Factor Mixture Model: an overview*

Catarina Marques

Instituto Universitário de Lisboa (ISCTE-IUL), BRU-IUL, catarina.marques@iscte.pt

Abstract

Factor analysis assumes that observations are sampled from a homogeneous population with parameters invariant across individuals. This assumption tends to be strong and in most empirical applications not very realistic thus parameter estimates turn out to be biased (Muthén, 1989; Jedidi et al., 1997; Yung, 1997; Wall et al., 2012). Unobserved heterogeneity may result from the existence of subgroups in population, which are characterised by distinct latent structures that varies when the underlying latent factors are not normally distributed with the same parameters or even when they have distinct distributions (Wall et al., 2012). Factor mixture models incorporate discrete unobserved heterogeneity (i.e. clustering structure) in covariance structure models. Consequently, the factor mixture model adds a categorical latent variable that captures the unobserved population heterogeneity to the factor model that imposes a low-dimensional structure on the covariance matrix of the observed variables within each component.

In this seminar it will be discussed the effects of population heterogeneity on factor analysis under the homogeneity assumption and potential biases resulting from aggregation. Specific characteristics of the model will be examined by synthetic data. The factor mixture model will also be illustrated using empirical data.

Keywords: Factor mixture models, Unobserved heterogeneity, Synthetic data, Outliers detection, Clustering.

* Joint work with José G. Dias – ISCTE-IUL