

Graduate Student Seminar Nov 2013

Limit Order Books and Machine Learning

Tadeu Ferreira

Most electronic markets nowadays are quote driven, where quotes (prices and volume that individuals are willing to buy and sell) collect in the Limit Order Book (LOB). There are several analysis and different approaches to understand the dynamics of the LOBs coming from different fields, such as, economics, statistics, physics and mathematics. In this talk, I will provide a brief description of LOBs along with an analysis of some aspects of its dynamics using selected models from Machine Learning such as Hidden Markov Models and Restricted Boltzmann Machines, where the probabilistic perspective is emphasized.

On Bias Adjustments for Web Surveys

Lingling Fan

With the rise of internet, web surveys have become an easy and inexpensive means to collect data for population studies, e.g. to gauge consumer attitudes or political opinions. Statistical inference based on web-survey samples often requires cautious treatment to account for potential biases when applying conventional statistical methods for survey data. In this talk, we focus on non-coverage bias, and discuss several approaches, including inverse-probability-weighting and GREG-type methods, for bias adjustment. We consider scenarios depending on the availability of information on the reference population, and propose a new approach based on a Horvitz-Thompson-type estimator. We present some simulation results where this proposed estimator performs well for all cases considered. Possible extensions of the approach will also be discussed.

Penalized dependence calibration in conditional copulas

Jialin Zou

We provide a nonparametric procedure, for both true and nonparametric estimated marginal, to detect whether the dependence structure is constant or not over certain region. The proposed method can recover the shape in flat region and estimate copula parameter at each grid simultaneously by combining group SCAD penalty with local polynomial regression. Moreover, the proposed method provides a tool to illustrate the dependence pattern in a more parsimonious manner. We derive asymptotic consistency, sparsity and normality result needed to construct asymptotic pointwise confidence interval for copula parameter at any grid.

Self-Exciting Event Data in Finance

Jason Ricci

When analyzing event data, many natural systems exhibit large amounts of clustering. Although there are cases when such behaviour can be explained by Poisson clustering, there are many others where self-excitation effects are more pronounced and quiet periods persist far longer. Other disciplines including biology, neuroscience, and seismology have modeled event data using the Classical Hawkes Process. In this talk, the Generalized Hawkes Process is introduced, fit to financial data, and some goodness of fit diagnostics are presented.

Effective dimension reduction for sparse functional data

Edwin Lei

We propose functional cumulative slicing as an effective means of achieving dimension reduction in the context of multiple index functional regression for a scalar response variable. In many real world scenarios, predictor trajectories are observed intermittently, contaminated with noise, and collected in the form of repeated measurements. These design issues force us to adopt the strategy of pooling data together from the entire sample for consistent estimation and inference. Our proposed method of functional cumulative slicing addresses these sparsely observed, as well as the easier case of densely observed, predictor trajectories. Compared to sliced inverse regression, the nature of cumulative slicing lends itself well to sparse designs because it avoids partitioning the data into an arbitrary number of slices -- thereby ensuring the maximum use of data. We demonstrate the effectiveness of our estimator compared to functional linear model and sliced inverse regression with simulations and data analyses, and provide a theoretical justification of our method.