Canterbury Statistics Open Day

Wednesday, 20 November 2013

9:00am – 4:30pm

HP Seminar Room, NZi3 Building

University of Canterbury
Programme

9:00 – 9:10am Welcome and Introduction by Professor Jennifer Brown, HoD, Department of Mathematics and Statistics, University Canterbury

9:10 – 9:20am Keynote Address by Professor Jan Evans-Freeman, PVC, College of Engineering, University of Canterbury

9:20 – 10:20am Plant and Food Keynote Speaker Session
   • Creating value through advanced analytics by Dr. Kevin Ross, Fonterra

10:20 – 10:50am Morning Tea

10:50am – 12:20pm Session I, Chaired by Dr. Alasdair Noble, Plant and Food Research
   • Doubly resolvable row – column designs by Dr. David Baird, VSN NZ
   • Bayesian methods in official statistics by Dr. Patrick Graham, Statistics NZ
   • Bayesian analysis of multivariate regression model through elliptical copulas and scale mixtures of normal with application to asset pricing model by Dr. Nuttanan Wichitaksorn, University of Canterbury

12:20 – 1:30pm Lunch

1:30 – 3:00pm Session II, Chaired by Dr. Richard Penny, Statistics NZ
   • Divining modes of speciation and extinction: Can statistics help us read the 'tree leaves'? by Professor Mike Steel, Department of Mathematics and Statistics, University Canterbury
   • A Bayesian approach to a binomial distribution with unknown p and unknown, but constrained, n by Dr. Alasdair Noble, Plant and Food Research
   • How can statisticians contribute towards agricultural science in New Zealand?: Examples from an AgResearch statistician by Dr. Chikako van Koten, AgResearch

3:00 – 3:30pm Afternoon Tea

3:30 – 4:30pm Session III, Chaired by Professor Jennifer Brown, Department of Mathematics and Statistics, University Canterbury
   • The New Zealand longitudinal census study by Dr. Robert Didham, Statistics NZ
   • Applying Real Options approach to determine the optimal vaccination timing and threshold by Dr. Elena Molchanova, Department of Mathematics and Statistics, University Canterbury
Abstracts

Creating value through advanced analytics

Kevin Ross
Fonterra

Fonterra is one of several organisations in New Zealand and around the world who are investing in the development of advanced analytics capability. In order to maximize the value created from milk, Fonterra needs to employ a combination of descriptive, predictive and prescriptive analytics. We have developed a range of models to support decisions from milk collection to product mix and asset investment. In 2013, we launched the New Zealand Analytics Forum, a group of professionals who meets regularly to share best practice principles. We see the potential for this to benefit a wide range of industries, and New Zealand as a whole. I will share some of the insights from both Fonterra’s efforts and the Analytics Forum, with a focus on what I believe will lead to the successful development of analytics capability.

Doubly resolvable row – column designs

David Baird
VSN NZ

These designs generalise Latin squares, with replicates made up of groups of rows and columns. Replicates in both directions are useful when spatial effects and trial operations can happen in both row and column directions. For example the trial may be sown row by row, and irrigation performed down the columns. The replicates may be made up of full rows and part rows, but are contiguous in both directions. The allocation of treatments is also made in a manner to optimize the normal row-column analysis of the trial. The advantage over the standard row-column designs is that, if only the row and column replicates are needed to control the spatial effects, then these are orthogonal to the treatments means, so no treatment information is lost and the raw means require no adjustment. A method of generating these designs will be given.

Bayesian methods in official statistics

Patrick Graham
Statistics NZ

Historically Bayesian methods have not been much used in official statistics. This may, in part, have been due to the assumed subjectivity of Bayesian methods, in contrast to the assumed objectivity of frequentist methods. However, Bayesian methods seem well-suited to some areas of official statistics which are problematic for the design-based approaches to inference traditionally used in official statistics agencies. These include small domain estimation and methods of estimation based on administrative data in which sampling is formally absent but numerous other sources of uncertainty are present. After briefly examining claims concerning subjectivity and objectivity of different modes of statistical inference, I will illustrate the potential of Bayesian methods to contribute to official statistics using examples drawn from problems in small domain
estimation, inference from administrative data and an area at the intersection of these issues, small domain population estimation.

---

**Bayesian analysis of multivariate regression model through elliptical copulas and scale mixtures of normal with application to asset pricing model**

*Nuttanan Wichitaksorn*

Department of Mathematics and Statistics, University of Canterbury

We propose a new method to implement the elliptical copula-based multivariate regression models through scale mixtures of normal. With the scale mixtures of normal simplification in both copula and margins, we can efficiently estimate the model parameters through the Markov Chain Monte Carlo method. We illustrate some examples from Gaussian and Student-t copulas through a simulation study and the results show the favorable performance. A real data example on the analysis of U.S. excess return through the capital asset pricing model is illustrated to show the applicability of the proposed method.

---

**Divining modes of speciation and extinction: Can statistics help us read the 'tree leaves'**?

*Mike Steel*

Department of Mathematics and Statistics, University of Canterbury

Biologists use genomic sequences to identify evolutionary relationships between species, and thereby to construct an evolutionary tree. Once we have many such trees we can then look at just their 'shape', ignoring which species are which. These shapes harbor a signal of the speciation-extinction process by which ancestral lineages split over time to create new lineages, or died out. I will describe some problems that arise when trying to fit stochastic models to observed data on tree shapes, along with our simple proposal that provides a consistent fit on a large data base of trees. Part of our approach can also be applied to better predict how much 'evolutionary heritage' (phylogenetic diversity) might be lost in the short term due to current high rates of extinction.

---

**A Bayesian approach to a binomial distribution with unknown p and unknown, but constrained, n**

*Alasdair Noble*

Plant and Food Research

During the work done when the Painted Apple Moth incursion occurred there was interest in finding female survival rates of eggs on various species of foliage. 50 eggs were placed on samples of foliage in 5 replicates, ie 10 eggs per replicate. However the sex of the eggs is unknown so assuming a binomial distribution the “n” ie number of trials from which we can estimate the survival rate is unknown. The number of female eggs was modeled with a binomial distribution and various priors for the sex ratio were considered.
How can statisticians contribute towards agricultural science in New Zealand?: Examples from an AgResearch statistician

Chikako van Koten
AgResearch

Agriculture is one of the most important industries in New Zealand and a key for our successful future. In this talk, I briefly introduce some examples of latest agricultural research at AgResearch, which I have been involved in, in order to demonstrate: 1) what kind of role a statistician can take at AgResearch, and 2) how statisticians, in collaboration with other scientists, can make significant contribution towards the progress of agricultural science and agriculture in New Zealand.

The New Zealand longitudinal census study

Robert Didham and Kirsten Nissen
Statistics NZ

Statistics New Zealand has developed a longitudinal database of linked censuses covering 1981 to 2006. The Census of Population and Dwellings has long been the core data source for social and economic research where population information has been needed. This has provided a set of cross-sectional snapshots of the New Zealand population, but until now it has only been possible to look at change over time using these independent data sources as synthetic cohorts. In contrast, the linking of censuses enables research to be carried out using real cohorts. Thus real value is added to the data and the analytical power has increased substantially. This presentation introduces the linked census data, explains some of the challenges faced in creating it, the limitations of the data and safe-guards necessary to protect confidentiality and privacy. For the present audience - predominantly mathematical statisticians - the focus will be on the value of the dataset as a genuine, real-world, set of information against which statistical models and measures can be tested or developed. As such it is an ideal testbed for theoretical statistical work on a range of topics such as multidimensional bias analysis, missingness and longitudinal data visualisation.

Applying Real Options approach to determine the optimal vaccination timing and threshold

Elena Moltchanova
Department of Mathematics and Statistics, University of Canterbury

Epidemics are an ever-present danger and although preventive vaccination helps reduce the disease pool considerably, the decision maker is often faced with the choice of buying the vaccine early and perhaps wasting money and loosing political credibility or buying late, when more information is available, but when the valuable initiative has been lost. Although guidelines exist, recognising an epidemic and deciding when to take action may still be problematic.

We apply a Real Options approach to the standard Susceptible-Infected-Recovered (SIR) epidemic model to develop a dynamic decision-making framework. Unlike the Present Value Optimisation, Real Options approach takes into account the cost of waiting and the knowledge
gained by waiting, making it an exciting and valuable tool for real-time strategic optimisation. The value of information is thus explicitly taken into account.

We demonstrate some simulations studies and discuss the potential challenges associated with the implementation of the approach. We also demonstrate the sensitivity of the analysis to exogenous parameters and present the possibility of a dynamically learning decision making system.

---

**Details**

**Date & Time:** Wednesday, 20 November 2013 from 9:00am to 4:30pm  
**Venue:** HP Seminar Room, NZi3 Building, University of Canterbury  
*(See campus map below.)*  
For information on parking at UC, please visit: [http://www.fm.canterbury.ac.nz/parking/](http://www.fm.canterbury.ac.nz/parking/)

**Contacts:**  
Dr. Nuttanan (Nate) Wichitaksorn  
Phone: +64 3 364 2987 ext 7677, E-mail: nuttanan.wichitaksorn@canterbury.ac.nz  
Professor Jennifer Brown  
Phone: +64 3 364 2987 ext 6696, E-mail: Jennifer.Brown@canterbury.ac.nz  
Dr. Alasdair Noble  
Phone: +64 3 325 9656, E-mail: alasdair.noble@plantandfood.co.nz

**Campus Map**

New Zealand ICT Innovation Institute (NZi3) Building  
Physical Location: 69 Creyke Road, Christchurch  
Or visit the online map at  
[http://maps.canterbury.ac.nz/home/browse/1160](http://maps.canterbury.ac.nz/home/browse/1160)