Statistics in Action: A Smörgåsbord

Chair: Wendy Lou (University of Toronto)

YIZHOU FANG, University of Waterloo
Fast and Flexible Non-linear Modeling of Financial Assets

The last few decades have featured numerous examples of delicately inter-related financial products. Stochastic models have gone very far in the financial modeling field. However, many of the complex multi-factor models used at present incur a large computational burden which increases dramatically with the size of the data. We propose a nonlinear framework including observable proxies for latent volatility components. Inference for the proposed framework is very fast and scalable to high frequency data, and therefore amenable to large-scale model comparisons and goodness-of-fit assessments. A number of case studies illustrating the methodology will be presented.

PHILIPPA SWARTZ, Simon Fraser University
The Quality of Pitches in Major League Baseball

We investigate the quality of pitches thrown in Major League Baseball (MLB). The approach is based on characterizing pitches based on pitch count and a descriptor. The descriptor may include information on the type of pitch, the speed of the pitch, the location of the pitch, etc. This is a big data problem with roughly 750,000 pitches thrown in a single season of MLB. If we are able to assess the quality of pitches, then a host of practical questions may be addressed. For example, can we detect when a pitcher's performance begins to deteriorate prior to bad results?

MICHAEL GUERZHOY, University of Toronto
Hierarchical Bayesian Models for Uncertainty-Quantified Ranking of Restaurant Chains by Food Safety Compliance

For an episode of CBC Marketplace (April 2014), data has been collected on the number of food safety violations found during inspections at hundreds of locations of 13 restaurant chains in 5 Canadian cities. We describe several methods to produce an uncertainty-quantified nationwide ranking of the chains by food safety compliance, addressing the issue of differing standards for what constitutes a violation in different cities. We demonstrate a ranking based on the coefficients of a Poisson regression. To obtain more precise rankings with less uncertainty, and gain insights about the data, we fit a series of Bayesian overdispersed Poisson models.

ANITA BROBBEY, Memorial University of Newfoundland
Analysis of Ordered Categorical Data from Designed Experiments

The responses of interest in many designed experiments are ordered categorical. Taguchi Accumulation analysis (AA), Nair's Scoring Scheme (1986) and Jeng's weighted Probability Scoring Scheme are existing approaches for analyzing ordered categorical data from industrial experiments for optimal factor settings in quality improvement studies. In this work, we discuss analyses of ordered categorical data using cumulative link models. This approach models both the location and dispersion effects and it is easy to implement. Relative Performances of the four approaches are compared through simulation studies and analysis of real data set.

THUVA VANNIYASINGAM, McMaster University
Determining Individual-Specific Terminal Half-Life Estimates and Optimal Sampling Time Points for Patients with Hemophilia B Administered rFIXfc

Individual-specific estimates have not been determined for patients with hemophilia B who are receiving prophylactic treatment for Alprolix. We aim to (i)determine individual-specific terminal half-life estimates and (ii)identify informative sampling points

to reduce the required number of blood samples, without loss of accuracy in pharmacokinetic estimates. Using nonlinear mixed-effects regression, we identified a three-compartment model to describe plasma concentrations after an IV bolus injection, while simultaneously estimating interindividual variability, intraindividual variability, and covariate effects. A limited sampling strategy will be used to determine optimal sampling time points to reduce the number of required concentration measures and improve overall prophylactic management.

CHARLES BORDET, Université Laval

A Stochastic Pella Tomlinson Model and its Maximum Sustainable Yield

This presentation investigates the biological reference points, such as the maximum sustainable yield, for the Pella Tomlinson surplus production model in the presence of a multiplicative environmental noise. This model is used in fisheries stock assessment as a firsthand tool for the elaboration of fishing strategies. The consideration of random noise leads to more conservative harvesting target than deterministic models.