Monday 22 July 2024

11:00-12:30 Invited Session 1 (Main Room)

Causal Inference and Machine Learning (Chair: Els Goetghebeur)

Efficient and robust machine-learning-based approaches for simple, cluster randomized, and sequential multiple assignment randomized trial analysis: Illustrations from HIV trials in East Africa.

Maya Petersen (University of California, Berkeley)

Covariate adjustment in randomized trials can improve precision and power while maintaining rigorous inference. Targeted maximum likelihood estimation (TMLE) is a general semi-parametric efficient approach for optimizing precision gains from covariate adjustment in trials by combining 1) flexible, machine learning-based outcome regression; and, 2) additional covariate adjustment in the treatment mechanism. This general method can be used across a range of trial designs, including those employing cluster, restricted, sequential, and adaptive randomization schemes, and can be extended to account for loss-to-follow-up, missing data, competing risks, and intercurrent events.

Using examples from completed HIV trials in Kenya and Uganda that used TMLE as the primary pre-specified analysis, we illustrate several design-specific TMLE estimators. We first discuss the SEARCH Study (NCT01864603), a trial in Uganda and Kenya of universal HIV testing and person-centered multi-disease treatment to reduce HIV incidence and improve community health. Using SEARCH, we illustrate implementation of TMLE for the analysis for cluster randomized data, with and without pair-matching. In particular, we illustrate how “adaptive-pre-specification”, an empirical efficiency maximization strategy, can be employed to achieve meaningful precision gains even in the context of cluster-randomized trials with few independent units.

We then use the ADAPT-R Study (NCT02338739), a trial of sequential behavioral intervention strategies to improve retention in HIV care among people with HIV in Kenya, applying TMLE to sequential multiple assignment randomized trials (SMARTs). Adaptive intervention strategies (or dynamic regimes), in particular those that modify treatment based on a participant's own response, are a core component of precision health approaches. SMART designs are growing in popularity and are specifically designed to facilitate the evaluation of sequential adaptive strategies. We present a robust and efficient approach using TMLE for estimating and contrasting expected outcomes under the dynamic regimes embedded in a SMART, together with generating simultaneous confidence intervals for the resulting estimates. We contrast this method with two alternatives (G-computation and inverse probability weighting estimators).

Selected References:

Balzer LB, van der Laan M, Ayieko J, Kamya M, Chamie G, Schwab J, Havlir DV, Petersen ML. Two-Stage TMLE to reduce bias and improve efficiency in cluster randomized trials. Biostatistics. 2023 Apr 14;24(2):502-517. doi: 10.1093/biostatistics/kxab043. PMID: 34939083; PMCID: PMC10102904.

Montoya LM, Kosorok MR, Geng EH, Schwab J, Odeny TA, Petersen ML. Efficient and robust approaches for analysis of sequential multiple assignment randomized trials: Illustration using the ADAPT-R trial. Biometrics. 2023 Sep;79(3):2577-2591. doi: 10.1111/biom.13808. Epub 2022 Dec 22. PMID: 36493463; PMCID: PMC10424093.