

SEM-Gibbs algorithm for Multilevel Cross-Classified Latent Class analysis of Binary Data

Nicola Piras and Silvia Columbu

University of Cagliari, Department of Mathematics and Computer Science

We present a Stochastic version of the EM algorithm for the estimation of Latent Classes for binary data in a multilevel cross-classified structure (Columbu & Vermunt (2021)). Data are cross-classified when first level units can be considered simultaneously nested within two (or even more) higher level units. The basic idea of the method is to include a stochastic step between the E and the M step in order to reconstruct the joint distribution of cross-classified latent classes conditionally to the observed data. This joint probability cannot be obtained as the product of the single latent classes probabilities. In particular, in the E-step we consider a Gibbs sampler by repeatedly simulating from the full conditional distributions of the two second level and first level component labels. In the M-step the mixing parameters and the Bernoulli probability parameter are updated. Numerical experiments with simulated data are presented in order to assess the performances of the method.

References

- J.K.Vermunt, *Latent class and finite mixture models for multilevel data sets*, Statistical Methods in Medical Research 2008; **17**: 33–51
- C.Keribin, V.Brault, G.Celeux, G.Govaert, *Estimation and Selection for the Latent Block Model on Categorical Data*, Statistics and Computing **25(6)** (2015), 1201-1216
- M.Selosse, J.Jacques, C.Biernacki, *ordinalClust: An R Package to Analyze Ordinal Data*, The R Journal (2021)
- S.Columbu, J.K.Vermunt “Multilevel Latent Class Models for cross classified data”, Book of Abstracts CFE-CMStatistics 2021, ISBN 978-9925-7812-5-6 (2021).
-