Estimation in the probit normal model for binary outcomes using the SAEM algorithm

Cristian Meza[1], Florence Jaffrézic[2] and Jean-Louis Foulley[2]

[1] Departamento de estadística, CIMFAV, Universidad de Valparaíso, Valparaíso, Chile
[2] INRA, UR337 Station de Génétique Quantitative et Appliquée, Jouy–en–Josas, F–78350, France

Generalized linear mixed models (GLMM) form a very general class of random effects models for discrete and continuous responses in the exponential family. They are useful in a variety of applications. The traditional likelihood approach for GLMM usually involves high dimensional integrations which are computationally intensive. In this work, we investigate the case of binary outcomes analyzed under a two stage probit normal model with random effects. First, it is showed how ML estimates of the fixed effects and variance components can be computed using a stochastic approximation of the EM algorithm (SAEM). The SAEM algorithm can be applied directly, or in conjunction with a parameter expansion version of EM to speed up the convergence. A procedure is also proposed to obtain REML estimates of variance components and REML–based estimates of fixed effects. Finally an application to a real data set involving a clinical trial is presented in which these techniques are compared to other procedures (PQL, ML, Bayes) already available in classical softwares (SAS Glimmix, SAS Nlmixed, WinBUGS). As observed with other mixed models, ML estimates were found to be biased downwards and this bias was large for complex models involving many coviariates. Strong discrepancies were observed for the PQL methods both with ML and REML versions. On the other hand, the SAEM–REML algorithm was found to provide estimates close to the Bayesian Mode a Posteriori.