Bayesian Variable Selection in Normal Regression Models

Gertraud Malsiner-Walli

A challenge in statistical analysis is variable selection in normal linear regression models, i.e. to decide which subset of covariables should be included in the final explanatory model. Many methods of Bayesian variable selection use spike and slab priors for the regression coefficients with the spike around zero to allow shrinkage of small effects to zero and a flat slab elsewhere. We consider different spike and slab priors where the spike is specified either as by Dirac measure at zero or by a continuous distribution with small variance. Variable selection is accomplished by MCMC methods with different sampling schemes depending on the spike types. In a simulation study with different settings (independent or correlated regressors, different scales of the prior parameters) we study the influence of the different prior approaches on the selection of the regressors. Application of the methods is illustrated on a data set on patients with psychopharmacological treatment, where the goal is to assess the effect of this treatment on parameters of lipid and glucose metabolism.