STOR-i Workshop 2016, 7th-8th January

Abstracts

Day 1

Routing, Logistics and Management of Recharging Infrastructure for Electric Vehicles
Pitu Mirchandani, School of Computing, Informatics and Decision Systems Engineering, Arizona State University

There is much reason to believe that many individuals and organizations will transform their vehicles to ones that utilize alternative fuels which are more sustainable. The electric vehicle (EV) is a good candidate for this transformation, especially which "refuels" by exchanging its spent batteries with charged ones. Unfortunately, although there is much research gone into technologies of EVs, little effort have gone into designing the battery charging-/ exchanging- infrastructure. This presentation discusses the issues that must be addressed, principally the issues related to the limited driving range of each electric vehicle's set of charged batteries and the possible detouring for battery exchanges. In particular, the talk will address the optimization and analysis of infrastructure design alternatives dealing with (1) the routing of vehicles from origins to destinations, (2) the optimum locations of battery-exchange stations, and (3) the recharging capacity and operations management of battery-exchange infrastructure. Some infrastructure design and optimization models, and some results, will be discussed in this lecture.

Lubrizol Internship: Statistical Predictions of Engine Test Results
Lisa Turner, STOR-i PhD student, Lancaster University

A key division in the Lubrizol Corporation is chemical additive performance packages and products for the use in passenger cars and heavy duty diesel engine oils. As well as developing packages that deliver engine oil performance requirements, they are also responsible for testing engine oils in customer refined base oils. These laboratory engine tests are expensive and time consuming. Therefore, it is in Lubrizol's best interest to optimise the information they extract from the engine tests and develop an understanding of the effects of the chemistry.

In this talk, I will give an introduction to how statistical methods, such as Bayesian Variable Assessment, are used to find important parameters, often from an extremely large number of possible predictor variables. I will also explain how the models are built to predict the engine test results, which are used by Lubrizol to help chemists and formulators select products for testing that are likely to pass the engine tests whilst minimising the formulation costs. For these models to be used, the chemists must have confidence in the model predictions. For example, the predicted value for an oxidation test at 72 hours should be less than the predicted value at 196 hours. I shall look in more detail at how constraints can be applied to the prior distribution of the parameters to ensure reasonable predictions for such properties.

Stratified Space-Time Infectious Disease Modeling
Jon Wakefield, Departments of Statistics and Biostatistics University of Washington

In this talk I will describe an extension to an existing modeling framework in order to analyze data on hand, foot and mouth disease (HFMD), collected in the central north region of China over 2009-2011.
An interesting class of space-time models for infectious disease data was proposed by Held et al (2005). In subsequent developments, random effects were added, with inference carried out via penalized quasi-likelihood. We provide a novel derivation of the model, which allows new insight into key assumptions and provides parameter interpretation. Subsequently, we extend the model class in two directions. First, we simultaneously model the

disease transmission between subgroups (i.e., age-gender strata), in addition to space and time. This is crucial in our application as infection varies considerably by age in particular and, as we illustrate, the temporal dynamics of disease incidence vary by age and gender. Second, we are able to use our model to make inference on informative summaries such as local reproductive numbers. Bayesian inference is carried out via Markov chain Monte Carlo. The results for the HFMD reveal interesting findings concerning the transmission mechanisms between different demographic groups. The local reproductive numbers show significant spatial and temporal variation within the study region. These findings provide vital information that can be used in public health initiatives such as the design of intervention programs. This is joint with Cici Chen, Brown University.

Reference:

Held, L., M. H□ohle, and M. Hofmann (2005). A statistical framework for the analysis of multivariate infectious disease surveillance counts. Statistical Modelling 5, 187-199.

Operational Research in Finance
Alexander Armstrong, Morgan Stanley

Working as a quantitative analyst (a quant) used to be all about pricing derivatives and other financial products with stochastic calculus. As the landscape of finance has changed since the global financial crisis, new problems and opportunities have arisen for operational researchers.

In this talk I will describe some of the projects that I have worked on since graduating from Lancaster University. Ranging from applying machine learning to over-the-counter markets, to examining risk on some of the world's largest hedge funds, to changing European securities regulation.

Multi-class queueing applications in automated warehousing systems
Ivo Adan, Department of Industrial Engineering and Innovation Sciences, University of Technology, Eindhoven

In this talk we discuss the use of multi-class queueing models for the design of automated warehousing systems, including sequential multi-segment zone-picking systems and compact picking systems. To estimate key performance statistics such as throughput capacity and recirculation, we propose various techniques, including product-form approximations, mean-value analysis and (dis-)aggregation.

Optimisation in the Energy Sector
Clym Stock-Williams, Eon

An overview will be given of the range of problems in the energy sector to which mathematical optimisation can be applied. Some examples of the work being undertaken within E.ON to solve these problems will be given, and we will conclude by highlighting some of the key areas where optimisation algorithms could have a big impact on the effectiveness of different energy technologies in the coming years.

STOR-i
excellence with impact

Lancaster University

Day 2

Exact Bayesian inference for change-point detection and comparison
Stéphane Robin, Applied Mathematics and Informatics, Agro Paris Tech

Change-point detection arises in many fields, including economics, earth sciences and biology. From a general point of view, the problem is to find changes in the distribution of data collected along 'time'. From a statistical point of view, change-point inference raises a series of specific problem due to the hybrid nature of the parameters to be inferred: the parameters of the emission distributions are typically continuous, whereas the change-point locations are discrete.

From a Bayesian point-of-view, such a setting requires to be able to integrate over both continuous and discrete sets. Using conjugate priors is convenient to solve the former problem but the later raise serious combinatorial issues due to the large number of possible segmentation of a series of observations.
We will present a series of results showing that the Bayesian inference of some change-point models can be achieved in an exact manner using some algebraic properties of the problem. A series of posterior distributions such as this of the change-point locations or of the total number of change-points can be computed exactly and efficiently. We will also show that these results can be extended to compare the location of change-points in series observed independently.

Inference from Noisy High Dimensional Acoustic Data
Tim Park, Shell

Distributed Acoustic Sensing (DAS) is a recently introduced sensing technique which has opened up many new possibilities. It is used to measure acoustic and thermal disturbances along the length of an oil producing well. The underlying physical phenomenon is elastic Rayleigh back-scattering of coherent light from inhomogeneities in the core of an optical fibre. The data is collected at a very high temporal frequency and regularly spaced depths. This allows for real time estimation of important quantities such as flow rates, flow composition, slug formation and the propagation of sound waves. These quantities, which previously could only be measured at the surface, can then be used to improve well management and ultimately increase production. The high volume of data collected can make analysis difficult, we therefore make use of efficient Fourier based methods. In this talk we present our methodology as well as applications to real wells.

Multi-armed bandit experiments in the online service economy
Steve Scott, Google

The modern service economy is substantively different from the agricultural and manufacturing economies that preceded it. In particular, the cost of experimenting is dominated by opportunity cost rather than the cost of obtaining experimental units. The different economics require a new class of experiments, in which stochastic models play an important role. This article briefly summarizes mulit-armed bandit experiments, where the experimental design is modified as the experiment progresses to make the experiment as inexpensive as possible.

Projection Pursuit for Clustering with Applications in Population Genomics
David Hofmeyr, STOR-i Research Fellow

Projection pursuit refers to a class of optimisation problems based on identifying subspaces within which data contain "interesting" structure. If "interestingness" can be defined relative to the objective of a data analytic task, then this framework is highly versatile. This talk will look briefly at some recently proposed projection pursuit formulations for clustering, and discuss some benefits of the projection pursuit framework for dimension reduction and clustering. Some early results from applications to some very high dimensional genetics data sets will be presented.