The Forum for Collaborative Research presents:

Causal Inference and Adaptive Trial Design Workshop

Wednesday, May 10, 2017

8:30am – 5:00pm

Hosted at:

1616 Rhode Island Avenue, NW, Washington, DC 20036

(Center for Strategic and International Studies building)

Instructors:

Mark van der Laan, PhD
Professor, Biostatistics and Statistics
University of California, Berkeley School of Public Health
laan@stat.berkeley.edu

Maya Petersen, MD, PhD
Associate Professor, Biostatistics and Epidemiology
University of California, Berkeley School of Public Health
mayaliv@berkeley.edu

SUMMARY

This course introduces a general framework for causal inference.

- Directed acyclic graphs and non-parametric structural equation models (NPSEM) are used to define the causal model.
- Target causal parameters are defined using counterfactuals and marginal structural models.
- Causal graphs are used to illustrate identification assumptions.
- G-computation and targeted maximum likelihood estimators for the resulting statistical estimation problem are introduced
- Super Learning, an ensemble machine learning approach, is introduced to estimate nuisance parameters.
- Targeted group sequential adaptive designs that learn the optimal individualized treatment rule, using targeted maximum likelihood estimation, while providing formal and assumption free inference.

We will introduce concepts focusing on causal effects of single time point treatments and then extend these concepts to longitudinal causal models, identifiability and of the joint effects of multiple interventions in the presence of time dependent confounding. We will also present estimators and designs for learning the optimal individualized treatment rule.

COURSE READINGS

- M.J. van der Laan and S. Rose. Targeted Learning: Causal Inference for Observational and Experimental Data. Springer, Berlin Heidelberg New York, 2011. Chapters 1-5. http://www.springerlink.com/content/978-1-4419-9782-1.

- Petersen ML, van Der Laan, MJ. Causal Models and Learning from Data: Integrating Causal Modeling and Statistical Estimation. *Epidemiology* 2014; 25(3): 418-26.

SOFTWARE RESOURCES

- SuperLearner R package: https://cran.r-project.org/web/packages/SuperLearner/index.html Implements the super learner ensemble machine learning prediction method and contains a library of prediction algorithms to be used in the super learner.

- LTMLE R package: https://cran.r-project.org/web/packages/ltmle/index.html Targeted Maximum Likelihood Estimation (TMLE) of treatment/censoring specific mean outcome or marginal structural model for point-treatment and longitudinal data.

- R labs with worked examples: http://www.ucbbiostat.com

SESSION I: CAUSAL INFERENCE

Part I: From causal questions to the statistical estimation problem: Introduction using single time point interventions.

- A General Roadmap for Tackling Causal Questions
- Introduction to Structural Causal Models (SCM)/Causal Graphs
- Defining Target Causal Quantities using Counterfactuals
- Identifying Causal Effects as Parameters of the Observed Data Distribution

Part II: Statistical estimation and interpretation: Introduction using single time point interventions.

- A Simple Substitution Estimator Based on the G-computation Formula
- Introduction to Super Learning
- Introduction to Targeted Maximum Likelihood Estimation

Part III: From causal questions to the statistical estimation problem: Extension to multiple/longitudinal interventions.

- Longitudinal Casual Models/Causal Graphs
- Counterfactual Casual Parameters Summarizing the Joint Effects of Multiple Interventions
- Identification in the Longitudinal Setting: The Challenge of Time-Dependent Confounding

Part IV: Statistical estimation and interpretation: Extension to multiple/longitudinal interventions.

- A Simple Substitution Estimator Based on the Longitudinal G-computation Formula
- An Alternative Sequential Regression Approach
- Longitudinal Targeted Maximum Likelihood Estimation

SESSION II: ADAPTIVE DESIGN

Part V: Targeted group sequential adaptive designs for learning the optimal individualized treatment rule

- Description of targeted group sequential adaptive designs that learns the optimal individualized rule from the past groups, and assign its best estimate to the next group.
- The targeted maximum likelihood estimator of the counterfactual mean outcome under optimal treatment rule (relative to control).

Part VI: Utilizing Seamless Adaptive Designs and Considering Multiplicity Adjustment for NASH Clinical Trials

Presenter:
Yeh-Fong Chen, PhD
Mathematical Statistician
Center for Drug Evaluation and Research
U.S. Food and Drug Administration

Presentation Abstract:

The prevalence of non-alcoholic fatty liver disease including non-alcoholic steatohepatitis (NASH) is increasing worldwide. NASH is the second most common indication for liver transplantation and is expected to be the leading indication by 2020. In light of increasing prevalence and burden of disease, it is imperative to develop therapeutic strategies for patients with NASH.

Although it is feasible to conduct clinical trials with liver transplantation or death as the clinical endpoint, it may take 10 to 20 years for NASH patients to develop cirrhosis or other liver-related morbidity and mortality. In addition, even with accelerated regulatory approval based on surrogate endpoints, studies need to be at least one or two years long to demonstrate clinically meaningful effectiveness in study drugs.

Seamless adaptive clinical designs that “roll over” patients from earlier phases can be adopted to shorten the duration of new drug development programs for NASH. To address the complex multiplicity problems resulting from multiple doses and multiple looks, these types of designs require flexible and powerful (or efficient) multiple testing procedures. Once the drug is partially approved (with accelerated approval), additional challenges include how to ensure later studies, including phase 4 studies, can be properly conducted and which controls groups are feasible.

In this presentation, we will focus on statistical considerations and the application of seamless design in NASH clinical trials. Thorough exploration, including different methods for efficiently controlling study-wise type I error rate as well as the potential use of historical controls will be illustrated and discussed.