Efficient MCMC Estimation for Binomial Binned Logit Models

Agnes Fussl, Sylvia Frühwirth-Schnatter[a] & Rudolf Frühwirth[b]

[a]*Department of Applied Statistics, Johannes Kepler University, Linz*
[b]*Institute of High Energy Physics, Austrian Academy of Sciences, Vienna*

Presenting author: agnes.fussl@jku.at

This talk considers the special case of binomial data where repeated measurements are available for identical covariate patterns.

Our method is based on various Bayesian approaches for estimating binary logit models. To carry out MCMC sampling with data augmentation, logit models are rewritten as random utility models (RUM) or difference RUM (dRUM). Following earlier papers (Frühwirth-Schnatter and Frühwirth, 2007, *Computational Statistics and Data Analysis*, **51:**3509–3528; Frühwirth-Schnatter and Frühwirth, 2010, *Statistical Modelling and Regression Structures - Festschrift in Honour of Ludwig Fahrmeir*, 111–132), the individual RUM or dRUM representations of the binomial logit models can be used to estimate the regression parameters. However, this approach leads to an unfeasibly high-dimensional latent variable, in particular if the group sizes of observations with identical covariate patterns are large. It is possible to reduce the dimension of the individual representation by introducing an aggregated RUM version of the binomial model (Frühwirth-Schnatter, Frühwirth, Held and Rue, 2009, *Statistics and Computing*, **19:**479–492).

To improve the sampler further, we suggest a new aggregated dRUM representation of the binomial binned logit model. The modifications lead to a considerable reduction of computing time and a remarkable gain in efficiency. The parameters appearing in the regression model are estimated by using various MCMC algorithms: a data-augmented MH sampler, an auxiliary mixture sampler and a new hybrid auxiliary mixture (HAM) sampler. To demonstrate the properties of the different methods, their performance is evaluated on some well-known data sets.

Keywords: Markov Chain Monte Carlo; Binomial Binned Logit Model; Repeated Measurements; Aggregated dRUM; HAM Sampler.