University of California, Irvine
Statistics Seminar

Covariate Adjustment in Multi-armed,

Possibly Factorial Experiments

Peng Ding
Associate Professor, Statistics
UC Berkeley

4 p.m., Thursday, Nov. 3, 2022
6011 Donald Bren Hall

Randomized experiments are the gold standard for causal inference, and justify simple comparisons across treatment groups. Regression adjustment provides a convenient way to incorporate covariate information for additional efficiency. This article provides a unified account of its utility for improving estimation efficiency in multi-armed experiments. We start with the commonly used additive and fully interacted models for regression adjustment, and clarify the trade-offs between the resulting ordinary least-squares (OLS) estimators for estimating average treatment effects in terms of finite-sample performance and asymptotic efficiency. We then move on to regression adjustment based on restricted least squares (RLS), and establish for the first time its properties for inferring average treatment effects from the design-based perspective. The resulting inference has multiple guarantees. First, it is asymptotically efficient when the restriction is correctly specified. Second, it remains consistent as long as the restriction on the coefficients of the treatment indicators, if any, is correctly specified and separate from that on the coefficients of the treatment-covariates interactions. Third, it can have better finite-sample performance than its unrestricted counterpart even if the restriction is moderately misspecified. It is thus our recommendation for covariate adjustment in multi-armed experiments when the OLS fit of the fully interacted regression risks large finite-sample variability in case of many covariates, many treatments, yet a moderate sample size. In addition, the proposed theory of RLS also provides a powerful tool for studying OLS-based inference from general regression specifications. As an illustration, we demonstrate its unique value for studying OLS-based regression adjustment in factorial experiments via both theory and simulation.