



Conference 2012

International Centre, Telford, UK

3-6 September 2012

Abstracts

Abstracts are ordered in session order for oral presentations followed by poster presentations

Index (authors alphabetical by first name)

A

<u>Adam Brentnall</u>	56
<u>Adam Butler</u>	97
<u>Adam M. Sykulski</u>	108
<u>Adrian Roberts</u>	87
<u>Alastair Young</u>	70
<u>Alberto Arcagni</u>	41
<u>Aleks Collingwood</u>	25
<u>Alex Hayman</u>	179
<u>Alexander Donev</u>	147
<u>Alexandra Hayman</u>	178
<u>Alexandre Santos</u>	90
<u>Alfred Akinsete</u>	195
<u>Alison Park</u>	12
<u>Amy Wilson</u>	141
<u>Andrew Epps</u>	44
<u>Andrew Oliver</u>	49
<u>Andrew Rouse</u>	37
<u>Andrew Smith</u>	136
<u>Andrew Titman</u>	46
<u>Ann Blake</u>	10, 120
<u>Ann Langford</u>	165
<u>Anthony Lee</u>	129
<u>Arief Gusnanto</u>	116

B

<u>B.N. NAVEEN KUMAR</u>	214
<u>Benson Ade Afere</u>	206
<u>Bothinah Altaf</u>	208
<u>Brendan Murphy</u>	65

C

<u>C. Chandramouli</u>	8
<u>Catrin Tudur Smith</u>	84
<u>Celia Chambers</u>	180
<u>Charlotte Greenan</u>	207
<u>Chris Wild</u>	19
<u>Colin Aitken</u>	93, 139

D

<u>D.R.Cox</u>	24
<u>Dan Jackson</u>	104
<u>Daniel Ayoubkhani</u>	122
<u>David Forrest</u>	113
<u>David Hand</u>	161
<u>David J. Hand</u>	24
<u>David Randell</u>	110
<u>David Walker</u>	25, 53

<u>Diana Tlupova</u>	31
<u>Dieter William Joenssen</u>	77
<u>Duncan Lee</u>	76

E

<u>Elisa Loza</u>	96
<u>Emma White</u>	78
<u>Evangelos Kontopantelis</u>	91

F

<u>Fanghong Zhang</u>	40
<u>Fatima Jichi</u>	7
<u>Fentaw Abegaz</u>	117
<u>Fiona Steele</u>	60
<u>Fiona Underwood</u>	33, 88
<u>Fola Ariyibi</u>	162

G

<u>Gabriela Czanner</u>	190
<u>Garnett Compton</u>	79
<u>Gavin Shaddick</u>	135
<u>Geoff Jones</u>	156
<u>George Ploubidis</u>	59
<u>Gerard Kuper</u>	127
<u>Gerton Lunter</u>	34
<u>Giancarlo Manzi</u>	177
<u>Gillian Lancaster</u>	29
<u>Golnaz Shahtahmassebi</u>	197
<u>Gordon Ross</u>	109
<u>Guy Rubin</u>	166

H

<u>Hal R. Varian</u>	111
<u>Heather Cordell</u>	36
<u>Heather Turner</u>	105
<u>Hossein Riazoshams</u>	200
<u>Howard Litwin</u>	100

I

<u>Ian Hunt</u>	20, 140
<u>Ian Nevison</u>	199
<u>Ian Plewis</u>	43
<u>Ian White</u>	55
<u>Idris Eckley</u>	114
<u>Ikhlaaq Ahmed</u>	155
<u>Ilaria Lucrezia Amerise</u>	198
<u>Irene Prete</u>	102

<u>Ivana Horova</u>	174
J	
<u>Jack Cuzick</u>	106
<u>Jacqueline Civil</u>	159
<u>Jacqueline Potts</u>	153
<u>Jake Ansell</u>	57
James Carpenter.....	95
James Grime	146
James Matcham.....	16
James Tucker	62
<u>Jan Kolacek</u>	183
<u>Jay Magidson</u>	18
<u>Jaya Satagopan</u>	196
<u>Jennifer Rogers</u>	67
Jenny Church	81
<u>Jenny Freeman</u>	17
<u>Jenny Wadsworth</u>	137
<u>Jens Mehrhoff</u>	22
Jil Matheson.....	50
<u>John Einmahl</u>	112
<u>Jonathan Blackledge</u>	71
<u>Jonathan Cook</u>	54
<u>Jonathan Marchini</u>	13
<u>Jose Pina</u>	21
<u>Joseph Frend</u>	138
<u>Jouni Kuha</u>	176
<u>Julia Brettschneider</u>	15
K	
<u>Karim Anaya-Izquierdo</u>	144
Kate Pickett.....	12
<u>Kishore Das</u>	212
L	
<u>Leonardo Trujillo</u>	203
<u>Lisha Wang</u>	184
<u>Louisa Blackwell</u>	121
<u>Louise Lloyd</u>	205
<u>Ludi Simpson</u>	52
M	
<u>M VENKATESWARAN</u>	216
<u>Mamusu Kamanda</u>	74
<u>Man Ying Edith Cheng</u>	171
<u>Margaret Eames</u>	42
<u>Margaret Hurley</u>	142
<u>Maria Vazquez-Montes</u>	143
<u>Marina Knight</u>	115
Markus Gesmann.....	66

<u>Marta Blangiardo</u>	134
<u>Martin Forster</u>	5
<u>Martina Aumeyr</u>	175
<u>Martine Durand</u>	11
<u>Matieyendou LAMBONI</u>	168
<u>Matthew Greenaway</u>	125
<u>Matthew Sperrin</u>	160
<u>Michael Imande</u>	182
<u>Michael Stumpf</u>	131
<u>Michail Papathomas</u>	38
<u>Muhannad Al-Saadony</u>	192
N	
<u>Nadine Chlass</u>	39
Naomi Givens	25
<u>Neil Spencer</u>	126
<u>Nick Bingham</u>	86
<u>Nick O'Donnell</u>	80
<u>Nicky Wright</u>	188
<u>Noorazrin Abdul Rajak</u>	181
<u>Nur Anisah Mohamed</u>	185
O	
<u>O. M. OLAYIWOLA</u>	172, 173
<u>OlaOluwa Yaya</u>	218
<u>Oluwafemi Oyamakin</u>	217
<u>Oluwole Nuga</u>	189
Omar Paccagnella	101
<u>Orlaith Burke</u>	69
P	
<u>Paola Zaninotto</u>	98
<u>Patrick Wolfe</u>	63
<u>Paul Askew</u>	123
<u>Paul Buckley</u>	209, 210
<u>Paul Clarke</u>	61
<u>Paul Fryers</u>	157
<u>Paul Johnson</u>	92
<u>Paul Kirk</u>	130
<u>Paul Marchant</u>	75
<u>Paul Noroski</u>	99
<u>Paul Northrop</u>	151
Paula Williamson.....	82
Penny Young	25
<u>Peter Goos</u>	148
<u>Peter Tankov</u>	85
<u>Peter Wilmshurst</u>	150
<u>Polycarp Chigbu</u>	202

R

<u>Rajen Shah</u>	68
<u>Rebecca Gillard</u>	164
<u>Rebecca Newell</u>	9
<u>Rebecca Pillinger</u>	119
<u>Rebecca Turner</u>	27
<u>Reem Al-Jarallah</u>	193
<u>Richard Alldritt</u>	51, 62
<u>Richard Hooper</u>	30
<u>Richard Parker</u>	47
<u>Riza Momin</u>	167
<u>Robert Aykroyd</u>	89
<u>Robert Dunn</u>	58
<u>Robert Grant</u>	194
<u>Robert M. Kunst</u>	204
<u>Robert Matthews</u>	6
<u>Roger Smith</u>	170
<u>Ronaldo Silva</u>	187
<u>Rosemary Bailey</u>	24
<u>Rosie O'Neil</u>	169

S

<u>Sabrina Juran</u>	8
<u>Sada Nand Dwivedi</u>	211
<u>Samuel Manda</u>	191
<u>Sandra De Francisci</u>	201
<u>Shah-Jalal Sarker</u>	186
<u>Shereen Hussein</u>	45
<u>Simon Tavaré</u>	14
<u>Simon White</u>	154
<u>Simon Wilson</u>	132
<u>Sir David Cox</u>	94
<u>Stephen D Walter</u>	28
<u>Stephen Senn</u>	103

<u>Steven Gilmour</u>	149
<u>Stuart Hillmansen</u>	37
<u>Sunil Mitra Kumar</u>	32
<u>Svetlana Tishkovskaya</u>	158

T

<u>Tatiana Xifara</u>	163
<u>Thomas Jaki</u>	83
<u>Thomas Laitila</u>	23
<u>Tim Paulden</u>	26
<u>Timothy Olabisi OLATAYO</u>	215
<u>Tom Fanshawe</u>	133
<u>Tom Sniijders</u>	64
<u>Tony Cox</u>	62
<u>Tony Lawrance</u>	124

U

<u>Ulrike Naumann</u>	35
-----------------------------	----

V

<u>Veronica Vinciotti</u>	72
<u>Victoria Harris</u>	152
<u>Vinayak Deshpande</u>	213

W

<u>Wendi Qian</u>	118
<u>William Browne</u>	48
<u>Wolfgang Maennig</u>	128

Y

<u>Yemisi Takwoingi</u>	107
<u>Yining Chen</u>	73

Z

<u>Zihua Yang</u>	106, 145
-------------------------	----------

Contributed 1 Tuesday 4 September, 9.20am - 10.20am

Bayesian sequential sampling for the economic evaluation of health technologies: opportunity or dead end?

Martin Forster¹, Paolo Pertile², Davide La Torre³

¹*University of York, York, UK*, ²*University of Verona, Verona, Italy*, ³*University of Milan, Milan, Italy*

Objectives

A Bayesian sequential sampling approach to the economic evaluation of health technologies allows a researcher to update a prior distribution for cost-effectiveness step-by-step, as new sample information arrives. In so doing, it offers the potential to improve how health care systems and regulators allocate funding for research into, and the provision of, health technologies. However, the demanding assumptions of such models have been criticised. Little research has been published which shows what an application of such a model might look like.

Method/Models

We present a Bayesian sequential sampling model applied to economic evaluation in which there exists flexibility over the timing of a decision to stop researching and recommend adoption of a new health technology or continued use of an existing one and which accounts for the costs and benefits of both research and adoption. We apply the model using in a case study, deriving dynamic thresholds defining optimal policies as a function of sample size, and comparing the results with decision rules from existing statistical approaches.

Results and Conclusions

We consider the insights that such models can offer, as well as the obstacles that might stand in the way of their use in practice. A full description of the model and its results, together with a wider discussion of the applicability of dynamic statistical methods to the economic evaluation of health technologies, is available in two discussion papers which presented preliminary results:

<http://ideas.repec.org/p/yor/yorken/10-24.html>

<http://ideas.repec.org/p/ver/wpaper/03-2011.html>

Using Bayes to assess the plausibility of new epidemiological findings

Robert Matthews

Aston University, Birmingham, UK

Objectives

There is mounting concern over the failure of results from observational studies to be replicated in subsequent studies. Many reasons for such failure have been put forward, such as inadequate power and publication bias. One issue that deserves closer scrutiny is the plausibility of the claim being made. Standard statistical measures such as p-values offer no obvious means of taking plausibility into account. Bayesian methods, in contrast, allow plausibility to be assessed quantitatively and transparently, and can show when "statistically significant" results should be treated with caution.

Method/Models

The assessment of plausibility requires that a new finding be set in the context of extant evidence. Bayesian methods are naturally suited to such a task, but require that we confront the well-known problem of establishing a suitable prior. The presentation will show how the data itself can be used to overcome this problem, via the concept of the "Credible Prior Value" (CPV). Its value is set by the condition that extant evidence for effect-sizes exceeding the CPV are required if the evidence is to be credible at the 95% level.

Results and Conclusions

When applied to real-life cases - which will be presented during the talk - the CPV method helps identify apparently convincing observational studies whose claims are unlikely to be replicated, despite being statistically significant.

The CPV method can also make sense of "out of the blue" findings, for which no extant prior knowledge exists. Currently, such studies stand or fall according to whether they are statistically significant. The CPV method goes further, allowing the evidential weight of such studies to be assessed. Specifically, if their central value does not exceed their calculated CPV, they lack the evidential weight to make their case. As such, they should be treated with caution. Again, real-life examples will be presented.

Growth Mixture Modeling of Child Behaviour in a Study of Children Receiving Multidimensional Treatment Foster Care in England.

Fatima Jichi¹, Sabine Landau¹, Stephen Scott², Alexandra Wretham²

¹King's College London, Institute of Psychiatry, Department of Biostatistics, London, UK, ²King's College London, Institute of Psychiatry, Department of Child Adolescent Psychiatry, London, UK

Objectives

The study aims to evaluate the response of children to a new treatment, Multidimensional Treatment Foster Care in England (MTFCE). The treatment is administered to children who have been previously abused or have received poor care from their parents. The Parent-Daily-Report (PDR) is a 37-part questionnaire completed by carers and is the outcome measure of the study. The questionnaire asks if any of 37 behaviours were present with the child, and of these, which were stressful to the carer. The objective of the analysis is to study trajectories of children's behaviours over time, and to find subgroups of treatment response.

Method/Models

Growth Mixture Modeling (GMM) was used to find subgroups in the data. A GMM describes longitudinal measures of a single outcome measure as being driven by a set of subject-varying continuous unobserved or latent variables - the so-called *growth factors*. The growth factors define the individual trajectories. GMM estimates mean growth curves for each class, and individual variation around these growth curves. This allows us to find clusters in the data. Starting characteristics of children were included into the GMM to see if these predicted class membership. Class membership was also checked to see if it predicted outcomes of interest.

Results and Conclusions

Several GMMs were run on the data to find the model with the best fit. Models with linear, quadratic and cubic growth factors were fit, as well as models with 1-4 classes. Based on several pre-specified criteria, the two-class model with a linear growth factor was chosen as the model which fits the data best. The analysis showed two distinct classes. The classes were interpreted as a responder group (class 1) and a non-responder group (class 2). Analysis to find starting characteristics which are predictive of class membership, and if class membership predicts outcomes of interest are underway.

Contributed 2 Tuesday 4 September, 9.20am - 10.20am

Counting the 7 Billion: The Future of Population and Housing Censuses

C. Chandramouli¹, Kris Oswalt², Sabrina Juran³

¹Office of the Registrar General & Census Commissioner, Delhi, India, ²DevInfo Initiative, New York, USA, ³United Nations Population Fund (UNFPA), New York, USA

Objectives

To create innovative and robust methods and procedures for conducting population and household censuses; and to harness information technology to disseminate census data for use by everyone. The 15th Indian National Census collected household and population data on 1.2 billion Indians by mobilizing modern technology to produce timely and accurate results.

Method/Models

The Registrar General's Office of India established cutting-edge technology to capture and process census data. The data were collected on paper forms that were scanned using high-speed image-based recognition scanners. The processes also harnessed GIS to digitize the boundaries of more than 600,000 villages, 8000 towns and 33 metropolitan areas. Training was conducted by 90 national trainers training 725 master training facilitators, who further trained 54,000 master trainers. The master trainers then trained 2.7 million enumerators and supervisors. ICT was used to enhance training methodology.

Results and Conclusions

Data were collected in 16 languages from 240 million households in February 2011. Within 100 days of the completion of the census, the preliminary results were released to the general public. The results were disseminated using an adaptation of the United Nation's CensusInfo database technology. The methods of data collection and tabulation ensured complete and unduplicated coverage, quality of data validation, reduction of overall cost of operations and efficient use of time from the date of collection to the date of data release. CensusInfo India is an effective open-source software platform for disseminating census data because of its robustness, ease of use, and versatility. The success of CensusInfo India, which manages the data collected in the 2nd largest census in the world, suggests that it would be an effective resource to manage and disseminate data for countries of all sizes and geographies.

Measuring Uncertainty in Local Authority Population Estimates for England and Wales

Rebecca Newell¹

¹Office for National Statistics, Titchfield, UK, ²University of Southampton (S3Ri), Southampton, UK

Objectives

The scope of the Improving Migration Statistics programme included recognition of a need to provide users with more detailed information regarding the uncertainty associated with local authority mid-year population estimates. In order to meet this need, the Office for National Statistics established the uncertainty project. One of the objectives for this project was to develop a methodology that would provide a statistical measure of uncertainty associated with the mid-year population estimates for local authorities in England and Wales. The development of this methodology is the focus of this presentation.

Method/Models

Statistical measures of uncertainty around local authority mid-year population estimates (using the cohort component method) are not straightforward to produce, as multiple data sources and a variety of estimation procedures are used within the process. Initial work identified the three components with the greatest impact on uncertainty; internal migration, international migration and the 2001 Census base. Simulation methodologies were developed to produce statistical measures of uncertainty around each identified component. Resultant simulated error distributions for each component are then combined to create an overall measure of uncertainty (such as the variance and any potential biases) for each local authority.

Results and Conclusions

Preliminary findings from the research are presented and discussed. This presentation also highlights some of the issues and challenges encountered during this innovative work in the area of quality measurement.

A conceptual framework for UK population and migration statistics

James Raymer², Phil Rees³, Ann Blake¹, Peter Boden⁴, James Brown², George Disney², Nikolas Lomax³, Paul Norman³, John Stillwell³

¹Office for National Statistics, Titchfield, UK, ²Southampton Statistical Sciences Research Institute (S3RI), University of Southampton, Southampton, UK, ³University of Leeds, Leeds, UK, ⁴Edge Analytics, Leeds, UK

Objectives

To develop a conceptual framework for UK population and migration statistics that that will facilitate communication with users of population and migration statistics through the development of a shared understanding of the underlying concepts, the available data and the methods used to produce key outputs.

Method/Models

High level review of current methods and approaches used to estimate population and migration statistics in the UK.

Results and Conclusions

Annual population and migration statistics are produced in the UK to meet a diverse range of requirements. They provide information on how populations are changing over time, which is influenced by wider social, cultural and natural environments. The data may come from censuses, administrative registers and surveys. Estimation is used to combine sources of information, infer missing patterns, or align the source to the particular migration or population concept required. To bring together the data, concepts, processing and outputs, a conceptual framework is required. In this presentation, we present a framework for UK population and migration statistics that has been developed as a collaborative project between Southampton Statistical Sciences Research Institute (S3RI), the University of Leeds and the Office for National Statistics. The framework provides a structured way to translate requirements to a coherent set of outputs. It is also useful as a basis for determining where to focus future improvements.

How's Life?

Martine Durand

OECD, Paris, France

Building on its long tradition of work, the OECD has recently taken important steps to provide direction to the international reflection on measuring well-being and progress. On the occasion of the OECD's 50th Anniversary, held under the theme "Better Policies for Better Lives", the Organization launched the OECD Better Life Initiative. The report *How's Life?*, which is part of this initiative, is a first attempt at the international level to go beyond the conceptual stage and to present a comprehensive set of comparable well-being indicators for OECD countries and, to the extent possible, other major economies. This set is based on a framework that distinguishes between current well-being and its sustainability, and which articulates the former through two main domains (material well-being, on one side; and quality of life, on the other) and 11 dimensions of life that have a claim to be relevant for all countries around the world. that This report aims to respond to the needs of citizens and policy-makers for better information on people's well-being and societal progress.

How's Life? highlights that average well-being has increased significantly in most countries over the past fifteen years: but also that there are large differences across countries and groups of the population, particularly less educated and low-income people. Beyond describing well-being patterns across countries and population groups, the *How's Life?* publication confronts the challenge of how to summarise information from a range of indicators into a synthetic index reflecting users' own views on the relative importance of different dimensions of life, and describes some of the priorities for future statistical work in this field.

Kate Pickett (University of York)

Understanding the effects of inequality in the UK: data from the Spirit Level

This paper will present a comparative analysis of the impact of income inequality on a wide range of health and social outcomes, and develop a theoretical explanation of a psychosocial pathway underpinning population-level health and wellbeing.

Alison Park (Head of Society and Social Change, NatCen)

What the British Social Attitudes Survey tells us about our attitudes to inequality

This talk will use data from the British Social Attitudes survey to explore public attitudes towards inequality and the government's role in alleviating it, focusing particularly on people's attitudes to the welfare state. It will explore how and why attitudes in this area have changed since the early 1980s, as well as how views vary between different sections of the population.

Linear complexity surrogate family haplotype estimation of whole chromosomes

Jonathan Marchini¹, Olivier Delaneau¹, Jean-Francois Zagury⁰

¹University of Oxford, Oxford, UK, ²Conservatoire National des Arts et Métiers, Paris, France

Objectives

Haplotype estimation is key step in the statistical analysis of many disease and population genetic studies of humans and other organisms. The literature on methods for haplotype estimation is large. Since datasets are becoming larger in terms of the number of samples and also the density of SNP markers new methods are needed that maintain high accuracy and are computationally feasible for routine use.

Method/Models

I will describe our recent work on a new MCMC that updates samples iteratively, conditional upon a carefully chosen subset of current haplotype estimates from other samples. We call this the surrogate family phasing approach as the conditioning haplotypes can be considered to be the ones most closely related in a genealogical sense to the individual being updated. We have combined this approach with the SHAPEIT1 method, which has the property that it scales linearly with both sample size, number of conditioning haplotypes and number of markers.

Results and Conclusions

To assess the performance of this new approach, called SHAPEIT2, we have used eight different whole chromosome datasets from low, medium and high density SNP chips, as well as genotype data derived from whole genome sequencing. For example, we analyze 1000 Genomes Project data from the Illumina Omni2.5M chip, as well as a dataset derived from the low-coverage sequencing carried out by the project. We also carry out an assessment of how different methods scale with much larger sample sizes that are becoming more routine, using data from the Wellcome Trust Case-Control Consortium with up to 9000 study samples. Results on all these datasets demonstrate that SHAPEIT2 produces substantial gains in phasing accuracy and computational speed when compared to other approaches, including our previous approach SHAPEIT1. We find that SHAPEIT2 reduces the switch error rate of SHAPEIT1 by as much as 45%, and by much more compared to other methods.

Assessing tumour heterogeneity: inference using the matrix-variate Normal distribution

Simon Tavaré¹, Anestis Touloumis^{1,2}

¹University of Cambridge, Cambridge, UK, ²European Bioinformatics Institute, Cambridge, UK

The nature of molecular variability within a tumour is of considerable interest to cancer researchers, as tumour heterogeneity has important implications for prognosis, treatment and prevention of cancer. In this talk we consider the problem of modeling gene expression levels in microarray experiments when measurements are made on multiple subsamples taken from the same tumour. The primary biological goal is to assess the heterogeneity of expression levels within each individual tumour and, subsequently, between a sample of individuals. From a statistical perspective our primary concern is the accurate estimation of the association structure of the correlated samples and of the genes. With high-dimensional data, naively calculating Pearson's or Spearman correlation coefficients might be problematic since they tend to overestimate the correlations. Instead, we utilize the matrix-variate normal distribution, which allows us to estimate separately the covariance/correlation matrix among the correlated subsamples and the genes. We derive estimators for the covariance matrices of the genes and the subsamples and we discuss how to test several biological hypotheses. Finally, we illustrate the methods by analyzing a dataset from a study of glioblastoma.

Subset based assessment of high-throughput gene expression data quality in time course experiments

Julia Brettschneider, G. Cohen Freue
University of Warwick, Coventry, UK

High throughput gene expression measurement is the basis for quantitative research in functional genomics. Microarrays have become standard laboratory equipment for conducting such assays in a genome-wide scale. For experiments that monitor a biological process over the course of time they are a cost efficient alternative to high-throughput sequencing. Such experiments are commonly used to understand the genetic basis of developmental (e.g. embryo growth) and cyclic (e.g. circadian rhythms) processes. The key to reproducible findings are accurate and precise measurements. Data quality assessment is specific to the platform. For the current (informal) industrial standard, short oligonucleotide arrays, a number of measures have been suggested and explored in a variety of situations [1], [2].

Among those, robust summaries (Median and IQR) of the the NUSE and RLE distributions have turned out to be particularly useful.

We aim to shed more light onto the appropriate use and the interpretation of these measures in the context of time course experiments. This includes a second look at the assumptions underlying the use of NUSE and RLE for quality assessment, the relationships between the different measures and their suitability for a range of biological experiments. Furthermore, as RLE can be calculated regardless which of the available preprocessing methods was used, we study the impact of that choice.

Quality assessment based on NUSE and RLE requires that a limited amount of genes are biologically active. Otherwise, biological activity may be mistaken for measurement error or technical artefacts. However, we found that the quality measures differ in their degree of sensitivity to increased gene activity. We propose an alternative method that restricts the data quality assessment to an appropriate subset of genes avoiding such effects. Regarding the impact of different preprocessing methods, IQR RLE is only mildly affected by that choice, whereas Median RLE is more. This is based on RMA with a range of settings for normalisation and background adjustment, Li/Wong and GCOS.

There are linear relationships between one-parameter summaries of RLE and NUSE, especially between IQR RLE and MedianNUSE, but only loose connections between these and the GCOS scores.

The data used in this study include different common types of microarray time course experiments, growth processes (embryo development) and cyclic processes (circadian clock), as well as non time course experiments comparing different mutants. The data sets are from fruit fly and arabidopsis. Perturbation studies and simulation studies complement the analysis of experimental data. Calculations were done in R using Bioconductor packages.

The results show that, quality assessment for microarray data can be tuned to appropriately address the special situation of time course experiments. We have developed experiment specific methods that allow to tailor high throughput quality assessment methods to such situations by restricting the quality assessment to appropriate subsets of genes. This method captures relevant quality patterns without being misguided by biological variation. Our method is very easy to implement for short oligonucleotide microarrays and can also be transferred other platforms.

Proposed Best Practice for Statisticians in the Reporting and Publication of Pharmaceutical Industry-Sponsored Clinical Trials

James Matcham (Amgen)

PSI offers the opportunity to experience something of the atmosphere of an RSS Ordinary meeting as we invite you to join us for a read paper session. James Matcham (Amgen) will present the paper "Proposed Best Practice for Statisticians in the Reporting and Publication of Pharmaceutical Industry-Sponsored Clinical Trials." (2011 *Pharmaceutical Statistics* 10:1;70-73). After the presentation, we will as is traditional have a proposer and seconded of a vote of thanks. There will then be the opportunity for attendees to make comment on the paper from the floor, or if they wish they can submit them in advance to the session chair, before James is given the opportunity to respond.

Copies of the article can be downloaded from the Pharmaceutical Statistics website at [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)1539-1612](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1539-1612)

Organised by Statisticians in the Pharmaceutical Industry (PSI)

Presenting at conferences

Tuesday 4 September, 12 noon

Jenny Freeman (University of Sheffield)

Improving your presentations: all you need to know in 80 minutes!

Have you ever felt nervous before a presentation; lost concentration during a presentation; wondered afterwards if you could have done better? If the answer to any or all of these then this workshop is for you. In less than an hour and a half we will look at what makes for an effective, memorable presentation and help you develop some strategies for improving what you do.

Regression Modelling with Many Correlated Predictors: A new approach to linear and logistic regression with high dimensional data

Jay Magidson^{1,2}, Gary Bennett^{4,1}

¹Statistical Innovations Inc, Belmont, MA, USA, ²Addinsoft, Paris, France, ³Tufts and Boston University, Boston, USA, ⁴Logit Research Limited, Kent, UK

Objectives

Recent advances in analysis of high dimensional data now allow reliable regression models to be developed even when the number of predictors exceeds the number of cases! With reference to real-life case studies from the marketing research industry and simulation studies we present a new approach - Correlated Component Regression - which addresses the limitations of traditional regression and outperforms penalised regression and other sparse data methods. The paper demonstrates an approach for screening out irrelevant predictors and recovering the population model even in very sparse conditions and in the presence of suppressor variables. Comparisons are drawn with alternative methods.

Method/Models

The main focus is on Correlated Components Regression (CCR), an exploratory predictive modelling methodology having four important components: (a) a unique approach to regularising model coefficients making use of correlated composite components that are decomposable into the original predictor terms ; (b) Model evaluation and selection based entirely on how well models predict out-of-sample using M-fold cross-validation (CV); (c) A stepping down algorithm for developing parsimonious models with optimal CV performance; (d) Ability to capture important suppressor variables. Comparisons in approach and performance are made with Ridge regression, LASSO, Principal Components Regression, PLS Regression and the R GLMNET package.

Results and Conclusions

High dimensional data pose unique problems when evaluating models purely on an in-sample basis, with over-fitting, perfect separation (Logistic regression) and spuriously high performance rife. Normal statistical inference is unfit for purpose.

CCR overcomes these issues by allowing the analyst to compare the performance of models with varying degrees of regularisation "out-of-sample".

The CCR step-down algorithm can help improve prediction and interpretation when extraneous variables are included among the candidate predictors.

Population parameters are recovered to a much greater degree of accuracy than with penalised methods (lasso, ridge regression) and other sparse methods.

When suppressor variables exist in data, they should be included in predictive models because they can improve prediction substantially. CCR has higher power for capturing effects of suppressor variables than stepwise regression or penalized regression.

CCR also performs well with larger samples. It is an important new method which is potentially generalizable to any predictive modelling family.

Semi-parametric methods for regression under two-phase sampling

Chris Wild, Alastair Scott, Gustavo Amorim
University of Auckland, Auckland, New Zealand

Objectives

This paper extends the work in Jiang, Scott & Wild (Int. Stat. Rev. 2011), Scott & Wild (Can. J. Stat. 2011) on fitting regression models with response-biased two-phase samples, that is, samples where some or all the covariates are missing for some units and the probability that this happens depends in part on the value of the response of that unit.

Method/Models

We look at a variety of methods based on estimating equations, at the relationship of these methods to semi-parametric efficient methods in cases where such methods exist.

The talk will concentrate on fitting linear models where the continuous response and some continuous covariates are available for all individuals but other covariates are observed only for subsets. Time permitting we will make connections with calibration and propensity scores.

Results and Conclusions

We show ways of obtaining efficiency gains that can sometimes be dramatic

Fisherian Testing for the 21st Century: Getting the Logic Straight and Multiple Testing

Ian Hunt

London School of Economics, London, UK

Objectives

Karl Popper's method of hypothesis conjecture and refutation (falsification) was a philosophical and practical failure. He couldn't show how falsification worked properly. RA Fisher promised falsification via statistical inference. I argue Fisher's falsification doesn't work properly either: Fisher's principal inferential output of "the hypothesis is rejected" is invalid; and his other inference, "either the hypothesis is false or the evidence is unlikely", is of little use in isolation. But a meaningful measure of evidential weight can be salvaged from Fisherian methods. This requires combining the logical implications from tests of multiple rival hypotheses. Such multiple testing seems Popperian in spirit.

Method/Models

(1.) Fisherian "rejection" is invalid in terms of falsification and useless otherwise. First, strict falsification is simply not what Fisher's tests can do - results are only probabilistic. Secondly, a reliable guide to belief about falsity is not provided - this would, at the least, require a low $p(\text{hypothesis/evidence})$, but Fisherian rejection only implies a low $p(\text{evidence/hypothesis})$. Anyway, there are logical objections to inducing categorical beliefs from probabilistic beliefs. Thirdly, a behavioural cue to set-aside/ignore/remove-from-consideration any single hypothesis is not implied by rejection. Fisher himself ruled this out because such cues only relate to a series of hypotheses in-the-long-run.

Results and Conclusions

(2.) Fisher's second inference "either the hypothesis is false or the evidence is unlikely" is logically valid (when p -values are extreme). But this disjunction is not useful on its own. It is no good for decision processes which require $p(\text{hypothesis/evidence})$ because only $p(\text{evidence/hypothesis})$ can be implied. And it does not reliably measure the evidential weight against a hypothesis because rival hypotheses are ignored.

(3.) A good measure of evidential weight can be salvaged from Fisherian methods. This entails testing multiple rival hypotheses and combining the logical implications of each result. Corollaries: (a.) Fisherian multiple testing is Popperian in spirit (hypotheses should be conjectured critically, rationally and creatively, and tested vigorously and openly); (b.) Fisherian testing should be more useful today than ever before because of an abundance of hypotheses and data with which to test them; (c.) Fisherian test outputs are in fact relevant to decision and belief processes (Bayesian or otherwise).

Measurement Error in Retrospective Reports of Unemployment

Jose Pina

University of Manchester, Manchester, Lancashire, UK

Objectives

To assess the prevalence of measurement error in the retrospective report of unemployment and to identify which are the main error-generating mechanisms.

Method/Models

I use three regression models:

a linear regression model specifying the underreport of spells of unemployment,

a logistic regression model specifying the probability of omitting spells of unemployment,

and a random intercepts logistic regression model specifying the probability of misclassifying work status.

Results and Conclusions

Retrospective reports of work histories are prone to different types of measurement errors: over/underreporting of the number of spells, over/underreporting of spells length, misdating of spells starts and misclassification of status.

In addition, these types of measurement error seem to be more highly associated with both features of the question used in the questionnaire and the nature of the topic to be reported than with characteristics of the interviewee. In particular the time-frame used for the recall of spells plus the level of saliency of unemployment (measured by the length and the number of spells) explain the occurrence of measurement error better than demographic variables (gender or age).

On the loss of accuracy of economic indicators released too early

Jens Mehrhoff, Karsten Webel

Deutsche Bundesbank, Frankfurt, Germany

Objectives

Virtually all short-term macroeconomic indicators are revised. On their first date of release the data contain estimates for missing values which will be updated by and by with actual figures. The analysis of these revisions is an integral part of business cycle analysis. Traditionally, the focus has been on some sort of mean revision or on the standard deviation of revisions measuring the expected correction and describing the uncertainty of economic indicators, respectively. We now enhance this concept as we consider the variance of the economic indicators themselves rather than their revisions over time.

Method/Models

To this end, the lower degree of "hard" data in the estimates released at early publication dates is treated theoretically and empirically from a sampling point of view. To bridge the gap between revision analysis and the current debate on the trade-off between timeliness and accuracy of short-term economic indicators, we suggest an aggregation scheme at European level that allows consideration of time-varying standard errors. Its key assumption is that national sample sizes evolve proportionally to the variance of the underlying firm level data and according to a growth process whose limit is approached at an inverse exponential rate.

Results and Conclusions

Using turnover data from the monthly German retail trade statistics, we demonstrate that national economic indicators are likely to suffer from a substantial loss of accuracy if they are released simultaneously with the respective European aggregate. More specifically, we find that 30 days after the end of any given reference month ($t + 30$ days), the width of the confidence interval of the European aggregate is just 40 per cent of the national one. Analogously, we demonstrate that the European aggregate reaches the national level of accuracy as of $t + 30$ days much faster. Hence, if national and European data are to be published simultaneously then it is necessary to wait for qualitatively sound data at the national level to avoid an asymmetry in the reliability of national and European data.

Heckman's sample selection model and calibration estimation under nonresponse

Thomas Laitila^{1,2}

¹*Örebro university, Örebro, Sweden*, ²*Statistics Sweden, Örebro, Sweden*

Objectives

The calibration estimator is by Lundström and Särndal (1999) suggested for estimation of population quantities under nonresponse in sample surveys. Heckman (1976, 1979) suggests a two-step estimator of regression models under sample selection. The objective of this paper is to study the relationship between these two estimators, suggest new estimators based on their relationship, and study the properties of the new estimators.

Method/Models

Adapting the two-step estimator for estimation of population totals yields a basis for deriving a mathematical relation between the two estimators. By specifying a synthetic population corresponding to the distribution obtained by weighting the population distribution with response probabilities yields a statistical relationship between the two estimators. Based on these relationships, new estimators are suggested. Asymptotic properties are studied and simulation studies are used for studying the finite sample properties of the estimators.

Results and Conclusions

Using the two-step estimator for estimation of population totals yields an estimator satisfying the calibration restriction fulfilled by the calibration estimator. The calibration weights obtained are only partly similar to those of the calibration estimator, however. Defining a calibration estimator with these weights yields consistent calibration estimates. Based on these results a new calibration estimator is suggested and a new type of instrument variable vector is developed for the calibration estimator. The new calibration estimator is shown consistent, and the results of the simulation study shows on good sample properties in terms of smaller bias and competitive RMSE estimates.

Fisher's Statistical Legacy *Tuesday 4 September, 2.20pm*

Organised by the Fisher Memorial Trust

Rosemary Bailey (QMW)

Randomisation 50 years after Fisher

One of the most important ideas that R. A. Fisher introduced into experimentation during his time at Rothamsted Experimental Station was randomisation. Most people agree with that. However, it turns out that they disagree about what is meant by randomisation: what it is, how you should do it, what its purpose is, whether or not it is desirable, and so on. I shall try to cover some of the different points of view.

D.R.Cox , Nuffield College, Oxford

Fisher and formal statistical inference

A brief outline is given of Fisher's contributions and of their relation with three complementary approaches. The present position is outlined.

David J. Hand, Imperial College, London

Multivariate analysis in the 21st Century

Fisher produced seminal work in several areas of multivariate analysis notably optimal scoring and discriminant analysis. Here I focus primarily on the second area, showing how the method has been generalised and extended in a wide variety of ways, but is still extensively used in the basic form that Fisher described.

Keeping the media honest in covering statistics Tuesday 4 September, 2.20pm

Stats has a media problem. Journalists can distort numbers, exaggerate risk and ignore basic rules of sampling and data handling. Things are improving, but the price of accuracy is constant vigilance. The session doesn't just look at how statisticians can protect themselves against misreporting. It's also about exploiting opportunities to get statistics into the public domain through the media, through more imaginative presentations and user-friendly releases. We've put together a panel of people who work with stats, who are in constant contact with the media. What do they recommend, both pro-actively and defensively? Can media organisations be pushed into greater care with stats? Are there new rules of the road for social media?

The panel will be moderated by David Walker, director of RSS getstats and former journalist with the Guardian, Times and BBC. Panel members are Penny Young, director of NatCen, the independent social research organisation, Aleks Collingwood, head of statistics for the Joseph Rowntree Foundation and Naomi Givens, statistics manager with GSK and promoter of the Science Media Centre's Behind the Headlines project.

Beating the odds: Successfully managing a rapidly-expanding statistical team

Dr Tim Paulden (ATASS Sports)

ATASS Sports is a successful Exeter-based statistical research company specialising in the forecasting of sports. Recently the company has undergone significant expansion, with the research team growing fivefold in less than five years. This talk will explore ATASS's experiences in managing its continuing expansion, and how critical processes such as recruitment, communications, project delivery, procurement, technology, and training have evolved (and continue to evolve). The presentation is designed to be of general interest to all professional statisticians, and of particular interest to decision-makers working in knowledge-based organisations.

The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews

Rebecca Turner, Sheila Bird, Julian Higgins
MRC Biostatistics Unit, Cambridge, UK

Objectives

Systematic reviews and meta-analyses of intervention studies aim to synthesise all available evidence on a specified research topic. The majority of meta-analyses include data from one or more small studies, which would not themselves have power to detect an intervention effect. The relative influence of adequately powered and underpowered studies in published meta-analyses has not previously been explored. Our objectives are to explore the distribution of power available in primary studies within meta-analyses published in the Cochrane Library, and to investigate the impact of underpowered studies on meta-analysis results.

Method/Models

Our analyses included 14,886 meta-analyses of binary outcomes from 1,991 Cochrane reviews. For each study in each meta-analysis, we calculated the power available to detect fixed levels of relative risk reduction. Associations between meta-analysis characteristics and power were examined. In a subset of 1107 meta-analyses including 5 or more studies, of which at least two were adequately powered and at least one underpowered, meta-analyses were repeated with underpowered studies excluded. Using random-effects models, we estimated the average differences in observed intervention effects between adequately powered and inadequately powered studies, overall and within medical specialties, outcome types and intervention comparison types.

Results and Conclusions

In 70% of 14,886 meta-analyses, all studies included were underpowered (power < 50%) to detect a 30% relative risk reduction. Only 17% of meta-analyses included at least two adequately powered studies (power ≥ 50%). In the subset examined, odds ratios in underpowered studies were on average 10% lower (95% CI 8% to 12%; $P < 0.0001$) than those in adequately powered studies, where lower odds ratios represent more extreme effects favouring the active treatment. The standard error of the intervention effect increased by a median of 12% (inter-quartile (IQ) range -1% to 35%) when underpowered studies were omitted. The between-study heterogeneity estimate decreased by a median of 27% (IQ range -100% to 14%). If at least two adequately powered studies are available, underpowered studies often contribute little information, and could be left out if a rapid review of the evidence is required. However, underpowered studies make up the entirety of the evidence in most areas.

Optimisation of the two-stage randomised trial design with some participants undecided about their treatment preference

Stephen D Walter¹, Robin M Turner²

¹McMaster University, Hamilton, Ontario, Canada, ²University of Sydney, Sydney, New South Wales, Australia

Objectives

In a two-stage randomised trial, participants are first randomly divided into two subgroups. In one subgroup, participants are randomly assigned to treatment, while in the other subgroup, participants may choose their own treatment. This permits estimation of the usual direct treatment effect, and also of the potentially important effects of patients' preferences between treatments (selection effects) and interactions between preferences and treatment received (preference effects). The objective of this work was to optimise the proportion of patients to allocate to the choice group when some proportion of study participants have no preference, and who are then re-randomised to treatment.

Method/Models

We extend earlier work on the case where all participants have a treatment preference (Walter *et al.*, Stat Med, in press) to develop more general expressions for the optimum proportion of participants who should be placed in the choice group. Optimisation is with respect to the standard errors of the selection, preference and treatment effects. The optimum depends on: the variances of the outcome in observable study subgroups; the proportions of participants who select each treatment in the choice group; and the selection, preference and treatment effects themselves, but it is independent of the total sample size.

Results and Conclusions

The optimum proportion in the choice group ranges between 40% and 50% for most reasonable scenarios. It is lower if preferences for one treatment dominate, or if the proportion of undecideds is low; otherwise the optimum is typically close to (but below) 50%. However, the variances of the selection and preference effects increase substantially if preference for one treatment dominates, or if the proportion of undecideds is large.

These ideas will be illustrated using a 2-stage randomised trial comparing medical and surgical management strategies for heavy menstrual bleeding. In this study, approximately 70% of participants in the choice group had no treatment preference, and were re-randomised.

We conclude that the two-stage design can be optimised to estimate selection, preference and treatment effects even when some participants have no treatment preference. However the absolute variation in these estimates becomes large if the proportion of undecided participants is large.

Feasibility/Pilot Studies in the design, conduct and evaluation of complex interventions

Gillian Lancaster

Lancaster University, Lancaster, UK

Objectives

Trials carried out in a health care setting typically involve complex interventions that require considerable planning if they are to be implemented successfully, as well as sensitivity to the local context in applying experimental methods. Issues in the design of complex interventions have been addressed over the past nine years by the Royal Statistical Society's Primary Health Care Study Group. With the aim of raising the profile of statistics and building research capability in this area with respect to methodological issues, the study group meetings have covered a wide range of topics - for overview see Lancaster et al. (2010).

Method/Models

The MRC's (2008) new guidance document on complex interventions and Campbell et al. (2007) emphasise the importance of thorough groundwork in designing and evaluating complex interventions and stress the importance of contextualising and conceptualising the problem at the development stage. Lancaster et al. (2004) and more recently Arain et al. (2010) have showed that there was and still is a dearth of pilot studies in the literature that stated they were specifically in preparation for a RCT, and that give a clear list of key objectives relating to the pilot phase.

Results and Conclusions

Feasibility and pilot studies are conducted to assess the feasibility and integrity of the study protocol, but the differences between the two are not clear-cut.

This talk will provide an overview of the use of feasibility and pilot studies in the literature, consider statistical issues in study design, and provide useful examples and references for future work.

Lancaster G.A., Campbell M.C., Eldridge S.E., Farrin A., Marchant M., Muller S., Perera R., Peters T.J., Prevost A.T., Rait G. (2010). Trials in Primary Care: statistical issues in the design, conduct and evaluation of complex interventions. *SMMR* 19: 349-377.

Arain M., Campbell M.J., Cooper C.L., Lancaster G.A. (2010) What is a pilot study? A review of current practice and editorial policy. *BMC Medical Research Methodology* 10:67.

Lancaster G.A., Dodd S.R., Williamson P.R. (2004). Design and analysis of pilot studies: recommendations for good practice. *Journal of Evaluation in Clinical Practice* **10** (2), 307-312.

A practical approach to sample size calculation using simulation

Richard Hooper

Queen Mary University of London, London, UK

Objectives

Choosing sample size on the basis of statistical power is an important ethical consideration when planning research in the life sciences. Statisticians resort to a variety of resources when calculating sample size, but new methodologies tend to outstrip new technologies for sample size calculation. Although the exact power achieved by a given sample size can always be estimated using Monte Carlo simulation, using simulation to determine the sample size required to achieve given power is more complicated, necessitating power to be estimated at different sample sizes. I discuss how this can be implemented in a practical and accessible way.

Method/Models

I describe an algorithm for determining sample size which searches heuristically and uses simulations with increasing numbers of replications: fewer replications in early iterations mean the algorithm finds a solution more quickly; more replications in later iterations mean that it estimates power with high precision to give a repeatable sample size. When implemented in general statistics software, the algorithm can invoke user-written code for generating and analysing a single data-set, and in this way calculate sample size for any method of analysis under any probability model that can be programmed. I present examples, and compare simulation with approximate, analytical alternatives.

Results and Conclusions

Implementing simulation on a widely available software platform makes it easy for investigators to share, and for others to validate, sample size calculations, for example in grant applications. It also allows a wider online community to own and extend the resource. Analytical approaches to sample size calculation can founder when there are a number of nuisance parameters, or in applications such as adaptive clinical trial design where the distribution of the test statistic under alternatives is intractable. The approach I have described is versatile and exact. Applications of simulation to sample size calculation have previously tended to use a fairly broad brush, but with sufficient replications simulation offers an exact solution that is repeatable, and efficient algorithms ease the burden of having to try different sample sizes to obtain a solution. With its non-analytical approach, simulation may seem mathematically unrefined, but this is precisely what gives it its versatility.

The Economic Benefits of Adult Apprenticeships: analysis of the Labour Force Survey.

Diana Tlupova, Joe Perkins, Martin Malinowski

National Audit Office, London, UK

Objectives

The Apprenticeship Programme in England has rapidly expanded in recent years and the government has announced substantial increases in spending. However, there is little quantitative evidence on the benefits of apprenticeships for the individual, employer and the Government. This paper aims to fill this gap by estimating the benefits to individuals in the form of wage and employment premiums and to use these estimates to conduct a rigorous cost-benefit analysis of the programme.

Method/Models

To estimate wage and employment premiums, we carried out multiple regression and probit regression analyses respectively using the Labour Force Survey (LFS). We employed data pooled from the successive Labour Force Surveys from 2004 to 2010. The wage and employment premium estimates from this analysis were then used in the cost-benefit model to produce an estimate of the return from apprenticeships to society and per pound of government investment.

Results and Conclusions

Our statistical analysis found that there are significant positive wage premiums associated with apprenticeships, although there is a high degree of variation depending on the sector of the apprentice's employment. Furthermore, there is indicative evidence that the wage premiums have declined over time for Intermediate apprenticeships, on average by about 2.4 percentage points per year. We also found evidence that apprentices are more likely to be in employment. The results of our cost-benefit analysis suggest that the returns from apprenticeships to economy as a whole are about £5 per pound of costs incurred; and around £18 per pound of government funding. This is substantially lower than the Department for Business, Innovation and Skills estimate of £40 per pound of government funding. These results were used to inform the NAO's conclusion on the Value for Money of the Apprenticeship Programme and were reported to the Parliament in March 2012.

Evaluating the impact of agricultural credit: A matching approach

Sunil Mitra Kumar

University of East Anglia, Norwich, UK

Objectives

To measure the change in the wealth status of farmers in India as a result of their having access to agricultural bank loans, where wealth status is measured using an index of consumer durable assets. Thereby, to provide a partial evaluation of government policy aimed at strengthening the provision of agricultural bank loans for farmers.

Method/Models

1. We use matching estimators based on the propensity score and Mahalanobis distance metric. We balance on pre-treatment values of relevant variables, i.e. before farmers received bank loans, and compare the post-treatment value of a wealth index for farmers who did and did not receive bank loans.
2. We pay close attention to the balance that results from matching, using quantile-quantile plots, their summary measures, and standardized bias plots to compare the distribution of variables between matched treatment and control groups.
3. Sensitivity analysis shows that the results are partially robust to the presence of unobservables.

Results and Conclusions

We find evidence of a small, positive treatment effect that is statistically significant. We check for robustness to unobservables and sensitivity to the matching specification and find that the treatment effect is quite robust to the latter and only mildly robust to the former.

The challenges of monitoring the illegal ivory trade to guide international decisions on elephant conservation.

Fiona Underwood, Robert Burn

University of Reading, Reading, UK

Objectives

The Elephant Trade Information System (ETIS) is one of two global monitoring systems for elephants under CITES (Convention for International Trade in Endangered Species). ETIS collects records of illegal ivory seizures and aims to: assess illicit trade in ivory; establish trends over time; determine whether or not such trends are related to CITES decisions. Three major challenges for analysis of these data are: (a) the opportunistic and interventionist nature of the data collection process (b) the complex, and poorly understood, inter-relationships between drivers of the illegal ivory trade (c) simple indicators are required to provide evidence to policy makers.

Method/Models

A Bayesian hierarchical latent variable model of the number of seizures reported by a country in a year captures the process by which shipments of illegal ivory may become records in ETIS. Potential proxy variables that affect the proportion of shipments that are seized and reported to ETIS are proposed. Posterior distributions of specific parameters in the model are combined with a separate model estimating the weight per seizure to obtain indicators of the illegal ivory trade. A conceptual framework for understanding causal relationships and drivers of the trade is outlined.

Results and Conclusions

The ETIS database holds over 17,000 records from 1989 to the present day. This analysis focuses on data from 1996 to 2011 to produce bias-adjusted indicators of time trends in the illegal ivory trade for different categories of illegal ivory shipments. Comparisons between countries are also made. Covariates that affect the latent variables of reporting rate and seizure rate are identified. The results form part of a report to be submitted to the CITES Conference of the Parties to be held in March 2013 in which further decisions on the regulation of trade in elephant products will be made.

Statistical challenges in high-throughput sequencing with examples in clinical resequencing

Gerton Lunter

University of Oxford, Oxford, UK

Objectives

Next-generation sequencing has provided the medical and biological sciences with a powerful new tool for clinical diagnosis and for research into biological function and evolution. An increasing range of 'library prep' protocols are available to prepare a sample for sequencing, allowing a range of different experiments to be carried out, from interrogating an entire genome sequence to estimating RNA expression levels and identifying protein-DNA interactions.

Method/Models

To extract meaningful information from the data, a range of tool-chains to analyze the often very large quantity of data is needed. Each combination of sequencing platform, library prep protocol and experimental design may require its own bespoke tool-chain. New sequencing platforms continue to be developed, which offer different advantages such as longer read length or lower cost, but do not always improve data quality, emphasizing the need for robust statistical tools.

Results and Conclusions

In this talk I will give an overview of the statistical and algorithmic challenges encountered in the analysis of next-generation sequencing data in the context of clinical resequencing applications, and I will try to sketch what I believe are the major challenges in this area today.

Statistical challenges in flow cytometry with application to cellular signatures for graft-versus-host-disease

Ulrike Naumann¹, George Luta², Matthew P. Wand³

¹*Kings College, London, UK*, ²*Georgetown University Medical Centre, Washington, DC, USA*,

³*University of Technology, Sydney, Australia*

Objectives

High-content flow cytometric screening (FC-HCS) is a 21st Century technology that combines robotic fluid handling, flow cytometric instrumentation, and bioinformatics software, and produces hundreds of large multivariate samples of cellular characteristics. A major component of the processing of these samples is a form of cell subsetting known as gating. Manual gating is time consuming and subjective. Our objective is therefore to develop an algorithm for automatic and semi-automatic gating, which is objective and less time-consuming.

We revisit an application of FC-HCS to the problem of cellular signature definition for acute graft-versus-host disease (GvHD).

Method/Models

We develop a statistical procedure, named curvHDR, for automatic and semi-automatic gating. The method combines the notions of significant high negative curvature regions and highest density regions. The underlying principles apply to dimension of arbitrary size, although we focus on up to three dimensions. We provide accompanying software.

Results and Conclusions

We apply the curvHDR method for cellular signature definition of GvHD data. The data were obtained from patients undergoing blood and marrow transplant, many of whom develop acute GvHD. The outcome of GvHD could be improved if it were treated as early as possible and if the diagnosis could be made more definitively.

Our methodology for obtaining cellular signatures of GvHD involves three phases in order to develop a test that would be useful 1-2 weeks prior to the onset of GvHD. Only a small number of tuning parameters need to be chosen.

We were able to develop an algorithm that is faster and possibly more objective than previous methods. The method is seen to adapt well to nuances in the data and to a reasonable extent, match human perception of useful gates.

Statistical challenges in Genome Wide Association Studies with an application to the detection of gene-gene interactions in Crohn's disease

Heather Cordell

Newcastle University, Newcastle upon Tyne, UK

Objectives

In this talk I will discuss possible strategies and statistical challenges involved in the analysis of data from genome-wide association studies (GWAS).

Method/Models

I will focus particularly on the issue of gene-gene interactions, and describe some recent work carried out by ourselves and others in relation to detecting gene-gene interactions using GWAS data.

Results and Conclusions

Genome-wide association studies (GWAS) have proved an enormously successful strategy for the detection of genomic regions harbouring genes associated with complex diseases and quantitative traits. However, moving from a replicable association signal to determination of the underlying causal genetic variant and an understanding of its mechanism has proved more challenging. In addition, the variants detected in GWAS generally confer modest effects (odds ratios of 1.5 or less), making them of limited use for risk prediction, and do not generally account for a large proportion of the observed heritability (the additive genetic contribution to the overall trait variance), suggesting there are many additional, possibly interacting, genetic contributors to be found. In this talk I will discuss some possible strategies and statistical challenges involved in addressing these issues.

Statistics at work in the Midlands: Cars, Trains and Flu Tuesday 4 September, 3.50pm

Organised by the RSS West Midlands Local Group

Stuart Hillmansen (University of Birmingham)

Energy saving strategies for the railway

Energy saving in the railway sector is a major consideration for current operators. This is driven by the continuing focus of environmental concerns, together with an increasing uncertainty about the availability and cost of liquid fuels. The extended cycle life of railway equipment means that decisions influencing traction systems can have lasting consequences of up to 50 years into the future. This presentation describes the fundamentals of railway vehicle kinematics and then analyses a number of notable recently implemented energy saving strategies. These include regeneration, use of energy storage, train control and optimisation, and operational effects.

Andrew Rouse (Heart of Birmingham teaching PCT)

NHS performance: How general is your General Practice?

Monitoring the quality of health services is notoriously difficult. In part this has been due to the fact that health services generated very performance data – and even this was incomplete and poor quality. Recently systems have been put in place which generate large, complete high quality data sets from over 8 thousand practices – but it is unclear how these should be used. In this talk I will demonstrate how a simple graphical processing technique can generate performance scatter-grams which can easily identify practices with extremely unusual activity patterns. The scatter-gram performance process can be used to process any very large data comprising many performance indicators collected from many providers.

Using Dirichlet process partitioning models to assist linear model determination: application to graphical log-linear models

Michail Papathomas¹, Sylvia Richardson²

¹*University of St Andrews, School of mathematics and Statistics, St Andrews, Scotland, UK,* ²*MRC Biostatistics Unit, Cambridge, UK*

Objectives

Detecting high order interactions between covariates in a linear model framework is not straightforward, due to the difficulty in investigating an unwieldy large space of competing models. Using simulated data we examine the relation between interactions terms in a log-linear model for a contingency table, and the profile of homogenous clusters created using flexible Bayesian partitioning.

Method/Models

We utilize the fact that variable selection within a clustering algorithm selects covariates that combine to form homogeneous groups of subjects, rather than covariates with a strong marginal signal, and propose a novel approach where a model search algorithm for a large space of log-linear models is informed by results from a Bayesian partitioning approach implemented in tandem with a variable selection procedure.

Results and Conclusions

We demonstrate that the combined effect of covariates on the partitioning of the subjects can sometimes be directly translated to high-order interactions in a log-linear model. Simulation and real data analyses show that for data with a complex underlying structure, this effect can be used to inform a conventional model search algorithm.

How Nonzero Ties distort Statistical Inference with Rank Tests

Nadine Chlass¹, Jens Krueger²

¹Friedrich-Schiller-University, Jena, Germany, ²Darmstadt University of Technology, Darmstadt, Germany

Objectives

We study how tied observations distort statistical inference with rank tests. We assess the performance of predominant tie breaking methods and study critical tie frequencies for statistical inference. We compare these critical frequencies to typical tie frequencies in seminal studies from economics. We discuss the central assumption of all tie breaking methods -- that tied observations occur independently of each other -- and illustrate that the assumption often contradicts the very existence of ties which may signal latent variables. Depending on how many values these latent variables can take on the resulting tie is 'big' (Dummy latent variable) or 'small'.

Method/Models

Monte Carlo study. We construct one sample of n minus the number of ties i.i.d draws. We induce ties by redrawing a controlled number of ties from this sample and append them to the sample to have n observations. We obtain a second sample by a linear projection whose slope controls the association between the samples (related vs. unrelated) and whose intercept controls the effect size. A small random component ensures that the second sample contains exactly the number of ties we control. We count how often the test rejects the null for different tie breaking methods and tie sizes.

Results and Conclusions

For all tie breaking methods, merely 10% of data tied into one big tie, and 30% of small ties distort the type-I errors of important tests such as the Wilcoxon Signed Rank and the Wilcoxon Rank Sum test to an extent that insignificant effects are judged to be highly significant at all conventional significance levels. Always, a true Null hypothesis is rejected more often than it should be at the given significance level. The weaker the actual underlying significance level, the more often the test overrejects the null. Small ties hardly affect test power, big ties can seriously decrease test power up to 40%. Theoretically, we show that the entropy of the data decreases faster if there is one big tie than when there are many small ties. In seminal studies, we regularly find tie fractions twice to six times as high as the critical fractions we report.

Probabilistic Index Mixed Models for Clustered Data

Fanghong Zhang¹, Stijn Vansteelandt^{1,2}, Jan De Neve¹, Olivier Thas^{1,3}

¹Ghent University, Ghent, Belgium, ²London School of Hygiene and Tropical Medicine, London, UK,

³Centre for Statistical and Survey Methodology, School of Mathematics and Applied Statistics, University of Wollongong, NSW, Australia

Objectives

The use of linear mixed models for continuous clustered data can be problematic when the data distribution is skewed in a way that is not easily accommodated by data transformation, or when the response is measured on an ordinal scale. The probabilistic index models (PIM) of Thas et al. (2012) form a flexible class of semi-parametric models for analyzing non-normal responses, but they assume mutual independence of all observations.

Method/Models

In view of this, we propose an extension of the PIMs to clustered data. Given two random vectors (Y_{ik}, X_{ik}) and (Y_{jl}, X_{jl}) , where Y_{ik} and X_{ik} denote the k th outcome and covariate measurement for subject i , we consider models of the form

$$P(Y_{ik} \leq Y_{jl} | X_{ik}, X_{jl}, b_i, b_j) = \Phi\{\beta(X_{ik} - X_{jl}) + (b_i - b_j)\},$$

Where b_i and b_j are i.i.d random effects with constant variance, β is a finite-dimensional parameter and $\Phi(\cdot)$ is the standard normal cumulative distribution function. Consistent and asymptotically normal estimators of the covariate effect β and of the random effects variance are established through semi-parametric estimation methods.

Results and Conclusions

The proposed framework extends non-parametric rank tests to deal with clustered data in settings that require covariate adjustment. In a simulation study, the finite-sample behavior of our estimators is evaluated and a data analysis example is provided.

Application of Zenga's distribution to a panel survey on household incomes of 15 European Member States

Michele Zenga, [Alberto Arcagni](#)

Dipartimento di Metodi Quantitativi per le Scienze Economiche e Aziendali - University of Milano-Bicocca - Piazza dell'Ateneo Nuovo, 1, Milano, Italy

Objectives

Our aim is to verify the fitting of Zenga's model on a large number of empirical income distributions, comparing different parameters estimation methods, both with and without restriction on the expected value.

Method/Models

Zenga's distribution is particularly suitable for modelling income distributions since it is positively skewed and it has Paretian right tail. This new density has three parameters: a scale parameter, which is equal to the expected value, and two shape parameters which affect the inequality.

Parameters estimates are obtained by the minimization of goodness of fit indexes and by the maximum likelihood method. These methods are implemented both without and with restriction on the expected value.

Results and Conclusions

A good fitting of the model is observed to empirical income distributions of States with very different economic conditions, such as: Germany, France, Luxembourg, Austria, Italy, Spain and Greece. This result shows that Zenga's model can be reasonably applied to describe income distributions.

What is Public Health Intelligence (PHI) ? Why is it critical to the best functioning of the NHS in 2012?

Margaret Eames

Imperial College, London, UK

Objectives

PHI is now the statistical branch of public health, which involves the analysis of medical, demographic and inequality data producing evidence to change public health policy. I will show how PHI has changed local public health policy in the past and give current examples of how PHI is the necessary bridge of evidence between public health and the new Clinical Commissioning Groups in the NHS. However, without wider knowledge of the statistical importance of PHI, the progress made in recent years, is in danger of being lost in the transition of Public Health from PCTs into the Local Authorities.

Method/Models

I will show how PHI has changed local public health policy in the past eg i) using smoking attributable mortality (SAM) rates-and regression models with recruitment to stop smoking clinics-as the dependent variable and SAM as the independent variable, to direct the relocation of stop smoking clinics- to places of most need ii) local childhood obesity analysis -showing gender and geographical variation to motivate community interventions to increase exercise and more healthy eating in Hertfordshire.

It will also give current examples of how PHI is the necessary bridge of evidence between public health and CCGs in the new NHS.

Results and Conclusions

There is need at this time for a professional "home", and training courses in PHI, to preserve standards and careers in this new NHS discipline. PHI has only just been established in the last 10 years. However, without wider knowledge of the statistical importance of PHI, and some academic vision for training - the progress made in recent years, is in danger of being lost in the transition.

It is important following the passing of the new "Health and Social care bill" on March 19th 2012 for PHI to be understood by the public and its function preserved in the transition of Public Health from Primary Care Trusts into the local authorities in April 2013. PHI as it measures and evaluates Public health outcomes is also becoming even more important for the accountability of commissioning health care in the brave new world of Clinical Commissioning Groups (CCGs).

Statistical Guidance on Optimal Strategies to Prevent Non-response in Longitudinal Studies

Ian Plewis

University of Manchester, Manchester, UK

Objectives

The organisational complexities of managing large-scale longitudinal studies require decisions to be made at the various stages of data collection. A number of these decisions relate to sample maintenance and to the need for the study to remain representative of its target population over time so that inferences about change are soundly based. Intelligent strategies do, however, need to be based on the efficient use of limited resources which take account of the emerging consensus that a higher response rate does not necessarily imply that the achieved sample generates unbiased estimates of change parameters.

Method/Models

The paper will show how to exploit the information contained in logistic regression models that generate estimates of response propensities, and the associated representativeness indicator and receiver operating characteristic curve to (a) examine the utility of reissuing to the field at wave t cases that did not respond at wave $t-1$, (b) determine an optimum threshold for allocating extra resources to respondents who might otherwise refuse. The ideas will be illustrated with data from waves one to four of the UK Millennium Cohort Study.

Results and Conclusions

The results suggest that, given the relatively low areas under the receiver operating characteristic curves, it might be difficult to target interventions effectively to reduce refusals. Some, but not all, reissuing strategies can improve representativeness.

Data.gov.uk: Opening the door on open data

Andrew Epps

National Audit office, London, UK

Objectives

The objective of the presentation is to reflect on the contents of data.gov.uk from the perspective of the National Audit Office (NAO), which uses some of the data on the site. The site contains around 8,000 datasets, but these are far from being homogenous. Many detail the spending of particular public organisations, whereas a few contain detailed information which can be useful for the NAO, as well as other analysts and application developers. The talk will describe and classify datasets based on the reason they have been released, the format of the data and how their data can be used.

Method/Models

The presentation will be based on an analysis of the metadata from data.gov.uk. This contains detailed information about each dataset linked from the site. The analysis will use descriptive statistics and qualitative analysis to examine the contents of the meta-data. In particular it will look at titles, producers, format of the data, national statistic status and geographic coverage.

Results and Conclusions

One of the main reasons for publishing data on data.gov.uk is transparency. For example, a large proportion of the datasets are regular releases containing itemised spending over £25,000 in public organisations. Other datasets contain summary level national statistics produced by government departments, which have always been available on their websites, but are now linked from this central point. The rarest type of dataset contains raw data, or samples of raw data, which enable statisticians and auditors to perform analysis and developers to produce web-based applications using the data. The release of this level data is the goal of the open government initiative, but such datasets are few and far between. This is where the economic value in open data lies, but is also where the most work still needs to be done.

Combining knowledge to estimate the percentage of care workers paid under the national minimum wage in the UK: a Bayesian approach

Shereen Hussein, Jill Manthorpe
King's College London, London, UK

Objectives

The care sector is estimated to employ about two million people, nearly 7% of the total UK labour force.. The Low Pay Commission reports continuous concern about low pay in the sector. Current estimates of the percentage of care workers paid under the National Minimum Wage (%UNMW) are likely to be inaccurate, due to lack of sensitivity of databases used and difficulties in calculating what constitutes 'duration' of work, among other reasons. The aim of this analysis is to present new estimates using adjusted sector-specific data while accounting for prior estimates.

Method/Models

We established pay distribution using the National Minimum Data Set for Social Care (NMDS-SC; $n=642,777$); we then adjusted hourly pay rates to account for a fraction of unpaid travel between clients and unpaid overtime using data from a large survey of care workers ($n=1205$). To account for prior knowledge we used a hierarchical Bayesian approach treating previous estimates as arising from a random process governed by the hyperparameters (α , β). We performed stochastic integrations via MCMC (Markov Chain Monte Carlo) approximation in order to estimate the densities. We implemented an iterative process of hyper Metropolis-Hastings/Gibbs sampling algorithm.

Results and Conclusions

The results show that the %UNMW of direct care workers has a 'stable' mean around 10 per cent, with different values of the gamma hyperprior parameters, indicating consistent and accurate results. 95 per cent credible intervals of the posterior inferences of %UNMW range from 9.2 per cent to 12.9 per cent. With very conservatively adjusted distributions of hourly pay rates (an average of 1% reduction), such percentage is higher than previous estimates. In addition, any further small adjustments in pay rates are likely to affect larger percentages of direct care workers due to their very narrow pay distribution. Using estimates of the workforce from the NMDS-SC and Labour Force Survey it is likely that payment UNMW affects 156,673 and 219,214 direct care jobs in the UK. This is of interest to policy makers, law enforcement bodies, employers and employees, as well as people needing care services.

Joint modelling of goals and bookings in association football matches

Andrew Titman, Deborah Costain, Gareth Ridall, Kris Gregory
Lancaster University, Lancaster, UK

Objectives

Using data collected from live match reports for Premier League and Championship football matches from the 2009/10 and 2010/11 seasons, a stochastic model is developed to determine the interdependence between match events. This can allow the effect of refereeing decisions on goal scoring rates to be quantified and also determine how booking rates are affected by the goal scoreline or bookings given to the opposing team. A particular aim is to have a model to allow dynamic prediction of final match outcomes given the current status of a game.

Method/Models

Match events are modelled through a multivariate counting process with intensities for the accrual of goals, yellow cards and red cards in continuous time. The intensities are assumed to be functions of time since kick-off, current match status and additional covariates such as relative team abilities and match referee. As a way of evaluating the model, a utility based live spread betting strategy using the model is developed and retrospectively applied to monitored prices from 93 matches in the 2011/12 season.

Results and Conclusions

Relative team ability, measured via bookmaker's prior match odds, have a significant effect on booking rates with a team's bookings peaking when the opposing side is slightly stronger. Yellow cards do not appear to have any direct effect on goal scoring rates. Red cards have a substantial effect, with the consequences being worse for an away team who are reduced to 10-men. Once a team receives a yellow card, the risk of the other players receiving a red card for a first bookable offence more than doubles, and the opposing team's booking rate increases by 25%. Referees' individual booking rates were generally more consistent in the Premier League than the Championship, with the exception of Mark Halsey who appears to be anomalously lenient.

The live betting strategy yielded a modest hypothetical profit, with particular success in the market relating to the total booking points occurring in a match.

Using interactive eBooks to teach and communicate statistical ideas

Richard Parker¹, Danilus Michaelides², Huanjia Yang², Christopher Charlton¹, Camille Szymaragd¹, Luc Moreau², William Browne¹

¹University of Bristol, Bristol, UK, ²University of Southampton, Southampton, UK

Objectives

The use of interactive eBooks, in which readers make choices which shape the content which appears, could be a very valuable addition to the toolkit available for teaching and communicating statistical ideas. To this end, an objective of the ESRC-funded e-Stat project (Digital Social Research programme) was the development of eBooks which would embed the executable statistical software packages, developed as part of the project, within traditional notebooks.

Method/Models

The software package Stat-JR has been developed as part of the e-Stat project; this incorporates a variety of open source templates, each of which takes users' input to perform particular functions with their data. Some templates fit models, whilst others concern data visualisation and summaries. The templates can inter-operate with externally-authored statistical packages, thus employing their engines as well. An eBook interface with Stat-JR has been developed which can utilise its functionality, with readers' input, and resulting output from the template executions, embedded within content written by an eBook author.

Results and Conclusions

The eBook tool, named DEEP, allows eBook authors to specify where, and what, inputs are required of the reader, the executions which take place as a result, and the format and location of the results returned; this is presented within content written by the eBook author, and displayed conditional on the users' input, thus allowing feedback to be tailored to it. The results presented can include plots, equations, data summaries, and model results, as well as the scripts / macros used to run externally-authored software. Thus readers can learn the general principles of analysing, plotting and summarising data, as well as the practicalities of how to do so via a variety of software packages. Readers can either upload their own data and/or use data pre-loaded by the eBook author. As well as a teaching tool, eBooks have a variety of other potential uses, such as a facility for multi-authored reports.

Using the STAT-JR software package for statistical analysis

William Browne¹, Christopher Charlton¹, Danus Michaelides², Richard Parker¹, Camille Szmaragd¹, Zhengzheng Zhang¹

¹University of Bristol, Bristol, UK, ²University of Southampton, Southampton, UK

Objectives

As part of a current ESRC funded grant within the Digital Social Research programme we have developed a new general purpose statistical software named STAT-JR. In this talk we will describe how the system works and how it allows users to perform statistical model fitting with a new MCMC estimation engine but also, through its interoperability features, to perform estimation in other commonly used packages.

Method/Models

The software package is designed to cater for three separate types of researcher. For novice researchers it offers a user-friendly interface and access to a variety of statistical operations, for example fitting particular model types or performing particular data manipulations. Each operation is constructed via a template and more advanced users can write their own templates and thus customise and expand the package. The software also produces custom computer code for model fitting that can be viewed and thus the third group of users, algorithm developers can take the code and customise it.

Results and Conclusions

We will be releasing the software to the user community in April 2012 and so by the conference we will be able to discuss initial user feedback. We will also describe how one can use the software to compare the capabilities and performance of existing software packages on the same problem. This feature allows the user to easily translate across packages thus teaching users new packages through packages they are already familiar with. We will finish by describing how we foresee its use in the future.

Explaining the significance of p: Presenting statistics to non-statistical audiences

Ah Mun Kuan, Andrew Epps, Lizzie Hogarth, Andrew Oliver, Caroline Harper
National Audit Office, London, UK

Objectives

The National Audit Office (NAO) scrutinises public spending on behalf of Parliament and reports to the Committee of Public Accounts on the economy, efficiency and effectiveness of government programmes and spending. The use of complex techniques in gathering and analysing evidence, and identifying relationships between inputs and outputs presents challenges. We need to balance our ambition to present core messages in simple language with assurance that we have used sound analytical techniques in drawing conclusions.

This presentation describes our current practice in communicating our methods and report findings to our stakeholders.

Method/Models

We provide training to help study teams interpret and understand the findings and the limitations of analysis, and how to present them. When presenting analysis in text, we avoid technical language and jargon, such as p-value or reference to particular statistical tests. Our analysis is often conveyed using tables, charts and diagrams, which are kept as simple as possible.

Prior to publication, draft reports are peer reviewed by analytical experts to ensure the graphics and text get the message across appropriately. This review also provides assurance to the teams that appropriate methods are being applied, and being applied rigorously.

Results and Conclusions

The use of simple graphics and plain English enables us to communicate our core messages in an appropriate format for our key audiences.

Our current practice is to produce a short report written, a detailed methodology appendix, and where appropriate a more technical report for other readers.

We find this creates a good balance between rigour and brevity, but are always looking for ways to improve the way we present our findings, and would welcome ideas from others.

Have the cutbacks in statistical budgets undermined the statistical base?

Wednesday 5 September, 9:00am - 10:00am

Organised by Official Statistics Section and General Applications Section

The National Statistician's assessment of the effect of the cuts across the whole of the Government Statistical Service

Jil Matheson, National Statistician, UK

The National Statistician will be talking about how statistical budgets, outputs and numbers of statisticians in the Civil Service have changed over the past few years.

When faced with budget reductions, what do you do?

How do you decide what the right programme of work is?

Have the cutbacks in statistical budgets undermined the statistical base?

Wednesday 5 September, 9:00am - 10:00am

Maximising Value From Statistical Budgets

Richard Alldritt

UK Statistics Authority, London, UK

Objectives

Around two hundred official bodies in the UK produce at least a few official statistics. Their budgets for this vary hugely and are set almost completely independently of each other. This restricts the scope for central planning, co-ordination or priority setting. But it does have some virtues as well.

Method/Models

We must start any discussion of value and budgets from where we are now. If cuts in budgets are made, their impact must be investigated and carefully assessed before deciding whether to challenge the budgetary decisions. It is not necessarily wrong to reduce the scale of statistical collections. That may allow value to be enhanced elsewhere within a fixed budget. The UK Statistics Authority has a mechanism to make a considered assessment via its Statistical Expenditure Report series.

To fully understand the value of statistical outputs and of the wider service provided by the statistical service, and thus assess the impact of any cuts, we need to identify the ways in which existing outputs are used; the nature of the public good (or beneficial impact) associated with those uses; and the scope to enhance that beneficial impact.

Results and Conclusions

Only then will we know what we are losing. Investigation of use, potential use, and the nature of the benefits that flow from those uses, are also steps on the path to maximising the value that can be obtained from whatever statistical budgets exist in the future. Maximising value needs to become a conscious focus of statistical work in government. And perhaps even more urgently, investigation of use and benefits is the path to maintaining and defending current public expenditure on official statistics. The case must be made convincingly that expenditure on statistics is not only good value, but better value than alternative uses of that money.

Have the cutbacks in statistical budgets undermined the statistical base?

Wednesday 5 September, 9:00am - 10:00am

The Radical Statistics view of the effect of the cuts on government statistics and their use

Ludi Simpson

Radical Statistics, -, UK

Objectives

Government research and the production of official statistics underpin several aspects of an effective and democratic state. They support the government's ability to appraise alternative policies and evaluate them once underway; they help make government accountable to civil society for its decisions; they allow others to conduct their own research for scientific or commercial ends; they support democratic debate in the public sphere. If research and statistics become out-of-date, irrelevant, or of lower quality and accuracy, all these functions are impeded.

It is difficult to quantify the scale of cuts to research capacity and statistical expenditure because there is no single standard measure of statistical and, more generally, research activity. It is even more difficult to assess accurately the effect of that activity or of its reduction on the care, conditions and culture of UK citizens.

Method/Models

Nonetheless, reductions in research and associated staff, in budgets and in activity have been evident since the current government's Comprehensive Spending Review of 2010; the curtailed expenditure follows the banking crisis in 2008, the subsequent recession, and their effects on public finances. Some major data-gathering exercises including the Citizenship Survey and the Tellus Survey have been ended. Budgets for commissioning research related to current policy have been severely curtailed. The abolition of Regional Assemblies has transferred planning and associated strategic research responsibilities to local authorities at the same time as their own budgets are reduced. The cuts, and their effects are ongoing. The quality of analytical services inside and outside government has deteriorated in parts. Where evidence is lacking, policy has developed based more on practical alliances and political expediency than on evidence of need, giving relatively more weight to convenience, political understanding, and financial priorities.

Results and Conclusions

The presentation arises from the 'Reduced Statistics' network of Radical Statistics' members and associates, who have sought to highlight changes to statistical output, and the implications of these changes in different fields.

We have to deal with media – to secure impact, to alert the world to new developments, to influence policy and render account to taxpayers and the public. But the media pose problems. Journalists may not be well informed; their levels of statistical literacy may not be adequate. So what should statisticians do to tell their story. This session offers practical advice. You can turn to press officers and media advisers, but the effectiveness of any story will depend on statisticians themselves having a grasp about how the media work.

Sessions themes are a) a brief guide to the media and the changing balance between the internet, print and broadcasting; b) a natural history of 'the story' and how journalists and editors construct and package information; c) what statisticians need to do in dealing with media – writing releases, making social media interventions, responding to error.

Discussion will be led by David Walker, director of RSS getstats and former journalist with the Guardian, Times and BBC and director of communications for the Audit Commission

A review of methods for specifying the target difference in randomised controlled trials (DELTA review)

Jonathan Cook¹, Jenni Hislop¹, Doug Altman², Craig Ramsay¹, Peter Fayers¹, Andrew Briggs⁴, John Norrie¹, Ian Harvey³, Luke Vale⁵

¹University of Aberdeen, Aberdeen, UK, ²University of Oxford, Oxford, UK, ³University of East Anglia, Norwich, UK, ⁴University of Glasgow, Glasgow, UK, ⁵University of Newcastle, Newcastle, UK

Objectives

When the sample size for a randomised controlled trial (RCT) is determined, a (target) difference is typically specified (following a Neyman-Pearson approach) which the RCT is designed to detect. This provides reassurance that the study will have the required statistical certainty and power. From both a scientific and ethical standpoint, selecting an appropriate target difference is crucial; too large or small a study is arguably unethical, wasteful and potentially misleading. Various methods exist for specifying a target difference, though their relative merits are uncertain. This project aimed to assess formal methods for specifying the target difference in a RCT sample size calculation.

Method/Models

This work had two main elements:

- a) Comprehensive systematic review based on electronic searches of a variety of biomedical and non-medical databases were performed to identify methods.
- b) Two surveys of trialists were conducted to assess current knowledge and practice: one of the International Society for Clinical Trials and one of UK and Ireland based trialists.

Results and Conclusions

Seven methods were identified - anchor, distribution, health economic, opinion-seeking, pilot study, review of evidence base and standardised effect size; each with important variations. Levels of awareness, use and willingness to recommend varied, with review of evidence base, pilot study and standardised effect size method most commonly used.

Some methods require strong assumptions to be used in the design of a RCT. The implications of each method from a statistical perspective will be highlighted. While no single method provides a perfect solution to a difficult problem, raising the standard of RCT sample size calculations and the corresponding reporting of them would be a step forward.

Sensitivity analysis for departures from the missing at random assumption in randomised trials with missing outcomes

Ian White¹, James Carpenter², Nicholas Horton³

¹MRC Biostatistics Unit, Cambridge, UK, ²London School of Hygiene and Tropical Medicine, London, UK, ³Smith College, Northampton, USA

Objectives

Randomised trials with incomplete outcomes are commonly analysed using a "missing at random" assumption. Less commonly, such trials with binary outcomes are analysed using a "missing=failure" assumption. The objectives here are (1) to propose a method for performing a sensitivity analysis for specified departures from these assumptions, (2) to show that the method agrees with standard results when standard results are available, and (3) to show that the method is suitable for practical use.

Method/Models

The analysis model is assumed to be a generalised linear model. We propose a mean score method in which the expectations of the missing outcomes are estimated using a second model, a pattern-mixture imputation model. Departure from "missing at random" or from "missing=failure" is quantified by a user-specified parameter δ in the linear predictor of the imputation model. The pair of estimating equations is solved with a sandwich variance, for which a small-sample correction is presented.

Results and Conclusions

The proposed method is illustrated using data from the QUATRO trial in mental health (with quantitative outcome, initially analysed under "missing at random") and from a smoking cessation trial (with binary outcome, initially analysed under "missing=failure"). In both sensitivity analyses, point estimates and standard errors vary with δ but agree exactly with standard results when δ corresponds to standard assumptions. Practical use of the method requires collaboration between statisticians and clinicians to identify a plausible range of values of δ : this is facilitated by the simple interpretation of δ in a pattern-mixture model.

Estimators for the effect of received treatment in a randomised trial with all-or-none compliance under a Poisson model

Adam Brentnall, Zi Hua Yang, Jack Cuzick, Peter Sasieni

Queen Mary University of London, London, UK

Objectives

In randomised trials a Poisson model is often used to report results when events are rare. A complication that sometimes occurs when a one-off treatment is offered is all-or-none compliance: some individuals do not comply with randomisation and take the other treatment. In this case the intention-to-treat hazard-ratio estimator may underestimate the effect of received treatment. The aim of this talk is to compare and contrast alternative estimators for a rate-ratio treatment effect for this setup, and as the Poisson distribution becomes a less reasonable assumption.

Method/Models

Estimators that respect randomisation are presented for the so-called conditional effect on an individual who would take the treatment if offered, including a non-iterative extension of an earlier estimator for binary risk ratios. Additionally, estimators are considered for the so-called marginal effect of treatment if everyone had adhered to randomisation allocation. The relationship between the new estimators, maximum likelihood and inverse probability weighting and g-estimation is discussed. Consistency is assessed, and relative efficiency is considered. Computer simulations compare the performance of the estimators against intention-to-treat and per-protocol approaches.

Results and Conclusions

The efficiency of a non-iterative conditional effect estimator relative to maximum likelihood decreases slightly as the level of censoring increases. The asymptotic efficiency of the non-iterative estimator relative to ITT depends on the level of censoring and the risk difference between potential compliance groups. Event-adjusted marginal-effect estimators are equivalent to the non-iterative estimator when exposure is fixed, but are biased when it is not. However, the bias might be small relative to that from ITT and simulation results were similar to exposure-and-event adjustments when the Poisson assumption was reasonable.

Modeling Credit Risk of Small and Medium-Sized Businesses: Evidence from the UK

Jake Ansell

The University of Edinburgh, Edinburgh, UK

Objectives

This presentation seeks to provide a comprehensive picture of the main approaches to modelling risk associated with lending to SMEs and issues related to it. Such risk shares the features of both retail and corporate sectors, and this has been recognised by Basel II and Basel III provisions. The dual nature of small business lending makes it possible to assess the credit risk using the approaches from both corporate and retail lending sectors. The corporate world relies mainly on structural market-based models for credit risk measurement, whilst retail lenders use empirical predictive models (credit scoring).

Method/Models

A number of risk-rating models are introduced for a UK sample of listed small businesses applying an accounting-based approach, which uses financial ratios to predict corporate bankruptcy. An enhancement to these models is considered through features typical to retail credit risk modelling. Furthermore, rather than using the traditional binary definition of survival and default the paper considers four categories: default, financial distress type I, financial distress type II and healthy SMEs. The regression models considered are logistic, survival and ordinal regression.

Results and Conclusions

One of the major problems related to SMEs consists in their informational opaqueness - financial statements are not necessarily audited, hence the information obtained from them is less reliable in comparison to large companies. Therefore, the predictive value of other type of information - transactional characteristics - is explored using another dataset obtained from one of the major UK lenders. Investigation of different stages of financial performance is continued through different clustering algorithms. Logistic and ordinal regressions are compared again.

Changes to the Compilation of UK National Accounts and Balance of Payments Statistics

Robert Dunn

Office for National Statistics, Newport, UK

Objectives

This paper sets out the main methodological and presentational changes which are going to be made to the UK National Accounts and Balance of Payments to become compliant with the new European System of National and Regional Accounts (ESA2010) and Balance of Payments and International Investment Position (BPM6) manuals. Both manuals will become the new basis for the compilation of National Accounts and Balance of Payments statistics for European Union countries from September 2014. These changes will appear in the UK National Accounts 'Blue Book' and UK Balance of Payments 'Pink Book' publications starting with the 2014 edition.

Method/Models

The revisions being made to the European System of Accounts manual are based on international guidance given in the United Nations' System of National Accounts (SNA) which was revised in 2008. All three manuals are periodically reviewed and updated so that National Accounts and Balance of Payments statistics continue to reflect economic reality in the most appropriate manner for monitoring, analysing and evaluating the performance of an economy. Therefore, it is important for users of these macroeconomic statistics to appreciate the forthcoming changes and what this means for the measurement of the UK economy.

Results and Conclusions

There is a large body of work confronting the Office of National Statistics over the next few years to move the compilation of the UK National Accounts and Balance of Payments statistics on to the new reporting basis consistent with the European and international guidance manuals. This body of work will mean changes across the statistical value chain, from data collection through to systems changes and data dissemination. The current plans are that these changes will appear in the UK National Accounts 'Blue Book' and UK Balance of Payments 'Pink Book' publications starting with the 2014 edition. Further changes will then be implemented in the years 2015-2017 with full compliance being achieved in the Blue Book 2017.

Lifelong Socio Economic Position and biomarkers of later life health: Testing the relative contribution of competing hypotheses.

George Ploubidis, Emily Grundy, Lenka Benova, Bianca DeStavola
London School of Hygiene and Tropical Medicine, London, UK

Objectives

The relative contribution of early and mid or later life Socio-Economic Position (SEP) to later life health is not fully understood and there are alternative hypotheses about the pathways through which they may influence health. One difficulty in the literature is that given the paucity of longitudinal studies with concurrent data collection spanning several decades, investigators often have to rely on retrospectively collected information on earlier life health and social circumstances. This is known to lead to bias as there is evidence that people's recall of this varies with current circumstances

Method/Models

In this paper we use data from the English Longitudinal Study of Ageing to address this issue and to investigate alternative hypotheses about life course influences on later life health, employing structural equation models as described within the causal mediation literature.

Results and Conclusions

Age group specific effects emerged, with current SEP dominating the effect on later life physical health and fibrinogen levels in participants under 65, while early life SEP had a more prominent role in explaining later life inequalities in physical health for men and women over 75. We extend previous findings on mid adulthood and early old age to old age and the beginnings of late old age. The complexity of our findings highlights the need for further research on the mechanisms that underlie the association between SEP and later life health.

Adjusting for Selection Bias in Longitudinal Analyses of the Relationship between Employment Transitions and Mental Health Using Simultaneous Equations Modelling

Fiona Steele¹, Robert French¹, Mel Bartley²

¹*University of Bristol, Bristol, UK*, ²*University College London, London, UK*

Objectives

There is substantial interest in understanding the association between labour force participation and mental health, and in particular the impact of unemployment on wellbeing. While panel data allow study of the dynamics of the relationship between changes in employment and mental health, selection bias remains a serious concern. We test for two types of selection effect: (i) direct selection (where prior health affects employment status), and (ii) indirect selection (due to unmeasured characteristics influencing both health and employment outcomes). We then examine the impact of adjusting for selection biases on estimates of the effect of employment transitions on mental health.

Method/Models

We investigate the relationship between men's employment transitions and mental health using data from the British Household Panel Survey, 1991-2009. We model the effect of a change in employment status between years $t-1$ and t on mental health at t , adjusting for mental health at $t-1$. Using a dynamic simultaneous equations model we allow explicitly for an effect of health at $t-1$ on employment transitions between $t-1$ and t to allow for direct selection. The health and employment equations include individual-specific random effects which are correlated across equations to allow for indirect selection due to shared unmeasured influences.

Results and Conclusions

There is strong evidence of indirect selection whereby men whose unmeasured characteristics place them at above-average propensity of anxiety and depression tend also to have above-average chances of being out of employment. There is rather less support for direct selection, with negligible effects of mental health at $t-1$ on employment transitions between $t-1$ and t . Ignoring indirect selection leads to an overstatement of the effects of employment transitions on mental health, but the effects remain substantial after adjustment. A move out of unemployment into employment is beneficial for mental health, while becoming unemployed is detrimental.

Adjusting for non-ignorable drop-out in a longitudinal analysis of residential mobility

Liz Washbrook, [Paul Clarke](#), Fiona Steele
University of Bristol, Bristol, UK

Objectives

We investigate the impact of drop-out on an analysis of residential mobility of men aged 20-59 using waves 6-18 of the British Household Panel Study (BHPS). Our concern is that drop-out depends on mobility itself and so is non-ignorable. However, models for non-ignorable drop-out can be highly sensitive to unverifiable modelling assumptions. We consider two approaches to non-ignorable drop-out from the statistics and econometrics literatures, and investigate a modelling strategy that uses 'exclusion restrictions' - predictors predicting mobility/drop-out but not drop-out/mobility - to reduce the sensitivity of our estimates to model mis-specification.

Method/Models

We consider two closely related but conceptually distinct drop-out models. Under the first model - 'direct dependence' - drop-out is assumed to depend directly on whether the person moved prior to that wave. Under the second model - 'bivariate probit' - drop-out is assumed to be due to wave-specific normally distributed omitted variables correlated with mobility. Both models are fitted using maximum likelihood and compared to the default complete-cases approach. Candidates for a 'drop-out instrument' (appearing in the drop-out but not mobility model) and 'mobility instrument' (vice versa) are proposed.

Results and Conclusions

We found that previous studies support the use of the statistical direct-dependence model over the econometric bivariate-probit model, but that a sensitivity analysis based on both models is to be recommended. In the application, the non-ignorable models indicate that the standard complete-cases analysis is biased, and the bias is in same direction for both. However, the direct-dependence corrections are slightly larger. More generally, we conclude that exclusion restrictions are most effective for the bivariate-probit model, but can reduce bias for both - especially in applications where the mobility and drop-out probabilities are not close to 0 or 1.

Who, What and Why? Understanding User Needs Wednesday 5 September, 10:10am - 11:30am

Organised by the Statistics User Forum

Every day, statistics are used by a wide range of individuals and organisations to make decisions. The value of statistics lies in their use, and an understanding of user needs is essential to ensure that statistics are of the greatest possible value. To this end, the UK Statistics Authority places great emphasis on improving user engagement for official statistics, and the RSS's Statistics User Forum was established to represent user needs.

This session is intended for anyone with an interest in improving the relevance and impact of their outputs and analyses. With the aid of a case study, it will give insight into the range of user requirements, explore the use of web technology, and highlight some of the challenges faced.

Richard Alldritt: The challenge of understanding the use of official statistics

James Tucker: The Statistics User Forum and StatsUserNet

Tony Cox: The experience of the RPI/CPI User Group

Null Models for Network Data

Patrick Wolfe

University College London, London, UK

Objectives

The analysis of datasets taking the form of simple, undirected graphs continues to gain in importance across a variety of disciplines. We address the issue of null models in this context. Two choices of null model, the logistic-linear model and the implicit log-linear model, have come into common use for analyzing such network data, in part because each accounts for the heterogeneity of network node degrees typically observed in practice.

Method/Models

Here we show how these both may be viewed as instances of a broader class of null models, with the property that all members of this class give rise to essentially the same likelihood-based estimates of link probabilities in sparse graph regimes.

Results and Conclusions

This facilitates likelihood-based computation and inference, and enables practitioners to choose the most appropriate null model from this family based on application context. Comparative model fits for a variety of network datasets demonstrate the practical implications of our results.

Models and estimators for longitudinal network data

Tom Snijders^{1,2}

¹University of Oxford, Oxford, UK, ²University of Groningen, Groningen, The Netherlands

Objectives

The main issue for statistical modelling of social networks (represented mathematically mainly by directed graphs) is how to express the dependencies between the ties in the network. This is less complicated for longitudinally than for cross-sectionally observed networks, because the time-ordering assists in the representation of these dependencies. Here the focus is on longitudinal network data collected in a panel design, i.e., with repeated measures of a network on a given group of nodes.

Method/Models

Stochastic actor-oriented models are a class of continuous-time Markov chain models for representing network dynamics. Tie changes here are modelled as the results of "choices" made by the nodes in the network, governed by multinomial logistic regression models. This allows much flexibility in the representation of dependencies between ties as well as dependence on covariates. Estimation procedures have been developed for network panel data, where the individual tie changes are not observed, and each next observation is the result of a series of tie changes.

Results and Conclusions

Models of network dynamics can also be used to study peer influence, i.e., ways in which the behaviour of individuals is influenced by that of their network partners ("peers"), while also the network ties can be changed depending on the behaviour of the individuals concerned. Examples are adolescent smoking and collaboration between firms. This is achieved by extending the outcome space with variables representing the behaviour of the individuals in the network. The presentation will give an overview of stochastic actor-oriented models and estimation procedures, illustrated by examples.

Role Analysis in Networks using a Mixture of Exponential Random Graph Models

Brendan Murphy, Michael Salter-Townshend

UCD, Dublin, Ireland

A novel and flexible framework for the investigating the roles of actors within a network is introduced.

Particular interest is in roles as defined by local network connectivity patterns, identified using the ego-centric networks extracted for each node in the network.

A mixture of Exponential-family Random Graph Models is developed for these ego-networks in order to cluster the nodes into roles.

An Expectation-Maximization algorithm is used to infer the unobserved cluster assignments and to estimate the mixture model parameters using a maximum pseudo-likelihood approximation. The flexibility and utility of the method are demonstrated on several examples of both simulated and real networks.

Creating interactive web graphs with R: Overview and googleVis tutorial

Wednesday 5 September, 10:10am - 11:30am

Markus Gesmann

The googleVis package [1] provides an interface between R [2] and the Google Visualisation API [3] to create interactive charts which can be embedded into web pages. The best known of these charts is probably the Motion Chart, popularised by Hans Rosling in his TED talks [4]. This talk will give an overview of the package and explain in detail how the functions can be applied to generate interactive charts for web pages and how it can be used as part of a reproducible research workflow.

[1] <http://code.google.com/p/google-motion-charts-with-r/>

[2] <http://www.r-project.org>

[3] <https://developers.google.com/chart/interactive/docs/gallery>

[4] http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html

Analysis of repeat event outcome data in clinical trials: examples in heart failure

Jennifer Rogers, Stuart Pocock

London School of Hygiene and Tropical Medicine, London, UK

Objectives

A composite endpoint (i.e. time to first if several types of disease event) is commonly used as a primary outcome in clinical trials, as it increases event rates and combines multiple outcomes into one, avoiding issues of insufficient power and multiplicity. Heart failure is characterised by repeat hospitalisations for worsening condition, rendering such endpoints suboptimal, as recurrent hospitalisations within individuals are ignored. Utilising all of the hospitalisations within individuals gives a more meaningful treatment effect on the true burden of disease. We will compare analyses of recurrent hospitalisations in clinical trials for heart failure with standard composite endpoints based methods.

Method/Models

This talk presents statistical analyses of two heart failure clinical trials: 'Eplerenone in Mild Patients Hospitalization and Survival Study in Heart Failure' (EMPHASIS-HF) and 'Candesartan in Heart failure Assessment of Reduction in Mortality and morbidity' (CHARM). The datasets will be analysed using the Cox proportional-hazards model for the composite of first hospitalisation and cardiovascular death, and using the Andersen-Gill, Poisson and Negative Binomial models for the repeat hospitalisations. The results of these analyses are compared and bootstrap simulations used to investigate statistical power of each method.

Results and Conclusions

In both datasets we observe that analysing repeat hospitalisations within individuals reduces the risk of heart failure hospitalisations to a greater extent than is captured by only studying the time to first. We also note that analysing all hospitalisations within an individual gives significant improvements in statistical power. Future clinical trials would benefit from incorporating repeat hospitalisations into their primary evaluation of treatment effects. A comparison of hospitalisation rates in heart failure can be confounded by the competing risk of death, as an increase in heart failure hospitalisations is associated with a worsening condition and a subsequent elevated risk of death. Analyses of recurrent events should take into consideration such informative censoring. Statistical methodology being developed to jointly model hospitalisations and death will briefly be presented.

Modelling interactions in high-dimensional data with Backtracking

Rajen Shah

University of Cambridge, Cambridge, UK

Objectives

In many scientific disciplines, it is now common to collect datasets where the number of variables p may be in the order of several thousands, and greatly exceed the number of observations. In these high-dimensional situations, even the number of first-order interactions, $p(p-1)/2$, can be overwhelming, and so most currently available regression methods only fit additive models. However, interactions may well be crucial for modelling the data generating process in a faithful way. Our work aims to provide a computationally feasible way to build regression models which include interactions, for high-dimensional data.

Method/Models

We introduce a new idea, called Backtracking, which can be incorporated into many existing high-dimensional methods which only fit additive models (e.g. the Lasso). It works by iteratively building up an increasing set of candidate interactions, which, along with the main effects, it makes available for selection by the base regression procedure.

Results and Conclusions

Backtracking is statistically motivated, and is computationally fast, making use of parallel processing. Since the method can be built into almost any existing high-dimensional regression technique, it is very flexible. Furthermore, in the special case where we use the Lasso for the linear model as the base procedure, we are able to give some theoretical support for our method.

Regional Variation in the Seasonal Behaviour of Indoor Radon Concentrations

Orlaith Burke¹, Patrick Murphy²

¹University of Oxford, Oxford, UK, ²University College Dublin, Dublin, Ireland

Objectives

Radon levels measured for any duration less than one year are subject to seasonal variation. Such radon measurements must therefore be adjusted using an appropriate seasonal correction factor in order to estimate an annual average radon level. Radon mapping has shown that radon concentrations vary between regions within a country but there has been little work examining regional variation in the seasonality of radon levels.

Method/Models

This paper investigates this regional variation in indoor radon seasonality in Ireland. Regional seasonal correction factors are estimated through Fourier decomposition.

Results and Conclusions

We find that there is significant regional variation in the computed seasonal correction factors. We conclude that the use of one national set of seasonal correction factors may not be appropriate in all countries.

Understanding Probability Matching Priors

Alastair Young¹, Thomas DiCiccio², Todd Kuffner³

¹Imperial College London, London, UK, ²Cornell University, Ithaca, USA, ³American University of Beirut, Beirut, Lebanon

Objectives

In Bayesian parametric inference, in the absence of subjective prior information about the parameter of interest, it is natural to consider use of an objective prior which leads to posterior probability quantiles that have the correct frequentist interpretation, at least to some higher-order approximation in terms of the sample size. Such priors are termed probability matching priors. The conventional analysis of probability matching expresses the prior as the solution of a system of partial differential equations. Our objective here is to provide a more useful, direct statistical interpretation of conditions for probability matching priors.

Method/Models

We consider problems concerning inference on a scalar interest parameter in the presence of a nuisance parameter and analysis of the Bayesian and frequentist distributional properties of the signed root likelihood ratio statistic. After consideration of general models, focus is given to models where the appropriate frequentist inference is a conditional one, in particular those admitting ancillary statistics, allowing development of issues relating to appropriate conditional probability matching. The theoretical analysis is combined with (conditional) simulation methodology to evaluate the usefulness of objective Bayes as a route to higher-order accurate frequentist inference.

Results and Conclusions

If we choose the prior distribution in a Bayesian analysis so that the frequentist and Bayesian posterior means of the signed root statistic match to first-order in the sample size, then the prior achieves first-order probability matching. Second-order probability matching is seen to correspond to matching of the frequentist and Bayesian variances of a very specific mean-adjusted version of the signed root statistic. A prior which is first-order probability matching in a marginal sense is necessarily first-order conditional probability matching. Second-order conditional probability matching is typically achieved only under very particular prior specification, though by this route exact conditional probability matching can sometimes be achieved. More generally, higher-order frequentist accuracy, conditional or marginal, by probability matching demands choice of prior from a class that can yield noticeable different accuracies of approximation in practice.

Financial Forecasting using the Generalised Kolmogorov-Feller Equation

Jonathan Blackledge

Dublin Institute of Technology, Dublin, Ireland

Objectives

The objective is to develop a model for a financial time series that is based on the Generalised Kolmogorov-Feller Equation (GKFE). Compared to the classical diffusion and fractional diffusion equations, the GKFE provides a fully consistent model for the 'density function'. However, the equation depends on two component functions: (i) the PDF used to describe the spatial evolution of the density function; (ii) a 'memory function' which determines the scale of the long term correlations associated with the time evolution of the density function. The aim is to characterize a financial time series in terms of this memory function.

Method/Models

By using the Green's function method to provide a general solution to the GKFE, expressions for the Impulse Response Function and the Correlation Function are derived. For a Levy distributed system characterized by a Mittag-Leffler (ML) memory function, solutions are considered that relate the density function to the ML index. The relationship between the time dependent characteristics of this index and the dynamic behaviour of a financial time series are then investigated using a standard moving window process. This process is extended to include an ML index analysis of Stochastic Volatility.

Results and Conclusions

The solution developed and the algorithm designed is applied to a range of financial time series including macroeconomic series such as the FTSE, currency pair exchange signals and energy commodities data. It is shown that an application of the ML index to a financial time series and its corresponding Stochastic Volatility provides a clear correlation between the start, direction and end of a trend with a time lag that depends upon the sampling rate of the time series and the look-back window that is used. In terms of a trading strategy, the model yields a forecasting efficiency of ~70% providing the data is well sampled and the trends are well correlated. Compared with existing algorithmic trading strategies, this result clearly demonstrates the statistical advantage of the method proposed which will be demonstrated for high frequency currency trading during the presentation using a system called *MetaTrader4*.

A Bayesian one-dimensional random field model for ChIP-seq data.

Veronica Vinciotti¹, Yanchun Bao¹, Ernst Wit², Peter 't Hoen³

¹*Brunel University, Uxbridge, UK*, ²*Groningen University, Groningen, The Netherlands*, ³*Leiden University Medical Centre, Leiden, The Netherlands*

Objectives

An important biological question is the one of detecting the regions in the genome bound by transcription factors, as these give insight into the mechanism of gene regulation. ChIP sequencing (ChIP-seq) is a biological method to detect these, by generating sequence reads at the positions bound by a transcription factor. In this talk, we present a statistical model to detect the enriched regions from ChIP-seq data.

Method/Models

In the absence of binding, reads would be expected to be scattered across the genome, following some non-specific binding pattern which generates a background signal. Statistically, the observed count data, summarised into bins across the genome, are assumed to come from either a background or a signal distribution. The interest is in inferring the state of the latent binary variable: "Enriched" or "Not Enriched". It is realistic in this application to assume a Markov property, whereby the probability of a region being enriched depends on the two neighbouring ones. We present a random field model that can capture these assumptions.

Results and Conclusions

The parameters of the model are estimated in a Bayesian framework, using either a Poisson or a Negative Binomial distribution for the counts. An extensive simulation study shows that the model is competitive with existing methods, which are either based on a simple mixture model (without latent states) or on a uni-directional Hidden Markov model. Finally, the approach is used to detect the binding sites of two transcription factors, p300 and CBP, for which six ChIP-seq datasets have been recently generated, as well as three further datasets from a previous study. We show how the model accounts for the different antibody efficiencies associated to the different antibodies used in the experiments. This aids the biological interpretation of the bound and differentially bound regions.

Log-concavity and its Application to Time Series Models

Yining Chen

University of Cambridge, Cambridge, UK

Objectives

In many existing time series models, it is assumed that the innovations are Gaussian. Statistical inference is then typically based on the idea of maximum likelihood estimation. However, time series with non-Gaussian innovations frequently occur in practice.

Motivated by recent developments in shape-constrained density estimation, we assume here that the innovations have a log-concave density. (i.e. the logarithm of the density function is concave) The aim is to develop a new framework which allows us to simultaneously estimate both the parameters and the density of the innovations (nonparametrically) for a wide class of time series models.

Method/Models

Denote a class of separable time series models by (f, θ) , where f is the density function of i.i.d innovations, and θ is its parametric part.

We propose the following semiparametric procedure to estimate both the density function and parameters:

1. For any given ϑ in the parameter space, compute its corresponding innovation estimates based on the observations. Then use the nonparametric log-concave maximum likelihood density estimator to estimate f and compute the likelihood.
2. Find out ϑ and f that maximizes the likelihood.

Results and Conclusions

Our method is free of tuning parameters and can be adapted to many popular existing time series models with only minor modifications.

- **Theoretical properties**
For ARMA, GARCH and ARMA-GARCH models, we show that if the true density is log-concave, then our estimators are consistent. Even if the true density is not log-concave (e.g., in the infinite variance ARMA), our estimators for the parameters are still consistent. The flexibility and robustness of this semiparametric procedure make it highly desirable in practice.
- **Simulation studies**
The simulation study we ran confirms the improved finite sample performance of our proposed estimator over the Gaussian quasi-maximum likelihood estimator in the setting of non-Gaussian innovations. (i.e. 50%-90% reduction in rooted mean squared errors)
- **Real data analysis**
We apply our framework to two real data sets concerning the streamflow of the Hirnant river and the FTSE daily return, where the Gaussianity assumption is strongly violated.

Education for All: Household and community level determinants of primary and secondary school attendance among children of official school going age in Sierra Leone

Mamusu Kamanda

University of Southampton, Southampton, Hampshire, UK

Objectives

The objectives of the research are to:

1. Explore the association between school attendance and household environment.
2. Investigate the association between school attendance and community environment.
3. Explore interaction effects between household and community environment, and between individual characteristics and community environment on school attendance.
4. Investigate whether the impact of community education and wealth operate above the household levels, and whether this association varies between communities.

Method/Models

Multilevel logistic regression was applied to the 2008 Sierra Leone DHS household members' datafile. Attending school at the official age (6-11 years for primary education, and 12-17 years for secondary education) was modelled as a function of individual, household, and community level predictors. Six models were fitted: (1) null model with random intercept at the household and community levels (2) individual level fixed effects are added to model 1 (3) household level fixed effects are added to model 2 (4) community level fixed effects are added to model 3 (5) interaction terms are explored between fixed effects in model 4 (6) Random coefficients are explored for education and wealth at community level.

Results and Conclusions

The environment in which a child lives is a significant predictor of whether he/she attends school. Household level predictors: education of the headship and wealth have the largest effects on school attendance. Both predictors are positively correlated with the outcome ($p < .01$). Having a female headship also has a significant positive association with attendance whilst living with many young children is inversely related to the outcome ($p < .05$). Community level predictors: the proportion of educated women has a positive effect on attendance; the effect is more pronounced among older children ($p < .01$), poorer households ($p < .01$), and in rural areas ($p < .01$). Living in a deprived community as well as living in a community in which the predominant livelihood is agriculture has an inverse relationship with school attendance. The association with community female education varied significantly between communities ($p < .05$); the random coefficient for community deprivation was not significant. Recommendations: efforts should be made to promote female education and develop community infrastructure, especially in rural areas.

Street lighting & safety

Paul Marchant

Leeds Metropolitan University, Leeds, UK

Objectives

To examine the effect of changing street lighting illumination, e.g. brightening & whitening or alternatively reducing it, on public safety, e.g. on occurrence of crime or road traffic accidents. Earlier work by the author on summated 'key' police reported crime at a local authority level suggests that new Private Finance Initiative (PFI) lighting has not reduced crime by the 20% that has been claimed (from a review of experiments) & indeed the new PFI lights may be associated with increased crime.

Method/Models

The presentation will extend the previous results to individual crime types, rather than just key crime, through modelling the joint time series of the number of crimes over several years. The variable of particular interest in the model is that of the proportion of a relighting programme completed at the time point. Additionally an analysis of road traffic accident data, done at the small scale of the vicinity of individual streetlights, will be presented.

Results and Conclusions

This is work in progress and the results have not yet been fully established. The results of previous work at local authority level suggest that the factor by which crime is increased has a 95% confidence interval of (0.87, 1.11). An analysis with a longer series gives a similar result but somewhat tighter CI, as might be expected. The work at the level of the vicinity of street lighting is in progress at the time of writing this abstract.

Localised spatial smoothing in Bayesian disease mapping

Duncan Lee, Richard Mitchell

University of Glasgow, Glasgow, UK

Objectives

Disease maps display the spatial pattern in disease risk, so that high-risk clusters can be identified. They are produced using Bayesian hierarchical models, where the spatial structure in disease risk is represented by random effects modeled with a conditional autoregressive (CAR) prior. Such models include a global spatial correlation parameter, whereas the risk surface will include areas of smooth evolution as well as discontinuities. The latter may occur where rich and poor communities meet, and thus have different risks despite being geographically adjacent. Therefore we propose an extension to CAR priors to capture such localized spatial smoothness.

Method/Models

CAR priors represent spatial dependence via a fixed neighbourhood matrix W , where $w_{ij} = 1$ if areas i and j share a common border and is zero otherwise. Here, we propose modeling the w_{ij} as random quantities, which allows for localized spatial smoothing because two areas sharing a common border can have random effects that are spatially correlated ($w_{ij} = 1$) or conditionally independent ($w_{ij} = 0$). The talk discusses different approaches for modeling the w_{ij} , including treating them as hyperparameters, which are updated conditional on the values of the remaining model parameters as part of a two-stage iterative estimation algorithm.

Results and Conclusions

We tested the efficacy of our methodology via simulation, which showed that it outperforms the existing CAR priors in terms of bias and RMSE. We then present a study mapping the risk of hospitalization due to respiratory disease in Greater Glasgow, Scotland. Greater Glasgow is an ideal study region to illustrate our methodology, as there are many locations where rich and poor communities live side by side. The results of this study suggest that respiratory disease risk is far from being spatially smooth, with 27% of borders between neighbouring areas not corresponding to a spatially smooth evolution in risk.

The Impact on Post-Imputation Parameter Estimation of Limiting Donor Usage in Hot Deck Methods

Dieter William Joenssen, Udo Bankhofer

Ilmenau University of Technology, Ilmenau, Germany

Objectives

Hot deck procedures rely on finding a "similar" object for supplying values to be imputed for other observations' missing data. Recommendations are to simply limit the donor usage frequency as this reduces the probability of using outliers or other values "too often". In contrast to this would be that, an "optimal" donor could not be used if its usage limit has been reached. Both points made are valid, suggesting an optimal donor limit based on a given situation. As systematic donor limit variation and its impact on parameter estimation has not been investigated it is prudent to do so.

Method/Models

To investigate the effect a donor limit or variation thereof has, data matrices with nine variables of mixed scale are generated using US Census Bureau data (2009 ACS) as a super population model. Using four imputation classes and varying the amount of observations in each class (50, 100, 250), the complete data is subjected to either 5, 10 or 15% missingness under MCAR, MAR, or MNAR mechanisms. The resulting data is imputed using six hot deck type procedures with different donor limits. Differences in the mean square error for the parameter estimates when using different donor limits are compared.

Results and Conclusions

Usage of donor limited hot deck imputation has a distinctive impact on the quality of post imputation parameter estimates in some situations. When a random hot deck method is used or the missingness mechanism is NMAR, a limit on donor usage does not impact parameter estimation noticeably. Under the MAR missingness type, two imputation procedures perform better without a donor limit, while the others exhibit no significant advantage for either donor limit. When data is missing MCAR, two methods perform better when donor usage is limited to once, while other methods show mixed results. These methods will estimate measures of association better when donor usage is not limited and univariate parameters better when donor usage is limited. Any advantages are intensified by an increasing number of observations per class and the amount of missing data.

Given these findings, clear recommendations as to when a donor limit is sensible can be made.

Making the most of the 2011 Census outputs for England and Wales

Wednesday 5 September, 12:00 - 1:20pm

2011 Census statistical releases

Emma White

Office for National Statistics, Fareham, Hampshire, UK

Objectives

The first results from 2011 Census are planned for release in July 2012. This presentation will highlight some key messages and trends from the results.

Due to the depth and breadth of the data, statistics on the usually resident population will be released in four stages. An overview of publication plans for each release will be given, including dissemination channels.

UK level outputs, and those for specialist products such as small populations tables, microdata, flow data and alternative population bases will also be discussed, in the context of the design and development of the full 2011 Census product suite.

Method/Models

Consultation and decision-making methodology in the face of competing requirements for census outputs will be discussed, including the detail of steps in user engagement, statistical disclosure control and seeking agreement across the UK on harmonisation and release policy. Outputs quality assurance activities will be described, as will the thinking behind statistical commentary and analysis plans.

Results and Conclusions

Published results from England and Wales 2011 Census will be described and discussed, while the content of the next statistical releases will be outlined, including where detailed estimates of population characteristics such as religion will be made freely available for the first time.

The range, scope and detail of the UK statistical releases and England and Wales specialist products, coupled with the four releases on the usually resident population base, mean that the product suite for 2011 Census includes more outputs on more topics and population bases in more detail than has ever before been made available as standard from a census in the United Kingdom.

The progress and plans for census outputs will be considered in the context of the National Statistics Code of Practice, the Statistics and Registration Services Act 2007, the Registrars' General Agreement 2006, and the fitness for purpose of these National Statistics.

Making the most of the 2011 Census outputs for England and Wales

Wednesday 5 September, 12:00 - 1:20pm

2011 Census - quality assuring the results

Garnett Compton

Office for National Statistics, Fareham, Hampshire, UK

Objectives

The first results from 2011 Census will be published in July 2012. The methods and processes for ensuring that the quality of the census results are fit for purpose has been a key strand of the census. This presentation will report results for the key quality targets for the 2011 census and the extensive quality assurance process that was undertaken as part of producing the census results.

Method/Models

This paper will present an overview of the demographic methods which were carried out to give confidence in the robustness of the estimates.

Results and Conclusions

The session will describe and discuss some of the early analysis of census quality and key processes for quality assuring the results of the 2011 Census. An overall assessment of census quality as measured against our quality indicators, including response rates and confidence intervals around the census population estimates, will be presented. In addition, the wide range of checks employed in the quality assurance process will be demonstrated. In particular the use of administrative and other survey sources to assure ourselves that the census population and household estimates are robust when comparing census estimates of total population; babies; school children; the elderly; resulting fertility and mortality rates; estimates of occupied households; and, household size. The presentation will draw on case study examples that demonstrate the work carried out and describe some of the adjustments which were made as a result of what was found.

Making the most of the 2011 Census outputs for England and Wales

Wednesday 5 September, 12:00 - 1:20pm

Improving access to Office for National Statistics (ONS) data

Francesca Kay, [Nick O'Donnell](#)

Office for National Statistics, Fareham, Hampshire, UK

Objectives

The release of the 2011 Census Outputs begins in July 2012. Because of the volume of data, outputs will be released in four stages.

This session will provide information about the improved tools being made by ONS to provide users with enhanced access to data, with a particular focus on the dissemination of outputs from the 2011 Census.

Planned developments include the provision of tools on the ONS website to find, download and explore ONS data online, as well as new opportunities for easier re-use of data by other systems through provision of an Application Programming Interface (API).

Method/Models

Extensive engagement with external organisations to develop the dissemination tools. In addition a widespread period of testing with a variety of users will be discussed.

Results and Conclusions

An outline of the census data and other ONS datasets being disseminated will be discussed. This is the first time that ONS will be using its website as the primary method of dissemination. The methods used will enable data to be disseminated in a way that maximises usage, is in line with government policy on transparency and meets the requirements of EU legislation.

The project being undertaken will deliver an enhanced website capability for organisations and individuals to use and explore ONS statistics more effectively. It will open up direct access to ONS data sources to power their own websites and applications.

Statistical nous

Wednesday 5 September, 12:00 - 1:20pm

Jenny Church

Statistical Consultant

Statistics and Computing are now indelibly linked, but we still need the human brain to check over outputs with the question - Does this make sense? This session, aimed at career young statisticians, will provide examples of publications where statistical nous wasn't applied, and some techniques to help you develop the skill of looking at an output and thinking - 'There's something wrong here.'

Introduction by Paula Williamson (University of Liverpool)

An overview of the MRC HTMR Network

The Hubs for Trials Methodology Research (HTMR) were established by the MRC to create a UK-wide regionally distributed research resource to improve the design, conduct, analysis, interpretation, and reporting of clinical trials.

The eight Hubs have a variety of methodological expertise and close links with clinical trials units and other methodological groups in universities, industry and relevant professional bodies and organisations. The Hubs were networked together to enhance their individual activities, strengthen the methodological research further, and aid dissemination and education relating to good methods in clinical trials.

The HTMR Network aims to promote and encourage collaborative methodological research relevant to trials and to enable implementation of the most effective and appropriate methods to improve the quality of trials and, ultimately, patient care. In order to achieve this, there are four interconnected areas in which the Network adds value:

- 1) Promoting high quality collaborative methodological research, both between Hubs and with other groups, UK-wide and internationally
- 2) Providing methodological advice to the clinical trials community
- 3) Encouraging the implementation of the most effective and appropriate methodological practice, for example by providing a coordinated package of education and training
- 4) Working with stakeholders, in particular to agree on shared priorities for research and guidance and to advocate for improvements in the conduct of clinical trials

The TAILoR study - an example of multi-arm multi-stage clinical trials

Thomas Jaki, Dominic Magirr, John Whitehead

Lancaster University, Lancaster, UK

Objectives

The TAILoR study aims to investigate the effectiveness of three different doses of an already licensed drug on insulin resistance in HIV positive patients. The aim is to efficiently identify all doses significantly reduce insulin resistance without assuming a dose response relationship.

Method/Models

A multi-arm multi-stage (MAMS) clinical trials design is used to evaluate the different doses of the treatment. Reduction in insulin resistance as measured by the HOMA-IR scale is modelled as a normally distributed variable.

Results and Conclusions

The developed approach is an extension of group sequential methodology that involves joint monitoring of the correlated, sufficient test statistics. We show the heavy computation involved can be dramatically reduced when the number of stages is small and the number of treatments is large. When designing the trial, sample size can be chosen to satisfy a power requirement based on the least favourable configuration of treatment effects. At each analysis, simultaneous confidence intervals are easily found such that the overall coverage is at the required level.

Clinical trial monitoring: towards establishing best practice?

Catrin Tudur Smith¹, Julie Bakobaki², Gill Booth³, Joanne Eatock¹, Carrol Gamble¹, Martin Landray⁴, Athene Lane⁵, Sarah Meredith², Paula Williamson¹

¹*University of Liverpool, Liverpool, UK*, ²*MRC Clinical Trials Unit, London, UK*, ³*University of Leeds, Leeds, UK*, ⁴*University of Oxford, Oxford, UK*, ⁵*University of Bristol, Bristol, UK*

Objectives

Clinical trial monitoring is a multi-disciplinary, multi-layered activity with substantial heterogeneity in approaches. Much needed progress has been made towards encouraging risk proportionate approaches to monitoring but the lack of evidence base to support the use of different methods presents challenges to clinical trial practitioners. The MRC Hubs for Trials Methodology Research (HTMR) funded a project to bring together the UK Clinical Research Collaboration (UKCRC) registered clinical trials units and the HTMR to discuss current practice, share ideas and experiences, and identify where, and how, the group can collaborate to facilitate advancement towards establishing best practice in clinical trial monitoring.

Method/Models

An on-line survey has been conducted to identify current practice across UK clinical trials units, followed by a one-day research workshop.

Results and Conclusions

A risk assessment to determine the level of trial monitoring, central monitoring and on-site monitoring is undertaken at most of the UKCRC registered trials units that responded to the survey. Within these broad categories there are a variety of approaches being adopted, some which were described at the workshop. There was positive support from the group to continue to share ideas and experiences and progress towards establishing best practice in clinical trial monitoring.

Asymptotic results and statistical procedures for time-changed Lévy processes sampled at hitting timesMathieu Rosenbaum², Peter Tankov¹¹Université Paris-Diderot, Paris, France, ²Université Pierre et Marie Curie, Paris, France**Objectives**

The goal of this study is to provide asymptotic results and develop statistical estimation procedures for time-changed Lévy processes sampled at random instants. The sampling times are given by first hitting times of symmetric barriers whose distance with respect to the starting point is equal to ε . Time-changed Lévy processes are often used to introduce a stochastic volatility component into a Lévy model and our sampling scheme can be seen as a representation of microstructure noise in such a model.

Method/Models

For a wide class of Lévy processes, we introduce a renormalization depending on ε , under which the Lévy process converges in law to an α -stable process as ε goes to 0. The convergence is then extended to moments of hitting times and overshoots.

Results and Conclusions

Our theoretical convergence results are used to build high frequency statistical procedures for our model. As examples we construct consistent estimators of the time change and, in the case of the CGMY process, of the Blumenthal-Gettoor index. Convergence rates and a central limit theorem for suitable functionals of the increments of the observed process are established under additional assumptions.

STATISTICAL ASPECTS OF MODELLING AND PREDICTION FOR FINANCIAL TIME SERIES

Nick Bingham

Imperial College, London, UK

1. Semi-parametric models
2. Multivariate elliptic processes
3. Prediction in general: Szeg\`o theory
4. ACF and PACF: Verblunsky's theorem
5. Szeg\`o's theorem
6. Discrete and continuous time
7. Stationarity v. non-stationarity

Improving penalised regression models for phenology by using thermal time

Adrian Roberts

Biomathematics and Statistics Scotland, Edinburgh, UK

Objectives

Phenology is the study of the timing of cyclical natural events, e.g. flowering, and how they are influenced by weather. Temperature is usually the key determinant. Recently Roberts (2008) has proposed the use of penalised signal regression (PSR) for phenology; this allows the use of daily temperatures, producing easily interpretable curves of regression coefficients, and outperforms the commonly-used stepwise regression (Roberts, 2012). Empirical models based on thermal time are sometimes used for plant first flowering and leafing. Here we aim to improve the performance of PSR by using the concept of thermal time.

Method/Models

PSR enables linear regression on many highly correlated covariates by penalising differences between adjacent covariates. We use a particular variant of PSR proposed by Marx and Eilers (1999), which uses a P -spline formulation to reduce dimensionality. In spring warming models, the response (timing) has a non-linear relationship with the temperature on a particular day. Temperatures below a threshold, α , have no effect, whereas temperatures above α have a linear effect. Using this function, we transformed temperature to give thermal time covariates for use in PSR. α was estimated by cross-validation. The performance of this modified PSR was assessed through simulation.

Results and Conclusions

Simulations from spring warming models showed that the modified PSR performed as well as or better than PSR on temperature and PSR on thermal time where α was set at 0°C . The improvement was greatest when the actual value of α was higher, reflecting the increased relevance of the threshold. With higher values of α , the fit of the PSR model became worse, using either temperature or thermal time. The estimates of α , found by cross-validation, were biased upwards. The modified PSR also performed better in simulations from sequential models (another type of empirical model with a chilling effect added).

The results from the simulation showed that the modified PSR using thermal time should be an improvement over PSR on temperature for studying phenological relationships.

A framework for adapting survey design through time for design-based wildlife population assessment

Fiona Underwood

University of Reading, Reading, UK

Objectives

Long-term wildlife conservation monitoring programs that aim to obtain design-based estimate of the status and trends in wildlife population abundance often face challenges in the survey design. At the start of the program it may be unclear what a good survey design would be and how auxiliary information can be used to determine strata or inclusion probabilities for sample design. After one or more surveys information about the species distribution becomes available and a map of species distribution produced. This information can then be used in future survey design.

Method/Models

After a first survey using simple random sampling (srs) a predicted map of expected abundance can be obtained and used to influence future design. A strategy is proposed in which part of the sample is sampled using srs and the rest with inclusion probability proportional to predicted abundance. Two-phase sampling ideas are used to obtain estimates of total abundance. The method is compared with an optimal allocation stratification strategy and a strategy where all of the units are selected with inclusion probability proportional to predicted abundance. Robustness of the methods to poor estimates of predicted abundance is investigated.

Results and Conclusions

The proposed strategy gives more precise estimates of the population total than standard design based methods. Stratification also performs well. The strategy is more robust than one in which all units are selected based on predicted abundance. This is particularly true if the design is repeated over time when each sample contributes both to design-based estimates of the total (and trend) and to future maps of predicted abundance. Diagnostics to determine when such designs may be more efficient than srs are given.

Data fusion modeling for archaeological geophysics

Robert Aykroyd¹, Salem Al-Gezeri²

¹University of Leeds, Leeds, UK, ²University of Garyounis, Benghazi, Libya

Objectives

Remote sensing methods are widely used in archaeological investigations to provide an initial description of an archaeological site. This will involve noninvasive surface measurements which give partial information about the subsurface. Traditional regularized least squares techniques require substantial, and possibly unrealistic, assumptions to be made. A key restriction is caused by the limited depth information in the surface data. Vertical cores provide an alternative source of depth information. The aim is to combine these two data sources to produce improved 3d reconstruction.

Method/Models

The analysis is divided into two stages: (i) the analysis of core data and (ii) the reconstruction of the subsurface using surface data and including output from the core data analysis. A layer model is proposed for the core data analysis with prior distributions for unknown parameters along with a noise model. The resulting posterior is investigated using MCMC methods. The posterior distribution is then included in the modeling of the subsurface to provide key depth information. The subsurface reconstruction remains an over-parameterized inverse problem which is regularized using a Mrf prior. Again estimation is via MCMC sampling.

Results and Conclusions

A preliminary analysis of the core data shows that not all parameters are separately estimable leading to an alternative re-parameterization. Reliable estimation of key parameters is now possible with the output summarized by corresponding marginal posterior distributions. These marginal distributions are used as prior models on depth allowing a full reconstruction of the site subsurface. The analysis is performed using a mixture of real archaeological data and simulated data based on site excavation records. Results are compared to previous site reconstructions indicating increased accuracy and easier interpretation.

Dynamic Bayesian Smooth Transition Autoregressive (DBSTAR) models

Alexandre Santos, Alvaro Faria

The Open University, Milton Keynes, UK

Objectives

Our main contribution in this research is an alternative formulation to the numerical Bayesian models that uses a polynomial approximation to the nonlinear transition function to obtain an algorithm for prior-to-posterior updating based on the Kalman filter, similar to a Dynamic Linear Model (DLM) of West and Harrison (1997).

Method/Models

In this research we describe Dynamic Bayesian Smooth Transition Autoregressive (DBSTAR) models which extend the classical Smooth Transition Autoregressive (STAR) models proposed by Chan and Tong (1986) for nonlinear time series processes. For parameters estimation, we use the Kalman filter with updating equations for recursive computations. This solution estimates the parameters in analytical closed form sequentially in time.

Results and Conclusions

The DBSTAR models do not require extensive historical data for parameter estimation and allow expert intervention via prior distribution assessment for model parameters, for being Bayesian. For being analytical, the DBSTAR models avoid computational problems associated with the numerical Bayesian formulations for STAR models by Lopes and Salazar (2005) and allow fast estimation of the dynamic parameters for real time applications. In addition, the DBSTAR models can be applied to time series data sets when stationarity, heteroscedasticity and/or seasonality are present. The application of a DBSTAR model to the Canadian Lynx data showed improved fitting performance when compared with both the classical and the non-sequential numerical Bayesian STAR models.

A short guide and a forest plot command (*ipdforest*) for one-stage meta-analysis

[Evangelos Kontopantelis](#), David Reeves

University of Manchester, Manchester, UK

Objectives

When individual patient data (IPD) are available, meta-analysts can avoid problems inherent to the two-stage approach and use a mixed-effects regression model to combine information across studies in a single stage, arguably the best meta-analysis approach.

However, one obvious advantage of two-stage meta-analysis is the ability to convey information graphically through a forest plot. Forest plots are informative, easy to follow and particularly useful for readers with little or no experience in meta-analysis methods.

Our aim was to develop a command in Stata that provides a forest plot for IPD meta-analyses.

Method/Models

The *ipdforest* command follows a random-effects IPD meta-analysis conducted using a linear logistic two-level regression, with patients nested within trials, and provides a meta-analysis summary table and a forest plot. Trial effects are calculated within *ipdforest* while the overall effect and variance estimates are extracted from the preceding regression.

The command estimates individual trial effects using one-level linear or logistic regression analyses. The command controls these regressions for fixed- or random-effects covariates that were specified. Patient weights are uniform and therefore each trial's weight is the ratio of its participants over the total number of participants across all studies.

Results and Conclusions

We provide a summary of available models and their implications and a practical guide for conducting IPD meta-analysis in Stata. Examples with the *ipdforest* command are provided, for both main and interaction effects.

The command aims to help meta-analysts in better communicating their results through the familiar and distinctive forest-plot, which cannot be obtained as standard in one-stage IPD meta-analysis.

Power analysis for generalised linear mixed models using simulations

Paul Johnson¹, Sarah Barry¹, Heather Ferguson², Carl de Wet³

¹Robertson Centre for Biostatistics, University of Glasgow, Glasgow, UK, ²Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK, ³NHS Education for Scotland, Glasgow, UK

Objectives

Generalised linear mixed models (GLMMs) are widely used to model binomial and count data where non-independence due to random effects or overdispersion is expected. Conventional power analysis for such models is challenging. The alternative approach of Monte Carlo simulation has several advantages, including flexibility, asymptotic exactness and the opportunity to validate analysis methods. The objective was to develop Monte Carlo methods for estimating power for GLMM analysis, and apply them to two areas of medicine where investigation is hampered by a lack of robust power analysis methods: testing interventions against disease-transmitting insects (malaria mosquitoes) and measuring preventable harm in general practice.

Method/Models

Data were simulated from GLMMs with binomial or Poisson responses using an R function written by the authors. Lognormal random effects and overdispersion were assumed. Two real study design scenarios were used to assess the methods: (1) field trials to evaluate the efficacy of tools to reduce the abundance (Poisson response) and increase the mortality (binomial response) of mosquitoes that spread malaria; (2) monitoring preventable harm in general practice (Poisson response). In both scenarios, clustering of events within hierarchical levels motivates GLMM analysis. Outcomes of interest were power, precision, bias, type I error rate and confidence interval coverage.

Results and Conclusions

At high event rates, and when overdispersion was low, power from binomial and Poisson GLMMs broadly matched analytical predictions, even when random effects were strong. Indeed, in some scenarios power was increased by strong random effects. However, strong overdispersion tended to reduce power substantially. The preventable harm study illustrated the difficulty of accounting for random effects (inter-practice variation) and overdispersion (inter-patient variation) when event rates are expected to be low. In both studies the simulation-based approach to power analysis enabled assessment of the sensitivity of the study designs to varying degrees of non-independence. In conclusion, Monte Carlo simulation provides a powerful and flexible tool for aiding study design, both by allowing accurate power estimates for GLMMs and by allowing the validity of the intended analysis methods to be assessed.

The evaluation of evidence in forensic science for discrete data

Colin Aitken¹, Erica Gold^{0,2}

¹University of Edinburgh, Edinburgh, UK, ²University of York, York, UK

Objectives

The objective is the evaluation of evidence for discrete data. The proposed method compares the likelihood of the evidence if the prosecution proposition is true with the likelihood of the evidence if the defence proposition is true. When the evidence is in the form of measurements, methods are well-developed for multivariate, hierarchical Bayesian multivariate random effects models. Methods are not so well developed for discrete data. The objective is to investigate the performance of possible models for such data

Method/Models

Data are provided from a project in forensic phonetics at the University of York in which the number of clicks per minute are recorded for each of 100 speakers, over a period of time ranging from four to six minutes.

The likelihood method will be investigated for multivariate, hierarchical discrete data.

Methods based on Dirichlet multinomial models, Poisson-gamma models, multivariate Poisson models and an empirical model, based on differences in counts between known and unknown source and on relative frequencies will be investigated.

Results and Conclusions

The evidence may be considered to be the number of clicks from a piece of speech from an unknown source and the number of clicks from a piece of speech from a known source, such as a suspect.

The support for the proposition these two pieces of speech were made by the same speaker versus the proposition they were made by different speakers is investigated.

Results obtained to date (early April 2012) provide intuitively reasonable results in that data for control and recovered sources that are close provide results which are supportive of speech from the same person and data that are far apart provide results which are supportive of the two pieces of speech coming from different people.

Organised by the RSS Research Section

Sir David Cox (University of Oxford)

Further Results on Tests of Separate Families of Hypotheses, JRSSB 24, 406-424, 1962

RSS Retrospective Read Papers recognise works which have made a seminal contribution to Statistics and were published in RSS journals, but were not originally presented as read papers. Previous retrospective read papers include the paper originally introducing the False Discovery Rate and the paper introducing the LASSO. This 3rd retrospective read paper is another prime example of such a paper. 50 years ago this year, Prof Sir David Cox published the paper "Further results on tests of separate families of hypotheses" in the JRSSB. This paper had a seminal influence in many areas of statistics and econometrics and is one of the most highly cited of all non-read RSS papers. This session will discuss the paper and its lasting contributions. In addition to the presentation by Prof Cox, Prof John Kent (Leeds) and Prof Grayham Mizon (Southampton) will propose and second the vote of thanks. The original 1962 paper can be downloaded at <http://stats.lse.ac.uk/fryzlewicz/RSS/>

James Carpenter (LSHTM)

Health Records are rich sources of medical information, but there are challenges in drawing casual inference

Routine, electronically collected, health record information is being increasingly used in medical research, complementing experimental and bespoke observational studies. While the volume of data available is enormous, and rapidly increasing, the fact it has been collected primarily for clinical purposes, rather than research, poses a range of statistical and associated computational challenges. In this talk, I will sketch some of these, and then discuss the potential of multiple imputation and related approaches to address them. I will illustrate with the application of multiple imputation to missing data and record linkage issues in this setting.

Studying Crop Disease and Soil: Evolutionary Analyses of Agricultural Genomic and Meta-genomic Data

Elisa Loza

Department of Computational & Systems Biology, Rothamsted Research, Harpenden, Herts., UK

Objectives

Rothamsted Research is the longest running agricultural research station in the world, with field experiments that have been operating, uninterrupted, for over 160 years. Our long-term experiments and the data collected from them are resources of immense value; analyses of both archived and new material continue to provide unique insights into the processes at the core of plant-based food production.

In this talk I will focus on the statistical methods that we employ at Rothamsted to answer scientific questions. I will concentrate on the statistical analysis of large genomic and meta-genomic (pertaining to one and several organisms' genomes, respectively) datasets.

Method/Models

In collaboration with Rothamsted's Molecular Microbial Ecology Group and the European Bioinformatics Institute, I have investigated the structure of soil microbial communities using molecular sequence data sampled from agricultural fields. We have applied simulation techniques, phylogenetic models and dynamic programming algorithms to identify the taxonomic and functional diversity of microbial communities in Rothamsted soils.

In a separate study, I have applied statistical methods to detect adaptive evolution of secreted effector proteins in *Mycosphaerella graminicola*, a fungal pathogen of wheat with huge economic importance due to its potential devastating effects. This is joint work with Rothamsted's Wheat Pathogenesis Group.

Results and Conclusions

I will present results from our on-going studies of soil metagenomics and fungal pathogen genomics at Rothamsted, and the insights these studies have produced.

Inferring seabird distribution and behaviour using data from electronic tags

Adam Butler¹, Ellie Owen³, Mark Bolton²

¹*Biomathematics and Statistics Scotland, Edinburgh, UK*, ²*Royal Society for the Protection of Birds, Sandy, Cambridgeshire, UK*, ³*Royal Society for the Protection of Birds, Inverness, UK*

Objectives

Seabirds are a familiar but elusive part of our natural environment, and a vital part of the marine ecosystem. We need to be able to understand key aspects of their behaviour - such as breeding and foraging - in order to assess the likely impacts of fisheries, offshore renewable projects or environmental legislation upon them. Behavioural data have historically been very difficult to collect, but the availability of relatively cheap electronic tags means that it has now become feasible to collect large quantities of behaviour-related data.

Method/Models

In this talk we will outline some of the statistical and computational challenges involved in using electronic tag data to study the distribution and behaviour of seabirds, and illustrate some of the statistical approaches that can be used in dealing with these using data collected during 2010 and 2011 within the "FAME" project.

Results and Conclusions

The first key methodological challenge lies in using data on speed, turning frequency and dive depth to draw inferences about the (unknown) behavioural state of each bird at each location, and we compare the classifications obtained from Hidden Markov models against those obtained using simpler rule-based approaches. The second main challenge lies in drawing inferences about the spatial distribution of foraging seabirds, and about the relationship between spatial distribution and environmental covariates. We present results from spatial point process models, as implemented using the integrated nested Laplace approximation (INLA), and contrast these against inferences derived from kernel density estimation and generalized linear mixed models.

The talk will focus upon discussing the uses and limitations of electronic tagging data in drawing inferences about behaviour and distribution, and will attempt to draw some wider conclusions regarding the analysis of large ecological datasets.

Cross-national differences in older adults physical functioning: results from HRS, ELSA and SHARE studies of ageing

Omar Paccagnella¹, Paola Zaninotto², Giovanni Boccuzzo¹

¹University of Padua, Padua, Italy, ²University College London, London, UK

Objectives

The purpose of this work is twofold. On the one hand, this paper aims at exploring and comparing the presence and severity of physical functioning among older adults living in the England, United States and mainland Europe and seeking to explain country differences. On the other hand, since data come from different surveys on ageing around the world, this work is an attempt to evaluate the comparability of these surveys analysing physical functioning issues.

Method/Models

Data come from the English Longitudinal Study of Ageing (ELSA), the Health and Retirement Study (HRS) and the Survey of Health and Retirement in Europe (SHARE) collected in 2006. The outcome measures are limitation with activities of daily living and instrumental activities of daily living, which are analysed separately. According to the idea of Heckman's models for treating sample selection bias, the statistical solution we adopt is written as a system of equations for two latent responses: the first analyses the presence of disability through a probit model; the second investigates the severity of disability through an ordinal probit model.

Results and Conclusions

Compared to the US, older people living in England, Spain and Portugal are more likely to have one or more limitations with ADLs, while Swedish and Belgians have a lower probability. Older people from England, Austria, Sweden, Netherlands, France, Greece, Belgium, and Czech are less likely to have limitations with IADLs than Americans, while Spanish older people are more likely to have limitations with IADLs, independently of other covariates. Country differences are more important comparing the presence of any disability rather than their severity. Findings also suggest a fair comparability across ELSA, HRS and SHARE surveys analysing physical functioning issues.

The Impact of Reporting Bias on the Measurement of Socioeconomic Health Inequalities in the U.S.

Kajal Lahiri, Paul Noroski

University At Albany, State University of New York, Albany, NY, USA

Objectives

The objectives of this paper are to estimate a model of self-assessed health among older Americans that is free of reporting bias, to compare these estimates to those of a traditional model that does not correct for reporting bias, and finally to perform a decomposition analysis to find the relative contributions of the determinants of health.

Method/Models

In order to achieve our objectives, we use anchoring vignettes (and other data from the Health and Retirement Study) along with a Hierarchical Ordered Probit model to estimate a health model that is free of reporting bias. A traditional Ordered Probit model is used as a baseline comparison.

Results and Conclusions

We find that many of the characteristics that affect health status are also found to affect perceptions of health. For example, we find that respondents in lower income groups tend to be more tolerant of mental health problems than the rest of the population. Overall, we find that failing to account for reporting differences ultimately leads to a distortion of the estimated relationships between health and socioeconomic status.

The measurement of social networks and their correlates in the older population: Lessons from SHARE and NSHAP

Howard Litwin

Hebrew University, Jerusalem, Israel

Objectives

The social networks in which older people are embedded are generally related to a wide range of well-being outcome measures. However, the measurement of the network phenomenon in large scale population surveys varies across studies and across populations. The aim of this paper is to consider two major parameters on which social network measurement differs: the source of the network information and its nature. The first parameter distinguishes between inferred and derived network information. The second parameter differentiates single and composite network indicators. Pairing the two parameters creates a property space with four analytical dimensions.

Method/Models

Drawing upon data from the Survey of Health, Ageing and Retirement in Europe (SHARE) and the National Social Life, Health and Aging Project (NSHAP) in the United States, the paper demonstrates each of the four analytical dimensions: 1) inferred network data with individual network indicators; 2) inferred network data with composite network indicators; 3) derived network data with individual network indicators; and 4) derived network data with composite network indicators. In each example, selected well-being outcome measures were regressed on the respective network variables within each analytical dimension.

Results and Conclusions

The respective analyses showed the following findings: 1) Inferred network data with individual network indicators-among younger-old respondents, aged 60-79, more significant associations were found between social network variables and well-being outcomes in comparison to older-old respondents, aged 80 or older. 2) Inferred network data with composite network indicators-respondents embedded in network types characterized by greater social capital exhibited better well-being in terms of less loneliness, less anxiety and greater happiness. 3) Derived network data with individual network indicators-satisfaction with one's personal network was the strongest predictor of well-being, followed by network size. 4) Derived network data with composite network indicators-network type retained a significant independent association with the CASP-12 well being measure after adjusting for age, gender, ADL, mobility limitations and country. The findings underscore the need to pursue comparative research to clarify which of the measurement approaches obtains the greatest accuracy and efficiency.

Childhood conditions, education and main job

Christelle Garrouste¹, Omar Paccagnella²

¹European Commission - Joint Research Centre, Ispra (VA), Italy, ²University of Padua, Padua (PD), Italy

Objectives

This paper aims at evaluating and comparing across many European countries the influence of early-life circumstances (like childhood disparities and educational attainments) on the main job of individual working career. This work gets into that part of the scientific literature which investigates intergenerational mobility (indeed, social mobility research has for a long time focused on the relationship between social origins, education and occupational destinations). The main objective of this study is to identify potential direct and indirect (through educational attainment) effects of early life circumstances on the quality of the main job in the working career.

Method/Models

Using life-history data from the SHARELIFE survey, country-specific structural relationships among childhood socio-economic status (at the age of 10: number of rooms per capita where respondent lived, availability of books in the respondent's accommodation, type of occupation of household's main breadwinner), educational attainments (number of years in full time education) and main job characteristics (incomes at the beginning and the end of the job spell, duration of the job spell) are investigated through the estimation of recursive models, controlling for several individual covariates and cohort effects.

Results and Conclusions

Poorer socio-economic conditions during childhood are associated with higher differentials in years of full-time education and higher income inequalities. Preliminary results confirm a strong direct effect of early-life circumstances on educational attainment in all countries and an indirect effect on the main job characteristics. The extent of these results however varies across countries, as well as the presence of cohort effects. Findings suggest that education policies may play a role in explaining the observed differences across countries. The introduction of support systems fostering access to education of students from disadvantaged households can weaken their financial dependence from parents and loosen the persistence in socio-economic conditions across generations.

Elderly Consumers and Cognitive Age Dimensions: The Influence of Hedonic/Utilitarian Contextual Factors on Perceived Youthfulness

Gianluigi Guido, Cesare Amatulli, Alessandro M. Peluso, M. Irene Prete

University of Salento, Lecce, Italy

Objectives

In affluent societies, the elderly market constitutes an expanding demographic segment with an increasing spending power. In this marketing scenario, understanding whether and to what extent both social and physical contexts affect elderly consumers' cognitive age (i.e., how old individuals perceive themselves to be) and youthfulness perception (i.e., how young individuals feel themselves to be in comparison with their actual age) is important for both marketers and researchers interested in improving the psychological well-being of this segment. The present study is the first investigation that explores how these two age-related constructs are affected by utilitarian versus hedonic circumstances.

Method/Models

This study explores the context-dependency of cognitive age and perceived youthfulness, measured on a sample of 600 elderly consumers (60 years and older). Cognitive age measures were collected by asking respondents to report the age they would feel to have in different physical and social contexts, previously identified through focus groups. For each context, either utilitarian (e.g., "senior center") or hedonic (e.g., "resort") one, perceived youthfulness was obtained by subtracting the actual age from the cognitive age of respondents.

Results and Conclusions

Results of both ANOVA and t-test showed that the cognitive age and the related perceived youthfulness of elderly consumers significantly change as a function of context. More specifically, respondents reported that they felt much younger when they imagined themselves to be in hedonic contexts than in utilitarian ones ($ps < .05$). These findings may be of help to marketers who want to advertise products targeting older people. Results suggest that marketing processes usually applied to middle-aged markets could be applied to approach the 'cognitively not-old consumer market'. Moreover, marketers can manage contextual factors - e.g., by using appropriate executional frameworks when developing advertising campaigns - in order to manipulate older consumers' perception of their age in a way that fits brand identity or to promote products or services by directing the focus to well-being determinants.

Ordered categorical predictors: with an application in pharmacogenetics

Stephen Senn

CRP Santé, Strassen, Luxembourg

Objectives

There are many models for dealing with ordered categorical outcomes. For ordered categorical predictors, however, there seem to be fewer tools available. A context in which ordered categorical predictors occur quite naturally, however, is that of pharmacogenetics. If patients are classified by the number of copies of an allele of a given sort at a given locus then they can be scored as having 0, 1 or 2 copies and it seems plausible in many contexts to expect that the response must be monotonic as a function of the number of copies. This work investigates possible approaches to modelling such predictors.

Method/Models

A general formulation is developed whereby an ordered categorical scale with three points can be linearised with the use of an additional parameter. The idea of a dose metameter is borrowed from the field of dose response. By this device, for an ordinal scale with three points, the scale of the predictor can be transformed to one which has a linear response. Likelihood solutions for this are explored and the connection to isotonic regression investigated.

Results and Conclusions

If estimation of the extra parameter is made in an unconstrained way then there is nothing to be gained compared to treating the predictor as nominal since two degrees of freedom have to be used. However if the parameter is constrained to produce monotonicity then there is a gain in efficiency. Of course, there is always the danger that a constrained model may not be correct. However, it is concluded that in many contexts it will be reasonable to accept the risk of some bias. Some implications for the design of pharmacogenetic trials are also considered.

Measuring the impact of heterogeneity in multivariate meta-analyses

Dan Jackson¹, Ian White¹, Richard Riley²

¹MRC Biostatistics Unit, Cambridge, UK, ²Birmingham University, Birmingham, UK

Objectives

Measures of the impact of heterogeneity are well established in the context of univariate meta-analysis. Our objectives are to motivate, develop and apply some multivariate generalisations of these statistics which are more appropriate when using multivariate meta-analysis.

Method/Models

We propose four new multivariate heterogeneity statistics, which are either generalisations or analogues of their established univariate counterparts. These multivariate heterogeneity statistics are easily obtained from standard output from the fixed and random effect multivariate meta-analysis models.

Results and Conclusions

We apply our methods to some contrasting examples. We find that our statistics provide additional insight when used in conjunction with the standard multivariate estimation methods and the established univariate heterogeneity statistics. In particular they illustrate the impact of multivariate heterogeneity better than the univariate statistics. We conclude that at least one of our statistics should accompany the usual results when fitting multivariate meta-analysis models, but the choice of heterogeneity statistic may depend on the aspect of the heterogeneity that the analyst wishes to quantify.

Modelling the Diluting Effect of Social Mobility on Health Inequality

Heather Turner^{1,2}, David Firth¹

¹University of Warwick, Coventry, UK, ²Freelance Statistical/R Consultant, Ramsgate, Kent, UK

Objectives

This presentation considers the incidence of limiting long-term illness (LLTI) in 2001 as recorded in the Office of National Statistics Longitudinal Study. The objective is to model the inequality in the incidence of LLTI across socio-economic positions (SEPs) and to quantify the diluting effect of social mobility on this inequality.

Method/Models

In public health studies, the effects of social mobility are commonly investigated by analysing all possible social trajectories, typically reducing to two classes for simplification. Rather than simplifying the data, we apply the diagonal reference model which provides a parsimonious representation of all possible social trajectories. The parameters of the model directly quantify the inequality across the classes and the diluting effect of social mobility. We include age as a covariate and model men and women separately. Given the binary outcome and the nonlinear form of the model, this is a particular case of a generalized nonlinear model.

Results and Conclusions

Inequality in the incidence of LLTI across SEPs is more acute amongst men than women. Compared to higher professionals, the odds ratio of LLTI for the intermediate class is 2.72 for men (95% CI: 2.40--3.08), 1.56 for women (95% CI: 1.34--1.81); for the routine labouring class this rises to 4.36 for men (95% CI: 3.95--4.81), 3.66 for women (95% CI: 3.15--4.25).

For a given destination SEP, the probability of LLTI decreases with a more favourable origin SEP and increases with a less favourable origin SEP, thus social mobility has a diluting effect on the health inequality. We find that the origin:destination ratio is 62:38 for men and 41:59 for women. Since their destination class is given more weight, social mobility has a greater impact for women.

Modelling Multiple Binary Events with Applications to Concurrences of Human Papillomavirus Infections

Jack Cuzick, Zihua Yang

Wolfson Institute, Queen Mary University of London, London, UK

Objectives

Excess relative risk is often observed in concurrent infections of human papillomavirus (HPV). A model is developed to capture the excess risks for such concurrences using a one-parameter gamma frailty model. The model is fitted to a dataset of 35 HPV infections types on 11,155 women from New Mexico with abnormal cervical cytology smears and age of less and equal to 30.

Method/Models

We present a general frailty model for modelling multiple binary events. The model yields simple closed-form expressions for all outcome probabilities. In particular, we consider a one-parameter gamma distribution for the frailty. By fixing the marginal probabilities to be the observed marginal prevalences, a consistent approximate maximum likelihood estimate is given for the frailty parameter. The estimator is efficient and the asymptotic variance is given. Goodness of fit is tested using the observed and estimated numbers of concurrences.

Results and Conclusions

Overall, the proposed model reduced dispersion in the concurrences significantly compared to the independence model for the New Mexico data but did not provide a good fit to the 35 types. This is mainly due to the heterogeneity in risk of concurrences associated with type 16 in particular. Further analysis suggests improved fits for the alpha-9 species (16-group) and alpha-7 species (18-group) of the 35 types.

Empirical assessment of the validity of uncontrolled comparisons of the accuracy of diagnostic tests

Yemisi Takwoingi¹, Mariska Leeflang², Jonathan Deeks¹

¹University of Birmingham, Birmingham, UK, ²University of Amsterdam, Amsterdam, The Netherlands

Objectives

Many studies of diagnostic test accuracy often assess the performance of only a single test evaluated against a reference standard, while the choice to be made is between alternative tests. Studies that directly compare tests within patients or between randomized groups are preferable but in the absence of such comparative studies, test accuracy reviews frequently rely on only indirect between-study (uncontrolled) comparisons. The aim of this study was to demonstrate the lack of comparative studies of diagnostic accuracy, and to investigate the magnitude of bias associated with test comparisons using non-comparative studies of test accuracy.

Method/Models

Systematic reviews of diagnostic test accuracy published between 1994 and February 2011 were identified in the Database of Reviews of Effects (DARE). For each pair of tests, meta-analysis was performed using the hierarchical summary ROC model. To compare test performance, a covariate for test type was included in the model; to compare comparative and non-comparative studies a second covariate for study type, together with its interaction with test type were added. We estimated relative diagnostic odds ratios (rDORs), and absolute differences in sensitivities and specificities, and compared estimates obtained from comparative studies with those from non-comparative studies.

Results and Conclusions

In 216 reviews where both study designs were eligible for inclusion, comparative studies for a specified pairwise comparison were sparse; many reviews (52%) had less than three. A subset of 27 reviews with sufficient data to conduct both direct and uncontrolled test comparisons contributed 38 pairwise test comparisons for meta-analyses. Large discrepancies were observed; DORs from meta-analyses of non-comparative studies of 52 of the 76 (68%) tests were greater than those from meta-analyses of comparative studies (sign test, $p=0.002$). There was more than a two-fold difference in ratio of rDORs for 18 test comparisons (47%). Direction of the bias was unpredictable.

In conclusion, between-study uncontrolled comparisons of tests may sometimes be the only or major source of evidence, but limitations of such comparisons should be highlighted when making inferences on the relative accuracy of competing tests. Funders should be encouraged to ensure future test accuracy studies address important comparative questions.

Statistical Modelling of Oceanographic Time Series

Adam M. Sykulski¹, Sofia C. Olhede¹, Jonathan M. Lilly², Eric Danioux²

¹University College London, London, UK, ²NorthWest Research Associates, Seattle, USA

Objectives

The collection and analysis of oceanographic data is central to understanding the global climate system. In particular an important feature of the ocean, known as the "ocean circulation", significantly contributes to decadal to-century-scale climate variability. The main sources of data on the ocean circulation are the time series obtained from free-drifting satellite instruments - more commonly known as surface drifters. The objective of our work is to build a complete statistical model for surface drifter time series, in order to gain a deeper understanding of the ocean circulation and how important properties might change over time and space.

Method/Models

The construction of our statistical model is motivated both from the observed data and from expected theoretical physical properties of the ocean. The underlying continuous-time physical processes are modelled as stochastic processes and we motivate and describe the applicability and key characteristics of the processes we use. Surface drifter time series, however, are obtained in discrete-time and we discuss the critical importance of correctly considering statistical sampling effects such as aliasing and leakage, which are often overlooked in the literature. The estimation of model parameters can then be efficiently and accurately performed in the frequency domain using the Whittle likelihood.

Results and Conclusions

We demonstrate that the statistical model can characterise the key features of surface drifter time series, in both the time and frequency domain. We also show results that demonstrate the significance of considering the sampling effects we highlight in this talk. Finally, we display results showing how the model can be used to efficiently generate new time series that mimic real drifter time series, which is of particular use in the analysis of many climate models. Overall, the model can be used to infer several important physical properties of surface drifter time series, and analyse how these properties might vary over time and space.

Nonparametric Monitoring of High Frequency Data Streams

Gordon Ross

University of Bristol, Bristol, UK

Objectives

The last decade has seen the emergence of data stream analysis as a key area within applied statistics. Data streams consist of a time-ordered sequence of observations which are received at a high frequency, with typical applications in fields such as finance often involving several thousand data points per second. This volume of data makes analysis using traditional time series methods infeasible.

Our focus is on detecting structural changes in the distribution of a data stream. We develop methods which can function in situations where computational efficiency is paramount, and where only limited information about the stream distribution is available.

Method/Models

As the stream distribution is often unknown, we use a nonparametric framework to minimise the number of assumptions that must be made. Working within a frequentist paradigm, we adapt nonparametric hypothesis testing to the data stream setting. By performing tests sequentially, the stream can be monitored while new observations are being received.

To increase computational efficiency, we use recursive update rules for all test statistics, and introduce discretisation to minimise the volume of information that must be processed. Unlike approaches which use sliding windows, our method does not incur any noticeable performance decrease as a result of pruning older data.

Results and Conclusions

Using a variety of simulated and real data sets, we show that our nonparametric approaches attain results that are competitive with their parametric counterparts, which assume full specification of the stream distribution. We consider change detection in both streams consisting of independent and identically distributed observations, along with time series of correlated data. The latter is illustrated using the example of detecting changes in the standard deviation of high frequency financial time series.

The computational burden of our approach is shown to be low, achieving $O(1)$ complexity and allowing several thousand points per second to be processed indefinitely. Additionally, we can exploit parallelisation in order to further increase efficiency.

All of our methods are implemented in an R package which we have released, and this will also be briefly described.

Locating and quantifying gas emission rates using remotely obtained concentration data

Bill Hirst¹, Fernando Gonzalez del Cueto¹, [David Randell](#)¹, Philip Jonathan¹, Oliver Kosut²

¹Shell Projects & Technology, UK, UK, ²MIT Laboratory for Information and Decision Systems, US, USA

Objectives

We describe a method for detecting, locating and quantifying gas emissions to the atmosphere using remotely obtained gas concentration data. The method is applicable to gases of environmental concern.

Method/Models

A Gaussian mixture model is used to model sources, a Markov random field for background concentrations, and a Gaussian plume eddy dispersion model to approximate the wind field between sources and measurement locations. Reversible jump Monte Carlo Markov chain inference provides estimated values and uncertainties for the number and characteristics of sources, background and other model parameters.

Results and Conclusions

We compare the performance of the mixture model with that of a compressive sensing model (i.e. least squares optimisation with L1 sparsity constraints) which estimates source intensities on a grid representing (a large number of) potential source locations. We apply these approaches to airborne methane concentration data to map emissions from a land-fill.

Plenary 2 - Champion (President's Invited) Lecture: Statistics at Google

Wednesday 5 September, 4.40pm - 5.40pm

Over the last decade private companies have invested in building real-time data warehouses. Now that they have all that data accessible, the challenge is to figure out what to do with it. As George Box has said "you can improve almost anything" and companies are trying use that great data to improve their operations. I will talk about how we have used statistics at Google in a number of areas, what kind of skills we find valuable, and what the future might hold for data analysis in technology-intensive industries.

Hal R. Varian is the Chief Economist at Google. He started in May 2002 as a consultant and has been involved in many aspects of the company, including auction design, econometric, finance, corporate strategy and public policy.

He is also an emeritus professor at the University of California, Berkeley in three departments: business, economics, and information management. He received his S.B. degree from MIT in 1969 and his MA and Ph.D. from UC Berkeley in 1973. Professor Varian has published numerous papers in economic theory, econometrics, industrial organization, public finance, and the economics of information technology.

Ultimate 100-m world record through extreme-value theory

John Einmahl

Tilburg University, Tilburg, The Netherlands

Objectives

We use extreme-value theory to estimate the ultimate world records for the 100-m running for men. For this aim we collected the fastest personal best times set between January 1991 and July 2012.

Method/Models

Estimators of the extreme-value index are based on a certain number of upper order statistics. To optimize this number of order statistics we minimize the asymptotic mean-squared error of the moment estimator. Using the thus obtained estimate for the extreme value index, the right endpoint of the speed distribution is estimated. The corresponding time can be interpreted as the estimated ultimate world record: the best possible time that could be run in the near future.

Results and Conclusions

We find the point estimate 9.36 seconds for the 100-m men. (This is joint work with Sander Smeets.)

David Forrest (University of Salford)

MODELLING MEDALS

The standard (Tobit random effects) regression model accounting for the distribution of Olympics medals between national teams emphasises the roles of living standards and population size in determining a country's medal share (or indeed whether it wins a medal at all). The presentation will report results from extending the basic model in a number of ways, including its application sport-by-sport rather than just to the overall medals table. Sporting disciplines are shown to be very different in terms of the degree of advantage exhibited by the richest countries.

Statistics and energy - challenges and opportunities

Idris Eckley

Lancaster University, Lancaster, UK

Energy infrastructure is undergoing significant change. In particular, with investment in new technology for renewable energy resources, carbon capture and smart metering, data collection is firmly embedded at the centre of the modern energy system. An opportunity therefore exists for statistics to play a major role in making our energy systems more efficient and sustainable. This talk will outline some of the challenges and opportunities for statistics in this area.

Forecasting nonstationary energy time series

Marina Knight¹, Rebecca Killick²

¹University of Bristol, Bristol, UK, ²University of Lancaster, Lancaster, UK

Within the energy sector forecasting is an important statistical tool. Each day many forecasts are made across a variety of time scales, such as production of renewables, consumer demand and trader pricing. Traditional statistical techniques assume stationarity of the past in order to produce accurate forecasts. For data arising from the energy sector, this stationarity assumption is often violated. There are a limited number of statistical models for nonstationary data which can be used to produce reliable forecasts. This talk will discuss appropriate forecasting models for energy data, highlighting potential issues. The talk will conclude with suggestions on overcoming these drawbacks.

Statistical analysis of copy number alterations in cancer samples using low-coverage next-generation sequence data

Arief Gusnanto¹, Stefano Berri², Henry Wood², Ibrahim Nafisah¹, Yudi Pawitan³, Pamela Rabbitts², Charles Taylor¹

¹University of Leeds, Leeds, UK, ²Leeds Institute of Molecular Medicine, Leeds, UK, ³Karolinska Institutet, Stockholm, Sweden

Objectives

Next-generation sequencing is a method to rapidly sequence biological samples' genomes. The method produces short sequences, usually called 'reads', that are mapped to the human reference genome. The number of reads in a genomic 'window' is utilised in the analysis of copy number alteration (CNA) by comparing the counts between normal and cancer samples. Our interest is to identify regions in the genome that 'deviate' from the normal copy number of two, either as amplifications or deletions. The locations at which the copy number changes indicate the position of important genes in the development of tumor.

Method/Models

In this talk, we will describe some challenges in the analysis, including estimating an optimal window size in the context of low-coverage sequence data, dealing with different coverage between normal and cancer samples, contamination of cancer cells with normal cells, and random experimental error. We will discuss some methods to deal with those challenges, including data-based histogram construction methods, linear models, and mixture models. We illustrate the concept with the analysis of CNA in lung cancer patients.

Results and Conclusions

The methods enable us to interpret the genomic copy numbers properly and identify the amplifications and deletions. The results from a simulation study indicates that our approach is relatively accurate to correct for contamination of normal cells up to 40 percent, and -- in overall -- is still better than previous methods in dealing with contamination for up to 80 percent. We estimate that an optimal window size for our data is approximately 120 reads per window, and window sizes between 30 - 180 reads per window can be considered near-optimal. Our works lay the foundation for further analysis, including testing for regions of CNA and survival analysis of cancer patients based on their CNA profiles.

Sparse Gaussian copula graphical model for a genomic breast cancer network

Ernst C. Wit, [Fentaw Abegaz](#)

University of Groningen, Groningen, The Netherlands

Objectives

Breast cancer is the leading cause of death among women in the world. In the last two decades knowledge of breast cancer grade determined by Nottingham prognostic index has been helpful to diagnose the most effective treatments. Nowadays other treatment predictive markers are incorporated that considerably improved breast cancer treatment decisions. To further tailor treatment for individual patients, identification of additional clinical and genetic markers is required. A comparative genomic experiment on breast cancer was performed to identify DNA amplification patterns to study the relationship between DNA amplification patterns and severity of breast cancer, as measured by several clinical indicators.

Method/Models

The vast literature on the study of association among mixed variables focuses on modeling with latent variables. Hoff (2007) and Dobra and Lenkoski (2011) implemented Bayesian approaches based on extended rank likelihood to study association among mixed variables. To account high dimensional mixed variables setting, we propose an L_1 penalized extended rank likelihood with Expectation-Maximization. This approach provides sparse estimates for Gaussian copula graphical models and much faster alternative to the Bayesian approaches. The proposed methodology is applied on breast cancer data to study conditional dependency and identify the underlying genes and clinical variables related to breast cancer aggressiveness.

Results and Conclusions

The breast cancer data includes 65 selected genes and 5 clinical variables obtained from 62 breast cancer patients. We applied the proposed L_1 penalized extended rank likelihood with EM estimation to the breast cancer data. We obtained a sparse representation of the conditional dependencies between the variables. The links between gene amplification/deletion and clinical information are summarized as follows. Breast cancer death is associated with breast tumor grade or NPI, tumor size and gene 23. Further NPI is associated with genes 19, 16, 21 and 46. Further more, breast cancer tumor size is associated with 56. A careful examination of these associations will generate several interesting hypotheses on the importance of these variables to the aggressiveness of breast cancer. This in turn be helpful to decide on the most effective treatments for breast cancer patients.

Detection of disease change from a biological marker: using CA125 in ovarian cancer as an example

Wendi Qian¹, Marcia Hall², Gordon Rustin², Diana Kornbrot³

¹University of Cambridge, Cambridge, UK, ²Mount Cernon Cancer Centre, London, UK, ³University of Hertfordshire, Hertfordshire, UK

Objectives

In Oncology drug development, the great majority of phase 3 trials are negative. New strategies are required to rapidly identify novel agents prior to large randomised trials. The CA125 doubling trial successfully showed that an effective drug could be identified efficiently by testing whether the rate of increase in the tumour marker CA125 decreased after starting the novel agent, at a point identified by CA125 rising to four times its nadir level. However efficiency could be improved, if more patients could be included. This work explores identifying an earlier effective starting point by analysing the time course of CA125 rise.

Method/Models

Tumour growth is measured by the slope of the linear regression of $\ln(\text{CA125})$ level on time. This slope is estimated for time points: 1 to n , 2 to $n+1, \dots$, $C-n+1$ to C , where C is the time patient transfers to new drug or leaves trial. A new potential transfer point to novel drug is identified as both CA125 level $>$ ULN and slope $>$.0121 (estimated minimum slope before transfer). The number of additional patients available for trial of the novel agent within 9 months is the end point, explored using $n=3$ and $n=4$.

Results and Conclusions

The number of patients who did not receive drug, but would have received drug with the new method, and the number of patients who did transfer to drug, but might have transferred earlier depends on the exact combination of slope and absolute CA125 level chosen, hence boundary estimates are given. Between 12 and 28 patients who did not receive drug with the existing protocol would have received drug with the new method. Between 12 and 25 patients who received drugs within 6 months would have received drug earlier using the new method. In addition, 2 patients who were transferred to drug after 9 months could have been transferred within 9 months. These preliminary analyses show that the proposed new monitoring approach could be applied in the selecting novel active agents suitable for larger randomised trials. Simulation of statistical performance of the approach will be presented.

Differential Heritability and Environmentality of Intelligence Test Subscales across Socioeconomic Status

Rebecca Pillinger

University of Bristol, Bristol, UK

Objectives

Two hypotheses suggest we should expect the magnitude of genetic effects on cognitive outcomes to differ depending on the environment an individual experiences. The bioecological model says some individuals have genes conferring the potential to achieve high scores, but that these genes can only fully operate in beneficial environments. The diathesis-stress model says some individuals have genes affording protection against the adverse effects of poor environments on scores. Similar hypotheses exist for unmeasured aspects of the environment. We examine whether each of these hypotheses is supported for the seven subscales of the WISC IQ test administered to 7 year olds.

Method/Models

Using the National Collaborative Perinatal Project (data from a large sample of children born in the US between 1959 and 1966) we fit multilevel genetic models and structural equation models to the ~1000 twins for whom data is available at age 7, decomposing variation between individuals in each subscale into a shared environmental, a non-shared environmental, and a genetic component. We allow the magnitude of each of these components to depend on parental income, occupation and education, measured at the time the children were born and when they were aged 7. Multilevel multiple imputation is used to handle missing data.

Results and Conclusions

The Comprehension subscale shows no differential heritability or environmentality across any of the family background measures. The other scales show differential heritability and/or environmentality for some or all of the background measures. However, the pattern of the differential components across the background measure varies between background measures as well as between subscales, with some combinations of subscale and background measure showing support for the bioecological model, some for the diathesis-stress model, and some showing support for each in different ranges of the variable.

Beyond 2011 - The future of population statistics

Alistair Calder, Andy Teague, [Ann Blake](#)

Office for National Statistics, England & Wales, UK

Objectives

The Office for National Statistics (ONS) is currently considering options for the future production of small-area population and small-area socio-demographic statistics in England and Wales. The 'Beyond 2011' programme has been set-up to take a fresh look at alternative approaches that will meet future user needs for 'Census-type' data.

Method/Models

While innovative ways of carrying out a traditional fieldwork based Census will be assessed most of the work is focussing on making better use of existing administrative and other data sources. These data sources are improving and maturing and Beyond 2011 aims to assess whether, properly combined, they will be able to provide a viable and cost effective means of producing the required statistics in future. A full understanding of user needs, quality, costs and the public acceptability of alternative approaches will all play a part in informing the recommendation on the way forward which will be made in 2014.

Results and Conclusions

Any recommendations from the Beyond 2011 work will have implications for all population based statistics in the UK and perhaps, in the longer term, for the statistical system as a whole. This presentation will describe the background to this important work and briefly outline the programme of research being taken forward. It will also explain how users and other interested parties can find out more and get involved.

Matching of administrative data to validate the 2011 Census in England and Wales

Louisa Blackwell, Nicola Rogers, Andrew Charlesworth

Office for National Statistics, Titchfield, UK

Objectives

This presentation describes the role of administrative data matching in the quality assurance process for the 2011 Census in England and Wales. Data sharing across government allowed ONS to use administrative microdata to check and understand census counts and estimates. The sources are described. The innovative data matching methods and systems designed for this task will be presented. A flexible, reactive approach applied analytic methods that were appropriate to the research questions that arose during the census quality assurance process. On their own and in combination, the administrative sources provided new insights, both empirical and theoretical, into population enumeration.

Method/Models

Data extracts from the NHS Patient Register, Valuation Office Agency, Electoral Registers, School Census and Higher Education Statistics Authority were matched at record level to create interacting matrices of linked information for dwellings and people. These were enriched with evidence from the Address Register History File. Matching ahead of census quality assurance focused on 37 local authorities representing a range of issues in administrative and census data. These data were mined and analysed at varying levels of detail, as required.

Results and Conclusions

The data architecture, systems and software used to support this matching and analysis had to be developed and built as data availability and requirements became clearer ahead of Census. The tension between a desire for hard-coded, rigorously tested and robust systems and the reality of operational imperatives and the need for early results was carefully reconciled through bespoke, ad-hoc solutions for urgent requests that then required careful quality scrutiny. Address matching was confounded by the different reporting of addresses to different administrative bodies. Name capture and differential recording posed additional challenges. Through matching people we learned about the behaviours of different types of administrative data in different types of area. There were characteristic lags that impacted the currency of the data for different age/sex groups. Comparability and match rates between administrative sources and census were highly geographically specific, but also typical.

Using Internet Search Data in the Production of Official Statistics

Daniel Ayoubkhani, Matthew Swannell

Office for National Statistics, Newport, South Wales, UK

Objectives

This paper investigates the potential for using Google Trends in the production of official statistics. The growing popularity of the internet as a virtual market place is amongst the most significant shifts in consumer behaviour of recent decades. Search engines are the mechanism that matches supply with demand and one search engine, *Google*, is dominant. The analysis reported in this presentation uses weekly data from *Google Insights for Search* in order to answer the question: "can the Office for National Statistics (ONS) make use of internet search query data to improve estimates of UK retail sales prior to their publication?"

Method/Models

This investigation makes several extensions to previous approaches to investigating Google Trends data in order to make conclusions more robust and to tailor the analysis to real-world conditions faced by ONS. These extensions include: (1) the identification of lagged relationships between internet search queries and retail sales; (2) the use of regARIMA (regression with ARIMA errors) models; (3) the inclusion of variables that account for the complex survey calendar of the UK Retail Sales Index.

Results and Conclusions

The results of the investigation are mixed. There appears to be some potential for using Google Trends data for predicting, and hence quality assuring, estimates of retail sales activity for a limited number of components of the UK Retail Sales Index. However, the results are less promising for other components. Google Trends terms appear to have only limited explanatory power for these series.

Several considerations, in addition to the empirical results obtained in this investigation, would need to be made before Google Trends will be utilised by ONS during routine production of the UK Retail Sales Index. These include: the fact that Google Trends data are not available for periods prior to 2004; the possibility of search terms being transient; and practical considerations regarding future accessibility, and cost, of Google Trends as a data source.

Driving Insight through developing Web Analytics in the not-for-profit sector

Paul Askew

Chartered Society of Physiotherapy, London, UK

Objectives

The shift to on-line services provides an opportunity, and increasing need, to evaluate effectiveness. This presentation illustrates the emerging value of web analytics in providing critical insight into the web usage, and specifically in the not-for-profit sector. This includes a focus on continuing professional development (CPD) in the health sector. This illustrates how the structured analysis of large quantities of dynamic data can provide tangible operational and strategic insight, while still being constrained by the nature of the data. This also illustrates the development and use of bespoke analysis to provide additional insight, and beyond the more typical commercial applications.

Method/Models

The approach builds on the increasingly established field of web analytics. While generally well recognised in the commercial sector, this is becoming an increasingly widespread analytical approach in other sectors. The not-for-profit sector examples illustrate the analytical use and development of the approach. There were three levels of analysis (1) Common web analytical methodologies provide the basic insight, (2) tailored analytical tools then provide the context specific analysis and insight and (3) the use of 3rd party web based applications - specifically with a professional educational application - provides additional insight through the more granular monitoring of use and behaviour.

Results and Conclusions

The use of web analytics, the specific context specific development of these, and the use of specialist applications, provided a range of new structured business intelligence. This enabled understanding, evaluation and development of new online educational resources.

The use of the three levels of analysis allowed tracking of specific new resources for continuing professional development initiatives, and the targeted at very specific audiences. This included a suite of tightly linked resources delivering detailed guidance, dynamic on-line tools and engaging multi-media case studies. The analysis enabled detailed insight into the patterns of use between and across these resources, both in aggregate and for individual pathways. This included access, exploration, and specifically the use and re-use of guidance, tools and case studies. This also highlighted the potential for the linkage between web services and other user data.

Volatility Graphics for Financial Time Series

Tony Lawrance

University of Warwick, Coventry, UK

Objectives

To develop exploratory graphical methods for revealing temporal and dependency aspects of volatility in financial time series, and to justify the proposed graphics by application to actual and simulated data.

Method/Models

A general time series model including a progression function and a volatility dependence function is proposed. It includes GARCH and ARCH models and a new comparable form of SV model. Individual volatilities in terms of absolutes are proposed on the basis of the general model. The smoothing of their temporal or dependency plots is designed so as to reveal volatility behaviour and the required aspects of bootstrap variability methodology is developed. The reliability of the graphics is probed via application to simulations of volatility models.

Results and Conclusions

The paper presents an applicable methodology for the exploration of volatility in financial time series. It focuses on using smoothed individual volatilities to give non-parametric versions of the volatility dependence function, together with bootstrap envelopes of its variability; these outcomes are compared to those of a fitted model in order to assess its provisional suitability. The methodology is demonstrated on recent FTST100 data, and validated by application to simulated series, both volatile and non-volatile.

Implementing bootstrap variance estimation on the ONS Business Register and Employment Survey

Matthew Greenaway

Office for National Statistics, South Wales, UK

Objectives

The ONS Business Register and Employment Survey is an annual survey of businesses which is used to produce employee estimates at detailed geographical and industrial levels. At levels at which the direct survey estimates are of insufficient quality a small domain estimation technique based on apportionment is used, but the analytic variance of these small domain estimates is difficult to derive. We aim to evaluate the use of bootstrap resampling for estimating the variance of these small domain estimates, considering two different bootstrap methods. A key challenge is that many strata have non-negligible sampling fractions.

Method/Models

We evaluate two bootstrap methods designed to account for non-negligible sampling fractions in terms of their practicality and the closeness of the bootstrap variances to analytic variances (where analytic variances can be derived). The first method is a standard with-replacement bootstrap where the re-sample sizes are adjusted to account for the sampling fraction within a stratum. The second is the Rao-Wu rescaling bootstrap, implemented as a with-replacement bootstrap with an adjustment to the survey weights (Rao & Wu, 1988; Girard, 2009). We implement these methods on 2009 and 2010 survey data.

Results and Conclusions

We found the bootstrap methods to be very computer intensive, as the survey calibration and small domain estimation must be re-run for each iteration of the bootstrap. This meant that the speed at which the variance estimates converged was a key practical consideration. The Rao-Wu rescaling bootstrap required a smaller number of iterations to converge, and each iteration took less time due to the smaller re-sample size. Once they converge, both bootstrap methods produce variance estimates that are fairly close to analytic variances, but there were some discrepancies. We conclude that the Rao-Wu rescaling bootstrap is the best option for small domain variance estimation on the Business Register and Employment Survey, and may be helpful on other ONS business surveys that use similar small domain estimation methods.

Item-by-item sampling for promotional purposes

Neil Spencer, Lindsey Kevan de Lopez

University of Hertfordshire, Hatfield, Hertfordshire, UK

Objectives

The objective of this paper is to present a method for sampling items that are checked on a pass/fail basis with a view to a claim being made about the success/failure rate for the purposes of promoting a company's product/service. Attention is paid to ensuring that the statistical methodology used is capable of giving supporting the language that the company wishes to use in its promotional message.

Method/Models

Commonly used statistical methodologies such as confidence intervals can make an assessment of the success/failure rate based on a sample but conclusions based on these cannot be correctly phrased in such a way that is useful for promotional purposes. This leads to the development of Bayesian credible intervals. The hypergeometric distribution is used to calculate successive stopping rules so that the costs of sampling can be minimised. Extensions to the sampling procedure are considered so as to allow the potential for stronger and weaker claims to be made as sampling progresses.

Results and Conclusions

It is shown that the use of Bayesian credible intervals and the hypergeometric distribution is able to produce a sampling scheme for item-by-item sampling that enables a company to make statistically valid claims about the success/failure rates of their products in such a way that is useful for promotional purposes. The issue of choosing appropriate prior probabilities is dealt with and the relationship between the true error rate and the probabilities of making correct claims is discussed.

Participation and Performance at the London 2012 Olympics

Gerard Kuper, Elmer Sterken

University of Groningen, Groningen, The Netherlands

Objectives

This paper presents a statistical model to predict national medal winnings at the Olympic Summer Games in London. The main focus is on the impact of economic, geographic and demographic determinants of Olympic participation and success. Moreover, we account for a home country advantage. Post-war studies estimate this advantage to be about two percentage points of the share in medals earned. Finally, we also include medal success at recent corresponding world cup events. We describe the model, its past performance, the data fitting and our forecasts for the 2012 Olympic Summer Games.

Method/Models

We use the country level as our unit of measure. This implies that we consider a country of holding a portfolio of likely medal winning athletes, each having idiosyncratic probabilities to win a medal.

We apply a two-step forecasting procedure. First, we analyze the decision to participate at the Olympic Summer Games. Secondly, we predict the number and medals each country is expected to win at the 2012 London Olympics, conditional on the participation. We distinguish individual events and team events. Individual events are separated in events for women and men (including mixed events).

Results and Conclusions

In our panel data model we include the last three editions of the Olympic Summer Games, and we include 126 countries. These 126 countries cover all medals, and 95-97% of total participation at the last three editions of the Olympic Games. The countries included in our database cover 99% of real World GDP, and 91% of the world population.

Both participation and medals are modeled in shares, and we compute the actual medal winnings by multiplying shares by the given total of medals to be awarded at the 2012 Olympics (302 gold and silver medals and 355 bronze medals).

The questions we address are:

1. Will China again lead the medal table, or will the USA overtake China?
2. Will Great Britain exploit its home advantage and maintain its Top-5 position?
3. Will the Netherlands re-enter the Top-10?

Olympic Games: Export Effect, Signal Effect?

Wolfgang Maennig, Felix Richter

University of Hamburg, Hamburg, Germany

Objectives

Recent studies find that Olympic Games host countries experience significant positive, lasting effects on exports. Some of the studies interpret their results as an indication that countries use the hosting of such events to signal openness and competitiveness. We challenge these empirical findings of significant Olympic export effects.

Method/Models

We employ the matching strategy of Rosenbaum and Rubin (1983) in order to systematically discriminate between Summer Olympic Games host countries (i.e., the treatment group) and non-host countries (i.e., the control group). Only countries that are otherwise structurally similar are included in the subsequent analysis.

Results and Conclusions

With appropriate matching, we are not able to isolate significant positive export effects.

Thursday 6 September, 11.10am

Harnessing many-core architectures using population-based Monte Carlo methods

Anthony Lee¹, Chris Holmes², Francois Caron³, Christopher Yau⁴, Arnaud Doucet², Michael Giles²

¹*University of Warwick, Coventry, UK*, ²*University of Oxford, Oxford, UK*, ³*INRIA, Bordeaux, France*,

⁴*Imperial College London, London, UK*

Objectives

1. To encourage the use of many-core architectures in computer intensive inference.
2. To highlight some challenges in doing so in the context of Monte Carlo methodology.
3. To present the general idea behind a number of population-based Monte Carlo methods, which are ideally suited to implementation in parallel on a many-core architecture.
4. To present some encouraging results on some problems.

Method/Models

Methods

1. Parallel tempering Markov chain Monte Carlo.
2. Sequential Monte Carlo.
3. Pseudo-marginal methods.

Models

1. State-space models, e.g. stochastic volatility.
2. Genome-wide association studies.

Results and Conclusions

The changing scientific computing landscape from single core to many-core processors motivates new algorithmic considerations for making the most out of available hardware. This is most acutely felt in computer intensive statistical methods, such as Monte Carlo based inference. Thankfully, a number of "population-based" methods are particularly well suited to implementation on emerging devices.

We present some encouraging results on some challenging canonical problems in Monte Carlo based inference, with speedups of between 20 and 550 in these cases. We also motivate algorithmic and theoretical attention to a restricted parallel model of computation.

Thursday 6 September, 11.10am

Accelerating integrative modelling using GP-GPU computing

Paul Kirk¹, Jim Griffin², Faiz Sayyid¹, Richard Savage¹, Zoubin Ghahramani³, David Wild¹

¹*University of Warwick, England, UK*, ²*University of Kent, England, UK*, ³*University of Cambridge, England, UK*

Objectives

The integration of multiple datasets remains a key challenge in systems biology and genomic medicine. Modern high-throughput technologies generate a broad array of different data types that provide complementary information (e.g. gene expression and ChIP-chip datasets). We have previously developed approaches for the integrative modelling of two different data types and demonstrated that these improve our ability to identify meaningful clusters in a variety of contexts (including gene module discovery and the identification of prognostic cancer subtypes). Given the abundance of different data types, we seek to extend these approaches to multiple (>2) datasets.

Method/Models

We have recently developed a new framework for the integrative modelling of diverse datasets, which we refer to as MDI (Multiple Dataset Integration). Each dataset is modelled using a Dirichlet-multinomial allocation (DMA) mixture model, with dependencies between these models captured via parameters that describe the levels of agreement among the datasets. The mixture models naturally capture the clustering structure of each dataset, while modelling the dependencies between models allows us to identify clusters that are common to several (or all) of the datasets. We perform inference in our model via Gibbs sampling.

Results and Conclusions

I shall present the model and some applications, and demonstrate that including more datasets can improve our ability to identify biologically meaningful groups of genes (as measured, for example, by the Biological Homogeneity Index). Given the massive numbers of entities (typically genes) that we seek to cluster, it is of vital importance to accelerate the inference procedure for our model in order to allow datasets of realistic sizes to be considered. GP-GPU computing (via Matlab and CUDA) provides one means by which to do this. By sampling (rather than integrating out) the component-specific parameters and mixture weights in the model, we can exploit conditional independence properties in order to allow computations to be conducted in parallel. This approach promises to enable us to tackle much larger datasets, thereby greatly increasing the range of problems that we can consider.

Thursday 6 September, 11.10am

Approximate Bayesian Computation with GPUs

Thomas Thorne, Juliane Liepe, Sarah Filippi, Michael Stumpf
Imperial College London, London, UK

Objectives

Real-world inference problems are frequently characterized by intractable likelihoods. In such cases so-called approximate Bayesian computation (ABC) schemes allow us to base inferences on simulations and comparisons between simulated and observed data (or meaningful statistical summaries of such data). However, even such approaches can be computationally cripplingly expensive. Here we show how we can employ GPUs in the context of ABC sequential Monte Carlo (SMC) schemes to study problems in molecular evolution and systems biology.

Method/Models

We illustrate the use of GPUs - and discuss the often non-trivial challenges arising from their use - in speeding up ABC SMC inferences in a number of large-scale biological applications. In particular we will study signal transduction systems and the evolution of biological networks. In the former problem we will show how we can use concepts from spectral graph theory to develop convenient descriptions of network data.

Results and Conclusions

Our ABC SMC approach can be easily implemented on GPUs and can lead to considerable speed-ups, which depending on the problem and the care and attention given to the implementation leads, can be approaching two orders of magnitude (compared to CPUs).

We will conclude by outlining some of the challenges inherent to the use of GPUs and how these can be tackled in practice.

Thursday 6 September, 11.10am

Experiences from the world of statistical parallel computing - large scale Bayesian factor analysis (Discussant)

Simon Wilson¹, Jason Wyse¹, Jiwon Yoon²

¹Trinity College Dublin, Dublin, Ireland, ²IBM Research, Dublin, Ireland

Objectives

In this work we describe a high dimensional factor analysis model that is being used for source separation of multi-spectral image data. For the data that we are considering, full implementation of a Bayesian solution would require the inversion of a very sparse matrix of dimension 10^8 by 10^8 . The objective of the work is to exploit the availability of multi-processor units to allow for tractable approximations to this matrix inversion.

Method/Models

First the problem is blocked into smaller units. Each unit can be done independently and in parallel, for which we use a multi-core CPU, with each core running a different block. The more interesting part comes from how to speed up the computation within a block, which is still a sparse matrix inversion. Here we look at GPU-based approaches and will discuss their viability.

Results and Conclusions

The original computation is completely intractable but, with some judicious exploitation of parallel processing, a tractable algorithm that can compute an approximation in a reasonable amount of time has been produced. This is illustrated with the data that we are considering, that come from all-sky microwave maps at several frequencies, collected by the Planck satellite. The data are 9 images, each of about 12 million pixels and the source separation task is to identify the signal from the cosmic microwave background (CMB) from all the other microwave sources. We demonstrate that approximate all-sky reconstructions of the posterior mean and variance of CMB are now possible thanks to these methods.

Detection of spatio-temporal disease clusters

Tom Fanshawe

Lancaster University, Lancaster, UK

Objectives

This talk gives an overview of methods for detecting disease clusters in spatial and spatio-temporal datasets. It highlights the need for an unambiguous definition of what is meant by a 'disease cluster', and the impact on choice of statistical method used to detect them. Approaches based on hypothesis testing of the existence or absence of a disease cluster are compared to those that aim at estimating risk of disease across space and time.

Method/Models

The primary methods compared are the spatial scan statistic method described by Kulldorff (1997), which aims to identify disease clusters through a systematic search through the study region, and the spatio-temporal surveillance method introduced by Diggle et al (2005). These methods are described both in terms of their theoretical justification and in their application to modelling spatio-temporal patterns in incidence of the gastro-intestinal infection giardiasis in north-west England between 2003 and 2011. The effect of risk factors linked to giardiasis risk, including water quality and exposure to livestock, are also assessed.

Results and Conclusions

Methods for detecting spatial and spatio-temporal disease clusters answer different scientific objectives, which should be determined by the context of the application considered. The term 'disease cluster' may not by itself be a precise enough term to capture these various objectives.

In the study of giardiasis, there was some evidence of spatial and temporal variation in incidence, although the association with proposed risk factors was weak. Further work is needed to understand the means of transmission of giardiasis in British populations, and its relation to similar conditions such as cryptosporidiosis.

Diggle, P., Rowlingson, B. and Su, T. (2005). Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics* 16:423-434.

Kulldorff M. (1997). A spatial scan statistic. *Communications in Statistics: Theory and Methods* 26:1481-1496.

A space-time model for investigating the effect of temperature on mortality: an England and Wales case study.

James Bennett, Marta Blangiardo, Daniela Fecht, Majid Ezzati
Imperial College, London, UK

Objectives

It is well established that extremes of temperature, even in a climate such as the United Kingdom's, are associated with raised mortality. However, there remains uncertainty in the temperature-response relationship.

Previous research in this area has concentrated on a two-stage approach, estimating the temperature response in each region (E+W, n=10) or city and then in a second stage combining them.

Here, we describe a district level (E+W, n=376) model from which we explore the temperature response relationship and identify environmental determinants of resilience or risk.

Method/Models

Data consist of postcoded cardio-respiratory deaths (ICD10 I, J) for England and Wales from 2001-2004 linked to mean daily temperatures obtained by interpolating over altitude adjusted Meteorological Office monitoring station data. We employ a case-crossover design in order to reduce confounding.

District level data are jointly modelled within a Bayesian hierarchical framework with implementation via the software WinBUGS. Confounders and possible response modifiers are added to the models.

Both spatially structured and unstructured random effects are included for each district. Temperature response functions focus on linear threshold models and consider both absolute and relative effects of temperature.

Results and Conclusions

There is considerable spatial heterogeneity in the temperature response function which remains even when confounders pollution are added to the model. We show the benefits of small area modelling by estimating the predicted mortality burden at a district level under various climate scenarios.

Ecological bias in studies of the short-term effects of air pollution on health

Gavin Shaddick

University of Bath, Bath, UK

Recently there has been increased interest in modelling the association between aggregate disease counts and environmental exposures measured, for example via air pollution monitors, at point locations. In such cases, the standard approach is to average the observed measurements at the monitoring locations, and use this in a log-linear health model. However, if the pollution surface is spatially variable this simple summary may not provide an accurate estimate of the average pollution concentration across the region, which may lead to bias in estimating the effects on health. In this talk, we investigate the incorporation of spatial variation in exposures within health studies, specifically in the context of the effects of air pollution on health. We investigate the potential for bias that may arise when using predicted concentrations from spatial models arising from underlying spatial variation, monitor placement, measurement error and where the monitoring network is sparse relative to the study area and the number of monitors is not sufficient to characterize the concentration surface. We also consider the problem of ecological bias that can occur when examining associations using aggregate level studies, i.e. examining the associations between aggregate disease counts and environmental exposures measured at point locations. Simulation studies are used to investigate the potential for bias arising from such issues, specifically within studies of air pollution on health. Models are proposed, including Bayesian spatial-temporal models and 'convolution' models, which incorporate spatial modelling of exposures and attempt to address a number of the issues discussed. In addition to being applied to simulated data, these models are implemented within examples of epidemiological studies of air pollution and respiratory health using data from Greater London.

Describing childhood diet with cluster analysis

Andrew Smith

University of Bristol, Bristol, UK

Objectives

Diet is notoriously complicated to record and quantify. However, it is a vital component in the development of many chronic diseases including cancer, cardiovascular disease and diabetes. Many studies of diet have focussed on the intake of individual nutrients, but it is becoming increasingly recognised that people eat foods rather than nutrients. Moreover, people eat foods in combination and therefore large correlations exist between these individual foods and nutrients. Analysing patterns of diet enable us to examine diet as a whole taking into account these correlations and similarities in foods eaten together.

Method/Models

This talk will give an introduction to dietary pattern analysis, using data collected from children in the Avon Longitudinal Study of Parents and Children (www.bristol.ac.uk/alspac) as an example. Dietary data has been collected via food frequency questionnaires and diet diaries, using over 90 food groups, so data reduction techniques such as principal components and cluster analysis are therefore appropriate methods to use to describe underlying dietary patterns in the data. We will focus on cluster analysis, in particular k-means clustering - the most popular method of cluster analysis in the dietary patterns literature.

Results and Conclusions

Cluster analysis of data from food frequency questionnaires yielded 3 clusters (describing a Processed diet, a 'Healthy' or Plant-based diet and a Traditional British diet) and cluster analysis of data from diet diaries yielded 4 clusters (the extra cluster describing a Packed lunch diet). There are a number of potential pitfalls in applying k-means cluster analysis: care must be taken when standardising the input variables, and the standard algorithm is not always reliable. The talk will explain these problems and offer appropriate solutions.

Modelling dependence between extreme values

Jenny Wadsworth^{1,2}, Jonathan Tawn²

¹EPFL, Lausanne, Switzerland, ²Lancaster University, Lancaster, UK

Objectives

Modelling extreme values requires special care due to the scarcity of data, and the requirement to estimate probabilities of events that have not yet happened (but still might). For multivariate data it is often of interest to assess whether the extremes of the different variables are likely occur simultaneously or not. Asymptotic theory motivates the possible class of limit models for multivariate extremes, but these are restrictive when the data are not sufficiently extreme to justify fitting a limiting model, and alternatives are needed.

Method/Models

We give an overview of some alternative characterizations which are sufficiently general to apply to a wide range of underlying distributions, but which do not assume that the limit model holds exactly.

Results and Conclusions

These characterizations can be shown to be more applicable for a variety of dependence structures, and the ensuing estimation methodology permits estimation of probabilities that would be inappropriately estimated by existing methodology.

Statistical Disclosure Control for Communal Establishments in the UK 2011 Census

Joseph Frend

Office for National Statistics, Fareham, UK

Objectives

On the 27th March the UK Census Offices conducted the 2011 Census. The objective of this presentation is to outline the Statistical Disclosure Control (SDC) methodology used to protect the confidentiality of respondents living in Communal Establishments (CEs) in the 2011 Census outputs.

Method/Models

Record swapping is a pre-tabular method of perturbation where the geographical variables of a percentage of households are swapped with those of other households matched on a set of key variables. Record swapping is a common SDC method for full population censuses and has been used in the past by both the UK Census Offices and the US Census Bureau. An expected weakness of the method was how to deal CE residents. This presentation shows how the household targeted record swapping methodology was adapted to cater for CE residents.

Results and Conclusions

We have been able to develop a bespoke disclosure control method for residents enumerated as living in CEs for the 2011 Census. The adapted record swapping methodology targets risky records and matches the sampled records efficiently in an attempt to minimise utility loss whilst providing the necessary protection that the UK Census Offices are legally obliged to provide. For more information on the disclosure control strategy for the 2011 Census, including the CE record swapping methodology, see Frend *et al.* (2012) [*Statistical Disclosure Control in the 2011 UK Census: Swapping Certainty for Safety*, ESSnet - Workshop on Statistical Disclosure Control of Census data, Luxembourg, 19-20 April 2012. Available at: http://neon.vb.cbs.nl/casc/ESSnet2/Frend_ESSnet%20Paper%2019Apr2012.pdf (Accessed: 28 August 2012)]

Communication and interpretation of statistical evidence in the administration of criminal justice.

Colin Aitken

University of Edinburgh, Edinburgh, UK

Objectives

The objective is to aid judges, lawyers, forensic scientists, expert witnesses and others involved in the criminal justice system in their understanding of statistical and probabilistic reasoning in the evaluation of evidence and to aid statisticians in their understanding of evidence in the administration of justice.

Method/Models

The Royal Statistical Society has a working group on Statistics and Law. The membership of the group includes judges, barristers, advocates and forensic scientists as well as statisticians. In early 2010 it received a grant from the Nuffield Foundation with Professor Colin Aitken (Professor of Forensic Statistics, Edinburgh University and Chairman of the working group) and Professor Paul Roberts (Professor of Criminal Jurisprudence, Nottingham University and a member of the working group) as Principal Investigators to produce four reports under the general heading of the title of this presentation.

Results and Conclusions

Two reports have been produced to date with a third due out in September 2012 and the last in January 2013. The first was on the fundamentals of probability and statistical evidence in criminal proceedings, the second on DNA profiling. The third is on the use of Bayesian networks in the evaluation of evidence and the fourth on case assessment and interpretation. There has been great interest amongst those involved in criminal justice. The talk will discuss the work that is being done amongst statisticians, the judiciary, barristers and advocates, and forensic scientists to develop the proper role for probabilistic reasoning in evidence evaluation and in which the Statistics and Law working group of the RSS plays a large part.

Statistical inference in the courtroom: experts getting it wrong?

Ian Hunt

London School of Economics, London, UK

Objectives

John Doe is in court accused of a crime. A finger-print trace of the culprit was found at the crime scene. Doe has a prior, so his finger-print sample is on the police database. Doe was identified as a suspect by a statistician who searched the database for a match between samples and the trace. What is the statistical evidence that links Doe with the crime scene? How can it be presented as a statistical inference? What inferential effect does the database search have? I answer these questions from Bayesian and stereotypical-frequentist perspectives. And I propose a novel modern-frequentist approach.

Method/Models

The court wants clear and justifiable testimony about Doe's presence at the crime scene. The evidence (E) is a "match" between Doe's sample and the trace. The hypothesis (H) is that Doe left the trace. A statistician delivers their "expert witness" testimony as a statistical inference about H and E. I assume Bayesians and frequentists concur about the key statistical models involved. But their inferences can clash. In particular, Bayesians claim stereotypical-frequentists get the inferential effects of database searches wrong. The Bayesians are correct. But a modern-frequentist approach puts things right (and identifies stereotypical Bayesian errors).

Results and Conclusions

Back to Doe's case. The Bayesian testimony will be based on a high likelihood ratio (posteriors being banished from UK courts). The stereotypical-frequentist will testify "either not-H is false or E is unlikely" if he naively assumes Doe was randomly selected from the population at large. But if he uses stereotypical adjustments for the effects of database searching he risks equivocating something like "not-H cannot be rejected".

I argue all this statistical testimony is useless or dangerous. The Bayesian's likelihood ratio says too little about other suspects. The stereotypical-frequentist's inferences are weak and use either false assumptions or logically invalid adjustments.

Getting the testimony right requires explicitly incorporating database search effects and analysing all "likely likelihood" ratios (high ratios from the search and ratios for alternative suspects). Bayesians will need their posteriors to do this properly. My modern-frequentist approach, using Fisherian logic and multiple tests, offers an alternative solution.

The Evaluation of Evidence Relating to Traces of Drugs on Banknotes

Amy Wilson¹, Colin Aitken¹, Richard Sleeman², Jim Carter³

¹University of Edinburgh, Edinburgh, UK, ²Mass Spec Analytical Ltd., Bristol, UK, ³Queensland Health Forensic and Scientific Services (formerly of Mass Spec Analytical Ltd.), Queensland, Australia

Objectives

Banknotes can be seized from crime scenes as evidence of illicit drug use or dealing. Mass Spec Analytical Ltd., an analytical chemistry company, has developed a technique to analyse the quantities of drugs on banknotes. Data are available from banknotes seized in criminal investigations, as well as from banknotes from the general circulation. This talk will describe some of the models used to analyse this data, with the aim of constructing likelihood ratios under competing hypotheses.

Method/Models

For each sample of banknotes tested, the analytical response over time is recorded for five different drugs. A peak detection algorithm used to convert these data into a measurement of the quantity of drug on each banknote will be presented.

There is evidence of autocorrelation between adjacent banknotes in samples. Three models have been developed to take this into account: an autoregressive process of order one, a hidden Markov model and a non-parametric model using conditional density estimates.

Results and Conclusions

Two questions are considered. The first focuses on the likelihood of the data under each of two propositions: that a set of seized banknotes is associated with drug crime, and that these banknotes are from the general circulation. The aim is to evaluate the associated likelihood ratio.

The second question involves the calculation of a likelihood ratio where data are available from bundles within samples, thus the within sample variation may be measured. The propositions are that two samples of banknotes have originated from the same source, and that they have originated from different sources.

The use of the above models in evaluating these likelihood ratios will be described, and the results presented.

Longitudinal quality of life in adults with cystic fibrosis.

Margaret Hurley, Janice Abbott

University of Central Lancashire, Preston, UK

Objectives

Measures of quality of life (QoL) are increasingly important in assessing the health status of adults with long term conditions but few studies have measured QoL longitudinally over many years. From cross-sectional studies involving adults with cystic fibrosis, it has not been clear whether declining lung function is associated with declining quality of life. The objective of this analysis is to investigate the association of lung function with QoL in a longitudinal study.

Method/Models

Our study followed adults with cystic fibrosis (CF) who attended for biannual assessments during a 12-year period. The study measured the nine domains of QoL together with patients' physical health and characteristics. QoL in CF are discrete measures with minimum domain count of zero and maximum domain counts of 15, 20, 30, 40 or 50 and modelling needs to respect the bounded and discrete nature of the data. We used random coefficient models with binomial distributions for QoL via MLwiN to investigate longitudinally the impact of lung function on QoL; an approach which has not been suggested previously.

Results and Conclusions

For each domain of QoL, the models estimated the percentage of CF patients whose QoL declined with declining lung function after correcting for confounders. For all domains except one domain the majority of CF patients were estimated to have a declining QoL with declining lung function in the long term. The random coefficients approach with a binomial model for QoL models the longitudinal nature of the data adequately and permits a more informative answer than would be obtained by a regression with fixed coefficients.

Control charts in medicine: An application to management of patients with bipolar disorder

Maria Vazquez-Montes, Richard Stevens, Joshua Wallace, John Geddes
University of Oxford, Oxford, UK

Objectives

Control charts are a tool widely used in industry for control of industrial processes. A visual display of a process over time is combined with algorithms called 'control rules' that are designed to distinguish systematic change in the underlying process from random noise. We investigate the ability of Shewhart's control rules to predict, and potentially prevent, manic or depressive episodes in patients with bipolar disorder, by applying Shewhart's control rules to weekly self-reported scores from the Altman Self-Rating Scale for Mania (ASRSM) and the Quick Inventory of Depressive Symptomatology (QIDS).

Method/Models

Manic and depressive episodes were defined as an ASRSM score ≥ 10 or a QIDS score ≥ 15 respectively. We restricted analysis to patients in whom a stable run-in period of 8 weeks without an episode could be identified. Mean and standard deviation of each score was calculated over the stable period. Shewhart's rules based on mean and standard deviation were then applied to data from subsequent weeks, and their sensitivity and positive predictive value (PPV) for predicting manic or depressive episodes within the next 4 weeks were calculated. Sensitivity analyses used global, instead of within-person, mean and standard deviation to define the rules.

Results and Conclusions

In 103 patients (37% male) included in the analysis, mean(SD) age was 47(13) years. During median(IQR) 65(26-154) weeks follow-up, 42 patients had manic episodes and 51 depressive episodes. Results show that for manic episodes, the sensitivity and PPV of Shewhart's control rules were 40%(95%CI=25 to 55%) and 5%(95%CI=3 to 7%), and for depressive episodes, 25%(95%CI=13 to 37%) and 7%(95%CI=5 to 9%) respectively. When control rules were defined using global instead of individual mean and standard deviation, sensitivity and PPV for manic episodes were 50%(95%CI=35 to 65%) and 8%(95%CI=6 to 10 %), and for depressive episodes, 31%(95%CI=18 to 44%) and 6%(95%CI=4 to 8%) respectively.

In conclusion, control rules have potential to predict episodes in time for clinical intervention and prevention, especially for manic episodes. Control rules performed best when defined using mean and standard deviation across all patients than when based on individual values during the stable run-in period.

Estimation of spatial spillover effects in cluster randomised trials

Karim Anaya-Izquierdo, Neal Alexander

London School of Hygiene and Tropical Medicine, London, UK

Objectives

In the context of trials where groups of subjects, rather than individuals, are randomised (commonly called cluster randomised trials) to estimate the extent of spillover between neighbouring clusters

Method/Models

A methodology is proposed to estimate the spillover effect, building on models from geostatistics and disease mapping. The spillover effect is represented by a parameter that controls the amount the intervention effect increases or decreases when individuals in the control arm are closer to neighbouring individuals in the intervention arm.

Results and Conclusions

We will show results using data from a paired matched cluster randomised trial against the dengue mosquito vector in Trujillo, Venezuela.

Estimating the treatment effect in a randomised trial with switching using a conditional Poisson model

Zihua Yang, Adam Brentnall, Jack Cuzick, Peter Sasieni

Wolfson Institute, Queen Mary University of London, London, UK

Objectives

The intention-to-treat (ITT) hazard ratio estimator is conservatively biased when some individuals cease to comply with treatment assignment during a two-arm randomised clinical trial. In an extension to all-or-nothing compliance (at the outset of the trial, we examine ways to deal with mid-trial switching by using a conditional Poisson model where the events are Poisson distributed conditional on principal compliance (compliers or non-compliers).

Method/Models

We estimate the complier average causal effect (CACE) by combining the ante-switch treatment effect and post-switch treatment effect amongst the compliers. In addition, we present a time-and-event adjusted estimator of the average marginal treatment by considering the ITT (overall rate ratio) in the absence of switching. The marginal approach may be viewed as a useful alternative to the CACE approach as it does not assume an equal number of potential switchers in both arms at the time of switching - this is not guaranteed if the potential switchers have a different rate to others ante-switching.

Results and Conclusions

Computer simulated examples compare the estimators for three models: (1) conditional Poisson events with random switching, (2) marginal Poisson events with informative switching and (3) piecewise conditional Poisson events with random switching. In general, the time-and-events adjusted estimator captured the marginal average effect quite well in mean but lacked efficiency for the piecewise exponential model. The CACE estimator was found to be generally poorer in terms of bias and efficiency than the TEA estimator for both the conditional and the marginal effect in all three models. This is mainly due to its sensitivity towards the bias in the (assumed equal) switching proportions as well as the unobservable compliance-effect for the ante-switch period.

James Grime

University of Cambridge's Millennium Mathematics Project

Alan Turing was one of our great 20th century mathematicians, and a pioneer of computer science. However, he may best be remembered as one of the leading code breakers of Bletchley Park during World War II. It was Turing's brilliant insights and mathematical mind that helped to break Enigma, the apparently unbreakable code used by the German military. We present a history of both Alan Turing and the Enigma, leading up to this fascinating battle of man against machine - including a full demonstration of an original WWII Enigma Machine!

Dr James Grime is a mathematician and public speaker who works for the Millennium Mathematics Project at the University of Cambridge. He can be mostly found touring the UK, and the world, giving talks about the history and mathematics of codes and code breaking as part of our Enigma Project.

A general class of models for data from mixture experiments

Alexander Donev, Liam Brown

University of Manchester, Manchester, UK

We propose a new class of models providing a powerful unification and extension of the existing statistical methodology for analysis of data obtained in mixture experiments. These models, which integrate the most commonly used models proposed by Scheffé (1958,1963), and the models of Becker (1968,1978), extend considerably the range of mixture component effects which may be described. Unlike previous generally applicable models, they are nonlinear, providing a novel, highly lucid and useful interpretation of the component effects. The novel method of estimation developed for this new class of models not only possesses application to any mixture experiment, but is likely to influence methods of model estimation in other areas. We aim to present the new class of models, their interpretation and estimation using several data sets, some from large industrial experiments carried out by scientists at Federal-Mogul Friction Products Ltd.

Designing real-life industrial experiments: design criteria and challenges

Peter Goos

Universiteit Antwerpen, Antwerp, Belgium

Many industrial experiments are much more complex than the textbooks on industrial design of experiments suggest. Experiments often involve several constraints on the randomization. The interest is often in several responses, and some of the responses may be non-normally distributed. In fact, it is not uncommon to encounter binary responses or ordered categorical responses. Finding a good experimental design is challenging in those cases. In this talk, I will describe some of the challenges and some potential design criteria, in the context of a polypropylene experiment.

Improved split-plot and multi-stratum designs

Steven Gilmour

University of Southampton, Southampton, UK

Many industrial experiments involve some factors whose levels are harder to set than others. The best way to deal with these is to plan the experiment carefully as a split-plot, or more generally a multi-stratum, design. Several different approaches for constructing split-plot type response surface designs have been proposed in the literature in the last 10 years or so, which has allowed experimenters to make better use of their resources by using more efficient designs than the classical balanced ones. One of these approaches, the stratum-by-stratum strategy, has been shown to produce designs that are less efficient than locally D-optimal designs. An improved stratum-by-stratum algorithm is given, which, though more computationally intensive than the old one, makes most use of the advantages of this approach, i.e. it can be used for any structure and does not depend on prior estimates of the variance components. This is shown to be almost as good as the locally optimal designs in terms of their own criteria and more robust across a range of criteria.

The English libel laws and other obstacles to honesty in medical research

Peter Wilmshurst

Consultant Cardiologist, -, UK

Misconduct in medical research, including failure to publish the results of negative trials, harms patients and wastes resources. Dishonesty is encouraged because publication of trials with positive outcomes in high impact journals are well rewarded, but publication of "negative" data is difficult and it brings no reward. Investigators, institutions, sponsors and journals have conflicts of interest that result in suppression of negative trials or converting them to a "positive" message. When misconduct is detected it often escapes effective punishment, but whistleblowers invariably suffer. As a result other investigators realise that it is best to turn a blind eye to misconduct by others or even to emulate the behaviour for similar gains.

I will use my experience of reporting more than 25 doctors to the General Medical Council, of being threatened and offered a financial inducement by a pharmaceutical company to omit data in sponsored research and of being sued three times for libel and slander by a medical device company, NMT Medical, to illustrate my beliefs that we have institutionalised misconduct in medical research and that the libel laws can be used to prevent scientific discussion.

Partitioning uncertainty in climate predictions using data from undesigned climate experiments.

Paul Northrop, Richard Chandler

University College London, London, UK

Objectives

To quantify the relative magnitudes of uncertainty in 21st century climate predictions arising from climate model, emissions scenario and internal climate model variability (i.e. over different simulation runs from a given model-scenario combination), globally and regionally, for surface air temperature and precipitation flux. To overcome problems (lack of balance, sparsity, a factor with a small observed number of levels) caused by the fact that the data arise from an undesigned experiment and to comment on how such experiments could be improved.

Method/Models

Random effects analysis-of-variance, fitted using restricted maximum likelihood and using a Bayesian approach with weakly-informative priors implemented using Markov Chain Monte Carlo.

Results and Conclusions

It is found that the relative contributions to climate uncertainty depend on climate variable, region and time horizon. As expected, scenario uncertainty increases from mid- to late- 21st century. However, for mid-century temperature, model internal variability makes a large contribution in northern regions. For mid-century precipitation, model internal variability is even more important and this persists in some regions into late-century. Thus, different climate variables and regions can have competing design requirements. For global temperature, where internal model variability is small, it is better to use multiple climate models and scenarios than multiple runs at single model-scenario combinations. For precipitation, where model internal variability more prominent, multiple runs for each model-scenario combination are indicated.

Multidecadal Oscillations in Multivariate Marine Ecological Datasets

Victoria Harris¹, Sofia Olhede¹, Martin Edwards²

¹University College London, London ²Sir Alister Hardy Foundation for Ocean Science, Plymouth, UK

Objectives

Changes in plankton communities are influenced by a mixture of natural climate oscillations and the recent warming trend in northern hemisphere temperature. The importance of different climate variables in influencing plankton abundance is heterogeneous in both space and across different species groups. Across the North Atlantic sea surface temperature is on average increasing but there are regions that are warming at a faster rate and other regions which are instead cooling. The aim of this study was to compare indicator species and plankton communities with climate trends.

Method/Models

A large scale statistical analysis is carried out on sea surface temperature data in order to determine how the influence of different climate variables varies spatially. The relationship between these climate variables and plankton abundance is explored by using spatial principal component analysis on various indicator species and species principal component analysis on two different species communities. The community analysis is novel in that it has been studied relatively little using this dataset. This facilitated the modelling of joint responses across functional groups to the various climate trends.

Results and Conclusions

The heterogeneous influence of climate oscillations on sea surface temperature can affect the local behaviour of marine pelagic communities and their spatial distribution. The detrended sea surface temperature in some spatial regions can be well understood using the Atlantic Multidecadal Oscillation (AMO) alone. Others, such as the North Sea, are better understood when the North Atlantic Oscillation is also included.

Although changes in temperature are the main influence on plankton abundance overall, other climate drivers also account for a large proportion of the variation. The AMO is a secondary driver of abundance for zooplankton species. For the Diatom subgroup the main driver is the AMO and not the warming trend. The community analysis can be used to determine which species are on average responding positively to the warming trend and others which are responding negatively.

Using the UKCP09 climate projections to assess uncertainty and variation in projected impacts of climate change on crop yields

Jacqueline Potts¹, David Elston¹, Adam Butler², Shibu Muhammad³, Mike Rivington³, Cairistiona Topp⁴, Helen Kettle², Nikki Baggaley³, Robert Rees⁴, Robin Matthews³

¹*Biomathematics and Statistics Scotland, Aberdeen, UK*, ²*Biomathematics and Statistics Scotland, Edinburgh, UK*, ³*The James Hutton Institute, Aberdeen, UK*, ⁴*Scottish Agricultural College, Edinburgh*

Objectives

Using outputs from climate projections to drive process models is a natural way to investigate many of the potential impacts of climate change. The UKCP09 projections combine climate model outputs with a stochastic weather generator that provides downscaled daily simulations of weather. However, the simulated weather data lack spatial coherence. UKCP09 applied complex methodology, including the use of emulators, to the outputs from climate models, eventually resulting in 10000 equiprobable climate samples. We demonstrate an efficient approach to selecting a subset of these for use as inputs to plant-environment models and explore the uncertainty and variation in the outputs.

Method/Models

We took a stratified random sampling approach to selecting a subset of the UKCP09 climate samples for selected 5km grid squares in Scotland. The stratification was based on projected spring and summer temperature and precipitation. For each specified grid square and selected climate sample, 30 years of daily weather were generated for two future time periods and used as inputs to three different process models: CropSyst for spring barley, DNDC for grass production, and PALM for short-rotation coppice. Variance components analysis of the simulated yields showed the relative contributions of uncertainty about the future climate and annual variation in weather.

Results and Conclusions

The estimates of variance components derived from fitting linear mixed models to these yields indicated that, after allowing for variation between grid squares, between 15% and 75% of the remaining variance in yields could be attributed to uncertainty due to climate projection. This means that stratification provides an efficient means of selecting a subset of the climate simulations to use as inputs to the process models. The fact that the stochastic weather generator used in UKCP09 operates on individual sites limits the usefulness of the UKCP09 projections for assessing the potential impacts of climate change on regional and national yields. The development of suitable multi-site weather generators would provide an obvious solution to this problem. However, in the meantime the use of copulas to impose a dependence structure on the predicted yields may be a possible way forward.

Injectors' initiation of others to injecting drug use in Scotland 2008–09: who initiates, how often, shared responsibility, and novice replacement-rate.

Simon White, Sheila Bird, Sharon J Hutchinson, Avril Taylor
MRC Biostatistics Unit, Cambridge, UK

Objectives

Over 2500 anonymous injecting drug user (IDU) participants were interviewed during 2008–09 in Scotland's Needle-Exchange Surveillance Initiative. We aim to estimate how many others, per annum, an IDU is fully responsible for initiating and which covariates determine whether a respondent initiated in the past year. Further, given the per annum initiation rate, to consider the IDU population dynamics in terms of a replacement-rate.

Method/Models

Interviewees were asked: i) how many times in the past-year have you been present when someone injected for the first time, and ii) how many other IDUs were co-present at the most recent initiation, as well as other demographic questions. Using bootstrap resampling we estimate a per annum initiation rate.

Respondents also reported their incarceration history, which can be used to estimate stepwise-constant incarceration rates per annum. These are combined with the initiation-rates to compute the IDUs' replacement-rate for each of three cessation-rate scenarios. Finally, we estimate how high initiates' quit-rate should be for injectors' replacement-rate to be one

Results and Conclusions

Newly estimated initiation rates identified IDU-locations (prisons, hostels, drug treatment centres) where interventions to reduce initiations could be deployed. Our methods can be used to track trends in initiation incidence. Estimated initiation rate was 0.26 (95% CI: 0.20–0.33), higher among early career IDUs. However, if all initiates persisted with injecting, the overall rate of initiations would only need to pertain for 5 years for IDUs to more than replace themselves. This does not happen, which implies a high initial quit-rate. To maintain a non-increasing number of IDUs requires a high immediate quit-rate for initiates. Further research on injector-initiation and novices' immediate disaffection with injecting is needed; a cessation rate of 15%, required a quit-rate as high as 55% (95% CI: 33%–70%).

Examining the performance of a predictive test using multiple studies: beyond meta-analysis

Richard Riley¹, [Ikhlaaq Ahmed](#)², Jon Deeks¹, Cindy Billingham², Pieter Noordzij³

¹University of Birmingham, Birmingham, UK, ²MRC Midlands Hub for Trials Methodology Research, University of Birmingham, Birmingham, UK, ³Boston University School of Medicine, Boston, USA

Objectives

A predictive test is a single factor that accurately predicts individual outcome risk for patients with a particular condition. For example, in patients with a thyroidectomy, parathyroid hormone (PTH) measured at between 1 to 6 hours post-surgery predicts which patients will become hypocalcemic within 48 hours. Often multiple studies examine the predictive accuracy of a particular test. In this talk, using individual patient data from 9 studies examining the predictive accuracy of PTH, we evaluate statistical approaches for examining a test's predictive ability given multiple studies.

Method/Models

The common approach is to 'lump' the data together and treat as if a single study. In contrast, a bivariate random-effects meta-analysis model accounts for clustering of patients within studies and accounts for any between-study heterogeneity in test accuracy. We compare these approaches in the PTH data, and then use them to evaluate how PTH's predictive performance (such sensitivity, specificity, and the ratio of predicted events (E) to observed events (O)) calibrates in all studies combined and in each study separately.

Results and Conclusions

The average PTH estimates of sensitivity and specificity barely change whether a 'lumping' approach or a bivariate meta-analysis is used. When these are combined with the average study prevalence of disease outcome, PTH is shown to give good predictive accuracy in the lumped data as a whole. However, when applying the average meta-analysis estimates back to each study separately, PTH's predictive accuracy varies dramatically (sometimes over-predicting, sometimes under-predicting), revealing that average predictive probabilities may not be helpful in clinical practice. We show that using study-specific prevalences considerably improves study-specific accuracy, leading to E/O estimates far closer to 1 in each study. Further, when applying PTH to a population where prevalence is unknown, we suggest calculating a 95% prediction interval for the potential E/O value. In summary, focusing on average test performance from meta-analysis can be naïve, and predictive test performance should be checked back in individual clinical settings.

Combining pooled and individual test data to estimate herd-level prevalence

Geoff Jones¹, Wes Johnson², Cris Verdugo¹, Cord Heuer¹

¹Massey University, Palmerston North, New Zealand, ²UC Irvine, Irvine, CA, USA

Objectives

The use of an expensive but sensitive diagnostic test with pooled samples can be a cost-effective way of monitoring herds for the presence of disease. For example the faecal culture test for Johne's disease can be applied to pooled faecal samples derived from random samples of animals in a deer herd, rather than to individual animals, to test the infection status of the herd. If a less accurate but cheaper test is available, it may be advantageous to also apply this to individual animals in a defined testing regime. We explore the estimation of such a regime's performance.

Method/Models

In the absence of a gold standard test, latent class analysis in a Bayesian framework is used to estimate the test sensitivities and specificities, along with the prevalence of disease, using data from two-stage sampling of animals within herds. The usual approach for combining the results of two tests is to condition on the true disease status of each individual animal, but here because of the combination of pooled and individual tests it is necessary instead to condition on the true status of the within-herd samples. It is also necessary to consider both herd-level and within-herd prevalences.

Results and Conclusions

In practice the model cannot be fitted without informative prior information on the distribution of within-herd prevalences. Modelling the herd-level and within-herd prevalences needs careful consideration.

Analysis of proportional admissions/mortality: proposed odds ratio methods to overcome flaws in standard methodologies.

Paul Fryers, David Jephson

East Midlands Public Health Observatory, Mansfield, UK

Objectives

We were asked to provide an indicator of ethnic inequalities based on emergency admissions. Because of intrinsic problems in relating ethnic coding of HES data to population estimates by ethnic group, we opted to use proportional admission ratios, which present emergency admissions as a proportion of all admissions, rather than emergency admission rates. There are several examples of this approach in the literature, eg proportional mortality ratios, but we discovered what we believe to be a fundamental flaw in the standard methodology. A literature search showed that the flaw has been flagged repeatedly, but is routinely overlooked.

Method/Models

We developed an indirectly standardised proportional admission ratio, whereby the number of emergency admissions is presented as a proportion of all admissions, standardised indirectly for age. The precise method does not appear to have been proposed before. It will be fully explained and justified in the conference paper: fundamentally, it uses odds within the calculation rather than proportions, maintaining the integrity of the underpinning binomial assumptions, ensuring that the final proportions are constrained within zero and one and are invertible (ie the proportion of non-emergencies can be calculated instead of the proportion of emergencies, with the same results).

Results and Conclusions

Authoritative sources going back to the 1970s (eg Breslow & Day, Miettinen) point out that it is only appropriate to use Poisson-based proportional methods with small proportions – otherwise binomial-based methods must be used. Miettinen and Wang proposed an age-adjusted odds ratio method in 1981. However, it appears that this guidance, and the odds ratio alternative, are routinely ignored, with mainstream publications (examples will be given) using the Poisson-based method for large proportions, resulting in distorted results including, in extremis, proportions that exceed one.

We have developed a second alternative measure for standardised proportional ratios, also based on odds ratios, which we believe should be used in specific, but not uncommon circumstances, alongside the Miettinen and Wang method.

We will present a clear analysis of the range of circumstances when each method should be used, which we believe constitute solid guidance for future analysis of proportional mortality/admissions analysis.

Bayesian Estimation of Animal Movements in Heterogeneous Environment Using Data Augmentation

Svetlana Tishkovskaya¹, Paul Blackwell²

¹University of Central Lancashire, Preston, UK, ²University of Sheffield, Sheffield, UK

Objectives

Advances in the field of wildlife radio-tracking, and in Global Positioning System telemetry in particular, have made it possible to acquire systematically scheduled data on many aspects of wildlife biology, including detailed information on animal movements and habitat use. These complex, spatially explicit ecological data require, in turn, advanced statistical and computational methods to characterise the details of movement and to relate the animal to its environment.

Method/Models

In this work, we develop a model for analysis of individual animal movements in a heterogeneous space.

Radio-tracking data are used to estimate both movement parameters and unobserved boundaries between landscape fragments. The data augmentation technique is employed to reconstruct the partition of the heterogeneous environment. In general, the process of multiple imputation carried out through data augmentation can be used to augment missing observations or to predict animal movements. The approach to inference is Bayesian using Markov chain Monte Carlo methods.

Results and Conclusions

The augmentation has provided a tool to diminish the impact of uncertainty about boundary crossing and potentially to extract some additional information from the given observations. The suggested methodology is illustrated on simulated data and applied to real movement data from a radio-tracking experiment.

Estimating the Oceanic Aircraft Separation Loss Distribution under Periodic Reporting

Jacqueline Civil

NATS, Whiteley, Fareham, Hampshire, UK

Objectives

A change in Oceanic air traffic control procedure is being tested to allow in-trail aircraft to fly closer together while transmitting position reports more frequently. To assess the impact on collision risk of the proposed changes, a constant variance AR(1) model for incremental aircraft speed error was previously used to estimate the separation loss distribution for a pair of aircraft over a fixed reporting period. This model has been found not to match the behaviour of real aircraft transit data, and an improved model is required.

Method/Models

The behaviour of the model was assessed by resampling the key factors of distance and forecast speed from a set of real aircraft Oceanic transit data and using them with the model to simulate the forecast transit time error. A process of trial and error was used to identify the functional changes to the original model necessary to replicate the key behaviours in the real data. An optimal parameter set for the new model was then identified using WinBUGS.

Results and Conclusions

The final model is an AR(1) model with non-constant variance which is a quadratic function of forecast speed multiplied by an additional unknown factor. The unknown factor is yet to be understood, but it is known that it must have a Rayleigh(1) distribution. This model replicates the observed behaviour of the variance of forecast transit time error by transit time, and the Normal-Laplace distribution of the forecast transit time error. The improvement to the model gives a much greater confidence in its ability to identify the appropriate reporting period required to safely allow reduced separation minima between aircraft.

Adult body mass index trends are levelling off in England: exploration of underlying patterns using a two-component model

Matthew Sperrin¹, Alan Marshall², Vanessa Higgins³, Iain Buchan⁴, Andrew Renehan⁵

¹Department of Mathematics and Statistics, Lancaster University, Lancaster, UK, ²School of Geography, University of Leeds, Leeds, UK, ³CCSR, School of Social Sciences, University of Manchester, Manchester, UK, ⁴School of Community Medicine, University of Manchester,, Manchester, UK, ⁵School of Cancer Studies and Enabling Sciences, Manchester, UK

Objectives

Excess body weight, commonly measured as body mass index (BMI) ≥ 25 kg/m², has become much more prevalent in many populations worldwide over the past three decades, but is levelling off in some western populations (including England), the reasons for which are unclear. We hypothesise that the levelling off reflects a resistant sub-population that tends to a normal BMI, and explore this using a two-component mixture model.

Method/Models

We used the Health Surveys for England (an annual cross-sectional household based survey of adults aged 20 to 75 years: 1992 to 2009), to derive a pseudo-panel. We described baseline trends in mean BMI using regression models with splines. Recognising the increasing variance in BMI with time, we generated a two sub-population model by gender using a latent class analysis approach. We compared fit of 2- versus 1-component models and performed sensitivity analyses to assess model assumptions.

Results and Conclusions

The 2-component model clustered normal and high BMI sub-populations with proportions for high BMI being 23.5% for men and 33.8% for women. The normal BMI sub-populations appear 'resistant' to weight gain with minimal increases over time. By age, we noted that high BMI population increased greatest between 20 and 34 years for men, then levelling off; for women, the increases were steady throughout ages.

In England, median BMI has increased since 1992 with levelling off between 2001 and 2009. The trends can be explained by a high BMI sub-population getting 'fatter' - levelling off being explained by a constant proportion 'resistant' normal BMI subpopulation. These findings suggest the need to consider targeted prevention of obesity alongside population wide public health initiatives.

Risk in business, and the business of risk

David Hand

Imperial College, London, UK

In business, risk is one side of a coin, the other side of which shows reward. Our aim, then, is not simply to minimise risk, but to optimise some appropriate combination of risk and reward. Statistical modelling is central to this optimisation task. This talk explores some aspects of the interaction between risk and statistical modelling, in the context of business applications, giving examples from several areas.

Subjective Well-Being: A beginner's Guide

Thomas Burnett, Isabella Saunders, [Fola Ariyibi](#)
Office for National Statistics, Newport, UK

Objectives

As part of the National Well-Being programme of work, ONS have started to investigate how to measure 'subjective well-being' through the inclusion of additional questions on the Opinions Survey. The Opinions survey is a monthly survey used to ask a variety of questions for different government departments. Using the Opinions survey for the experimental subjective well-being questions meant that data could be gathered on a range of different questions relatively quickly.

Method/Models

In December 2011, ONS published a report on this experimental data for subjective well-being. This included some basic estimates for subjective well-being from the four "headline" questions included on the Opinions survey, as well as discussion about interview mode effects and the additional questions that were asked.

Results and Conclusions

The poster will focus on the methodology behind the publication, detailing how the estimates were generated, and the techniques used in investigating the effect of interview mode.

A coupled hidden Markov model for disease interactions

Tatiana Xifara, Chris Sherlock

Lancaster University, Lancaster, UK

Objectives

In natural populations, animals are likely to be infected by different diseases either simultaneously or successively. Ecologists are interested in the interactions between parasite species. A longitudinal study was designed for that purpose by recording the sequences of infections events in a population of field voles. As the dataset contains many missing observations, a simple analysis is not preferred as it would entail discarding much of the dataset. We offer a more powerful alternative.

Method/Models

We use a discrete-time hidden Markov model for each disease with transition probabilities dependent on covariates via a set of logistic regressions. For each disease we wish to evaluate the way in which the presence or absence of each of the other diseases affects the probability of contracting it.

Inference is performed via a Gibbs sampler, one iteration of which cycles through each of the diseases, first sampling from the conditional posterior for the parameters for that particular disease using an adaptive Metropolis-Hastings step and then sampling from the hidden disease states using the Forward-Backward algorithm.

Results and Conclusions

The Markov model offers a more detailed description than the existing modelling approach and the inference methodology that we introduce is able to use more of the data than the existing standard inference methodology. This enabled us to identify interactions among the six diseases in our dataset.

Welsh Index of Multiple Deprivation - An Overview

Rebecca Gillard

Welsh Government, Wales, UK

Objectives

The Welsh Index of Multiple Deprivation (WIMD) is the official measure of deprivation for small areas in Wales. The most recent index was published in 2011.

Method/Models

Deprivation is a wider concept than poverty alone - it refers to a general lack of resources and opportunities. Currently WIMD is calculated from 40 separate deprivation indicators. These indicators range from income and employment benefits to measures of health, education and crime. An area has a higher overall deprivation score than another area if the proportion of people living there who are classed as deprived is higher.

Results and Conclusions

This poster provides an overview of the Welsh Index of Multiple Deprivation and outlines how the index is calculated. Some graphical illustrations of the types of analyses that can be performed on WIMD and its underlying indicators are demonstrated. Details of recent and future developments are given along with examples of how WIMD is used to inform policy making.

New approaches to Response Chasing in ONS

Ann Langford

ONS, Newport, UK

Objectives

ONS is currently investigating a number of different approaches to response chasing in stratified business surveys. The first approach is based on the examination of key results at different levels of stratification during the response chasing period and prioritising those strata where results change substantially. The second approach is based on prioritising the strata that have relatively high sampling fractions. The final approach is based on allocation of priority scores to individual businesses. Initial results from this project are presented and discussed.

Method/Models

The effect of the three approaches on the bias and variance of the final estimate are examined by means of a simulation study on a super-population derived from the ONS Monthly Business Survey.

Results and Conclusions

Initial results from this project are presented and discussed.

Dental Patients, Public and the Dental Professionals: Attitudes to Regulation

Guy Rubin

General Dental Council, London, UK

Objectives

To present key statistical findings from major opinion surveys carried out in 2011 and 2012 with the General Public, Dental Patients and Dental Professionals (Dentists and Dental Care Professionals) carried out by independent research organisations for the General Dental Council (GDC). The submission will summarise, compare and contrast professional and public attitudes and opinions on a range of subjects relating to the regulation of dentistry and dental standards, and key policy and topical issues in Dentistry. The presentation will explain its role in evidence based policy making at the GDC and will feature examples where the findings have influenced key policies.

Method/Models

The patient and public survey involved a representative sample of patients and the general public across England, Wales, Scotland and Northern Ireland. Face to face interviews were carried out in spring 2011. The surveys of Dental Professionals were carried out with a representative sample of Dentists and Dental Professionals and involved the sample completing on-line questionnaires. A core set of questions were replicated in each survey to allow for comparison. The presentation will consider issues of comparability created as a result of different sampling methodologies being used.

Results and Conclusions

There is strong support for Dental Regulation among the General Public, Patients and Dental Professionals.

Though patients and the public are aware that Dentists is regulated, knowledge and awareness of responsibility for regulation is limited.

There is a broad consensus between Dentists, Patients and the Public are key priorities for Dental Standards

There is support and satisfaction with the priorities and performance of the GDC, though there are some areas where performance needs to be improved.

Investigation of the effects of life course events and decisions over the long term on the physical and mental health of older people.

Riza Momin

Kingston University, Kingston upon Thames, Surrey, UK

Objectives

Increased life expectancy in the UK and the current Government's policy of extending the retirement age has implications for the country's ageing population, including the importance of their mental and physical well-being. This project aims to investigate publicly available longitudinal data to determine the long-term effects of life course events and decisions on mental and physical health in later life, as well as to apply appropriate statistical methodologies to evaluate the effects of circumstances, behaviours and choices over a period of time on the mental and physical well-being of the older person.

Method/Models

This study will investigate several approaches to examining the effects of life course events and decisions on the health of older people. The work involves the application and development of existing and emerging statistical methodologies that deal with longitudinal studies. For this, publicly accessible longitudinal datasets will be examined, the most important ones being the English Longitudinal Study of Ageing (ELSA) and the National Child Development Study (NCDS).

Results and Conclusions

There is increasing evidence that the effects of life course events and decisions over the long term can affect mental and physical health and general well-being in older age. This project uses longitudinal datasets that include measurements on engagement with social activities (such as organisational memberships of sports clubs, social clubs, charitable associations, political parties and educational classes), taking into account age, sex, self-rated health and NS-SEC, as well as measurements on leisure-time activity levels by occupational class, age and sex. This poster presents some initial results from the investigation.

Global Sensitivity Analysis: local polynomial estimation of Total Sensitivity Indices

Matieyendou LAMBONI

Institute for Environment and Sustainability, Ispra, Italy

Objectives

Global Sensitivity Analysis (GSA) methods [1], [2] and Multivariate Sensitivity Analysis (MSA) methods [3], which aim to apportion model output variability into input variables and their interactions, are an objective way to evaluate the impact of the uncertainty in input variables on the model output. In particular, The Total Sensitivity Index (TSI) gives for each input its overall contribution, including the effects of its interactions with all the other inputs, in the variance of the model output. The computation of TSI requires a large number of model evaluations [2]. We investigate a methodology to compute TSI with few model evaluations.

Method/Models

Under the assumption of independent input factors Z_1, \dots, Z_d , the model output $Y(Z_1, \dots, Z_d)$ can be expanded like a sum of three components [4]: the mean of the model output; the function $g()$ that contains all the information about Z_i , with $i = 1, 2, \dots, d$, and the function $h()$ that does not include Z_i . TSI is the square integral of $g()$ [4]. In practice, function $g()$ is not known and should be estimated. Based on the Taylor expansion, we show that the conditional expectation of $Y(Z_1, \dots, r^*Z_i, \dots, Z_d) - Y(Z_1, \dots, Z_d)$ is, locally, proportional to the function $g()$, where $r \sim N(1, \sigma)$.

Results and Conclusions

We propose a local polynomial estimator of TSI of Z_i , with $i = 1, 2, \dots, d$, and we compute the TSI in two steps. We, firstly, fit the multivariate local polynomial for each input, and secondly, use the fitted function to estimate the TSI. Multivariate local linear or quadratic lifting, theoretically, requires a reasonable number of observations ([5]), and this should result in few model evaluations for the TSI estimation. Some numerical tests of the estimation of the TSI for Sobol's function show the reduction, up to factor 10, of the number of model evaluations. The reduction becomes more significant when the model output is controlled by a small number of input variables and when the model does not include much interaction among input variables. For a high dimensional model, the upper bound of the TSI ([4]) is used for preliminary factor screening.

Longitudinal modelling for monitoring progressive chronic diseases in computerised GP records : a case study in Chronic Kidney Disease(CKD)

Rosie O'Neil¹, Penelope Bidgood¹, Simon deLusignan², Zalihe Yarkiner¹, Sybel Joseph¹

¹Kingston University, Kingston upon Thames, UK, ²University of Surrey, Surrey, UK

Objectives

This project is aimed at identifying, developing and applying a longitudinal modelling framework to investigate the natural history of chronic disease, using routinely collected GP records. Here the application is to Chronic Kidney Disease (CKD) but the methodologies should be applicable to many other progressive chronic disorders where long term management of a deteriorating condition is required.

Aims include development of methodologies for, identifying and evaluating factors related to the diagnosis and management of chronic diseases; and mechanisms for identifying early indicators of chronic disease which may be used in early diagnosis and preventative treatment.

Method/Models

This project is aimed at identifying, developing and applying a longitudinal modelling framework to investigate the natural history of chronic disease, using routinely collected GP records. Here the application is to Chronic Kidney Disease (CKD) but the methodologies should be applicable to many other progressive chronic disorders where long term management of a deteriorating condition is required.

Aims include development of methodologies for, identifying and evaluating factors related to the diagnosis and management of chronic diseases; and mechanisms for identifying early indicators of chronic disease which may be used in early diagnosis and preventative treatment.

Results and Conclusions

The application here is the natural history of Chronic Kidney Disease (CKD), a relatively recently recognised condition that affects up to 10% of the population. It is a multi-stage, progressive disorder classified between stage 1 (mild) and stage 5 (renal failure), but most commonly diagnosed between stages 3 and 4. CKD is associated with increased all cause mortality and co-morbidities, however the natural history of CKD and its complex associations with other conditions is currently unknown. Interventions which can be delivered in primary care can slow or stop its progression but the current lack of knowledge is recognised to be an issue in the management of CKD. Early results on a sample of more than 80000 patients indicate that identification of trends in routine laboratory tests are linked to earlier diagnosis and management of associated co-morbidities may help to slow progression.

Measuring the number of international visits to UK during London 2012: how the International Passenger Survey is adapting its design

Roger Smith, Paul Smith

ONS, Newport, UK

Objectives

To present to users the challenges of producing accurate estimates of visits to UK by overseas residents during London 2012, and to outline how the International Passenger Survey has adapted its design to address this special event.

Method/Models

Sample design.

Weighting.

Results and Conclusions

By making changes to the IPS sample design during August 2012 it has been possible to increase robustness of estimates of visits and earnings in the period around London 2012, in response to anticipated fluctuations in passenger flows.

Investigate the pattern of Chlamydia incidence in England using space-time models

Man Ying Edith Cheng

Southampton University, Southampton, UK

Objectives

The objectives of this study are (i) to explore whether the incidence and risk pattern of Chlamydia varies over regions and time in England between 2009 and 2011, (ii) to investigate how the time element within the disease mapping models can be used to improve and strengthen the interpretation and knowledge about the overall disease risk pattern, (iii) to identify, if there is variation over time or regions or both, the unstable time period or regions so that the limited healthcare resources can be planned ahead the potential benefits for the patients can be maximised.

Method/Models

Chlamydia is the most commonly diagnosed sexually transmitted infectious disease in the UK. The National Chlamydia Screening Programme is free of charge for everyone who is aged under 25 in England. The incident cases were reported periodically by the Health Protection Agency.

Bayesian, hierarchical, spatial temporal models were applied to explore the spatial and time variation in risk patterns. The time, spatial and space-time interaction were included in the model. Different assumptions were applied to classify regions and time periods as stable and unstable disease patterns in this study based on the posterior distribution of the space-time interaction.

Results and Conclusions

The results highlighted some regions had higher risk e.g. the Midlands, South West England and some other major cities in the south. The incidence rate varied across the seasons. This may be associated with some underlying risk factor (e.g. deprivation). By targeting the higher risk seasons and regions we may help to prevent further infections. The results strengthen the knowledge about Chlamydia disease patterns in England and the outcomes may potentially be used to detect the local hotspots and high risk seasons. This information allows the healthcare services to plan their limited resources and to reduce misallocation. The results were displayed on maps, which can be used to communicate with health care professionals as well as the general public. When there are spatial, time or both variations in England then different screening policies can be considered to improve the efficiency of screening.

DETERMINATION OF BROILER CHICKENS' MAXIMUM WEIGHT WITH RESPECT TO AGE AND FEEDING RATE

O. M. OLAYIWOLA¹, G.N. AMAHIA², O. A. OYEGUNLE³, A. A. ADEWARA⁴

¹STATISTICS DEPARTMENT, FEDERAL UNIVERSITY OF AGRICULTURE, ABEOKUTA, OGUN STATE, Nigeria, ²STATISTICS DEPARTMENT, UNIVERSITY OF IBADAN, IBADAN, OYO STATE, Nigeria,

³STATISTICS DEPARTMENT, UNIVERSITY OF IBADAN, IBADAN, OYO STATE, Nigeria, ⁴STATISTICS DEPARTMENT, UNIVERSITY OF ILORIN, ILORIN, KWARA STATE, Nigeria

Objectives

Livestock industry is ridden with myriad of problems, which resulted to a gross shortage of meat and other animal-products. The growth rate of agriculture sector is still below the potentials of the natural and human resources due to high cost of agricultural inputs. To increase protein intake, there is urgent need to increase broiler production at household and commercial holdings. This research was conducted to assess the weight gained by the broiler chickens served with fluted pumpkin leaves extract, to examine the possible combination of number of weeks and quantity of fluted pumpkin leaves extract that result in maximum weight.

Method/Models

Experiment was conducted to assess the weight gained by the broiler-chickens served with fluted pumpkin leaves-extract (FPLE). Forty day-old broiler-chicks were randomly distributed to 5 treatments which contained 0, 30, 60, 90, 120 ml of FPLE per litre of water in a completely randomized design. Each treatment was replicated four times with two birds per replicate. The feed and water were served. Factorial-design and Response surface methodology were used to study the effects of number of weeks and quantity of FPLE on the weight of broiler-chickens and to capture the combination of weeks and quantity of FPLE that brings maximum-weight.

Results and Conclusions

The results showed that the average body weight gained was significant ($P < 0.05$). It was least in control compared to the birds served with 30-120 ml of FPLE. Factorial Design revealed that birds served with FPLE gained more weight than those in control. The use of FPLE in broiler chickens production is most effective from five weeks of age. The fitted Response Surface Model indicated that number of weeks and quantity of FPLE together with their mutual interaction significantly ($P < 0.05$) determined the weight of broiler chickens. The maximum weight was achieved when number of week was ten with 100ml of FPLE.

Number of weeks, FPLE and their mutual interaction play a key role in obtaining maximum weight of broiler chickens. These factors should be put into consideration in making of feed for broiler chickens.

STATISTICAL ANALYSIS OF PATIENTS' WAITING TIME IN THE HOSPITAL

O. M. OLAYIWOLA¹, O.J. OGUNDIPE², A. A. ADEWARA³, G. N. AMAHIA⁴

¹STATISTICS DEPARTMENT, FEDERAL UNIVERSITY OF AGRICULTURE, ABEOKUTA, OGUN STATE,

Nigeria, ²STATISTICS DEPARTMENT, FEDERAL UNIVERSITY OF AGRICULTURE, ABEOKUTA, OGUN

STATE, Nigeria, ³STATISTICS DEPARTMENT, UNIVERSITY OF ILORIN,, ILORIN, KWARA STATE, Nigeria,

⁴STATISTICS DEPARTMENT, UNIVERSITY OF IBADAN, IBADAN, OYO STATE, Nigeria

Objectives

Long waiting time of the patients in the hospital has been identified as the major of patients failure to present themselves for the care in the hospitals. This research examined the waiting time activities of Ear Nose and Throat (ENT) out-patients clinic, University College Hospital, Ibadan, Nigeria. The research is to determine the extent and causes of long waiting time by the patients, when trying to see the doctor.

Method/Models

Direct observation method was used to obtain the data and they were collected between the hours of 7.00 am and 4.00 pm from Monday to Friday. An interval of hourly bases was taken as a unit of time. Causes of long waiting time, the traffic intensity, queue length and waiting time were determine.

Results and Conclusions

Inavailability of facilities and equipments, Lack of sufficient human resources, and patients registration process were the major causes of long waiting time of patients. Tuesday was found to has the greatest traffic intensity of 0.9722 with queue length of 34 patients and waiting time of 4.22 minutes. More patients visit hospital tuesday than any other days. We recommend that more personnel be made available to increase the efficiency of the personnel and to reduce the queue length and the service time per patient.

Bandwidth matrix selectors for multivariate kernel density estimation

Ivana Horova¹, Jan Kolacek¹, Kamila Vopatova²

¹Masaryk University, Brno, Czech Republic, ²University of Defence, Brno, Czech Republic

Objectives

The most important factor in a multivariate kernel density estimation is a choice of a bandwidth matrix. Because of its role in controlling both the amount and the direction of multivariate smoothing, this choice is a particularly important. This contribution aims to develop data-driven selectors of full bandwidth matrices for a density and its gradient.

Method/Models

The proposed method is based on an optimally balanced relation between integrated variance and integrated squared bias. The analysis of statistical properties shows the rationale of the proposed method.

Results and Conclusions

In order to compare this method with cross-validation and plug-in methods the relative rate of convergence is determined. The utility of the method is illustrated through a simulation study and application to real data. The advantage of the proposed method consists in the fact that it does not need any pre-transformation of the data.

Effectiveness of Keep-In-Touch-Exercises on Response Evidence from the Wealth and Assets Survey (WAS)

Martina Aumeyr

Office for National Statistics, Newport, Gwent, UK

Objectives

The Wealth And Assets Survey (WAS) was subject to an experiment over a 12 month period. Three different types of Keep-In-Touch-Exercises (KITE) were used as treatments on sample members: a five minute long telephone interview to up-date respondent contact details, and a newsletter which was posted to respondents informing them about the use WAS data, and both of these treatments together. The aim was to establish the effectiveness and impact of these treatments on the response rates of the next wave. Furthermore, the experiment should provide information on how well the treatments support the tracing process of original sample members.

Method/Models

Detailed analysis of the experiment data is planned including hypothesis testing and statistical modelling to identify differences in the response rates at the following interviewing wave, and establish any differences in the characteristics of responders and movers between all four experiment groups. The number of ineligible addresses identified with the KITE will also be explored further. An interim report with indicative results of the first six months of the experiment year was produced in March 2012. A final report presenting results of the complete experiment year will be produced in August 2012.

Results and Conclusions

At this stage only the first six months of the experiment year are available for analysis, therefore only descriptive analysis has been conducted so far. The results show that there are only marginal differences in response between all four experiment groups. However, it can be observed that treatment groups who were sampled for a telephone interview, or a telephone interview and a newsletter have slightly higher response rates in the next wave compared to the control group. The treatment group who only received a newsletter has a slightly smaller response rate than the control group. Focusing only on the treatment groups that were sent the newsletter, results suggest that there is a difference in response depending on the length of time between receipt of the newsletter and the next wave interview. Furthermore, the number of movers identified within each experiment group appears to differ.

Cross-national measurement equivalence in latent variable modelling: Sensitivity and model selection

Jouni Kuha, Irini Moustaki, Sally Stares, Jonathan Jackson

London School of Economics, London, UK

Objectives

In latent variable modelling of cross-national survey data, substantive interest usually focuses on comparisons of the distributions of latent variables across countries. Measurement models of the observed variables are of lesser interest, but nevertheless need to be specified them appropriately before cross-national comparisons on the latent constructs can be made. A potentially serious problem is lack of measurement equivalence, where some survey items do not function similarly across all the countries. This talk examines issues in the analysis of measurement equivalence, focusing in particular on model selection and the sensitivity of conclusions to ignoring lack of equivalence.

Method/Models

We consider latent variable models for categorical observed variables, i.e. latent class and latent trait models. Under a range of true models, we use numerical methods to examine which values parameter estimates converge to for different correctly and incorrectly specified models, and which models are selected well different methods of model selection.

Results and Conclusions

Results are reported on how sensitive are conclusions on comparisons of latent distributions to different incorrect assumptions about measurement equivalence, and how different methods of model selection perform in this context.

What makes for a good missing value imputation? New algorithmic approaches in PCA and random forests-based techniques

Nadia Solaro², Alessandro Barbiero¹, [Giancarlo Manzi](#)¹, Pier Alda Ferrari¹

¹Department of Economics, Business and Statistics - University of Milan, Milan, Italy, ²Department of Statistics - University of Milan-Bicocca, Milan, Italy

Objectives

The increasing availability of data often characterized by missing values has paved the way for the development of new powerful algorithmic imputation methods for handling missing data, among which two recent proposals seem most promising: Stekhoven and Bühlmann's method (*missForest*), a nonparametric technique based on a random forest, and Josse, Pagès, and Husson's imputation method (*missMDA*) based on the EM-PCA algorithm. In this work, a new PCA-based procedure is developed by drawing on the forward-imputation approach introduced by Ferrari, Annoni, Barbiero, and Manzi in the context of ordinal data (*ForImp*). Comparisons with the two methods above are then considered.

Method/Models

The proposed method (*ForImpPCA*) is designed as a forward-imputation technique for quantitative data. Missing values are recovered by averaging donors' values detected through a sequential procedure, which involves applying both PCA and NNI methods on the complete data submatrices derived from each imputation step. Comparisons with *missForest* and *missMDA* are made through a simulation study based on random generating, firstly, complete data matrices from different multivariate distributions (multivariate normal, Azzalini's skew-normal, multivariate exponential power distributions) and under different settings (number of variables, correlation or association structures, skewness or kurtosis parameters), and then MCAR missing data in different percentages.

Results and Conclusions

Simulations carried out up to now reveal that *missForest* has generally a slightly less satisfactory performance than *missMDA* and *ForImpPCA*. *missMDA* shows to perform better than *ForImpPCA* in presence of multivariate normal and exponential power distributions as the number of variables and correlation coefficients increase, regardless of the percentage of missingness. *missMDA* and *ForImpPCA* have a similar, overlapping performance when the variables are few, *ForImpPCA* indeed performing better when variables are correlated to a different extent, or when data are skew-normally distributed. The simulation study involving skew-normality is however still tentative, a small part only of the great variety of skewness patterns having been considered. Finally, our simulation results have evidenced the existence of turning points, related to the different experimental settings, where a given technique should be preferred to the others. To this issue therefore our attention will be specifically addressed.

How many subjects do I need in a pilot RCT to estimate the variance of the outcome with a reasonable degree of precision?

Neil Shephard, Munya Dimairo, Alexandra Hayman, Dawn Teare, Stephen Walters
University of Sheffield, South Yorkshire, UK

Objectives

Pilot studies are conducted before a main study in order to answer the question "Can this study be done?". They are sometimes used to estimate important parameters (e.g standard deviation (SD) of the outcome) that are needed to inform the design of the main study. This work aims to answer the question: "how many subjects are needed in a pilot randomised controlled trial (RCT) to estimate the critical parameter (SD) needed to design the full definitive RCT with a reasonable degree of precision?" We limit consideration to pilots of two arm RCTs with a continuous outcome.

Method/Models

Two treatment arms were simulated with outcome taking $N \sim (0,1)$ in one arm and varying standard deviation (SD) in the second arm between 0.4 and 2.0. Sample size varied in the range 2 to 70. Pooled and treatment specific empirical SDs were calculated and deviation from theoretical parameters investigated. Skewed data was simulated from $\Gamma \sim (2,2)$ in one arm and varying the scale and shape parameters in the second arm between 2 and 3 and again empirical SDs were compared to theoretical values. $N = 10000$ simulations were performed under each simulation.

Results and Conclusions

For normally distributed variables at low sample sizes the SD is consistently underestimated, but as sample size increased the accuracy and precision of the estimated SD increases. However above $N = 40$ there is little improvement in the precision of estimates. Smaller variances are estimated more accurately and precisely, but in the absence of evidence to suggest otherwise it is inappropriate to assume that the variance will be identical in both treatment arms. It is recommended that the treatment arms are not pooled to estimate the standard deviation and a minimum of 40 individuals are recruited to each arm of a pilot studies. In such cases pooling results in poor estimation of the true variation in either arm. Similar results are presented and commented on when the underlying variable is skewed.

How big should a pilot trial be? A case for binary outcomes in Randomised Controlled Trials.

Munyaradzi Dimairo^{1,2}, [Alex Hayman](#)^{1,2}, Neil Shephard^{1,2}, Dawn Teare¹, Stephen Walters^{1,2}

¹ScHARR, University of Sheffield, Sheffield, UK, ²NIHR RDS for York and Humber, Sheffield, UK

Objectives

There is great debate over what sample size should be used for pilot trials to inform the design of definitive randomised controlled trials. Some recommendations have been developed with no consensus. The majority of the recommendations focus on estimating the variation of a continuous outcome and relatively little attention is paid to binary outcomes. The aim of this research was to investigate the sample size required to accurately and precisely estimate parameters such as event, consent and attrition rates. We assume that the pilot study (and the full RCT) will be a two group treatment versus control or placebo trial.

Method/Models

We simulated binary outcome data from binomial distributions for varying sample sizes of 10 to 200 in steps of 5 and outcome probabilities of 0.1 to 0.5 in steps of 0.1.

We assessed the precision of the estimates by calculating the mean absolute error, the gain in precision through the width of the 95% exact confidence interval (CI) and the percentage gain in precision per increase in sample size of 5. The proportion of times the sample estimate was smaller than the true population parameter was also investigated.

Results and Conclusions

The mean absolute error suggested that a sample size of around 75-100 was required to produce an accurate event rate for all proportions. We observed similar results using the width of the CI. In addition, there was little gain in precision after a sample size of 70 with only a percentage gain in precision of less than 4% thereafter by increasing the sample size by 5 subjects.

We conclude the required sample size for pilot trials to accurately estimate binary outcome parameters should be greater than 70. This is substantially larger than the generally accepted recommendations for continuous outcomes.

Have the more restrictive abortion laws in the Republic of Ireland and Northern Ireland improved the health of Irish women since 1968? Have the more restrictive abortion laws in the Republic of Ireland and Northern Ireland improved the health of Irish women since 1968?

Celia Chambers

Objectives

To compare Irish incidence of maternal deaths, stillbirths and premature and low weight births, Cerebral Palsy reported by the Register in Northern Ireland, mental health disorders, breast cancer and female to male ratio for certain diseases of the immune system with other countries against abortion rates. For the RSS conference this year it is proposed to develop the epidemiological chapters in the report Ireland's Gain[1] published in December 2011 by extending the international comparisons such as with Sweden and the USA where abortion rates have been high and the health conditions regarded as abortion sequelae are also common.

Method/Models

There is detailed age specific information in Abortion Statistics published by ONS and the Department of Health since 1968 for England & Wales and the corresponding Abortion Statistics for Scotland for Nulliparous and Parous abortions. British abortion rates can be computed. Some data on Irish resident women coming to England and Scotland and the Netherlands for abortions is available so that, age specific abortion rates for Ireland and Northern Ireland can be estimated. The association between the abortion rate and Maternal death rates, Still births, and Premature birth rates, will be investigated and the relationship demonstrated for different countries.

Results and Conclusions

It is also found that Irish women have proportionately fewer maternal deaths, give birth to fewer stillbirths and premature and low weight births, with a low incidence of Cerebral Palsy reported by the Register in Northern Ireland, and enjoy better mental health and experience less breast cancer and have a lower female to male ratio for certain diseases of the immune system. These findings point to health benefits accruing to Irish women as a result of higher birth rates and lower abortion rates.

Reference

1. *Ireland's Gain: The Demographic Impact and Consequences for the Health of Women of the Abortion Laws in the Republic of Ireland and Northern Ireland since 1968.* Patrick Carroll. PAPRI 2011. www.papriresearch.org under Published Papers.

Imprecision and Robustness in Bayesian Experimental Design

Noorazrin Abdul Rajak, Malcolm Farrow

Newcastle University, Newcastle upon Tyne, UK

Objectives

It focus on experimental design within the Bayesian framework using a decision theoretic approach, in particular cases where the specifications which are normally required to solve a Bayesian experimental design problem are not made precisely. This happens due to the inability or unwillingness of an individual or expertise to specify the exact risk trade-offs or to make a mutual agreement on a precise values for the trade-offs between the various attributes. We are looking at the case where the observations are vector. The approach is illustrated by the design to an experiment related to Oral Glucose Tolerance Test (OGTT).

Method/Models

We develop methods for analysing multi-attribute utilities which do not require the specification of precise trade-offs between different risks, for choosing designs which are robust against such imprecision. In this study, we will consider the forms of imprecision in expectation (previsions) and the probabilities, and the shape of the utility function of the imprecision, we do not have to scale the function because the scale of a function is arbitrary scale. Our analysis is based on whatever limited collection of preferences we may assert between attribute collections. Similarly, we will take into consideration the forms of robustness to imprecision.

Results and Conclusions

We identify the Pareto optimal designs under the imprecise specification and suggest a criterion for selecting between such candidate designs. We will show how to reduce the designs further by combining rules which are almost equivalent and introduce general principles appropriate to selecting decisions in an imprecise hierarchy. Further work is still going through on the selecting designs to find the robustness in Bayesian experimental design problem.

Improving the adequacy of multilevel models in the face of outliers and error-prone random coefficient predictors

Michael Imande

Benue State University, Makurdi, Makurdi, Benue, Nigeria

Objectives

The paper intends to establish that the efficiency of outlier contaminated and/or error-prone multilevel models can be enhanced by systematically accommodating the outliers and adjusting for the incidence of measurement errors in the predictors using measurement error variance estimated via a combination of bootstrapping and Gibbs Sampling techniques and to show that this efficiency enhancing approach for a model based on data that is known to be associated with predictors having very high measurement error variance (or very low reliability) is less effective than what obtains in the case of predictors associated with lower measurement error variance.

Method/Models

Using the framework of a K-level model of the (compact) form $Y = X\gamma + ZU + Z^{(1)}e$, with $ZU + Z^{(1)}e$ constituting the random segment while $X\gamma$ the fixed segment, a 2-level model based on educational data was considered. Non-parametric bootstrapping and multivariate normal distribution modelling were employed to simulate additional levels 1 and 2 units which allowed for possible incidence of anomalous values. The model efficiency was examined by accommodating a couple of outliers and then systematically adjusting for the errors in a predictor using estimated measurement error variances estimated via bootstrapping and Gibbs Sampling techniques.

Results and Conclusions

The predictor, STM score per student in JSS1 subjects (NJS1avg), was perceived error-prone and estimating the model via IGLS gave the NJS1avg coefficient estimate, standard error and coefficient of variation (CV) as 0.680(0.017), 0.025 while 0.671(0.020), 0.030 and 0.744(0.018), 0.024 were obtained in the face of simulated additional level 1 (scenario 1) and levels 2 (scenario 2) units, respectively. Outlier accommodated estimates were 0.740(0.017), 0.023 while 0.872(0.023), 0.026; 0.857(0.021), 0.025 and 0.911(0.020), 0.022 were obtained after adjusting for errors using measurement error variance of 0.25 where as 0.906(0.011), 0.012 were estimates on accommodating outliers and adjusting for errors for scenario 2. A measurement error variance of 0.64, however, gave estimates 0.901(0.015), 0.017; being slightly worse than the 0.25 case. Model efficiency appears to be better and more realistically done with the systematically estimated 0.25 measurement error variance than with the value (0.64) estimated from the more contaminated data even after accommodating outliers in the model.

Iterative bandwidth selector - extension to kernel regression.

Jan Kolacek, Ivana Horova

Masaryk University, Brno, Czech Republic

Objectives

The aim of the contribution is to extend the idea of an iterative method known for a kernel density estimate to kernel regression.

Method/Models

The method is based on a suitable estimate of the mean integrated square error. This approach leads to an iterative quadratically convergent process.

Results and Conclusions

We conduct a simulation study comparing the proposed method with other data-driven bandwidth selector methods. Results are implemented in a user-friendly interface in Matlab.

Calibration estimation under non-response and missing auxiliary information

Thomas Laitila, [Lisha Wang](#)

Örebro University, Örebro, Sweden

Objectives

Non-response is undesirable, but inevitable in surveys. Techniques are required to promote the accuracy of estimation when missing data is non-ignorable. The calibration approach (Deville and Särndal, 1992) is suggested for estimation in sample surveys under non-response given access to suitable auxiliary information. Auxiliary variables from register system frequently contain missing values for some units in the response set and/or in the population, which can be substituted with values derived from imputation rules defined by information at the response, sample or population levels. In this essay, the properties of the calibration estimator under imputation of missing auxiliary information are considered.

Method/Models

The effects of imputation, and the level of information used for deriving the imputation rule, on the calibration estimator bias is studied by considering asymptotic properties. Probability limits are derived for the calibration estimator under imputation. Finite sample bias is studied by means of simulations.

Results and Conclusions

Results show that imputation does not add bias to the calibration estimator while it hampers the potentials of using auxiliary information for reducing non-response bias. The simulation study shows auxiliary information at the response level gives slightly better results compared with those at the sample level in the calibration estimation. The bias of the calibration estimator increases when the non-response is not random, whereas bias scarcely changes when missing values in auxiliary information are non-random.

Myopic decision rules using the regret-regression method with application to Warfarin data

Nur Anisah Mohamed, Robin Henderson

Newcastle University, Newcastle Upon Tyne, UK

Objectives

A dynamic treatment regime is defined as a set of decision rules on the treatments which is based on the history of the patients at the certain time point. Warfarin is an anticoagulant. It is a medication to stop blood from clotting. The warfarin data is taken from Rosthoj et al (2006). The study period is from February 1995 to August 2000 involving 350 patients. The treatment time varied from 16 days to almost five years which involved 2 to 124 clinic visits. The data is concentrate only on the first 14 clinic visits for 303 patients.

Method/Models

Henderson et al (2009) introduced the regret-regression method by combining the regret function and regression modelling to obtain the optimal dynamic treatment regimes from observational data. This method is applied to warfarin data to help choose the best potentially time-varying dose for a patient on a long-term treatment. This approach is also found suitable to estimate myopic decision rules which can be achieved by fitting the model at the end of the j^{th} interval, for $j=1,2,\dots,K$.

Results and Conclusions

The percentage time in range (PTR) is used as an overall measure of control of INR (International normalised ratio). The results for comparing between long-term and myopic decision rules using the PTR response are possible to obtain. It shows that the myopic models is valid and proportional to the long-term model.

NIH Stroke Scale as a Common Predictor of Short Term functional Status, Length of Hospital Stay and Survival - Cohort Study Within the South London Stroke Register (SLSR)

Shah-Jalal Sarker^{1,3}, Salma Ayis^{1,2}, Anthony Rudd^{1,2}, Kashfia Chowdhury³, Charles Wolfe^{1,2}

¹*Division of Health and Social Care Research, King's College London, London, UK,* ²*NIHR*

Comprehensive Biomedical Research Centre, Guy's & St. Thomas' NHS Foundation Trust and King's

College London, London, UK, ³*Centre for Experimental Cancer Medicine, Barts Cancer Institute, Queen Mary University of London, London, UK*

Objectives

Early prediction of future functional status, length of hospital stay and patient survival may be very useful for efficient use of health care resources. This study explores the link between neurologic deficit as measured by the National Institute of Health Stroke Scale (NIHSS), and its prognostic ability of short-term functional status, length of hospital stay and patient survival.

Method/Models

395 patients from the population-based South London Stroke Register were assessed with Nottingham Extended Activities of Daily Living (NEADL) 3-months after first-ever stroke. Their lengths of hospital stay as well as survival status were also ascertained. The baseline predictors included socio-demography, case severity: acute stage NIHSS and pre-stroke BI, co-morbidities and acute treatments. The statistical methods included negative binomial regression, multiple linear regression, Cox's proportional hazards model, Kaplan-Meier curves, and fractional polynomials.

Results and Conclusions

Results - Of the 395 patients, 44% had low NIHSS score (≤ 5), 40% had medium score (6 - 13) and 16% had high score (> 13). Compared to patients with a low NIHSS score, patients with a medium score had 22% lower ($p=0.02$) and those with a high score had 67% lower ($p<0.001$) NEADL score. Patients with medium NIHSS score had a 73% increase in average length of hospital stay (138% increase for those with high score) compared to patients with low score. Optimum length of hospital stay of about 33 days occurred at NIHSS score of 8, given NEADL at 3- months. One unit increase in NIHSS score was associated with a 10% higher risk of death (HR: 1.10, CI 1.05 to 1.16).

Conclusion -Acute stage NIHSS score is a common predictor for short-term functional status, length of hospital stay and patient survival, adjusting for other variables.

Pseudo-Bayesian optimum designs for nonlinear models using standard functions in R

Ronaldo Silva, Luzia Trinca

UNESP Univ Estadual Paulista, Botucatu, São Paulo, Brazil

Objectives

For obtaining informative and efficient designs in nonlinear models it is necessary to incorporate prior knowledge of the unknown parameters being estimated. This dependence on unknown values of the parameters lead us to use a pseudo-Bayesian method which introduces uncertainty about the parameters providing robustness to this uncertainty, justified for a substantial improvement of the design in terms of efficiency. We aim at proposing a general purpose optimization algorithm and an adaptive multidimensional integration (cubature) method in order to obtain efficient designs for some popular nonlinear models.

Method/Models

We use the optimization "L-BFGS-B", a quasi-Newton method implemented in the function `optim` and an adaptive multidimensional integration (cubature) method implemented in the package `cubature` of R in order to obtain efficient experimental designs for some popular nonlinear models. This method was applied for two classes of pharmacokinetics models and chemical kinetics, the exponential decay model with one parameter and a compartmental model with three parameters. We use the most popular design criterion, the D-optimality which maximizes the generalized variance, or its logarithm, of the parameter estimates. Designs were compared through relative D-efficiency on sampled priors.

Results and Conclusions

In this work we illustrate the application of approximate integration by adaptive quadrature/cubature method for higher dimensional problems for finding pseudo-Bayesian designs for nonlinear models. The examples for exponential decay (one parameter) and compartmental (three parameters) models shown that the proposed methods produced, in general, more efficient designs than those found by other authors. An additional advantage of our approach is that it uses general purpose routines for integration and optimization readily available in R. It was not possible to make a comparison between computational time because the other authors do not give such details, but our designs was found in a acceptable time for the practice. We

note that a transformation in the design space could be used in order to search in an unconstrained design space. That is under development and we believe will facilitate the optimization process.

Adaptive clinical trials - A review of statistical methodology and current status

Nicky Wright

BioSigma Ltd, Winchester, Hampshire, UK

Objectives

Over recent decades, increased spending in clinical research has not been accompanied by an increase in success in pharmaceutical development. Both the European Medicines Agency and the Food and Drug Administration have produced discussion papers and guidelines concerning so-called adaptive design methodology as a means of enabling more efficient clinical studies. These adaptive trials use data that is gathered during the trial to change trial parameters, such as dose levels or sample size, while the study is in progress. The objective of this paper is to review adaptive trial designs from a statistical viewpoint, and to identify potential problems.

Method/Models

Papers and reports of clinical studies employing adaptive designs will be described that were identified from searches conducted using Google Scholar, PubMed and Scopus, with the keywords 'adaptive trial design' and 'statistics'. Different methods for adapting the trial parameters will be discussed, and examples and case studies presented in which approaches to statistical issues such as maintaining overall type I error rates are addressed.

Results and Conclusions

Types of adaptive trial designs include: adaptive randomization; group sequential design; sample size re-estimation; drop-the-losers design; adaptive dose finding; and biomarker-adaptive design (Chow and Chang, 2008). Statistical issues associated with adaptive designs include the introduction of bias in data collection, shifts in the target population thus affecting statistical inference, and consequential modifications to the hypothesis tested by the study. Adaptive design trials can be difficult to interpret and appropriate statistical methodology is required to be developed to achieve unbiased and accurate estimates of treatment effects. The conditions under which regulators would consider an adaptive trial to be adequately controlled will be discussed.

Reference Chow S-C, Chang M (2008) Adaptive design methods in clinical trials - a review. *Orphanet Journal of Rare Diseases*, 3, 11.

Optimal Designs for the estimation of five variance components in a linear random effect model

Oluwole Nuga¹, G.N⁰

¹*Bells University Of Technology, Ota, Ogun/Southwest, Nigeria,* ²*University of Ibadan, Ibadan, Oyo/Southwest, Nigeria*

Objectives

To provide experimental designs procedure for the estimation of five variance components of the maximum likelihood estimators arising from a model that consists of two random crossed factors and two nested factors within the treatment combination of the crossed factor.

To determine the number of designs that can be generated for any fixed total sample size.

To obtain the optimal design for the estimation of the five variance components from the candidate design generated for a fixed total sample size

Method/Models

the linear random effect model used has two crossed factor A and B, the interaction effects, a nested factor effect and the random error term. the total sample size $N=abcm$. Candidates designs for a fixed total sample size were generated in such a way that all five variance components can be estimated

Information matrix of the maximum likelihood estimators are determined by the negative matrix of second order derivative of the log likelihood for individual candidate design. The A optimality and D-optimality were obtained for a particular configuration of the five variance components.

Results and Conclusions

For the estimation of the five variance components, it was observed that the design that has the largest number of observation in the second stage nesting of the nested designs within each treatment combination are A- optimal. Designs with the largest concentration of observation for the crossed factor are D- optimal.

Overall the D-optimal designs should be used to obtain the optimal balanced design for a fixed total sample size for the model above because of the concentration of observation at the level of the crossed factor where the true variance components are larger.

Statistical perspective on measuring the signal-to-noise ratio for history event data in neural systems

Gabriela Czanner¹, Emery Brown^{2,3}

¹University of Liverpool, Liverpool, UK, ²Harvard Medical School, Boston, USA, ³Massachusetts Institute of Technology, Boston, USA

Objectives

The signal-to-noise ratio (SNR) is a commonly used measure of system fidelity, is defined as the ratio of the squared amplitude of a signal to the variance of the noise. This definition is widely used in analysis of physical systems however it is not appropriate for history event data in neural systems. One example of such data are the times of electrical discharges from a neuron, this data are not Gaussian, not independent and can be well characterized within a framework of point processes. Our goal was to propose a new SNR measure for neural event history data.

Method/Models

We extended the existing measure of SNR to account for the properties of neural data. Our measure is based on deviances and can be viewed as an estimate of improvement in expected prediction error in Kullback-Leibler sense due signal. We investigated the properties of this measure in general dynamic event history regression models i.e. in models where the data are not independent. We studied the bias correction of deviance-based measure via simulations to show small sample properties of the measure.

Results and Conclusions

We illustrate the ideas in real recordings of time of electrical discharge from a neuron and also in a computer simulated data. We show that the bias correction is important in short datasets and/or in datasets where neural signal is very low. Our methods finally allow explicit measuring of SNR in neurons. We show that SNR in neurons are low of order -3 dB or lower. Our SNR measure allows measuring the signal content in neurons on same scale as in physical systems.

MODELLING AND ESTIMATION OF A JOINT LONGITUDINAL MODEL FOR MULTIPLE REPEATED MEASURES AND MULTIPLE RECURRENT EVENTS PROCESSES

Samuel Manda^{1,2}

¹South African Medical Research Council, Pretoria, South Africa, ²University of Kwazulu-Natal, Pietermaritzburg, South Africa

Objectives

In biomedical studies, a number of analytical estimating procedures have been developed for the analyses of joint longitudinal modeling of two processes: a single repeated measurement variable and a single recurrent time-to-event process, which are assumed to be correlated. However, in many studies, the observations may involve multiple repeated measures processes and multiple recurrent time-to-events processes. These several processes may be correlated with each other and within subjects.

Method/Models

We propose a process-specific and shared random effects model (*shared component model*) for the multiple processes. The expectation-maximization (EM) algorithm is used for estimation and inference of the model. The method is tested using simulation studies, and applied to an analysis of national pharmacovigilance surveillance system of HIV-infected patients on antiretroviral (ARV) in South Africa. The results are compared to those obtained using a Bayesian estimation approach.

Results and Conclusions

The EM algorithm performed well on the simulation studies. The results from applying the method to the national ARV pharmacovigilance data provide important data for public health programmes relating to the safety and morbidity profiles and impact of ARV medicines in adults and adolescents in the country

Inference for the Fractional Heston Model using the Auxiliary Particle Filter

Muhannad Al-Saadony

University of Plymouth, Plymouth, UK

Objectives

The aim of our work is to sample the underlying volatility process and update the posterior distribution of the parameters of the fractional Heston Model sequentially as data arrive over time using the auxiliary particle filter.

Method/Models

The fractional Heston Model is a generalization of the Heston Model obtained by replacing Brownian motion by fractional Brownian motion in the two defining equations. We present the fractional Heston Model as a Dynamic state space model. We perform inference about the fractional Heston Model in the Bayesian framework using the auxiliary particle filter.

Results and Conclusions

We apply the auxiliary particle filter to make inference about the underlying volatility process and the parameters of the fractional Heston Model using real and simulated data. In general, the results are successful.

Bayesian Non-Parametric Modeling in Measurement Error using A Polya Tree Distribution

Reem Al-Jarallah, David Stephens
Kuwait University, Kuwait, Kuwait

Objectives

The objective in this paper is to formulate the (ME) problems in Bayesian framework by fitting a Bayesian linear hierarchical model with measurement error and nonconstant variance to a real data problem. In measurement error models, Gaussian distribution is commonly used for modelling measurement error, however this choice places a strong constraint on the shape of ME distribution.

Method/Models

We consider adding a non parametric aspects to the calibration function in the ME distribution by introducing a Polya tree as an alternative to the parametric approach. A comparison is conducted between the parametric and the non-parametric approach in a Bayesian framework using MCMC methods.

Results and Conclusions

Bayesian non-parametric methods are appealing in measurement Error(ME) modeling. Interest in modeling using rich classes of nonparametric methods have been grown rapidly. Among the many models proposed for nonparametric Bayesian modeling in recent literature is the Polya tree. Recently, Polya tree distributions have received much attention and have been shown to uncover new insights, that can not be captured by usual parametric methods.

Two forms of attrition in a longitudinal health study involving ageing participants, and their predictors: analysis of Whitehall II.

Robert Grant¹, Gill Mein¹, Suneeta Johal¹, Clive Seale³, Richard Ashcroft³, Anthea Tinker²

¹Kingston University & St George's, University of London, London, UK, ²King's College London, London, UK, ³Queen Mary, University of London, London, UK

Objectives

We sought to describe patterns of attrition in the Whitehall II longitudinal study around the age of retirement. Known predictors from previous research include older age, impaired cognitive function, lower educational level, living alone, fewer social activities, and lower socio-economic status. We tested whether these held for Whitehall II and whether they differed between non-response and actively requesting to withdraw.

Method/Models

Data were extracted on 10,308 participants over 6 phases of the study from 1985 to 2001, including social, economic, demographic and health covariates, and the date of any request to withdraw from the study. Non-response was modelled by mixed-effects logistic regression and withdrawal by proportional hazards Cox regression. Inter-dependency between the two processes was investigated by experimenting with including predictions (modal BLUPs) of individual random intercepts for non-response as a predictor for withdrawal, as well as a joint model with a shared latent variable. We will also discuss some of the experiences of interviewing participants and their reasons for withdrawing.

Results and Conclusions

The two forms of attrition have some shared and some unique predictors. For non-response, older age, male gender, lower job grade, not owning one's home, no long-standing illness, higher levels of education, and not being retired were significant predictors. Being married was associated with higher odds in women and lower in men. Withdrawal shared the first five predictors above, as well as lower SF-36 physical function and mental health scores, fewer social activities, and the individual's BLUP for non-response.

We found a strong gender effect which had not previously been reported. The effect of health is complex: poorer quality of life is associated with greater hazard of withdrawal, and self-reported 'long-standing illness' with less attrition of either form, possibly an effect of the health screening in this study. Longitudinal researchers need to ensure attrition does not cause under-representation of older people, particularly those who are more socially isolated.

Statistical Properties of the Convoluted beta-Weibull Distribution

Alfred Akinsete, Jianan Sun

Marshall University, Huntington, West Virginia, USA

Objectives

This work discusses some statistical properties of the convoluted beta-Weibull distribution (CBWD). The beta-Weibull distribution, defined and studied extensively by Famoye et al. (2005), arose as a result of the logit of the beta distribution with the Weibull distribution. Various statistics of CBWD are obtained, including for example, moment, moment and characteristic generating functions, hazard function, and the entropy. The method of Maximum Likelihood Estimation (MLE) is proposed for estimating the parameters of the distribution. A simulation of the distribution is obtained using the Markov Chain Monte Carlo procedure.

Method/Models

A generalized class of the beta distribution was introduced by taking the logit of the beta distribution. Under this scheme, the cumulative distribution function (cdf) for the generalized class of distribution for the random variable X is generated by applying the inverse cdf of X to a beta distributed random variable. The probability density function of the generalized distribution then becomes,

$$g(x) = [1/B(\alpha, \beta)] [F(x)]^{(\alpha-1)} [1-F(x)]^{(\beta-1)} f(x)$$

Substituting the cdf of the convoluted beta-Weibull distribution in the above equation, we have the required density of the CBWD.

Results and Conclusions

Various statistical properties of this distribution are obtained, including for example, moment, moment and characteristic generating functions, hazard function, and the entropy. The method of Maximum Likelihood Estimation (MLE) is proposed for estimating the parameters of the distribution. A simulation of the distribution is obtained using the Markov Chain Monte Carlo (MCMC) procedure.

Evaluation of removable statistical interaction for binary traits

Jaya Satagopan¹, Robert Elston²

¹Memorial Sloan-Kettering Cancer Center, New York, NY, USA, ²Case Western Reserve University, Cleveland, OH, USA

Objectives

This work evaluates the null hypothesis of no interaction between two sets of risk factors for a binary trait against the alternative that an interaction exists and is removable. Statisticians define the term "interaction" as a departure from additivity in a linear model on a specific scale on which the data are measured. Certain interactions may be eliminated via a transformation of the outcome such that the model is additive on the transformed scale. Such interactions are known as removable interactions. We develop a novel test statistic for evaluating a removable interaction in case-control studies.

Method/Models

Including interaction terms in a model is equivalent to including higher degree polynomial terms, and the resulting model is non-additive in the risk factors. Large interactions may induce curvature effects. We use these concepts to develop a novel logistic regression model for disease risk, which is a polynomial of degree two in the risk factors. When the model satisfies certain monotonicity conditions, the interaction terms can be approximated in a parsimonious form such that they depend upon a single curvature parameter. Testing this single term is equivalent to testing for the presence of a removable interaction.

Results and Conclusions

We develop a test statistic that has an asymptotic chi-squared distribution with one degree of freedom. To calculate this test, we estimate the main effects as unweighted averages of the outcomes based on orthogonal contrasts. We apply this test to two published case-control studies on smoking and NAT2 acetylation in relation to bladder cancer and advanced colorectal adenoma. We fit logistic regression models to these data and show that the interaction between smoking and NAT2 acetylation is removable in both these studies, suggesting that there exists an alternative transformation (i.e., link function) under which additive models may be fit to these data. We consider the Guerrero and Johnson (GJ) link function, develop an iteratively reweighted least squares algorithm to estimate the model parameters, and show that additive models based on the GJ link provide more precise estimates of the log odds of the disease than models based on logistic regression.

The Analysis of Ultra-High Frequency Financial Data via Particle Filters

Golnaz Shahtahmassebi, Rana Moyeed

Plymouth University, Plymouth, UK

Objectives

The availability of *ultra-high frequency* (UHF) data on transactions has revolutionised data processing and statistical modelling techniques in finance. Unique characteristics of such data, e.g. discrete structure of price change and unequally spaced time intervals have introduced new theoretical and computational challenges. Our objective in this talk is to develop an appropriate methodology for modelling price change as an integer-valued variable.

Method/Models

We introduce a class of models, called *dynamic zero inflated Poisson difference* (DZPD) models, which enable us to model integer-valued variables taking both negative and positive values. Using *sequential Monte Carlo* methods, we develop a strategy to sequentially learn about the states and parameters of the DZPD model. Furthermore, the sequential version of the deviance information criterion is used for the purpose of model comparison.

Results and Conclusions

The application of the DZPD model is illustrated by two sets of UHF data: Western Texas Intermediate oil price change and FTSE100 index change. Finally, we show that our model is flexible enough to adapt to the structure of the data and able to characterise the behaviour of price/index change well.

A treatment of ties in rank correlation

Ilaria Lucrezia Amerise⁰

¹*University of Calabria, Department of Economics and Statistics, Via Bucci, 87036, Rende (CS), Italy,*

²*Royal Holloway, University of London, Egham Hill Egham, Surrey TW20 0EX, UK*

Objectives

A popular method for resolving ties is the midrank method. An alternative solution is the max-min tie-breaking algorithm which will enable us to compute rank correlations by the usual definition in the absence of ties. In essence, the algorithm replaces the tied values by the ranks they would have if they were not tied. Two orderings are considered for computing a rank correlation: one that maximizes the positive correlation and one that maximizes, in absolute terms, the negative correlation. The two values are averaged to arrive at a unique result. We propose a weighted average estimator.

Method/Models

In this work, we have used modified version of the algorithm found in [Gideon & Hollister, 1987] and [Gideon & Rothan, 2011] that suggests a randomization method and a computation of the minimum and maximum possible correlation values when ties are present. This technique had already been discussed by [Gini, 1939]; he found a weighted average solution based on unknown parameter, which could not be determined reliably because of lack of computing resources. To overcome this problem, information from a range of possible randomizations may be incorporated into the weights. This idea is illustrated with an example involving several rank correlation coefficients.

Results and Conclusions

Equal values are fairly common when rank methods are applied to rounded data or integer-valued observations. The midrank method assigns to tied observations the average of the ranks they would have had if there were no ties. The mean rank of the permutation remains unaltered, but the variance is reduced and changes according to the number and location of ties. Moreover, the usual extremes of many coefficients are not fixed and should be determined on a case-by-case basis. We propose a weighted average of the values of rank correlations for the two permutations that most favor positive and negative correlation. If it is found that the difference between these values is large, then the conclusion should be drawn that there is little information in the data set. On the other hand, in case of tied ranks, a weighed average of the minimum and maximum possible correlation values yield a significant improvement in the approximations.

Observations on the design of sensory profiling trials for a back-cross population of potato varieties

Ian Nevison¹, John Bradshaw², Carol Shankland³, Glenn Bryan²

¹*Biomathematics and Statistics Scotland, Edinburgh, UK*, ²*James Hutton Institute, Dundee, UK*,

³*Sensory Scotland Ltd, Ayr, UK*

Objectives

The statistical aspects of designing sensory profiling trials for the comparison of numerous hot samples present challenges that do not arise in the majority of sensory studies. We consider alternative design strategies to optimise efficiency while satisfying practical constraints, the respective trade-offs involved and make recommendations for future designs. Our work was motivated by a large study of potato varieties from a back-crossed population, where the samples were boiled potatoes from individual field plots.

Method/Models

A design for profiling in the order of 130 samples in 35 tasting sessions was required. Within individual sessions, 12 assessors scored the same five samples; the order of presentation of these samples was balanced across assessors as far as possible based on Williams Latin square designs. Most samples were allocated to a single session but there were some samples common to more than one session, including one control in all sessions. Linear mixed models were fitted to the data using REML and the variance components used subsequently to assess the benefits of the chosen design relative to alternatives.

Results and Conclusions

Balancing for the demonstrated effects of order of presentation of samples to assessors in the design avoids potential problems of interpretation due to confounding. Comparison of analyses including and excluding the control variety scored throughout the study demonstrated the benefit of its inclusion over replacing it with additional samples instead. The substantially enhanced connectivity between sessions arising from such an approach enabled more accurate comparisons between samples assessed in different sessions. It also led to more precise estimates of the sample by session variance component, and thus to standard errors for comparing samples that can be larger but are more appropriate.

Robustifying the Least Squares estimate of parameters of variance model function in nonlinear regression with heteroscedastic variance

Hossein Riazoshams

Statistics Department, Stockholm university, Stockholm, Sweden

Objectives

This Article deal with robust nonlinear regression with heteroscedastic error when the errors follows a unknown parametric function of predictors. A robustified form of multistage estimate is proposed.

Firstly, the Classic Least Squares estimate for parameters of the model function is robustified and replaced by robust MM-estimators, and secondly the Least Square estimates for variance function model parameter which calculate the sum of errors between sample variance and variance model, is robustified by generalized M-estimates.

The second part of above process is proposed first time by this Article.

Method/Models

Stage1; The Least Squares estimate for the parameters of regression model is replaced by robust MM-estimate.

Stage 2: For estimating the parameters of variance model, firstly the sample variance and errors are calculated. Then either the Least Square estimate or Maximum Likelihood estimate of chi-square distribution can be applied. In order to robustify this form, firstly we robustify the sample variance and calculate the errors, secondly robustify the Least Square estimate by Generalized M-estimates.

Stage 3: Using estimates of stage 1 and 2, the function model parameter can be estimated by Generalized robust estimates.

Results and Conclusions

The simulation is shown the proposed methods are more robust than Classic forms, and also is more robust than the Robust Multistage estimate when it uses robustified chi-square likelihood estimate of estimating parameters of variance model.

This method is faster than old robust methods and numerically is easier. The easier calculation and faster computation of this method can be extended to the methods which needs high dimension computation, such as GARCH error models.

The proposed method can be extended to identify outliers.

Construction and validation of the Physiological Emotions Friendship Scale (PhEFS) through a multidisciplinary approach.

Emma Zavarrone, Sandra De Francisci

IULM University, Milano, Lombardia, Italy

Objectives

This paper aims at creating and validating a Physiological Emotions Friendship Scale (PhEFS) combining social network analysis, psychometric characteristics and emotional/physical internal states. The advantages of each approach are: measuring different aspects of friendship (e.g. influence, connection, etc.), detecting the items related to the latent construct, encoding emotional reactions to stimuli. The study of the neurophysiological signals detected on emotion stimulation will permit to link the items related to the friendship construct and the interaction among individuals. PhEFS will be created on the basis of the most items-related neurophysiological parameters and will be validated on a random sample of students.

Method/Models

We intend to create an experimental biofeedback scaling on four random sample of IULM's students. Each group will fill in the specific questionnaire (e.g. Friendship Qualities Scale, Friendship Scale) under neurophysiologic stimuli detection. A set of sensors, connected to a Computer and to each student, will record physiological activations caused by questionnaire submission. The collected data will refer to neurophysiological parameters (e.g. skin conductance, EEG, indices derived from eye-tracking, etc.), psychometric measures on friendship construct and network information. The use of multiway Anova approach on neurophysiological parameters, psychometric responses (treatment 1) and network measures (treatment 2) will allow to discriminate items and subjects.

Results and Conclusions

The creation and validation of PhEFS through the use of a multidisciplinary approach is aiming at overcoming the different approach related limits: social network analysis does not allow to estimate the latent construct whereas psychological approach does not take into account interpersonal aspects. The neurophysiological discipline states that patterns of peripheral physiological activity are correlated to different encodable emotions. Through the classification of specific arousal, detected by electronic devices (FlexComp Infiniti™ System, Thought), affective computing offers a substantial explanation of human attitudes towards friendship. We assume that a set of items, exploring friendship construct, administered in a network of students, generate specific emotions and internal states variation. In this perspective, items become the stimuli aimed at measuring the construct in the network. As a result, detected neurophysiological parameters will confirm the presence/absence of the friendship construct while parameters variations will allow us to determine the bond friendship intensity within the network itself.

A typical quasi-semi-Latin square design: Combinatorial Structures and Application

Polycarp Chigbu, Abimbola Oladugba

University of Nigeria, Nsukka, Nsukka, Enugu, Nigeria

Objectives

Quasi-semi-Latin square design is a combinatorial object whose design layout has the semi-Latin square formation except for the subsequent definition of its blocking system. In this work, we discuss the combinatorial structure (randomization, and treatment and block structures) of a typical quasi-semi-Latin square design. The bases of this work are to obtain a null randomization model for this quasi-semi-Latin square design as a step to obtaining its statistical model, null analysis of variance table showing the quadratic breakdown of the yield, covariance matrix and expected value associated with the breakdown of the yields.

Method/Models

The combinatorial structure (block structure, treatment structure and randomization) were used to derive the null analysis of variance, covariance matrices and expectations of mean, which formed the bases for obtaining its statistical model and full analysis of variance table. Simulated data were used to illustrate this design numerically.

Results and Conclusions

The statistical model and full analysis of variance for the quasi-semi-Latin square design was obtained using its combinatorial structures. From the simulated data used, the 'Positions' occupied by any particular 'Cleaner' do not have effect on the response when tested by each 'Housewife' at a particular time.

Logistic Regression estimator for Warner's randomized response model

Claudia Rivera, Leonardo Trujillo

National University of Colombia, Bogotá, Cundinamarca, Colombia

Objectives

- Consider the use of auxiliary information to find two estimators for the total when randomized response is use.
- Get an approximation for variance of the estimators and determine under which circumstances each estimator has a lower variance.

Method/Models

For avoiding bias in sensitive questions, (Warner, 1965)proposed a method called Randomized Response (RR). The respondent should choose randomly (with a success probability p) between two questions,

1. Do you belong to group A?
2. Do you belong to group A^c?

This answer is called z and represents the answer listened by the pollster. Mathematically,

$z_k = \{y_k \text{ with probability } p ; 1-y_k \text{ with probability } 1-p\}$

The logistic regression estimator (Lehtonen, 1998), is used for developing the two estimators. The coefficients of the model are found through the likelihood of z and accuracy

Results and Conclusions

The logistic regression estimator is applied to Warner's randomized response(RR) model. The variable y , which indicates if an individual belongs to a sensible group **A** is modeled through a logistic model ζ and an auxiliary information vector $[x_k]$, this is done theoretically because the variable is unknown.

Two different estimators are presented, one is asymptotically unbiased. The estimators are applied to a simulated population of size 10000, five sample sizes and 10 values of p . These are compared with their MSE estimations, which turned out to be lower for the biased estimator, although the asymptotically unbiased estimator turned out to be better in regard to precision. The two estimators produce good estimations for $p > 0.7$ and $p < 0.3$, although the symmetry in the behavior does not maintain for the biased estimator.

Jittered phase diagrams for seasonal patterns in time series

Robert M. Kunst^{1,2}

¹*Institute for Advanced Studies, Vienna, Austria,* ²*University of Vienna, Vienna, Austria*

Objectives

In the analysis of seasonal time series, an issue of main concern is the discrimination of deterministic patterns (shape reversion) and seasonal unit roots (permanent shape transition). The suggested method assists in this discrimination.

Method/Models

We suggest complementing the customary hypothesis tests with jittered phase diagrams on a discretized set of basic seasonal shapes that represent the dynamic transition between the patterns. These jittered phase diagrams provide a convenient visualization that supports the discrimination among the main classes of potential seasonal data-generating processes.

Results and Conclusions

For several classes of interest, characteristic patterns in the phase diagrams are determined. We also consider a nonparametric hypothesis test constructed from the charts. Variants for quarterly and for monthly data are studied. Exemplary applications to economics data and to some other variables are presented.

What's going on with the road accident fatal trend?

Louise Lloyd^{1,2}, Jon Forster²

¹*Transport Research Laboratory, Wokingham, UK*, ²*University of Southampton, Southampton, UK*

Objectives

Recently there has been a major reduction in the number of fatalities on British roads. Road accidents are now treated as a public health issue and it is important to understand the reasons for these changes and to judge whether further reductions are likely in order to assess the impact on health care resources. Our overall aim is to determine the major influences on this changing trend, and to answer specific questions such as is the rise in the number of 4x4s and people carrier type vehicles involved in accidents due totally to the rise in the number of these vehicles using the roads.

Method/Models

We take detailed exposure data from sources such as national traffic estimates, vehicle registrations and the National Travel Survey and evaluate changes in accident rates by year, vehicle type, road use and driver type using a Bayesian Graphical model made up of a series of MCMC GLM models. An interesting evaluation of the influence of the recession on the trend is included.

Results and Conclusions

The models are not yet finalised, however initial results suggest that the rise in accidents involving 4x4 type cars can be attributed to the increase in the number of these vehicles and that the recession appears to be influencing the number of fatal accidents occurring. It has been shown that this is not due to the decrease in the number of miles driven but it is not clear what other reasons there might be for this influence.

On the innovation of the family of multivariate higher-order hybrid polynomial kernels

Benson Ade Afere¹, Sunday Martins Ogbonmwan^{0,2}

¹Adekunle Ajasin University, Akungba Akoko, Ondo State, Nigeria, ²University of Benin, Benin City, Edo State, Nigeria

Objectives

The objectives of this study are:

1. to obtain the generalized higher-order asymptotic optimal bandwidth of the proposed family of multivariate higher-order hybrid polynomial kernels.
2. to obtain a reduction in global error of the proposed family of multivariate higher-order hybrid polynomial kernels by using asymptotic mean integrated squared error (AMISE).

Method/Models

The method behind this work lies on taking the derivative of $K(t, p) = \{2^{2p+1}B(p+1, p+1)\}^{-1}(1-t^2)^p$, $|t| \leq 1$ and sum it with the family of polynomial kernels. This is done by using the augmented version of the work of Jones & Foster (1993).

Also, the derivation of the schemes for AMISE lies mainly in the application of multivariate Taylor's series expansion of $f(\mathbf{x})$ up to $2m+2$ (where m is a positive integer). These ideas are applied to the proposed family of higher-order hybrid kernel density estimation.

Results and Conclusions

The findings of this work are as follows:

1. A new family of higher-order polynomial kernels is obtained.
2. The multivariate global error has faster convergence rate than those provided in the literature.
3. The multivariate optimal bandwidth has faster convergence rate than those provided in the literature.

Conclusively, in the kernel density estimation setting, the smaller the error, the better the method or the kernel function and vice versa. We have seen that the embedded kernels in the family of the proposed higher-order hybrid polynomial kernels have been able to provide a reduced global error when compared with their counterparts in the existing families of higher-order polynomial kernels and classical polynomial kernels. This however prompts high efficiency for all the proposed higher-order hybrid kernels. With these, we therefore recommend the use of the higher-order hybrid polynomial kernels proposed in this work.

Diffusion of Innovations in Dynamic Networks

Charlotte Greenan

University of Oxford, Oxford, UK

Objectives

To simultaneously model the evolution of a dynamic social network with the diffusion of an innovation, as the novel behaviour, practice or attitude is appropriated by the network's members. The aims are to provide a more realistic model than could be attained whilst assuming a static network, and to enable us to disentangle social influence and homophilous peer selection.

Method/Models

The innovation process follows a proportional hazards model, and includes time-varying social contagion measures as explanatory variables; co-evolving is the dynamic network, which follows a stochastic actor-oriented model first developed by Snijders (2001). We estimate the parameters of the model using the Method of Moments. The model is implemented in the R package RSiena.

Results and Conclusions

We apply the method to an existing dataset to analyse the initiation of smoking amongst a group of 129 Glaswegian teenagers, and find the presence of both selection and influence. We explore the properties of the estimator of the model parameters using both theory and simulation.

Monitoring of an industrial process using adaptive multi-block partial least squares

Bothinah Altaf, Elaine Martin OBE, Gary Montague

Newcastle University, Newcastle upon Tyne, UK

Objectives

- To remain competitive and to maximise production, one approach is to continuously monitor the performance of the industrial process and to identify and take corrective action when non-conforming performance is identified
- One tool for developing a monitoring scheme is Partial Least Squares (PLS). PLS is appropriate for processes that operate at steady state. In practice most industrial processes are time varying hence the development of Adaptive PLS
- To simplify the monitoring task, monitoring charts for the whole process as well as individual 'operating units' based on an adaptive PLS model that is robust to outliers is proposed

Method/Models

This poster describes an extension to adaptive PLS, robust adaptive multi-block PLS (RAMBPLS). The proposed approach uses a combined index to calculate a weighting that reduces the impact of the outlying observation thereby retaining it in the updating process. A multi-block approach is then used to monitor the whole process as well as the individual 'operational units'. The performance of the proposed method is assessed using a benchmark data set from a time varying process. The capability of the monitoring charts is statistically evaluated using average run length and the root mean square error is used to assess model performance.

Results and Conclusions

In this poster, the monitoring of time varying processes is discussed. It is shown that conventional PLS is not appropriate for monitoring time varying process as it is only applicable for processes operating at steady-state. To overcome this limitation, RAMBPLS (Robust Adaptive Multi-block PLS) is proposed. RAMBPLS combines adaptive PLS and multi-block analysis thereby enabling the development of monitoring charts for both the whole process and individual process units. It is based on a model that is adaptive and robust to outliers. The proposed method shows satisfactory performance in terms of process monitoring. Compared to PLS, the application of RAMBPLS to a simulation of a time varying process demonstrates

- An improvement in model prediction
- A reduction in the false alarm rate.
- A simplification of the monitoring task by isolating those parts of the process that are most affected by the fault.
- A reduction in the effect of the outlying observations.

The Prevalence rate of Work related Stress, anxiety and depression in Great Britain 2001/02-2010/11

Paul Buckley, Paul Cushion

Health and Safety Executive, Liverpool, UK

Objectives

Work-related stress is defined as a harmful reaction people have to undue pressures and demands placed on them at work. Work related stress remains the largest single cause of work related sickness absence in Great Britain. The Health and Safety Executive (HSE) has sought to assess the prevalence of work related stress in the working population of Great Britain through the self-reported Labour Force Survey (LFS) carried out annually by the UK Office for National Statistics (ONS)

Method/Models

The Labour Force Survey (LFS) is a national survey of households (N = 50'000) living at private addresses in the UK. The survey is managed by the Office for National Statistics in Great Britain and by the Department of Finance and Personnel in Northern Ireland on behalf of the Department of Enterprise, Trade and Investment (DETINI). The HSE commission's questions in the LFS, to gain a view of work-related illness and workplace injury based on individual's perceptions. The HSE questions are included in two survey modules - '**The Workplace Injury survey**' module and the '**Self-reported Work-related Illness (SWI) survey**' module.

Results and Conclusions

Analysis of work related stress, depression and anxiety from this LFS data available between 2001/02 and 2010/11 demonstrated that the prevalence of work related stress in Great Britain was at a rate of 1330 cases per 100'000 for people employed in the last 12 months (95% CI's 1210-1440 case rate, $P < 0.05$) in 2010/11 and was statistically significantly lower than the prevalence rate reported in 2001/02 of 1560 cases per 100'000 for people employed in the last 12 months (95% CI's 1460-1660 case rate, $P < 0.05$). It is considered that greater awareness of work place stress, anxiety and depression and work related interventions by the health and safety community in Great Britain may have contributed to the reduction in the observed prevalence rate.

The prevalence rate of work related musculoskeletal disorders in Great Britain 2001/02-2010/11

Paul Buckley, Paul Cushion

Health and Safety Executive, Liverpool, UK

Objectives

The term MSD covers any injury, damage or disorder of the joints or other tissues in the upper/lower limbs or the back. For the employee, they cause personal suffering and loss of income; for the employer, they reduce business efficiency; and for government, they increase social security costs. The Health and Safety Executive (HSE) has sought to assess the prevalence of work related musculoskeletal disorders in the working population of Great Britain through the self-reported Labour Force Survey (LFS) carried out annually by the UK Office for National Statistics (ONS).

Method/Models

The Labour Force Survey (LFS) is a national survey of households (N = 50'000) living at private addresses in the UK. The survey is managed by the Office for National Statistics in Great Britain and by the Department of Finance and Personnel in Northern Ireland on behalf of the Department of Enterprise, Trade and Investment (DETINI). The HSE commission's questions in the LFS, to gain a view of work-related illness and workplace injury based on individual's perceptions. The HSE questions are included in two survey modules - '**The Workplace Injury survey**' module and the '**Self-reported Work-related Illness (SWI) survey**' module.

Results and Conclusions

Analysis of work related musculoskeletal disorders 2001/02 and 2010/11 from the LFS modul demonstrated that the prevalence rate of work related musculoskeletal disorders in Great Britain was at a rate of 1690 cases per 100'000 for people employed in the last 12 months (95% CI's 1560-1820 case rate, $P < 0.05$) in 2010/11 and was statistically significantly lower than the prevalence rate reported in 2001/02 of 2140 cases per 100'000 for people employed in the last 12 months (95% CI's 2020-2260 case rate, $P < 0.05$). Reductions in the prevalence rate of upper limb disorders and back disorders accounted for the greatest portion of the rate reduction. It is considered that government regulations, greater awareness of musculoskeletal disorders in the workplace and interventions by employers and the health and safety community in Great Britain may have contributed to the reduction in the prevalence rate.

Community effects on Infant Mortality Rate in Rural India: a multilevel model

Sada Nand Dwivedi¹, Shahina Begum¹, Alok Kumar Dwivedi¹, Arvind Pandey^{0,2}

¹All India Institute of Medical Sciences (AIIMS), Ansari Nagar, New Delhi, Delhi, India, ²Indian Council of Medical Research, Ansari Nagar, New Delhi, Delhi, India

Objectives

The data on any aspect of public health including that on infant mortality generally has inbuilt hierarchical structure. Traditional regression approach, i.e., ignoring hierarchical structure, does not maintain desired assumption related to independence of records. Due to this, results may get distorted due to probable underestimation of standard error of the regression coefficients. To be more precise, from public health program point of view, an irrelevant covariate may emerge as an important covariate leading to inappropriate public health implications. The objective of the present work was to deal with multilevel analysis of the data on infant mortality from rural area.

Method/Models

The multilevel analysis of the data on infant mortality from rural area, available under second round of National family Health Survey, 1998-99, was used. This method is expected to provide more accurate results leading to meaningful public health implications. In addition, under this method, estimation of variability at different levels and their covariance may also be obtained. Basically, focus of this presentation is to present the detail findings of the multilevel analysis that takes into account hierarchical structure of above mentioned dataset; and notify changes in results under traditional regression analysis that ignores hierarchical structure of data.

Results and Conclusions

The results indicate that the community (e.g., state) level characteristics still have major role regarding infant mortality in rural India. Thus, if computational facilities are available, multilevel analysis may be preferred in dealing with data involving hierarchical structure leading to accurate results having meaningful public health implications.

Modelling fatigue-life failure distribution

Kishore Das, Tanusree Deb Roy

Gauhati University, Guwahati, Assam, India

Objectives

An attempt has been made to model fatigue-life failure data using Weibull distribution considering weight in the Weibull distribution. Various properties of the newly developed distribution has been studied. Characterization of the weighted Weibull distribution has been considered.

Method/Models

Modelling on the fatigue-life data has been done using Weibull distribution. Further, weighted Weibull distribution and Weibull distribution are fitted for determining the best fitting distribution using Newby's method and method of moments. The performances of the distribution are evaluated using the chi-square goodness of fit test.

Results and Conclusions

Survival probabilities and hazard function of the weighted Weibull distribution has been obtained. The distributions have been fitted with the fatigue-life data. It is seen from the two tabulations of expected frequencies, which provides an excellent fit of the observed data. However, the fit by the weighted Weibull distribution is found to be better fit than the Weibull distribution.

Bootstrapping method to confirm presence or absence of outlier subject in a standard bioequivalence study

Vinayak Deshpande

Sankhya Analytical Research Pvt Ltd, MUMBAI, India

Objectives

In a standard bioequivalence study, equivalence is established if 90% confidence interval for the ratio of least square means of C_{max} , AUC_t and AUC_{inf} lie between 0.80 and 1.25. The outcome of the study is influenced by 'outlier' subjects. It is necessary to evaluate the impact of outlier subject on the outcome of the study.

The objective of this paper is to demonstrate the use of **bootstrapping method** to confirm the presence or absence of outlier subject.

Method/Models

Analysis of variance was performed using sequence, treatment and period as fixed factors and subjects nested within sequence as a random factor.

The pharmacokinetic parameters were log transformed prior to analysis.

To check for the presence of outlier subject, 3 methods were used:

- 1) Lund's test based on studentized residuals
- 2) Estimated Distance Method (ED Method)
- 3) Bootstrapping Method

Bootstrapping was performed using sampling with replacement.

5000 samples of size equal to original sample were generated.

For each sample, ratio of least square means and 90% confidence intervals were obtained.

Percentage of samples in which bioequivalence was established was used to derive final conclusion.

Results and Conclusions

Bioequivalence was not established as 90% confidence interval for ' C_{max} ' was outside 0.80-1.25.

Graphical presentation of the data suggested presence of outlier subject.

Although, the absolute studentized residual was maximum for suspected subject outlier, it was smaller than the critical value. Therefore, Lund's test could not confirm the presence of outlier subject. As a confirmatory test, Estimated Distance method was used.

ED method identified the subject as an outlier. Since the two tests did not agree on conclusion, it was decided to use bootstrapping method.

The bootstrapping results provided a very clear conclusion and the subject was confirmed as statistical outlier.

Finally, results were presented including and excluding the outlier subject.

Medical Decision Support Systems by Statistical Scale Measures through Stochastic Modeling

B.N. NAVEEN KUMAR¹, P. TIRUPATHI RAO², P.R.S. REDDY¹

¹SRI VENKATESWARA UNIVERSITY, TIRUPATI, ANDHRA PRADESH, India, ²PONDICHERRY UNIVERSITY, PUDUCHERRY, India

Objectives

The objective of the paper is to develop the Decision support systems for Medical treatments by developing suitable stochastic models. Scaling measures for Psychometric parameters have been measured with the constructed Stochastic the formulation of probability distributions with the available empirical data sets. The Statistical measures like Location, Dispersion, shaping and Consistency parameters are derived by using the method of moments. Inferential aspects on testing several heterogeneous hypothesis and verification of reliability of the tool can also be done with the proposed study.

Method/Models

This study has proposed 'n >0' number of decision items each is rated on 'k>0' points scale and extracted the procedures of obtaining the Probability distributions from the empirical data sets. The model is developed by considering two situations such as (i) The decision items are considered to be equally weighted and (ii) the decision items are having different weights as per their importance. Bivariate probability distributions were developed by identifying the proper data sets. Computations templates were developed using MS Excel.

Results and Conclusions

Model behaviour is observed with some hypothetical data sets for having the sensitivity analysis and other related issues. The joint Probability distribution for Decision Items and Decision score; the marginal distributions for both Decision Items and Decision score; The Mean, Standard deviation, Coefficient of Variation, Coefficient of Skewness and Kurtosis; Product Moments along with correlation coefficients were calculated with some empirical data. It is observed that the Prioritized decision items have results that are more consistent, whereas the non-prioritized items have exhibited the symmetric phenomena. Observing the results.

Comparative health-education output and economic growth in Nigeria

Timothy Olabisi OLATAYO¹, Moses C Ekperiware²

¹, *Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria,, Ago-Iwoye, Nigeria,* ²*Obafemi Awolowo University, Ile- Ife, Nigeria, Ile-Ife, Nigeria*

Objectives

The study examined the effect of education and health sectors of the government to economic development in Nigeria. Education and health sectors are the main human capital component of any economy. The development of human capital is a connection point of the domestic and the foreign scene for maximizing the gains from abroad. In Nigeria different uprising like; education standard, quality of graduates, tertiary institution ranking, standard of medical facilities, rising low life expectancy, access to medical service, rate of infant mortality etc. have seek to question the effort of government in contributing to transform these sectors

Method/Models

Annual data of education, health and economic output with the dynamics of both descriptive and econometric methods were used.

Results and Conclusions

The descriptive analysis showed that health and education effort by the government were inconsistent until the nations' democratic dispensation where a progressive efforts were evident in human capital development. Innovating the endogenous growth model, the OLS regression showed that the education and health sectors contributed 7% and 5% to economic growth in Nigeria respectively. Put differently, for every unit change in GDP, education and health sectors contributes 0.9 and 0.5 to economic output in Nigeria respectively. Conclusively, human capital is germane to economic development in the country and no effort exerted will be too much. Government and private sector of less developed countries are therefore recommended to invest in education and health related areas for healthy economic transformation in Nigeria.

Dam Models for Expected Amount of Overflow allowing any Number of Emptiness

M VENKATESWARAN¹, P.R. VITTAL², P.R.S. REDDY¹

¹SRI VENKATESWARA UNIVERSITY, TIRUPATI, ANDHRA PRADESH, India, ²UNIVERSITY OF MADRAS, CHENNAI, TAMIL NADU, India

Objectives

A dam model with linear release rule policy together with random inouts and random outputs is considered. The random inflow-outflow process assumed to be Poisson and the amounts of inouts and outputs are assumed to be separable densities.

Method/Models

An integro-differential equation is arrived at for expected amount of overflow allowing any number of emptiness by reaching the barrier and crossing the barrier

Results and Conclusions

The master equation in this model is expressed in terms of separable kernel and this enables to arrive at a closed form solution by using imbedding technique of Richard Bellman.

Sensitivity of Simultaneous Equation Techniques for application in forest growth modeling

Oluwafemi Oyamakin^{1,2}, Adedayo Adepoju²

¹Forestry Research Institute of Nigeria, Ibadan, Oyo State, Nigeria, ²Dept. of Statistics, University of Ibadan, Nigeria, Ibadan, Oyo State, Nigeria

Objectives

1. Determine the performance of the parameter estimates across upper and lower triangular matrices.
2. Rank simultaneous estimators to know the best estimator for serious consideration as an estimation procedure for structural parameters.
3. Determine the performance of simultaneous estimation methods as the sample size varies.
4. Know the effect of the varying correlation coefficients among the random normal deviates in simultaneous equation techniques.

Method/Models

The followings estimation methods used are;

1. Ordinary least squares(OLS)
2. Indirect least squares or Reduced form (1LS)
3. Two stage least squares estimation or instrumental variable (2SLS).
4. Limited information maximum likelihood (LIML).
5. Three stage least squares estimation method (3SLS)
6. Full information maximum likelihood (FIML)

Results and Conclusions

Using the Average of parameter estimates criterion, **2 3SLIM** are consistent and ranked best estimators followed by **FIML** and by **OLS** for the three cases studied. Absolute bias of parameter estimate showed a decrease in the bias as the value of N changes for P_1 and P_2 in case I. the same pattern was noticed in the second and third case. Case I was better than case II and III in terms of the bias. Hence, OLS generated the least results in all cases. In addition, across the result of the lower triangular matrix were better. For the Root Mean Square Error criterion, OLS still performed the best. There is no clear pattern on the asymptotic behavior of the estimators. The performance of OLS using the Root Mean Square Error showed that OLS is not sufficiently inferior to other estimators.

STATISTICAL ANALYSIS OF PRE-MENOPAUSAL AND MENOPAUSAL AGE OF WOMEN WITH DIETARY INTAKE AND BMI AS PREDICTING FACTORS

OlaOluwa Yaya¹, ModupeOluwa OlaOluwa², Angela Chukwu¹

¹University of Ibadan, Ibadan, Nigeria, ²University College Hospital, Ibadan, Nigeria

Objectives

Different studies have considered premenopausal and menopausal timings of women. A recent study conducted in South western Nigerian on age at menarche concluded that menopause age is set at confidence interval 13.66 1.82 years, with 6.1% of the respondents at 11 years, 49.3% between 13 and 14 years (Raji, et al, 2006). This study is motivated by the work of Raji, et al. (2006). We then consider the effect of dietary intake, Body Mass Index (BMI) on the pre-menopausal and menopausal age of women in Southwestern Nigeria.

Method/Models

After the initial frequency analysis, we will apply the regression method with the pre-menopausal and menopausal age as the dependent variable. The dietary intake and BMI are the independent components. Correlation analysis is also used to check for possible relationship between some of the variables.

Results and Conclusions

The result obtained shows most of the women started menstruation at the average age of 13 years according to Rsji et al (2006).