Bayesian Analysis (2009) **4**, Number 2, pp. 213–216

Comments on Article by Yin

Ciprian M. Crainiceanu*

Inferential methods for Generalized Linear Mixed Models (GLMMs) are under intense methodological development because they: 1) are widely applicable; and 2) raise non-trivial technical and inferential challenges. The Generalized Method of Moments (GMM) (Hansen (1982); Newey and West (1987)) provides a powerful and robust set of inferential tools for GLMMs, especially when the likelihood formulation is difficult and interest is centered on the fixed effects parameters.

The paper by Yin (2009) is an important contribution to this literature. The main idea of the paper is to provide a simple Bayesian framework for what I considered to be a frequentist method, par excellence. I found the paper thought provoking, fresh and definitely worthy of discussion. Below I summarize my reactions and comments and provide a set problems that could be, but are not currently, addressed by this methodology.

1 Why?

The most important question in my mind after reading the paper was "Why should we use Bayesian GMM instead of GMM?" Simulations seem to indicate that both methods produce similar results, with the Bayesian methodology requiring more computational effort. One answer that I do not particularly like is "Because we can". Another possible answer could be that in some data sets with a smaller number of clusters the posterior distribution $\widetilde{\pi}(\boldsymbol{\beta}|\mathbf{y}) \propto \widetilde{L}(\mathbf{y}|\boldsymbol{\beta})\pi(\boldsymbol{\beta})$ might not be well approximated by a normal. In such a context, the next natural step would be to consider the sampling variability of the data by conducting a nonparametric bootstrap of the clusters. Pooled analyses using Bayesian GMM and GMM could then be compared. Some applications and simulations supporting these ideas would add credibility to the proposed methods.

2 What?

The approach proposed by Yin is to treat the quadratic objective function

$$Q_n(\boldsymbol{\beta}) = \mathbf{U}_n^T(\boldsymbol{\beta})\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\beta})\mathbf{U}_n^T(\boldsymbol{\beta})$$

as an approximation of minus twice the log of the conditional likelihood $L(\mathbf{y}|\boldsymbol{\beta})$. More precisely, the author replaced the unknown $L(\mathbf{y}|\boldsymbol{\beta})$ by the approximate likelihood $\widetilde{L}(\mathbf{y}|\boldsymbol{\beta}) = \exp\{-Q_n(\boldsymbol{\beta})/2\}$. When observations are not clustered $\mathbf{U}_n(\boldsymbol{\beta}) = \sum_{i=1}^n U_i(\boldsymbol{\beta})/n$, where $U_i(\boldsymbol{\beta}) = \mathbf{D}_i v_i^{-1}(y_i - \mu_i)$, $\mathbf{D}_i = \partial\mu_i/\partial\boldsymbol{\beta}$ is the vector of derivatives of the subject i-specific mean with respect to the model parameters, and $v_i = \mathrm{var}(y_i|\mathbf{Z}_i)$ is the conditional vari-

*Department of Biostatistics, Johns Hopkins University, Baltimore, MD, mailto:ccrainic@jhsph.edu

 DOI:10.1214/09-BA407B