]

Institute of Social and Preventive Medicine

Problems of INLA in generalized linear mixed models for binary responses

11. September 2013

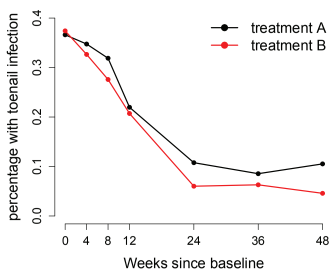
Rafael Sauter and Leonhard Held

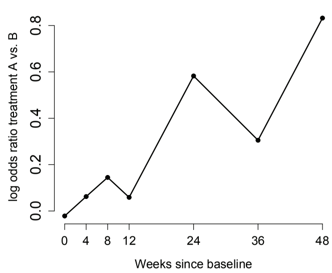
University of Zurich[UZH]

Institute of Social and Preventive Medicine

Toenail infection data (De Backer et al., 1998)

- Randomized, double-blinded study comparing two oral treatments for toenail infection.
- Binary outcome: absent, mild (0) or moderate, severe (1).
- N =289 patients with 1903 observations.





GLMM for toenail infection data

- Patients $i = 1, \ldots, N$, sessions $j = 1, \ldots, n_i$.
  $y_{ij} \mid \mathbf{x}_{ij}, \mathbf{z}_{ij}, \mathbf{b}_i \sim \text{Bernoulli}(\mu_{ij})$
  $\text{logit}(\mu_{ij}) = \mathbf{x}_{ij}^\top \beta + \mathbf{z}_{ij}^\top \mathbf{b}_i$
  $\mathbf{b}_i \sim \mathcal{N}(0, \mathbf{D}(\boldsymbol{\theta}))$

University of Zurich[UZH]
Institute of Social and Preventive Medicine

GLMM for toenail infection data

- Patients $i = 1, \ldots, N$, sessions $j = 1, \ldots, n_i$.
  $y_{ij} \mid \mathbf{x}_{ij}, \mathbf{z}_{ij}, \mathbf{b}_i \sim \text{Bernoulli}(\mu_{ij})$
  $\text{logit}(\mu_{ij}) = \mathbf{x}_{ij}^\top \beta + \mathbf{z}_{ij}^\top \mathbf{b}_i$
  $\mathbf{b}_i \sim \mathcal{N}(0, \mathbf{D}(\boldsymbol{\theta}))$

  $\mathbf{x}_{ij} = (\text{intercept}, \text{time}_{ij}, \text{time}_{ij} : \text{treatment}_i)$
- time_{ij} is measured in months since baseline.
- Fixed effect for treatment is omitted because treatment and y_{ij} are independent at baseline due to randomization.

11.9.2013 ROeS 2013 Page 3

GLMM for toenail infection data

- Patients $i = 1, \dots, N$, sessions $j = 1, \dots, n_i$.
  $y_{ij} \mid \mathbf{x}_{ij}, \mathbf{z}_{ij}, \mathbf{b}_i \sim \text{Bernoulli}(\mu_{ij})$
  $\text{logit}(\mu_{ij}) = \mathbf{x}_{ij}^\top \beta + \mathbf{z}_{ij}^\top \mathbf{b}_i$
  $\mathbf{b}_i \sim \mathcal{N}(0, \mathbf{D}(\boldsymbol{\theta}))$
  $\mathbf{x}_{ij} = (\text{intercept}, \text{time}_{ij}, \text{time}_{ij} : \text{treatment}_i)$
- time_{ij} is measured in months since baseline.
- Fixed effect for treatment is omitted because treatment and y_{ij} are independent at baseline due to randomization.
- Random intercept, random slope: $\mathbf{z}_{ij} = (1, t_{ij})^\top$
  covariance matrix $\mathbf{D} = \begin{pmatrix} \sigma_{z_1}^2 & \rho\sigma_{z_1}\sigma_{z_2} \\ \rho\sigma_{z_2}\sigma_{z_1} & \sigma_{z_2}^2 \end{pmatrix}$.

ML-estimation with Laplace Approximation

Marginal likelihood contribution for patient i:

$$L_i(\mathbf{y}_i \mid \beta, \mathbf{D}) = \int \pi(\mathbf{y}_i \mid \beta, \mathbf{b}_i)\pi(\mathbf{b}_i \mid \mathbf{D}) d\mathbf{b}_i$$

ML-estimation with Laplace Approximation

Marginal likelihood contribution for patient i:

$$L_i(\mathbf{y}_i \mid \beta, \mathbf{D}) = \int \pi(\mathbf{y}_i \mid \beta, \mathbf{b}_i)\pi(\mathbf{b}_i \mid \mathbf{D}) d\mathbf{b}_i$$

Iterate:

1. Find the mode $\hat{\mathbf{b}}_i = \arg\max_{\mathbf{b}_i} \pi(\mathbf{y}_i \mid \hat{\beta}, \mathbf{b}_i)\pi(\mathbf{b}_i \mid \hat{\mathbf{D}})$ with penalized iterative reweighted least squares.
2. Maximize Laplace-approximation of $L_i(\mathbf{y}_i \mid \beta, \mathbf{D})$ around $\hat{\mathbf{b}}_i$ and obtain $\hat{\beta}, \hat{\mathbf{D}}$.

ML-estimation with Gauss Hermite (GH) Approximation

– Adaptive GH-approximation for GLMM (Pinheiro and Bates, 1995):

ML-estimation with Gauss Hermite (GH) Approximation

- Adaptive GH-approximation for GLMM (Pinheiro and Bates, 1995):
  - The accuracy of the GH approximation increases with increasing number of quadrature points (Q).
  - The adaptive GH approximation with Q=1 is equal to the Laplace approximation.

University of Zurich[UZH]
Institute of Social and Preventive Medicine

ML-estimation with Gauss Hermite (GH) Approximation

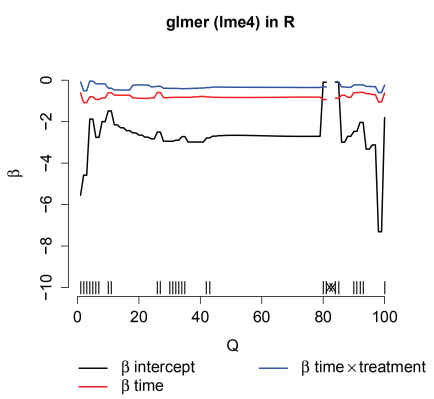
- Adaptive GH-approximation for GLMM (Pinheiro and Bates, 1995):
  - The accuracy of the GH approximation increases with increasing number of quadrature points (Q).
  - The adaptive GH approximation with Q=1 is equal to the Laplace approximation.
- Lesaffre and Spiessens (2001) compared non-adaptive and adaptive GH approximation for the toenail data.

11.9.2013 ROeS 2013 Page 5

ML-estimation with Gauss Hermite (GH) Approximation

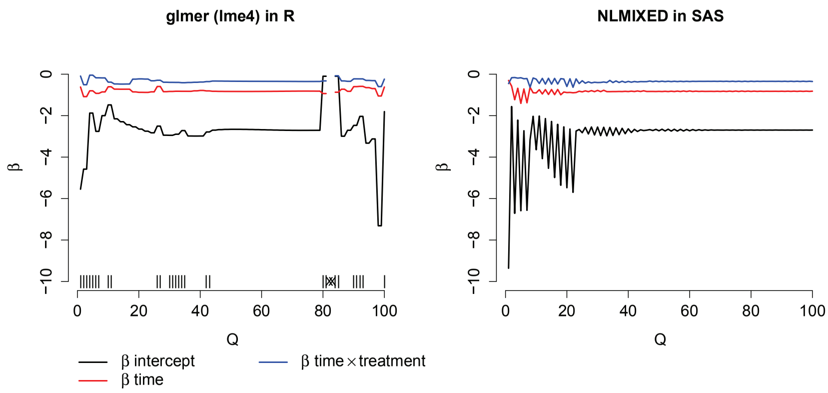
- Adaptive GH-approximation for GLMM (Pinheiro and Bates, 1995):
  - The accuracy of the GH approximation increases with increasing number of quadrature points (Q).
  - The adaptive GH approximation with Q=1 is equal to the Laplace approximation.
- Lesaffre and Spiessens (2001) compared non-adaptive and adaptive GH approximation for the toenail data.
- Zhang et al. (2011) simulation study comparing statistical packages for GLMM with binary response.

Fixed effects estimates with varying Q



- **X**: "Error message: Downdated X'X is not positive definite, 1."
- |: "Warning message:In mer_finalize(ans) : false convergence (8)"

Fixed effects estimates with varying Q



– **X**: "Error message: Downdated X'X is not positive definite, 1."

– |: "Warning message:In mer_finalize(ans) : false convergence (8)"

Approximate Bayesian estimation with integrated nested Laplace approximation (INLA)

– Data $\mathbf{y}$, parameters $\gamma = (\beta, \mathbf{b}_i)^\top$, hyperparameters $\boldsymbol{\theta}$ define a GMRF (Rue and Held, 2005).

Approximate Bayesian estimation with integrated nested Laplace approximation (INLA)

– Data $\mathbf{y}$, parameters $\boldsymbol{\gamma} = (\boldsymbol{\beta}, \mathbf{b}_i)^\top$, hyperparameters $\boldsymbol{\theta}$ define a GMRF (Rue and Held, 2005).
$\rightarrow$ Requires a prior $\pi(\boldsymbol{\theta})$.

University of Zurich^UZH
Institute of Social and Preventive Medicine

Approximate Bayesian estimation with integrated nested Laplace approximation (INLA)

– Data $\mathbf{y}$, parameters $\boldsymbol{\gamma} = (\boldsymbol{\beta}, \mathbf{b}_i)^\top$, hyperparameters $\boldsymbol{\theta}$ define a GMRF (Rue and Held, 2005).
$\rightarrow$ Requires a prior $\pi(\boldsymbol{\theta})$.

– Marginal posterior:

$$\pi(\gamma_k \mid \mathbf{y}) = \int \pi(\gamma_k \mid \boldsymbol{\theta}, \mathbf{y}) \pi(\boldsymbol{\theta} \mid \mathbf{y}) d\boldsymbol{\theta}$$

Approximate Bayesian estimation with integrated nested Laplace approximation (INLA)

- Data $\mathbf{y}$, parameters $\boldsymbol{\gamma} = (\boldsymbol{\beta}, \mathbf{b}_i)^\top$, hyperparameters $\boldsymbol{\theta}$ define a GMRF (Rue and Held, 2005).
  $\rightarrow$ Requires a prior $\pi(\boldsymbol{\theta})$.
- Marginal posterior:

$$\pi(\gamma_k \mid \mathbf{y}) = \int \pi(\gamma_k \mid \boldsymbol{\theta}, \mathbf{y}) \pi(\boldsymbol{\theta} \mid \mathbf{y}) d\boldsymbol{\theta}$$

- INLA (Rue et al., 2009):

$$\pi(\gamma_k \mid \mathbf{y}) \approx \sum_u \tilde{\pi}(\gamma_k \mid \boldsymbol{\theta}_u, \mathbf{y}) \tilde{\pi}(\boldsymbol{\theta}_u \mid \mathbf{y}) \Delta_u$$

University of Zurich[UZH]
Institute of Social and Preventive Medicine

Approximate Bayesian estimation with integrated nested Laplace approximation (INLA)

- Data $\mathbf{y}$, parameters $\boldsymbol{\gamma} = (\boldsymbol{\beta}, \mathbf{b}_i)^\top$, hyperparameters $\boldsymbol{\theta}$ define a GMRF (Rue and Held, 2005).
  $\rightarrow$ Requires a prior $\pi(\boldsymbol{\theta})$.
- Marginal posterior:

$$\pi(\gamma_k \mid \mathbf{y}) = \int \pi(\gamma_k \mid \boldsymbol{\theta}, \mathbf{y}) \pi(\boldsymbol{\theta} \mid \mathbf{y}) d\boldsymbol{\theta}$$

- INLA (Rue et al., 2009):

$$\pi(\gamma_k \mid \mathbf{y}) \approx \sum_u \tilde{\pi}(\gamma_k \mid \boldsymbol{\theta}_u, \mathbf{y}) \tilde{\pi}(\boldsymbol{\theta}_u \mid \mathbf{y}) \Delta_u$$

  1. $\tilde{\pi}(\boldsymbol{\theta}_u|\mathbf{y})$ Laplace approximation for $\pi(\boldsymbol{\theta}_u|\mathbf{y})$.

11.9.2013 ROeS 2013 Page 7

Approximate Bayesian estimation with integrated nested Laplace approximation (INLA)

- Data $\mathbf{y}$, parameters $\boldsymbol{\gamma} = (\beta, \mathbf{b}_i)^\top$, hyperparameters $\boldsymbol{\theta}$ define a GMRF (Rue and Held, 2005).
  $\longrightarrow$ Requires a prior $\pi(\boldsymbol{\theta})$.
- Marginal posterior:

$$\pi(\gamma_k \mid \mathbf{y}) = \int \pi(\gamma_k \mid \boldsymbol{\theta}, \mathbf{y}) \pi(\boldsymbol{\theta} \mid \mathbf{y}) d\boldsymbol{\theta}$$

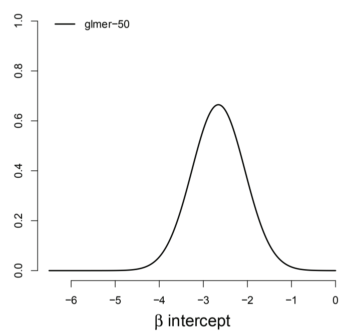
- INLA (Rue et al., 2009):

$$\pi(\gamma_k \mid \mathbf{y}) \approx \sum_u \tilde{\pi}(\gamma_k \mid \boldsymbol{\theta}_u, \mathbf{y}) \tilde{\pi}(\boldsymbol{\theta}_u \mid \mathbf{y}) \Delta_u$$

  1. $\tilde{\pi}(\boldsymbol{\theta}_u|\mathbf{y})$ Laplace approximation for $\pi(\boldsymbol{\theta}_u|\mathbf{y})$.
  2. $\tilde{\pi}(\gamma_k|\boldsymbol{\theta}_u, \mathbf{y})$ Gaussian, simplified or full Laplace approximation for $\pi(\gamma_k|\boldsymbol{\theta}_u, \mathbf{y})$.

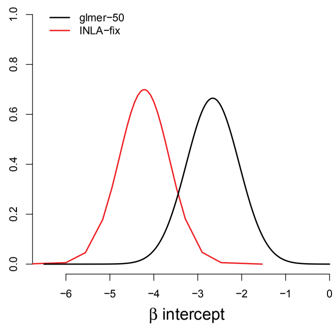
Intercept for GLMER, INLA, MCMC fixed θ

- **glmer-50**: ML-estimation with GH approximation using `glmer` in package `lme4` by Bates et al. (2011) with Q=50.



Intercept for GLMER, INLA, MCMC fixed θ

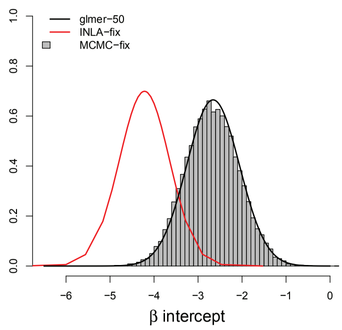
- **glmer-50**: ML-estimation with GH approximation using `glmer` in package `lme4` by Bates et al. (2011) with Q=50.
- **INLA-fix**: INLA-estimation with fixed $\boldsymbol{\theta}$ at ML-estimates.



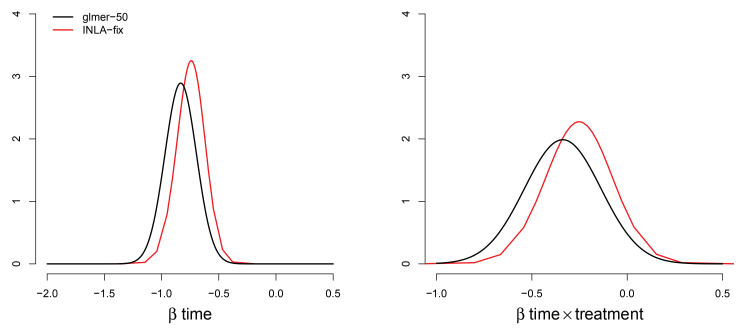
University of Zurich[UZH]
Institute of Social and Preventive Medicine

Intercept for GLMER, INLA, MCMC fixed θ

- **glmer-50**: ML-estimation with GH approximation using `glmer` in package `lme4` by Bates et al. (2011) with Q=50.
- **INLA-fix**: INLA-estimation with fixed $\boldsymbol{\theta}$ at ML-estimates.
- **MCMC**: MCMC sampling with fixed $\boldsymbol{\theta}$ at ML-estimates.

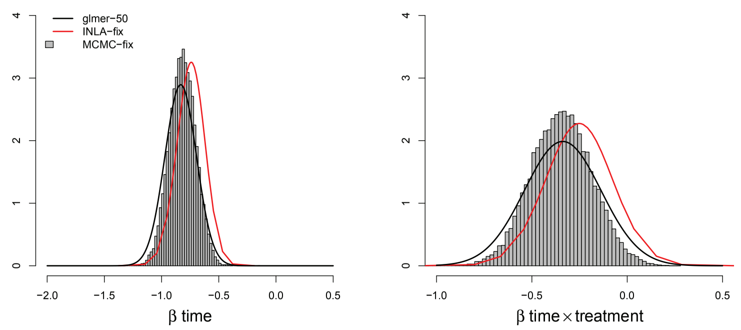


Time, time:treatment for GLMER, INLA, MCMC fixed θ



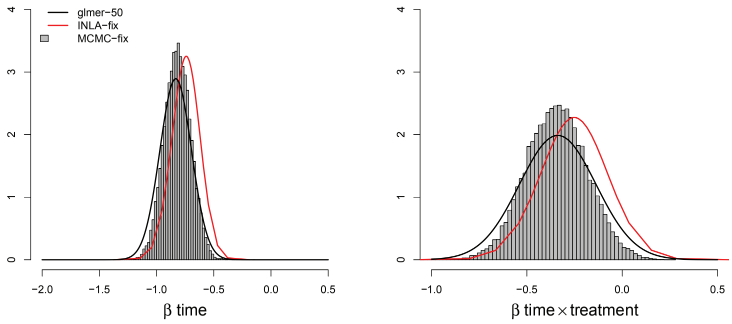
University of Zurich[UZH]
Institute of Social and Preventive Medicine

Time, time:treatment for GLMER, INLA, MCMC fixed θ



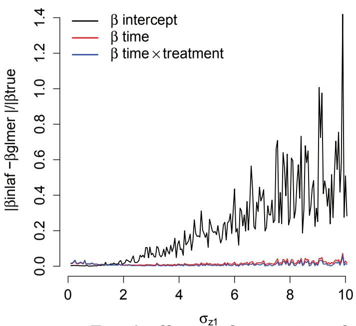
11.9.2013 ROeS 2013
Page 9

Time, time:treatment for GLMER, INLA, MCMC fixed θ



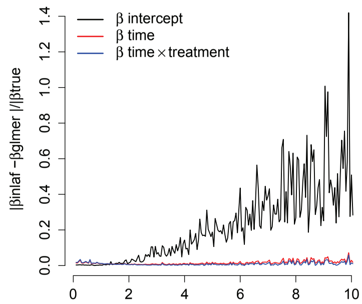
- 95%-CI GLMER: [-0.73, 0.06]
- 95%-CI INLA: [-0.60, 0.09]
- 95%-CI MCMC: [-0.66, -0.03]

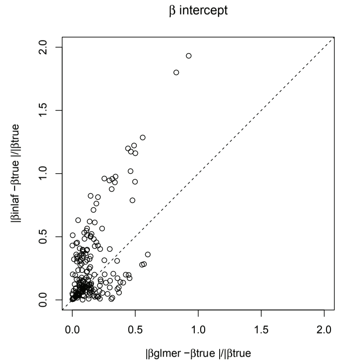
Simulation for a RI model



- Fixed effects: $\beta_0 = -2.5$, $\beta_1 = -0.8$, $\beta_2 = 0.8$.
- Random effect standard deviation σ_{z1}: 0.1 to 10 by 0.05.
- 290 patients for 7 visits at times 0, 1, 2, 3, 6, 9, 12.
- $Q = 50$ for `glmer`.

Simulation for a RI model





- Fixed effects: $\beta_0 = -2.5$, $\beta_1 = -0.8$, $\beta_2 = 0.8$.
- Random effect standard deviation σ_{z1}: 0.1 to 10 by 0.05.
- 290 patients for 7 visits at times 0, 1, 2, 3, 6, 9, 12.
- $Q = 50$ for `glmer`.

Summary

- ML-estimation of GLMM parameters for toenail data show variation depending on Q in GH approximation.

Summary

- ML-estimation of GLMM parameters for toenail data show variation depending on Q in GH approximation.
- Is there a better way to approximate the likelihood?

University of Zurich[UZH]
Institute of Social and Preventive Medicine

Summary

- ML-estimation of GLMM parameters for toenail data show variation depending on Q in GH approximation.
- Is there a better way to approximate the likelihood?
- Use INLA with fixed hyperparameters and compare... ...with GLMER and MCMC sampling.

Summary

- ML-estimation of GLMM parameters for toenail data show variation depending on Q in GH approximation.
- Is there a better way to approximate the likelihood?
- Use INLA with fixed hyperparameters and compare... ...with GLMER and MCMC sampling.
- Approximation of the Likelihood with INLA resulted in shifted marginal posterior distributions.

Bates, D., Maechler, M. and Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. .

De Backer, M., De Vroey, C., Lesaffre, E., Scheys, I. and De Keyser, P. (1998). Twelve weeks of continuous oral therapy for toenail onychomycosis caused by dermatophytes: A double-blind comparative trial of terbinafine 250 mg/day versus itraconazole 200 mg/day. *Journal of the American Academy of Dermatology* **38** S57 – S63.
URL http://www.sciencedirect.com/science/article/pii/S0190962298704864

Lesaffre, E. and Spiessens, B. (2001). On the effect of the number of quadrature points in a logistic random effects model: an example. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **50** 325–335.

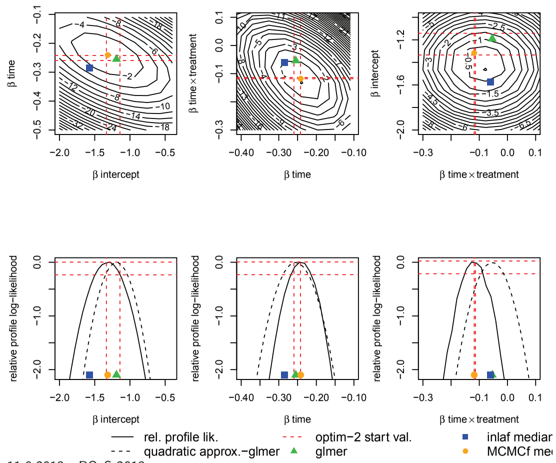
Pinheiro, J. C. and Bates, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics* **4** 12–35.

Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications*. Chapman & Hall/CRC Press, London.

Rue, H., Martino, S. and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society - Series B* **71** 319–392.

Zhang, H., Lu, N., Feng, C., Thurston, S. W., Xia, Y., Zhu, L. and Tu, X. M. (2011). On fitting generalized linear mixed-effects models for binary responses using different statistical packages. *Statistics in Medicine* **30** 2562–2572.
URL http://dx.doi.org/10.1002/sim.4265

Profiled relative log-likelihood surface



Computational time for GLMER and INLA