Testing of Parametric Models fitted to High-dimensional Contingency Tables using Complex Survey Data

Irini Moustaki
London School of Economics and Political Science, London, United Kingdom
i.moustaki@lse.ac.uk

Chris Skinner*
London School of Economics and Political Science, London, United Kingdom
c.j.skinner@lse.ac.uk

Data on multiple categorical variables (items) are often collected in surveys in order to measure a smaller number of underlying dimensions. Latent variable models, analogous to those in factor analysis, may then be fitted to the resulting contingency table in order to capture these dimensions. Maximum likelihood type estimation methods can be computationally burdensome because of the need to compute multi-dimensional integrals and we shall consider composite likelihood methods which reduce this computation and can be adapted to allow for weights and other features of complex survey sampling schemes. We focus on the problem of testing the goodness-of-fit of the model or of testing associated nested hypotheses. Because the contingency table can be sparse when the number of items is large, even for large sample surveys, we consider methods which focus on lower-order margins of the table. Such methods are sometimes called limited information tests. We consider their extension to complex survey data, including the use of Rao-Scott tests.

Key Words: complex surveys; composite likelihood; latent variable models; survey weights.