SPPH681A: CAUSAL INFERENCE IN PUBLIC HEALTH SCIENCES

INSTRUCTORS

Boris Sobolev, office hours: Tuesdays, 9 to 11am at VGH, or by appointment boris.sobolev@ubc.ca

PLACE AND TIME

- Term 1, Weekly 3-hour class on Mondays, 1 to 4 pm, TBD

COURSE DESCRIPTION

Drawing causal conclusions from observational data is a common task in the public health sciences. The goal of this 3-credit course is to develop knowledge, skills, and competency in causal inference methodology. The course offers in-depth coverage of methods developed over the past three decades. We will look at probabilistic causality, causal diagrams, counterfactuals, mediation analysis, and methods for evaluating treatment effects.

This course is divided into 12 modules. In each module you will find reading material, tutorials, and videos. By the design, they complement each other. The readings introduce the topic, the videos discuss ideas, and the tutorials help you put them into practice. I will also share a wealth of material that you can use both during the course and long after you've completed it. The tutorials will use Stata; alternatively, you can apply your R programming knowledge from previous courses. Additional tutorials will teach you how to create and analyze cause and effect diagrams.

RATIONALE FOR TAKING THIS COURSE

This offering fulfills a recognized need for a graduate course that connects concepts and practical skills in drawing causal conclusions. Whether you are studying epidemiology, health services research, occupational or environmental health, you will benefit from the lectures, videos, tutorials and readings in 681A! We begin by outlining the framework of causal reasoning. You will get a full exposition to causality and causal effects. Then you will learn about conditions to classify observed associations as causal. After that, the course invites you to master causality by practical application to real-world data.

After taking the course, you will be able to create causal diagrams for your thesis projects, refine your research questions, find variables to adjust, detail your analysis plan, and attempt to estimate causal effects using data from your projects. The course will be useful for those who analyze data from patient registries, routine care records, hospital discharges, or study cohorts.

LEARNING OBJECTIVES

On completion of the course, students will be able to

(1) explain the causal inference framework;
(2) develop directed acyclic graphs and identify a minimal sufficient adjustment set;
(3) estimate total, direct and indirect treatment effects from observational data; and
(4) express causal queries through counterfactuals quantities.

PREREQUISITES

This 600-level course builds on the knowledge and skills acquired in SPPH500, SPPH503, and SPPH548. Working knowledge of statistical software is recommended.

COURSE STRUCTURE

The course combines lectures, discussions and tutorials, in-class and homework assignments, and relies on course reading materials and short video lectures.

EVALUATION OF LEARNING

Evaluation will be conducted using a combination of marks for 12 in-class assessments (maximum 24% of the total mark), six homework assignments (36%), and final homework assignment (40%).

In-class assignments

Each class involves assessment of the learning progress in class. The assessment tools will include a combination of multiple choice tests and short writing exercises covering the content from lectures, in-class discussions and required readings. The assessments provide an opportunity for students to evaluate their own progress through the course and help strengthen their understanding of core concepts and methods. Each test is worth 2 points (maximum 24 points for 12 tests).

Homework assignments

Bi-weekly homework assignments will involve various aspects of data analysis and preparing short reports. The quality of each report will be judged by clarity of presentation, suitability of methods, and interpretation of results. Each assignment is worth 6 points (maximum 36 points for 6 assignments).

A final homework assignment will involve data analysis and paper preparation. The paper should contain the following sections: Introduction, Methods, Results, Discussion, and References. Results should be presented in an organized fashion, such as in table or graphical formats. Computer outputs should be edited to eliminate irrelevant or redundant material. The quality of the report will be judged by the suitability of methods, correct computing, interpretation of results, and clarity of presentation (maximum 40 points).

PREPARATION

The student is expected to be prepared for topics discussed in class. Sufficient time should be allocated for reading of required and assigned texts, watching video lectures.

COURSE MATERIALS

- Video lectures
- Lecture slides and handouts of lab tutorials
- Selected articles and book chapters:

1. Susser M. Glossary: Causality in Public Health Science. Journal of Epidemiology and Community Health 2001;55(6):376-378.
2. Cattaneo. Journal of Econometrics 2010, 155:147
3. Almond D. The Quarterly Journal of Economics 2005:1031-83
4. Berkson J. Limitations of the applications of fourfold table analysis to hospital data. Int. J Epidemiol 2014: 1-5
5. Maldonado G, Greenland S. Estimating causal effects. Int J Epidemiol 2002;31(2):422-429.
6. Austin PC. Absolute risk reductions, relative risks, relative risk reductions, and numbers needed to treat can be obtained from a logistic regression model. J Clin Epidemiol 2010;63(1):2-6.
7. Greenland S. Causation and causal inference. In: Lorvic M, editor. International encyclopedia of statistical science. Berlin Heidelberg: Springer; 2011. p. 216-221.

8. Vittinghoff E, Glidden DV, McCulloch CE, Shiboski SC. Chapter 9: Strengthening Causal Inference. Regression Methods in Biostatistics. 2nd ed.: Springer; 2012. p. 331-394.
9. Elwert F. Graphical Causal Models. In: Morgan SL (ed.), Handbook of Causal Analysis for Social Research: Springer Netherlands; 2013. p. 245-273.
10. Greenland S, Pearl J. Causal diagrams. In: Lorvic M, editor. International encyclopedia of statistical science. Berlin Heidelberg: Springer; 2011. p. 208-216.
11. Greenland S, Pearl J, Robins JM. Causal Diagrams for Epidemiologic Research. Epidemiology 1999;10(1):37-48.
12. Hernán MA, Robins JM. Chapter 1: A Definition of Causal Effect. Causal Inference; 2013. p. 3-12. www.hsph.harvard.edu/faculty/miguel-hernan/causal-inference-book
13. Hernán MA, Robins JM. Chapter 3: Observational Studies. Causal Inference; 2013. p. 25-39. www.hsph.harvard.edu/faculty/miguel-hernan/causal-inference-book
14. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. J Epidemiol Community Health 2006;60(7):578-586.
15. Holland PW. Statistics and Causal Inference. Journal of the American Statistical Association 1986;81(396):945-960.
16. Kaufman J, MacLehose R, Kaufman S. A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. Epidemiologic Perspectives & Innovations 2004;1(1):4-4.
17. Sobolev B, Kuramoto L. Section 2.1 and 2.2 - Waiting-time data used in this book. Analysis of waiting-time data in health services research. New York: Springer New York; 2008. p. 17-29.
18. Suzuki E, Yamamoto E, Tsuda T. On the relations between excess fraction, attributable fraction, and etiologic fraction. Am J Epidemiol 2012;175(6):567-575.
19. Textor J, Hardt J, Knüppel S. DAGitty: A Graphical Tool for Analyzing Causal Diagrams. Epidemiology 2011;22(5):745-745.
20. Textor J, Liskiewicz M. Adjustment Criteria in Causal Diagrams: An Algorithmic Perspective. 2012.
21. Vittinghoff E, Glidden DV, McCulloch CE, Shiboski SC. Chapter 10: Predictor Selection. Regression Methods in Biostatistics: Springer; 2012. p. 395-429.
22. Westreich D, Greenland S. The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients. Am J Epidemiol. 2013;177(4):292–298
23. Miettinen O. Stratification by a multivariate confounder score. Am J Epidemiol 1976; 106: 609-620