

## Latent Markov models with covariates for longitudinal categorical data

Francesco Bartolucci

Department of Economics, Finance and Statistics

University of Perugia - Italy

email: bart@stat.unipg.it

The latent Markov model of Wiggins (1973) is a model for longitudinal data which relies on an unobservable Markov chain to model the dependence between the observations at the different time-occasions. In the case of categorical data, two different extensions for the inclusion of covariates are compared. In the first the covariates affect the initial and transition probabilities of the latent Markov process whereas, in the second, the covariates affect the conditional distribution of the observable variables given this process. The choice between the two approaches is driven by the context of application. The first extension is suitable when the latent process represents the evolution of a certain latent characteristic (e.g. quality of life) which is measured by the observable responses, typically corresponding to responses to a questionnaire. Then we want to evaluate how the evolution of this characteristic depends on the covariates. In the second extension, the latent process captures the unobserved heterogeneity between subjects so that we want to obtain estimates for the direct effect of the covariates on the response variables. With respect to traditional random effect models, the advantage is that the heterogeneity between subjects is modelled in a more flexible way. We show that, in both cases, an EM algorithm may be used to obtain parameter estimates which is based on recursions taken from the hidden Markov literature. We also discuss the problem of testing hypotheses of interest on the parameters through the likelihood ratio statistic.

The two approaches are illustrated by two applications. The first concerns the evaluation of the performance of the nursing homes operating in an Italian Region for what concerns the quality of life they ensure to elderly people. The second concerns the relation between fertility and female participation to the labour Market.

### Relevant papers on the topic

- Langeheine, R. (1994), Latent variables Markov models, in *Latent Variables Analysis: Applications for Development Research*, A. von Eye and C. C. Clogg, Sage (Eds.), Thousand Oaks, pp. 373-395.
- MacDonald, I. L. and Zucchini, W. (1997), *Hidden Markov and Other Models for Discrete-valued Time Series*, Chapman & Hall.
- Vermunt, J. K., Langeheine, R., and Böckenholt, U. (1999), Discrete-time discrete-state latent Markov models with time-constant and time-varying covariates, *Journal of Educational and Behavioral Statistics*, **24**, pp. 179-207.
- Wiggins, L. M. (1973), *Panel Analysis: Latent Probability Models for Attitude and Behavior Processes*, Elsevier.

### Papers of the author on the topic

- Bartolucci, F. (2006), Likelihood inference for a class of latent Markov models under linear hypotheses on the transition probabilities, *Journal of the Royal Statistical Society, series B*, **68**, pp. 155-178.
- Bartolucci, F. and Farcomeni, A. (2008), A multivariate extension of the dynamic logit model for longitudinal data based on a latent Markov heterogeneity structure, *under review*.
- Bartolucci, F. and Nigro, V. (2007), Maximum likelihood estimation of an extended latent Markov model for clustered binary panel data, *Computational Statistics and Data Analysis*, **51**, pp. 3470-3483.
- Bartolucci, F. and Pennoni, F. (2007), A class of latent Markov models for capture-recapture data allowing for time, heterogeneity and behavior effects, *Biometrics*, **63**, pp. 568-578.
- Bartolucci, F., Lupporelli, M. and Montanari, G. E. (2008), Latent Markov model for binary longitudinal data: an application to the performance evaluation of nursing homes, *under review*.
- Bartolucci, F., Pennoni, F. and Francis, B. (2007), A latent Markov model for detecting patterns of criminal activity, *Journal of the Royal Statistical Society, series A*, **170**, pp. 115-132.
- Bartolucci, F., Pennoni, F. and Lupporelli, M. (2008), Likelihood inference for the latent Markov Rasch model, in *Mathematical Methods for Survival Analysis, Reliability and Quality of Life*, C. Huber, N. Limnios, M. Mesbah and M. Nikulin (Eds.), Wiley, pp. 239-254.