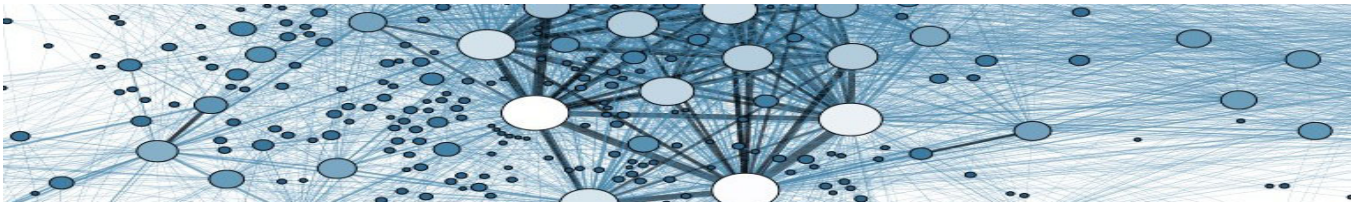


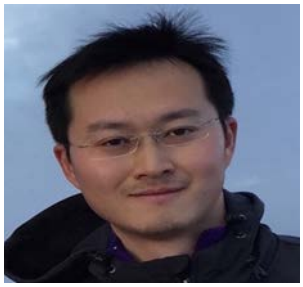
Department of Statistics Seminar

Tuesday, May 30, 2017

3:45 – 4:45 p.m., Room 420, Olmsted Hall

Reception in Olmsted 1331 at 3:15 p.m.

Principal Stratification Analysis Using Principal Scores



Peng Ding, Ph.D.

Assistant Professor

Department of Statistics

University of California, Berkeley

Abstract:

Practitioners are interested in not only the average causal effect of the treatment on the outcome but also the underlying causal mechanism in the presence of an intermediate variable between the treatment and outcome. However, in many cases we cannot randomize the intermediate variable, resulting in sample selection problems even in randomized experiments. Therefore, we view randomized experiments with intermediate variables as semi-observational studies. In parallel with the analysis of observational studies, we provide a theoretical foundation for conducting objective causal inference with an intermediate variable under the principal stratification framework, with principal strata defined as the joint potential values of the intermediate variable. Our strategy constructs weighted samples based on principal scores, defined as the conditional probabilities of the latent principal strata given covariates, without access to any outcome data. This principal stratification analysis yields robust causal inference without relying on any model assumptions on the outcome distributions. We also propose approaches to conducting sensitivity analysis for violations of the ignorability and monotonicity assumptions, the very crucial but untestable identification assumptions in our theory. When the assumptions required by the classical instrumental variable analysis cannot be justified by background knowledge or cannot be made because of scientific questions of interest, our strategy serves as a useful alternative tool to deal with intermediate variables. We illustrate our methodologies by using two real data examples, and find scientifically meaningful conclusions.

Biography:

Dr. Ding obtained his Ph.D. from the Department of Statistics at Harvard University. His primary research area is in Applied & Theoretical Statistics with sub-focuses in the following: (I) Applied Statistics, (II) Bioinformatics/Biostatistics, (III) Bayesian Statistics, (IV) Statistics in Social Sciences. Additional research interests include: (I) Casual inference in experiments and observational studies, (II) Missing data.