Symposium on Recent Advances in Methods for the Analysis of Panel Data
15-16 June 2011
ISCTE-IUL, Lisbon University Institute, Portugal

Title

Graphical chain models and their value for panel data

Ronja Foraita and Iris Pigeot

Bremen Institute for Prevention Research and Social Medicine (BIPS), University Bremen

Abstract

Graphical models are probability models for multivariate observations to analyze and visualize conditional associations between random variables encoded by a conditional independence graph. In contrast to regression models, graphical modelling is concerned with identifying association structures for all study variables, including those which are usually regarded as explanatory. In particular, they allow differentiating between direct and indirect associations. They are therefore appropriate in situations where complex associations have to be dealt with. Due to the visualization in graphs, these models make it easier to display complex dependence structures. Furthermore, they can handle simultaneously categorical and continuous variables. Additionally to these characteristics, graphical chain models (GCM) are suitable to account for prior substantial knowledge of an underlying dependence structure by forming a dependence chain where all variables are subdivided into an ordered sequence of disjoint subsets.
The dynamic aspect of panel studies may be considered being reflected by GCM as the chain is designed by the natural temporal order of the examinations and life events.

Panel analyses aim to describe the influence of past and present covariate information on the development of risk factors over time and/or on primary events. For this purpose we need the conditional distribution of a status at time t given the development of covariates for an individual who has not experienced any events before time t. We will illustrate the GCM approach to analyze longitudinal data by means of a study on atopic diseases in infants.

However, as with any statistical analysis on the one hand the obtained results partly reflect the uncertainty being inherent in any type of data. On the other hand, the results heavily depend on the selected variables to be included in the analysis, the coding of these variables and the selection strategy used to fit graphical models to the data. Sensitivity analyses might be recommended to assess the stability of the obtained results.