**Pattern-Mixture Models for Categorical Outcomes with Non-Monotone Missingness**

Ivy Jansen [1] and Geert Molenberghs[1]

[1] Hasselt University, Center for Statistics, Belgium

When analyzing incomplete longitudinal data, several modelling frameworks can be considered, as there are selection models, pattern-mixture models and shared parameter models. In the past, selection models were very popular, but more recently, pattern-mixture models have gained considerable interest. Even several authors have contrasted selection models and pattern-mixture models, to either compare the answer to the same scientific question, such as marginal treatment effect or time evolution, as a form of sensitivity analysis, or to gain additional insight by supplementing the results from a selection model analysis with those from a pattern-mixture approach. Examples can be found in [6] for continuous outcomes, while categorical outcomes have been treated by [3]. However, they all restrict to the case of monotone missingness. Models to deal with non-monotone missing data ([5, 2, 1]) all belong to the selection model framework. Our focus will be on pattern-mixture models to analyze non-monotone missing categorical data. One of the major differences with work for continuous data, is that, unlike in the normal case, marginalization of pattern-specific effects is less than straightforward, owing to the presence of a non-linear link function in most models commonly considered. In particular, such marginalizations do not merely follow as simple weighted averages of the pattern-specific parameters. Attention is given to the derivation of a marginal covariate effect in pattern-mixture models for non-monotone categorical data.
Another important issue is that pattern-mixture models are by construction under-identified. This problem is addressed through the use of identifying restrictions: inestimable parameters of the incomplete patterns are set equal to (functions of) the parameters describing the distribution of the completers. In this way, the conditional distribution of the unobserved measurements, given the observed ones in a specific pattern, is identified. [4] proposed a particular set of restrictions for the monotone case which corresponds to MAR. We focus on the assumptions needed when intermittent missingness is present.
The techniques developed are used to analyze data from a clinical study in psychiatry.

## References

[1] Jansen, I. and Molenberghs, G. (2008) A flexible marginal modeling strategy for non-monotone missing data. *Journal of the Royal Statistical Society – Series A*, Accepted.

[2] Jansen, I., Molenberghs, G., Aerts, M., Thijs, H. and Van Steen, K. (2003) A local influence approach applied to binary data from a psychiatric study. *Biometrics*, **59**, 409–418.

[3] Michiels, B., Molenberghs, G. and Lipsitz, S. R. (1999) Selection models and pattern-mixture models for incomplete categorical data with covariates. *Biometrics*, **55**, 978–983.

[4] Molenberghs, G., Michiels, B., Kenward, M. G. and Diggle, P. J. (1998) Monotone missing data and pattern-mixture models. *Statistica Neerlandica*, **52**, 153–161.

[5] Troxel, A. B., Harrington, D. P. and Lipsitz, S. R. (1998) Analysis of longitudinal data with non-ignorable non-monotone missing values. *Applied Statistics*, **47**, 425–438.

[6] Verbeke, G., Lesaffre, E. and Spiessens, B. (2001) The practical use of different strategies to handle dropout in longitudinal studies. *Drug Information Journal*, **35**, 419–439.