# Modeling DHS data using dynamic mixture models

### José G. Dias

Lisbon University Institute, Portugal

#### Abstract

Finite mixture models have become very popular tools in modeling unobserved heterogeneity (*vide*, *e.g.*, McLachlan and Peel (2000), Dias and Willekens (2005)). This research extends the application of these models to sequential discrete data by incorporating misclassification error. The specified model is estimated by the Baum-Welch algorithm (Baum *et al.*, 1970) and the selection of the number of components in the mixture is based on the Bayesian information criterion of Schwarz (Schwarz, 1978).

Data in the illustration come from the 1996 Brazil Demographic and Health Survey (BDHS). This data set includes a Life History Calendar of the reproductive career of the women surveyed, which identifies monthly state occupancies (*e.g.*, *being pregnant*). We selected the Northeast region of the Brazil as it is the most determinant in the evolution of the Brazilian total fertility rate (TFR). The results show three subpopulations with different dynamics.

## Keywords

Finite mixture models, Latent class models, Markov chains, DHS – Demographic and Health Surveys.

## References

Baum, L.E., Petrie, T., Soules, G., and Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.* 41, 164–171.

Dias, J.G. and Willekens F. (2005). Model-based clustering of sequential data with an application to contraceptive use dynamics. *Math Popul. Stud.* 12(3), 135–157.

McLachlan, G.J. and Peel, D. (2000). *Finite Mixture Models*. New York: John Wiley & Sons.

Schwarz, G. (1978). Estimating the dimension of a model. Ann Statist. 6, 461–464.