

## Statistics Workshop in Commemoration of Professor Chris Heyde

### Abstracts

#### **A Model for the Spread of Cholera**

*Professor Joe Gani*

Mathematical Sciences Institute, Australian National University

Email: [gani@maths.anu.edu.au](mailto:gani@maths.anu.edu.au)

In 1854, a cholera epidemic broke out in London, which was recorded by the British physician John Snow. He showed that its cause was water contaminated by faecal matter from the Broad St pump; once this had been put out of action, the epidemic subsided. A simple model is proposed for the spread of cholera.

#### **An Introduction to Probability**

*Professor Peter Taylor*

Australian Mathematics Trust, University of Canberra

Email: [pjt@olympiad.org](mailto:pjt@olympiad.org)

Chris Heyde was one of the founding members of the Board of the Australian Mathematics Trust and always available to the founding Executive Director Peter O'Halloran and myself. To honour his memory I would like to discuss some of the nicest probability problems and related issues from the Australian Mathematics Competition. In their way these problems helped introduce talented students for the first time to the structure and thinking in probability and hopefully inspired them to further learning.

## **Seasonality in Chlamydia Pneumoniae Infection in Canberra**

*Dr Alice Richardson*

Faculty of Information Sciences and Engineering, University of Canberra

Email: [Alice.Richardson@canberra.edu.au](mailto:Alice.Richardson@canberra.edu.au)

ACT Pathology at The Canberra Hospital stores large amounts of data relating to the serology tests it carries out. To date the patterns in this data have not been systematically analysed. Nonetheless pathologists are interested in many aspects of the test data, including quantification of seasonal patterns in testing load.

Knowledge of peak testing time would allow the lab to prepare for these busy times of year.

In this talk I will introduce the data set relating to the infection Chlamydia pneumoniae. I will describe how a straightforward application of time series analysis techniques was able to satisfy the pathologists' request for numerical measures of seasonality. I will also describe the broader data mining project in which this analysis sits.

This talk describes joint work by Simon Hawkins, Brett Lidbury, Fariba Shadabi and Dharmendra Sharma (all of UC) and John Fulcher of the University of Wollongong.

## **Simultaneous Prediction of Toxicity of Multiple Chemicals to Multiple Species**

*Dr Richard Morton*

CSIRO Mathematical and Information Sciences, GPO Box 664, Canberra

Email: [Richard.Morton@csiro.au](mailto:Richard.Morton@csiro.au)

A theoretical model is developed for the estimation of the toxicity of a large number of chemicals to a group of species for which the log-toxicity values have strong linear relationships. The model is a multivariate Functional Relationship (FR) that allows for many missing values. It depends on the assumption that there is an unobserved variate (denoted by  $\xi$  in the text), with a value for each chemical; for each species, the expectations of the log-toxicity values are each linearly related to  $\xi$ . Theory of Estimating Equations leads to an algorithm for estimating the parameters in the presence of incidental parameters  $\xi$ . Simulations show that there are worthwhile gains in multivariate over the bivariate FRs hitherto used. The model is applied to data on the toxicity of 51 chemicals to 10 species. It allows for species to have unequal variance about their expectations.

The results are very encouraging, and show that good prediction of log-toxicity from measurements on other species where only 30% of the combinations were measured is possible even when the taxa are not closely related. Among the uses for this technique are (i) to predict toxicity for combinations of species and chemical that have not been measured (ii) to suggest a suitable range of concentrations for a new bioassay (iii) identifying specificity i.e. chemicals that have a different relative toxicity for a particular species.

The number of incidental parameters to be estimated can be reduced by finding their relationship to  $K_{ow}$  (a ratio of uptake of solvent) and the MOA (Mode of Action) of the toxicant. This should enable application of the methodology to data sets with even higher proportions of missing observations.

Keywords: Estimating equations, Inter-species relationships, Missing values, Multivariate functional relationship, Toxicity.

Reference: Morton, R., Warne, M.St.J. & Correll, R.L. (2008) Environmetrics (In press).

## **Causal Integral Equations**

*Dr Bob Anderssen*

CSIRO Mathematical and Information Sciences, GPO Box 664, Canberra

Email: [Bob.Anderssen@csiro.au](mailto:Bob.Anderssen@csiro.au)

The first rigorous formulation of a convolutional causal integral equation appears to be due to Boltzmann. He proposed it as a model for linear viscoelasticity. Clearly, the simplest version of a causal integral equation is the fundamental theorem of calculus. Such equations are now known as first kind Volterra integral equations. For the linear viscoelasticity equation, there are interesting connections back to statistics, probability and financial mathematics, some of which will be explored in the talk.

## **A Survey of Numerical Methods for Exponential Lévy Process Models**

*Dr David Heath*

Mathematical Sciences Institute & School of Finance and Applied Statistics  
Australian National University

Email: [heath@maths.anu.edu.au](mailto:heath@maths.anu.edu.au)

In recent years more and more attention has focused on financial models that depart from the standard Black-Scholes model. Observations of large, sudden movements in asset prices have led to the use of stochastic processes with jumps. Some of the most popular of these are exponential Lévy processes. However use of these models has been limited partly due to a relative lack or scarcity of suitable quantitative methods that can be employed for this class of models. This talk will survey the numerical methods that have been used for the pricing and hedging of vanilla and exotic options. The main methods are Monte Carlo techniques, partial integro-differential equations and lattice or tree type methods. For European instruments other techniques include valuation via the transition densities, characteristic function and corresponding Fourier inversion, and integration using the subordination property of a Lévy process, expressed as a time-changed Brownian motion. Some commentary on how these techniques can be tailored or harnessed for Chris Heyde's FATGBM, to improve the tractability of this model, will be given.

## **A Generalised Skewness Statistic with Reference to Financial Applications**

*Professor Ross Maller*

Mathematical Sciences Institute & School of Finance and Applied  
Statistics  
Australian National University

Email: [Ross.Maller@anu.edu.au](mailto:Ross.Maller@anu.edu.au)

I will report on an investigation into the skewness of stock returns, originally begun as a collaboration between Chris Heyde, PhD student Ben Kaehler, and myself. We suggest a class of generalised skewness statistics, depending on a parameter  $\beta$ , that reduces to the usual skewness statistic when it is applicable, but may also be used when a sample's distribution is believed to be leptokurtic. In order to use the (studentised) statistics for hypothesis testing, their asymptotic distributions are derived in the context of stationary ergodic martingale differences, rather than under the usual assumption of independent identically distributed samples, making the theory particularly well suited for use in financial applications. The new statistics are applied to a large data set of 614 stocks listed on the Australian share market. Significant positive (but no negative) skewness was found in 18% of the stocks.