Statistical Sciences
UNIVERSITY OF TORONTO

SEMINAR

February 23, 2017 at 3:30 pm
Refreshments will be provided at 3:15pm
Sidney Smith Hall, Room 2108
Speaker: Alexander Franks, University of Washington
Host: Mike Evans

Information-sharing schemes for complex data analysis: examples from high-throughput biology and professional basketball

Modeling complex data usually involves sharing information between models for smaller subsets of the data. I will discuss examples of this in a discussion on data analyses from two distinct domains. In the first part of the talk, I will describe a model-based method for evaluating heterogeneity among several p x p covariance matrices in the large p, small n setting. Here, we regularize estimates of the covariance matrices by exploiting the fact that data from similar sources share similar structure. I will illustrate the utility of the method for exploratory analyses of high-dimensional multivariate gene expression data. In the second half of the talk, I describe the analysis of a spatio-temporal dataset of professional basketball players. Efficient inference requires models which pool information both between players and across space. By blending Bayesian hierarchical models with geography inspired mapping tools, we will shed light on previously unidentified aspects of skill. In particular, our results describe spatial variation of defensive ability among NBA players. Although we apply these methods to professional basketball data, the models are widely applicable in other domains.