An Overview of Clustered Observational Studies and Multilevel Matching

September 30, 2019

This abstract describes the proposed lead presentation in the proposed symposium "Causal Inference in Clustered Observational Studies using Optimal Multilevel Matching." The lead presentation will provide background for a set of studies that were conducted using this framework. The rest of this abstract describes the material covered in the lead presentation.

Many interventions in education occur in settings where treatments are applied to groups. For example, a reading intervention may be applied to all the students in some schools and withheld from the students in other schools. When such treatments are non-randomly allocated, outcomes across the treated and control groups may differ due to the treatment or due to baseline differences in the groups. We classify such study designs as the observational analogue to clustered randomized trials or clustered observational studies (COSs). Clustered observational studies (COSs) are a critical analytic tool in educational effectiveness research. In recent work, we developed a design framework for COSs. The framework we develop is built on the counterfactual model for causal inference and promotes the concept of designing COSs that emulate the targeted randomized trial that would have been conducted, were it feasible.

Target trial emulation calls for using design principles from randomized trials for the analysis of observational data. Under the target trial approach, the investigator explicitly ties the analysis of the observational study to the trial it is emulating, and causal estimands of interest are derived from the hypothetical target trial. The purpose of target trial emulation is to improve the quality of observational studies through the application of trial design principles. For example, in an experimental study, the sample and study design are clearly delineated to enable randomization. In contrast, observational studies, particularly those conducted after the fact, often necessitate some level of investigation to inform decisions about sample construction and study design. Using the concept of target trial emulation, we develop a framework thats provides a set of guidelines for both the development and critique of clustered observational studies (COSs). We also emphasize the key role of the understanding the assignment mechanism. We review how treatment assignment at the group level is advantageous in observational studies. In the context of a CRT, group-level treatment assignment typically reduces power relative to unit-level assignment but confers an advantage in terms of bias. In a COS, group-level treatment assignment also can reduce bias.

In a COS, researchers use statistical adjustment to make treated and control groups similar in terms of observed characteristics. Multilevel regression models are commonly used for this purpose. Recent work in statistics has developed matching methods designed for contexts where treatments are clustered. This form of matching conducts separate matches at the student and school level. It seeks to not only balance covariates but also to ensure schools of equal size are matched to avoid sample size loss. Covariates that are deemed to be of key importance can be prioritized for better balance. This form of matching method can also accommodate trimming of the data to ensure overlap in the treated and control distributions. We review how regression models can be profitably combined with matching and note best practice for estimates of statistical uncertainty.

We evaluate the performance of multilevel matching methods in two ways. First, we use these matching methods to recover treatment effect estimates from three clustered randomized trials. Second, we conduct a simulation study. We find that the best analytic approach to statistical adjustment in clustered observational studies is first to apply matching and then to use regression models with the matched data for further bias reduction. Finally, we review how sensitivity analysis can determine whether conclusions are sensitive to bias from potential unobserved confounders.