samsi
NSF·Duke·NCSU·UNC

Causal Inference Program Opening Workshop
December 9-11, 2019

SPEAKER TITLES/ABSTRACTS

Edward Kennedy
Carnegie Mellon University

"Optimal Causal Inference with High-Dimensional Discrete Data"

In this work we consider estimation of causal effects in a setting where the covariates required for confounding adjustment are discrete but arbitrarily high-dimensional. Contrary to classical results that consider dimension d fixed and sample size n tending to infinity, we give non-asymptotic risk bounds for arbitrary n and d, and consider non-classical regimes where d can grow with or exceed n. We study the plug-in estimator of the average treatment effect, showing that this estimator is only consistent in the regime where the dimension grows slower than the sample size. We also consider a propensity-weighted effect estimator, which has some surprising ability to adapt to both exact and approximate effect homogeneity. Then we go on to characterize the minimax lower bounds for the average effect under heterogeneity and homogeneity, indicating the fundamental limits of causal inference in high dimensions as well as when an estimator can no longer be improved upon.