EMORY COLLEGE OF ARTS AND SCIENCES | Department of Quantitative Theory and Methods

Policy Learning in Adaptive Experiments

Learning optimal treatment assignment policies from historical data enables personalization gains in a wide variety of applications including healthcare, digital recommendations, and online education. The growing policy learning literature focuses on settings where the data collection rule stays fixed throughout the experiment. However, adaptive data collection is becoming more common in practice, from two primary sources: 1) data collected from adaptive experiments that are designed to improve inferential efficiency; 2) data collected from production systems that are adaptively evolving an operational policy to improve performance over time (e.g. contextual bandits). Adaptivity yet complicates the ex post optimal policy identification, since samples are dependent and each treatment may not receive enough observations for each type of individual. In this talk, we aim to address the challenges of learning the optimal policy with adaptively collected data and make initial research inquiries into this problem. We propose an algorithm based on generalized augmented inverse propensity weighted (AIPW) estimators, which non-uniformly reweight the elements of a standard AIPW estimator to control worst-case estimation variance. We establish a finite-sample regret upper bound for our algorithm and complement it with a regret lower bound that quantifies the fundamental difficulty of policy learning with adaptive data. When equipped with the best weighting scheme, our algorithm achieves a minimax rate optimal regret guarantee even with diminishing exploration. Finally, we demonstrate our algorithm's effectiveness on both synthetic data and public benchmark datasets. This is joint work with Zhimei Ren, Susan Athey, and Zhengyuan Zhou.

Ruohan Zhan

Stanford Graduate School of Business
Stanford University

11.2.2022 | 1pm-2:15pm
PAIS 561

Fall 2022 QTM Speaker Series