Hyunseung Kang
Department of Statistics
University of Wisconsin-Madison

A Robust, Differentially Private Randomized Experiment for Evaluating Online Educational Programs with Sensitive Student Data

Wednesday, September 21, 2022
11:45 AM

110 Frelinghuysen Road, Hill Center, Room 552

Zoom Meeting: Meeting ID: 99075124232
Password: 952486

https://rutgers.zoom.us/j/99075124232?pwd=UDdPVjRncXZFcXpvbFE0OWJyMVdSUT09

Light refreshments will be served

Abstract: Randomized control trials (RCTs) have been the gold standard to evaluate the effectiveness of a program, policy, or treatment on an outcome of interest. However, many RCTs assume that study participants are willing to share their (potentially sensitive) data, specifically their response to treatment. This assumption, while trivial at first, is becoming difficult to satisfy in the modern era, especially in online settings where there are more regulations to protect individuals' data. The paper presents a new, simple experimental design that is differentially private, one of the strongest notions of data privacy. Also, using works on noncompliance in experimental psychology, we show that our design is robust against "adversarial" participants who may distrust investigators with their personal data and provide contaminated responses to intentionally bias the results of the experiment. Under our new design, we propose unbiased and asymptotically Normal estimators for the average treatment effect. We also present a doubly robust, covariate-adjusted estimator that uses pre-treatment covariates (if available) to improve efficiency. We conclude by using the proposed experimental design to evaluate the effectiveness of online statistics courses at the University of Wisconsin-Madison during the Spring 2021 semester, where many classes were online due to COVID-19.

Bio: Hyunseung (pronounced Hun-Sung) is an Assistant Professor in the Department of Statistics at the University of Wisconsin-Madison. His research is focused on developing theory and methods to analyze causal relationships by using instrumental variables, econometrics, semi/nonparametric methods, network analysis, and machine learning. He is interested in applications to genetics, epidemiology, infectious diseases, health policy, education, and applied microeconomics.