

STATISTICS COLLOQUIUM

MONDAY, MARCH 21, 2011

TALK: 4:00 PM - SCIENCE CENTER RM. 309

RECEPTION: 5:15 PM - SCIENCE CENTER, 7TH FLOOR

“A Bayesian Semiparametric Approach to Intermediate Variables in Causal Inference”

Fabrizia Mealli
Università di Firenze

ABSTRACT

In causal inference studies, treatment comparisons often need to adjust for confounded post treatment (intermediate) variables. Principal stratification (PS) is a framework to deal with such variables within the potential outcomes approach to causal inference. Continuous intermediate variables introduce inferential challenges to PS analysis. Existing methods either dichotomize the intermediate variable, or assume a fully parametric model for the joint distribution of the potential intermediate variables. However, the former is subject to information loss and the latter is often inadequate to represent complex distributional features. We propose a Bayesian semi parametric approach that consists of a flexible parametric model for the potential outcomes and a Bayesian nonparametric model for the potential intermediate outcomes using a Dirichlet process mixture (DPM) model. The DPM approach provides flexibility in modeling the possibly complex joint distribution of the potential intermediate outcomes and offers better interpretability of results via its clustering feature. The Gibbs sampling based posterior inference is developed. We illustrate the method by two applications: one concerning partial compliance in a randomized clinical trial, and one concerning the causal mechanism between physical activity, body mass index and cardiovascular disease in the observational Swedish National March Cohort study.

This is joint work with Scott L. Schwartz and Fan Li.