Modern Statistical Methods for Complex Health Data

Chair: Kevin He (University of Michigan)
Organizers: Kevin He and Yi Li (University of Michigan)

RICHARD COOK, University of Waterloo
Statistical Challenges Arising in Transfusion Research

There are many important scientific questions in transfusion medicine that remain to be addressed. There is currently considerable interest in i) exploring the relation between the duration blood has been in storage prior to transfusion and adverse health outcomes in recipients, ii) the effect of transfusing blood to patients with a different blood type, and iii) the effect of transfusing gender mismatched blood (i.e. blood from female donors to male recipients). The statistical challenges and approaches to analysing data from a large registry of transfused patients are discussed with a view to exploring these issues.

JOHN KALBFLEISCH, University of Michigan
Optimization in a Kidney Paired Donation Program

Kidney transplant is the best available treatment for patients with end stage kidney disease. Often, however, patients have a willing living donor, but that donor is incompatible in blood type and/or histology. A kidney paired donation (KPD) program consists of transplant candidates and their incompatible living donors as well as non-directed or altruistic donors (ADs). Exchanges of donors among candidates and chains created by ADs overcome the incompatibilities. A problem of importance is how best to arrange exchanges and chains in order to achieve as many transplants as possible. We develop allocation schemes that account for uncertainties. These methods are compared through simulations with methods currently in use and illustrate substantial gains in transplants achieved.

YI LI, University of Michigan
Gateaux Differential-Based Boosting for Time-Varying Effects in Survival Analysis

Survival models with time-varying effects provide a flexible framework for modeling the effects of covariates on event times. However, the difficulty of model construction increases dramatically as the number of variable grows. Existing constrained optimization and boosting methods suffer from computational complexity. We propose a new Gateaux differential-based boosting procedure for simultaneously selecting and automatically determining the functional form of covariates. The proposed method is flexible in that it extends the gradient boosting to functional differentials in general parameter space. In each boosting learning step of this procedure, only the best-fitting base-learner (and therefore the most informative covariate) is added to the predictor, which consequently encourages sparsity. In addition, the method controls smoothness, which is crucial for improving predictive performance.

ROSS PRENTICE, Fred Hutchison Cancer Research Center
Nonparametric and Semiparametric Analysis of Bivariate Failure Time Data

The nonparametric maximum likelihood estimation problem is re-defined to entail maximization over marginal hazard rates, and over double failure hazard rates at all grid points formed by uncensored times where there is a possibility of double failure given preceding data. The resulting unique NPMLE can be calculated using a two-step procedure. The fist involves setting aside all doubly censored observations that are interior to the risk region, leading to a Dabrowska-type NPMLE from the remaining data. A corresponding NPMLE arises also for the censoring 'survivor' function. The second brings in the omitted data using self-consistency, yielding non-iterative NPMLEs for both the failure and censoring distribution functions. Simulation studies demonstrate modest efficiency improvements. Regression generalizations will be briefly mentioned.