

Statistical Methods for Causal Inference in Observational and Randomized Studies

3-day Executive-Level Course

September 26-28, 2011
Washington, DC

Forum for
Collaborative HIV Research



School of
Public Health

UNIVERSITY OF CALIFORNIA, BERKELEY

Target Audience

The course is designed for senior statisticians from regulatory agencies, industry and academic settings.

This course concerns statistical methods for causal inference using observational and experimental longitudinal data. The course will focus on the application of methodological advances in statistical and causal research to improve the design and interpretation of safety analyses. These analyses will become increasingly important in the post-marketing safety environment for new drugs.

About the Course

A review and scientific critique of current estimation methods will be provided, including an introduction to semi-parametric targeted learning. Structural causal models (causal graphs) and working marginal structural models will be introduced as tools for translating a policy question and background knowledge into a target statistical quantity and model. The course will emphasize understanding and responding to the challenges posed by safety data in randomized controlled trials as well as observational cohorts, including informative drop out/censoring, missing data, time-dependent confounding, non-compliance, and high dimensional covariates. Examples from the fields of HIV and cardiovascular disease, together with other fields, will be used to illustrate the methods and to provide practical experience with analytic design and accurate interpretation of results.

About the Instructors

Mark J. van der Laan, PhD is a Hsu/Peace Professor of Biostatistics at the University of California, Berkeley School of Public Health. He is the recipient of the 2005 COPSS Presidents' and Snedecor Awards, as well as the 2004

Spiegelman Award, and is a Founding Editor for the International Journal of Biostatistics. Mark has co-authored various books, and his most recent book is *Targeted Learning: Causal Inference for Observational and Experimental Data*, van der Laan, Rose (2011), Springer: New York.

Maya Petersen, MD, PhD is Assistant Professor of Biostatistics and Epidemiology at the University of California, Berkeley School of Public Health. She received her MD from the University of California, San Francisco and her PhD in Biostatistics from UC Berkeley. Her doctoral work was funded by a fellowship from the Howard Hughes Medical Institute and was honored by the Evelyn Fix prize. Her research focuses on the development and application of novel causal inference methods to problems in health. Maya has a strong interest in and has published on the interface between biostatistics, epidemiology, and clinical medicine.

Sherri Rose, PhD is an NSF Mathematical Sciences Postdoctoral Research Fellow in the Department of Biostatistics at Johns Hopkins Bloomberg School of Public Health. She received her PhD in Biostatistics from the University of California, Berkeley School of Public Health and is co-author of *Targeted Learning: Causal Inference for Observational and Experimental Data*. Sherri's research interests include methodology for causal inference and prediction in rare diseases.

The course will employ a two-tiered pricing system with early and late registration fees. Cost of applicable software and textbook is included in the course fee.

Registration and Fees

Early Registration (beginning June 1, 2011):

\$500.00 (USD) Academic/Government
\$2000.00 (USD) Industry

Late Registration (after August 1, 2011):

\$750.00 (USD) Academic/Government
\$2500.00 (USD) Industry

*Upon completion of
the course students
will be able to...*

- 1)** Translate a scientific question and background knowledge into a formal causal model and target causal parameter using the causal graphs and counterfactual (potential outcome) frameworks, including specifying a working marginal structural model.
- 2)** Understand the assumptions needed to make causal inferences from observational data, and how causal graphs can facilitate the application of background knowledge to evaluate these assumptions.
- 3)** Translate the causal model and target parameter into a parameter of the observed data distribution (estimand) and statistical model.
- 4)** Understand the shortcomings of traditional parametric regression-based techniques for the estimation of causal effects.
- 5)** Be familiar with and able to apply basic data adaptive (machine learning) approaches as a tool for confronting the curse of dimensionality in the context of causal effect estimation.
- 6)** Understand the properties of and be able to apply three distinct estimators of the causal effect of a point treatment (an exposure or intervention that occurs at a single time point): Maximum Likelihood, Inverse Probability of Treatment Weighted, and Targeted Maximum Likelihood.

- 7) Be able to implement estimators of a point treatment estimator in R using available packages (tmle and Superlearner).
- 8) Understand the challenges raised by estimation of the causal effect of longitudinal exposures, and the particular vulnerabilities of traditional regression-based adjustment in this setting.
- 9) Understand how to estimate the causal effect of a longitudinal exposure in the context of right censoring and time-dependent confounding.

Course Materials

The course fee will cover the necessary course materials for attendees. Book: *Targeted Learning: Causal Inference for Observational and Experimental Data*, van der Laan, Rose (2011), Springer: New York. Hand-outs: Copies of PowerPoint presentations and computer lab material including R-code

8:30 – 10:00	Lecture part 1
10:00 – 10:30	Coffee break
10:30 – 12:00	Lecture part 2
12:00 – 1:30	Lunch
1:30 – 3:30	Lab
3:30 – 4:00	Coffee break
4:00 – 5:00	Review and recap

Preliminary Daily Schedule

For more information and to register

Registration will be available starting June 1, 2011.

Please email Rob Besaw (rbesaw@hivforum.org) with any questions.