Randomized Trials of Educational Technology: What (and How) Can Log Data Tell Us About Causal Effects?

As Educational Technology (EdTech) products proliferate, so do questions about their effectiveness. Researchers have used randomized controlled trials (RCTs) to estimate the overall effects of a number of prominent EdTech products, but have not fully exploited their potential. EdTech RCTs often produce rich and granular student log data—electronic records of what EdTech users actually did during the RCT. Log data could help address more nuanced causal questions about when, how, and for whom the products work best. But causal modeling of log data can be tricky; when only the treatment group has access to the EdTech product, log data do not exist for control students. Further, log data can be messy and complex. This talk will discuss a “principal stratification” approach to modeling log data from an RCT of the Cognitive Tutor Algebra I program. Principal stratification is a causal framework in which treatment effects are modeled as a function of “potential usage,” or how students would use the technology if assigned to treatment. The talk will describe conventional Bayesian principal stratification models alongside a novel model that includes a latent variable component to address the complexity of log data.

Thursday, December 5, 2019
10:00AM- 11:00AM
Stratton Hall, room 309