

On the Identifiability and Estimability of Latent Class Models: A Bayesian Analysis

Ernesto San Martín
Department of Statistics
P. Universidad Católica de Chile
esanmart@mat.puc.cl

Paul De Boeck
Department of Psychology
K. U. Leuven
Paul.DeBoeck@psy.kuleuven.ac.be

Keywords: Identifiability, Minimal Sufficient Statistics, Minimal Sufficient Parameter, Posterior Expectation.

Abstract

Latent Class Models (LCMs) assume that the population, from which the observed sample is taken, is composed of m mutually exclusive latent classes. The parameters of interest are the probabilities with which a randomly chosen subject belongs to each of the latent classes. For each person, L dichotomous measurements are made. In LCM it is assumed that, for each item, every class has a specific probability of positive responses: these are the parameters of interest in the conditional model.

To ensure a coherent inference on the structural parameters, their identifiability is needed. Different solutions to this problem can be found in the psychometric literature. A first group can be found in the early works of Andersen (1954), McHugh (1956), Madansky (1960), Goodman (1974) and Clogg and Goodman (1984). These works are all related since they are based on linear equation systems defined on the marginal probabilities of response patterns of the individuals from the population as a whole.

Another approach to analyze this identification problem is motivated by the hierarchical structure underlying LCMs. The key idea is to relate the identifiability of the statistical model with the identifiability of both the conditional model and the marginal latent model. A heuristic analysis in this line can be found in Maris (1999).

This paper is motivated by the last approach. At the identification level, we establish that, under prior independence of the parameters of interest, the identification of the marginal model is a necessary condition for the identification of the statistical model. A second result we explore in this paper deals with the consistency of a Bayesian estimator of the probabilities of correct responses given the latent class, denoted by P . More specifically, we show that the conditional expectation of the P given the observations and the latent class probabilities converges to P . A similar result is established for the probabilities of latent classes.

References

- Anderson, T. W. (1954), On Estimation of Parameters in Latent Structure Analysis. *Psychometrika* **19**, 1–10.
- Clogg, C. C. and Goodman, L. A. (1984), Latent Structure Analysis of a Set of Multidimensional Contingency Tables. *Journal of the American Statistical Association* **79**, 762–771.
- Goodman, L. A. (1974), Exploratory latent structural analysis using both identifiable and unidentifiable models. *Biometrika* **61**, 215–231.
- Madansky, A. (1960), Determinantal Methods in Latent Class Analysis. *Psychometrika* **25**, 183–198.

Maris, E. (1999), Estimating Multiple Classification Latent Class Models. *Psychometrika* **64**, 187–212.

McHugh, R. B. (1956), Efficient Estimation and Local Identification in Latent Class Models. *Psychometrika* **21**, 331–347.

Autor1, Autor2 (year). Title etc..