LSE Department of Statistics

Recent advances in the analysis of categorical and count data

Thursday 12 December 2013
Vera Anstey Room (Old Building)

10.00 – 10.30	Coffee
10:30 – 11:15	**Ardo van den Hout** (UCL): Joint models for discrete longitudinal outcomes in ageing research
11.15 – 12.00	**Peter W. F. Smith** (University of Southampton): Statistical Modelling of Population Processes
12.00 – 13.00	Lunch – 5th floor Old Building, Director's Dining Room
13.00 – 13.45	**Maria Kateri** (RWTH Aachen University): Generalized odds ratio structures in ordinal response analysis
13.45 – 14.30	**Jeroen K. Vermunt** (Tilburg University): Multilevel Latent Markov Modeling in Continuous Time with an Application to the Analysis of Ambulatory Mood Assessment Data
14.30 – 14.45	Coffee

Abstracts

Ardo van den Hout — Joint models for discrete longitudinal outcomes in ageing research

Given the ageing population in the UK, statistical modelling of cognitive function in the older population is of interest. Joint models are formulated for survival and change in cognitive function in the older population. Because tests of cognitive function often result in discrete outcomes, binomial and beta-binomial mixed-effects regression models are applied to analyse longitudinal sum scores. Dropout due to death is accounted for by parametric survival models, where the choice of a Gompertz baseline hazard and the specification of the random-effects structure is of specific interest. The measurement model and the survival model are combined in a shared-parameter joint model. Estimation is by marginal likelihood. The methods are illustrated by analysing data from the Cambridge City over-75s Cohort Study and the English Longitudinal Study of Ageing.

Peter W. F. Smith — Statistical Modelling of Population Processes

Some recent work on the modelling of population processes using Bayesian methods is described. After the rationale for using Bayesian methods is discussed, two examples are presented. The first example explores the consequences of choosing different specifications of age-specific fertility and mortality in a closed cohort projection model in terms of its forecasted populations and measures of uncertainty. For illustration, a historical time series of fertility and mortality from England and Wales are used. The second example provides estimates of international migration flows amongst 31 countries in the European Union and European Free Trade Association from 2002 to 2008, based on data collected by individual countries with separate collection systems and designs. As a result, the reported data are inconsistent in availability, definition and quality. Covariate information and information provided by experts on the effects of undercount, measurement and accuracy of data collection systems are incorporated, and a synthetic data base with measures of uncertainty for the migration flows and other model parameters is produced.

Maria Kateri — Generalized odds ratio structures in ordinal response analysis

The role of generalized odds ratios and their structures in analyzing ordinal response variables is investigated, considering three different setups, starting with the basic problem of comparing two independent groups on an ordinal categorical scale. Most inferential methods addressing this problem are designed to detect a location effect, such as by focusing a single-degree-of-freedom test on an effect parameter. Often it is of interest to consider whether a stronger conclusion can be made, as for example, whether the population distributions are stochastically ordered. The Bayesian approach seems to be more straightforward for handling such problems. This is discussed in the context of stochastic ordering and other types of ordinal odds ratio structure, for the two-group comparison and for more general contexts. In the sequel, the role of odds ratio structures in analyzing correlated multinomial responses (ordinal or nominal) by the generalized estimating equations (GEE) approach is discussed. Finally, the classical quasi symmetry model (QS) for square contingency tables with commensurable classification variables is revisited and studied from the odds ratio structure point of view.

Jeroen K. Vermunt — Multilevel latent Markov modeling in continuous time, with an application to the analysis of ambulatory mood assessment data

Latent Markov modeling, also referred to as hidden Markov, Markov switching, regime switching, and latent transition modeling, is becoming a quite popular longitudinal data analysis tools in social science research. In this talk I will illustrate recent contributions to this field as well as ongoing research using a complex but very modern application. I will also show some developments in the Latent GOLD 5.0 software, which was released recently.

The application concerns the analysis is of a longitudinal data set in which there are multiple ordinal responses (mood items), measurement error, unobserved heterogeneity, a (longitudinal) nesting structure, and an unequal spacing of the measurements. The proposed continuous-time multilevel latent Markov modeling can deal with all these issues. I will also talk about parameter estimation and the assessment of model fit.