The Hong Kong University of Science and Technology

Seminar on Business Data Science

Department of ISOM

Powerful randomization tests for subgroup analysis

By

Professor Zijun GAO, USC Marshall Business School

Date: 26 Jun 2025 (Thursday)

Time: 11:00am – 12:00nn

Venue: Room 4047 (LSK Business Building)

Abstract

Randomization tests are widely used to generate valid p-values for testing sharp null hypotheses in finite-population causal inference. This article extends their application to subgroup analysis. We show that directly testing subgroup null hypotheses may lack power due to small subgroup sizes. Incorporating an estimator of the conditional average treatment effect (CATE) can substantially improve power but requires splitting the treatment variables between estimation and testing to preserve finite-sample validity.To this end, we propose BaR-learner, a Bayesian extension of the popular method R-learner for CATE estimation. BaR-learner imputes the treatment variables reserved for randomization tests, reducing information loss due to sample-splitting. Furthermore,we show that the treatment variables most informative for training BaR-learner are different from those most valuable for increasing test power. Motivated by this insight, we introduce AdaSplit, a sample-splitting procedure that adaptively allocates units between estimation and testing. Simulation studies demonstrate that our method yields more powerful randomization tests than baselines that omit CATE estimation or rely on random sample-splitting. We also apply our method to a blood pressure intervention trial, identifying patient subgroups with significant treatment effects.

Bio

Zijun Gao is a tenure-track assistant professor in the Department of Data Sciences and Operations at USC Marshall Business School. She received her Ph.D. in Statistics from Stanford University in 2022 supervised by Professor Trevor Hastie. She served as a research associate in the Statistical Lab at the University of Cambridge from 2022 to 2023 hosted by Professor Qingyuan Zhao. Her research focuses on the estimation and inference problems in causal inference with heterogeneity, with side interests in distribution learning, selective inference, and model evaluation. She also works on real-world data motivated topics, with a specific emphasis on the applications in adaptive clinical trial and personalized medication.

All interested are welcome!

For details, please contact ISOM Department.