Referee's report

100112M: Bayesian analysis for nonlinear mixed-effects models under heavy-tailed distributions

Authors: Rolando De la Cruz

Journal: *Biometrics*

Date: *March 5, 2010*

The author presents a nonlinear mixed model in which both random effects and within-subject errors follow a scale mixture of multivariate normal distributions (SMMN) instead of conventionally used multivariate normal distribution. Bayesian inference based on MCMC is used in the paper. This is certainly a useful contribution in the field of mixed models. Nevertheless, I have some remarks which should be considered before the paper may get accepted. Especially, I do not like too much the "automatic" style of the paper: this is the model, these are the distributions, these are full conditional distributions etc. without providing more information concerning strong and weak points and sharing the particular experience with the proposed model.

Particular remarks

- ✧ In expression (7), why don't you allow for covariates working on the level of random effects? That is, why not to replace expression (7) by $\boldsymbol{\phi}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i$ (like in Lindstrom and Bates, 1990, expression (2.2)). In the most standard situations with $\boldsymbol{Z}_i$ matrix being a submatrix of $\boldsymbol{X}_i$, the random effects have then easily understandable interpretation as subject-specific modifiers of the mean values $\boldsymbol{\beta}$.

- ✧ You state on bottom of page 10 that "*In the absence of good prior information, a convenient strategy to avoid improper posterior distribution is to use diffuse proper priors so that the analysis is dominated by the data likelihood.*" However, neither here nor in the application part (Section 6), there is indication on how to choose these diffuce priors in particular applications. Certainly in Section 6, there must be a list of particular hyperparameter values used for that particular application. Otherwise, the results shown in the paper are not reproducible by others.

- ✧ On several places (pp. 12–15) you state that you employ the Metropolis-Hastings (MH) algorithm to sample from the full conditional distribution which is not analytically available without being specific about the details of the MH algorithm. That is, how you choose proposal distribution, is it necessary to tune the proposal, if so, how etc. This does not allow the reader to repeat your analyses, especially when a good proposal distribution is rather crucial for the success of the MH algorithm.

- ✧ I do not understand how you derived the formula for $\widehat{\mathrm{CPO}}_i(\mathcal{M}_l)$ in the middle of page 17. Is this the correct expression?

- ✧ Discussion in Section 8 is more a summary of the paper than the discussion. I think that this section should discuss much more the strong and weak points of the proposed methods. The last paragraph of the discussion states that the current approach does not account for possible skewness and that further generalizations are needed to allow for skewed distributions. As far

as I know, the author himself has already done something in this direction (De la Cruz, 2008). Why is this not mentioned in the paper (and maybe compared to the proposed method)?

- ✧ In the paper, there is no indication on how to run the MCMC simulation in practice without the need to implement everything again when the reader decides to analyze his/her data using the methods presented in the paper. I strongly recommend that the author accompanies the paper by some well documented software with the example code (certainly at least for the analysis of publicly available Theophylline Kinetics data shown in the paper). Personally, I feel that the contemporary statistical literature is full of methods whose usage is hampered by too high time requirements to implement them which cannot be taken care of by potential users of the method.

I am aware of the fact that I require to add some technical details: guidelines on how to choose weakly informative prior distributions, more details on Metropolis-Hastings algorithm, which could inflate the length of the paper beyond acceptable size. Nevertheless, I think that many technical details could be placed in a separate document which would be only on-line available (maybe as a part of the documentation of the software that I also mention above). Also some other pieces from current document (e.g., all full conditional distributions) could be placed in on-line supplement to break down the "automatic" style of the paper and to improve its readability and appeal.

Typographical and other minor remarks

- ✧ **P. 4 (+4)** "Let $\boldsymbol{Y} \sim$" should be "Let $\boldsymbol{Z} \sim$".
- ✧ **P. 4 (−6)** $\boldsymbol{\Lambda}$ in "$SMMN_d(\boldsymbol{\mu}, \boldsymbol{\Lambda}; H)$" should probably be replaced by $\boldsymbol{\Sigma}$ (also on p. 5 (+3)) to correspond to expressions (2), (3), (4).
- ✧ **P. 5 (+7)** Twice the word "multivariate".
- ✧ **P. 18 (+11)** Reference van der Linde (2009) should be van der Linde (2005).
- ✧ **P. 19 (+9)** I would prefer to replace the word "non-informative" by the word "weakly informative" (or similarly).

References

R. De la Cruz. Bayesian non-linear regression models with skew-elliptical errors: Application to the classification of longitudinal profiles. *Computational Statistics and Data Analysis*, 53(2):436–449, 2008.

M. J. Lindstrom and D. M. Bates. Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46(3):673–887, 1990.

A. van der Linde. DIC in variable selection. *Statistica Neerlandica*, 59:45–56, 2005.