CAPTURING CORRELATED CLUSTERS USING MIXTURES OF LATENT CLASS MODELS

Gertraud Malsiner-Walli¹, Bettina Grün¹ and Sylvia Frühwirth-Schnatter¹

 $^{\rm 1}$ Institute of Statistics and Mathematics, Vienny University of Economics and Business

(gertraud.malsiner-walli@wu.ac.at, bettina.gruen@wu.ac.at, sylvia.fruehwirth-schnatter@wu.ac.at,)

ABSTRACT: Latent class models rely on the conditional independence assumption, i.e., it is assumed that the categorical variables are independent given the cluster memberships. Within the Bayesian framework, we propose a suitable specification of priors for the latent class model to identify the clusters in multivariate categorical data where the independence assumption is not fulfilled. Each cluster distribution is approximated by a latent class model, leading overall to a two-layer mixture of latent class models. By carefully specifying the priors on the model parameters, the Bayesian approach allows to identify the clusters and fit their cluster distributions using MCMC sampling. We provide suitable estimation and inference methods for the mixture of latent class models and illustrate the performance of this approach on a real data set containing patients suffering frome one of three types of low back pain.

KEYWORDS: Bayesian inference, model-based clustering, prior on the number of components, telescoping sampler

References

- FOP, MICHAEL, SMART, KEITH M, & MURPHY, THOMAS BRENDAN. 2017. Variable selection for latent class analysis with application to low back pain diagnosis. *The Annals of Applied Statistics*, **11**(4), 2080–2110.
- FRÜHWIRTH-SCHNATTER, SYLVIA, MALSINER-WALLI, GERTRAUD, & GRÜN, BETTINA. 2021. Generalized Mixtures of Finite Mixtures and Telescoping Sampling. *Bayesian Analysis*, 16(4), 1279–1307.
- MALSINER-WALLI, GERTRAUD, FRÜHWIRTH-SCHNATTER, SYLVIA, & GRÜN, BETTINA. 2017. Identifying mixtures of mixtures using Bayesian estimation. *Journal of Computational and Graphical Statistics*, **26**(2), 285–295.