



Conference on Social Statistics Honouring the Scientific Contributions of Professor emeritus David Bartholomew: Clement House, LSE, 12 and 13 December 2011

Programme timetable

| Monday 12 December 2011 | | | |
|--------------------------|---------------|--|---|
| | 13:00 – 14:00 | Registration and coffee | |
| | 14:00 – 14:15 | Opening introduction | |
| | 14:15 – 15:00 | <i>Strategic Design for Generalization: Experiments and Surveys</i> | Colm O'Muircheartaigh |
| | 15:00 – 15:45 | <i>The development of the use of social statistics in the conduct of the Census</i> | Ian Diamond |
| | 15:45 – 16:15 | Coffee break | |
| | 16:15 – 17:00 | <i>Mixture latent autoregressive models for longitudinal data</i> | Francesco Bartolucci |
| | 17:00 – 17:30 | <i>Assessing students' achievement under survival analysis methods and Markov models with fuzzy states</i> | Aglaia G Kalamatianou |
| | 17:30 – 18:00 | <i>Educational measurement under latent variable models</i> | Shing On Leung |
| | 19:00 – 21:00 | Buffet reception | Senior Common Room, 5 th Floor, Old Building |
| Tuesday 13 December 2011 | | | |
| | 09:30 – 10:00 | Coffee | |
| | 10:00 – 10:45 | <i>Modelling household decisions with individual-level longitudinal data</i> | Fiona Steele |
| | 10:45 – 11:30 | <i>Social statistics and statistical physics: the chicken and the egg</i> | David Hand |
| | 11:30 - 12:15 | <i>Group means as explanatory variables in multilevel models</i> | Jouni Kuha |
| | 12:15 – 13:30 | Lunch | |
| | 13:30 – 14:15 | <i>Accuracy indicators and measurement error</i> | Chris Skinner |
| | 14:15 – 15:00 | <i>Francis Galton, measurement, psychometrics and social progress</i> | Harvey Goldstein |
| | 15:00 – 15:45 | <i>Godfrey H. Thomson: David Bartholomew's part in his reconstruction and resurgence</i> | Ian Deary |
| | 15:45 – 16:00 | Closing | |



Professor Francesco Bartolucci

University of Perugia

bart@stat.unipg.it

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Mixture latent autoregressive models for longitudinal data

Many relevant statistical and econometric models for the analysis of longitudinal data include a latent process to account for the unobserved heterogeneity between subjects in a dynamic fashion. Such a process may be continuous (typically an AR(1)) or discrete (typically a Markov chain). In this paper, we propose a model for longitudinal data which is based on a mixture of AR(1) processes with different means and correlation coefficients, but with equal variances. This model belongs to the class of models based on a continuous latent process and then it has a natural interpretation in many contexts of application. Moreover, it is more flexible than other models in this class, reaching a goodness-of-fit similar to that of a discrete latent process model, with a reduced number of parameters. We show how to perform maximum likelihood estimation of the proposed model by the joint use of an Expectation-Maximisation algorithm and a Newton-Raphson algorithm, implemented by means of recursions developed in the hidden Markov literature. We also introduce a simple method to obtain standard errors for the parameter estimates and suggest a strategy to choose the number of mixture components. The proposed approach is illustrated by an application to a longitudinal dataset, coming from the Health and Retirement Study, about self-evaluation of the health status by a sample of subjects. In this application, the response variable is ordinal and time-constant and time-varying individual covariates are available.



Professor Ian J Deary

University of Edinburgh

i.deary@ed.ac.uk

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Godfrey H. Thomson: David Bartholomew's part in his reconstruction and resurgence

David Bartholomew was a co-investigator with me and Martin Lawn in a recent ESRC-funded project entitled, 'Reconstructing the Scottish School of Educational Research, 1925-1950'. As the project progressed, it quickly became clear that the scientific star during that quarter century—during which Scotland punched well above its weight in Educational research—was Professor Sir Godfrey H. Thomson (1881-1955), who held the Bell Chair in Education at the University of Edinburgh from 1925 until 1951. The problem we found was that Thomson and his work have almost disappeared from scientific discourse since his death. In his working lifetime, Thomson contributed to: the theory of human intelligence differences; the statistical procedures used to analyse intelligence test data; constructing, validating and distributing intelligence tests on an industrial scale; the population-wide Scottish Mental Surveys of 1932 and 1947; testing the possibility of a dysgenic trend in UK intelligence; and other aspects of educational research and reform more generally. This presentation outlines Thomson's life and work. It describes the re-discovery of many of Thomson's materials. It describes his contributions, especially to intelligence testing, theory and analysis. In particular, the presentation describes the essential contributions that were made by David Bartholomew to re-evaluating Thomson's work on statistics and on the theory of how intelligence differences arise. Largely thanks to David Bartholomew's work, Thomson has resurged as a 'statistical pioneer' and the deviser of a theory of intelligence differences that now has 'a new lease of life'.



Professor Ian Diamond
University of Aberdeen
ian.diamond@abdn.ac.uk

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

The development of the use of social statistics in the conduct of the Census

Historically, censuses were seen as accurate counts of the population, and not subject to error. This paper will first argue that as censuses became more central to the allocation of finance then census accuracy became much more scrutinised. As a result methods needed to be developed to estimate census accuracy. The paper will then describe some of the methods used to provide estimates of census counts in the UK round of censuses and comment on the future of population counts in the UK.



Professor Harvey Goldstein
University of Bristol
h.goldstein@bristol.ac.uk

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Francis Galton, measurement, psychometrics and social progress

Modern cultures are deeply imbued with notions of measurement. Nearly all scientific disciplines depend heavily on mathematical measurement and arguably the social sciences have, relatively, seen the most rapid recent development of quantitative methodology and accompanying measurement regimes. Political debate also involves the use of measurements of all kinds and it is often assumed that the introduction of new measurements or the expansion of existing ones needs little justification. Sometimes, but rather rarely outside of the statistical and other scientific professions, discussion will revolve around the accuracy of measurements and whether they could be refined, but the default assumption is that more measurement is a good thing.

Yet measurement has not always been so popular, and 100 years after the death of Francis Galton (and also incidentally Alfred Binet) is not a bad time to reflect on how we got here, who helped us to get here and how much of it was worthwhile. In the space of a relatively short presentation, little more can be done than to provide a sketch and since it is the Galton centennial, his contribution will be dwelt on. Using Galton and his views as a starting point I shall attempt to describe some of the ways in which psychometric test theory and attitudes towards measurement have tended to develop over the last 100 years.



Professor David Hand
Imperial College London
d.j.hand@imperial.ac.uk

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Social statistics and statistical physics: the chicken and the egg

The natural and the social sciences are often presented as having contrasting properties, with each side being suspicious of the claims and merits of the other. However, the fact is that each has exerted a powerful influence on the way the other has developed, with ideas and developments in the one leading to major advances in the other. In this talk I track one strand of this historical development, showing how each side of the divide has been immeasurably enriched by developments on the other side.



Professor Aglaia G Kalamatianou
Panteion University, Athens
AglaiaKalamatianou@gmail.com

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Assessing students' achievement under survival analysis methods and Markov models with fuzzy states

This talk considers results of a research carried out to respond to issues arisen in the Greek university system and concern the students' likelihood and timescale of graduation. In general, the case emerges when studies are not assumed to be completed in certain duration of time i.e. there is a minimum prescribed time for graduation but there is no corresponding upper limit. Under these circumstances duration of studies data are lifetime data and have to be analysed in a framework of survival analysis. For handling the distribution of duration of studies, survival models are proposed. Fitting these models to real data reveals that students can graduate at the minimum prescribed time, can have a long period of graduation or may remain perpetual students. This reality gives rise to a further question about how such a distribution can be predicted at an early stage. For this purpose, the theory of non-homogeneous Markov systems with fuzzy states is used in order to describe students' academic progress through academic years which gives light to the previous results. The methodology developed could be useful to analyse similar data from other type or level of education in Greece and elsewhere.



Dr Jouni Kuha
LSE
j.kuha@lse.ac.uk

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Group means as explanatory variables in multilevel models

Research questions for models for clustered data often concern the effects of cluster-level averages of individual-level variables. For example, data from a social survey might characterise neighbourhoods in terms of average income, ethnic composition etc. of people within each neighbourhood. Unless the true values of such averages are known from some other source, they are typically estimated by within-cluster sample estimates, using data on the subjects in the observed data. This incurs a measurement error bias when these estimates are used as explanatory variables in subsequent modelling, even if the individual observations are measured without error. The measurement error variance can, however, be estimated from within-cluster variation, using knowledge of the sampling design within each cluster, and we can then apply measurement error modelling methods to adjust for the error.

This talk considers such estimation for generalised linear mixed models.

The methods are illustrated with models for political attitudes and behaviour, using data from the 2010 British Election Study.



Professor Shing-On Leung
University of Macau
soleung@umac.mo

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Educational measurement under latent variable models

Latent variable models provide a unified framework for dimension reduction, quantifying variables and achieving conditional independences, and are applicable in many areas in social sciences and education. With recent development in Rasch modelling and item response theories, educational measurement addresses many issues in teaching and learning. These range from measuring students' abilities, estimating item difficulties and discrimination power, investigating average abilities of those students choosing a particular category of a specific item, etc. Further, from an instructional point of view, it is important to identify those areas which students need help most in their learning process, and these areas are termed as zone of proximal development (ZPD). It turns out that all these tasks in educational measurement mentioned can be worked under the framework of latent variable models. This talk shall discuss details involved in applying latent variable models to various tasks in educational measurement. However, there are unresolved problems under *latent variable models* need further research, such as new reliability measures, items selection, confirmatory factor analysis, goodness-of-fit, person and item in-fit and out-fit statistics etc.



Professor Colm O'Muircheartaigh
University of Chicago
colm@uchicago.edu

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Strategic Design for Generalization: Experiments and Surveys

The purpose of social experiments is to make inferences about treatment effects in policy-relevant populations, but they seldom utilize probability samples of the populations of those populations. External information from a probability sample survey of the inference population can be used to obtain an estimate the mean causal effect of treatment in a sample that matches the composition of the population on a set of covariates measured in both the experimental sample and the survey sample. Conditions are given in which this procedure provides an estimate of the average causal effect of the treatment in the inference population.

This methodology permits the design of a program of research that will maximize the value of subsequent experiments in establishing effects for the whole population and its significant subpopulations.



Professor Chris Skinner
LSE
c.j.skinner@lse.ac.uk

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Accuracy indicators and measurement error

Survey measurement may be accompanied by the recording of information relevant to measurement accuracy. Examples include: interviewer assessments of respondent accuracy; respondent utterances of uncertainty about their responses; records of whether respondents refers to documents when answering factual questions. This talk will discuss the question of how to make use of such information to compensate for measurement error in survey data analysis. A particular challenge arises from the fact that an accuracy indicator may be directly related to the true value of a variable as well as measurement error.



Professor Fiona Steele
University of Bristol
fiona.steele@bristol.ac.uk

CONFERENCE
HONOURING THE WORK OF
PROFESSOR EMERITUS
DAVID J. BARTHOLOMEW

Modelling household decisions with individual-level longitudinal data

Many events occurring to married or cohabiting individuals are the result of decisions that are made jointly by both partners. However, studies of life course events usually take an individual or head of household perspective and so do not explicitly reflect the joint nature of these decisions. Household panel studies and population registers are a rich resource for studying household events, but the analysis of such data presents major analytical challenges. Allowing for the influence of both partners in a couple's decision is complicated by that fact that individuals can change partner and can have periods when they are not in a co-residential union. This talk presents a general non-hierarchical multilevel model in which the probability of an event is specified as a function of individual and partner covariates and a weighted combination of random effects for each decision-maker in the household. This approach avoids the need to choose individuals or households as the unit of analysis, while allowing for changes in household composition over time. The proposed method is compared with previously used approaches via a simulation study, and illustrated in an analysis of residential mobility using data from the British Household Panel Study.

Conference website:

<http://www2.lse.ac.uk/statistics/events/SpecialEventsandConferences/2011conference.aspx>

Enquiries:

Statistics.events@lse.ac.uk