Bayesian model selection using INLA with application to longitudinal count data

Christine Adrion[1] & Ulrich Mansmannn[1]

[1] *University of Munich, Institute for Medical Information Sciences, Biometry and Epidemiology, Germany.*
{adrion, mansmann}@ibe.med.uni-muenchen.de

An important aspect in modeling longitudinal count data is to select among a series of distributional assumptions accommodating overdispersion or zero-inflation. Additionally, clustering is handled by introducing a set of random effects. We used the INLA approach for fitting generalized linear mixed models (GLMMs) for clinical trial data on vertigo attacks within a Bayesian setting. The INLA methodology enables the direct computation of leave-one-out predictive distributions which are crucial for Bayesian model assessment. Competing GLMMs for longitudinal counts were evaluated according to the DIC, PIT, and the logarithmic score (LS). A simulation study was performed to investigate the discriminatory power of these Bayesian tools for model criticism in case of a longitudinal, negative binomial data generating process, by assuming varying degrees of overdispersion and sample sizes. We found that Bayesian methods are not only appealing for inference but notably provide a sophisticated insight into different aspects of model performance, such as forecast verification or calibration checks, and can be applied within the model selection process. The mean of LS provides a robust tool for model ranking which is not sensitive to sample size. For negative binomial models a non-vague prior should be assigned for the dispersion parameter to minimize the risk of a biased posterior mean which may affect the LS.

Keywords: Bayesian Analysis; INLAs; GLMMs for count data; Model Assessment; Model Comparison.