Causal Inference with Observational Data

Two-day workshop, University of Bergen Dr. Thomas Däubler

This two-day workshop provides an introduction to statistical methods for causal inference with observational data. We begin with more general considerations, familiarizing ourselves with the Potential Outcomes Framework and Directed Acyclic Graphs. The specific methods to be covered then are matching, instrumental variables, difference-in-differences, regression discontinuity and synthetic control. Finally, we briefly discuss some recent applications of machine learning approaches to problems of causal inference.

Upon successful completion of this course the participants should be able to:

- Use Directed Acyclic Graphs and the Potential Outcomes Framework to derive strategies for causal identification
- Choose among various methods of causal inference, apply them to real data and interpret the findings
- Explain and (where possible) test the underlying assumptions of these techniques

Prerequisites

The course assumes familiarity with the fundamentals of inferential statistics (e.g., hypothesis testing, interpretation of p-values) and basic applied regression modeling (i.e. using the linear model with continuous and categorical predictors).

The course is taught using the R statistical package. People not familiar with R could team up with a coursemate for the exercises. Many (but not necessarily all) the exercises can in principle also be done with other statistical software, such as Stata.

Schedule

Day 1 morning (2.5 hours)	Day 2 morning (2.5 hours)
• Introduction to causality	• Difference-in-differences
Potential Outcomes Framework	• Panel data approaches
Directed Acyclic Graphs	
Day 1 afternoon (2.5 hours)	Day 2 afternoon (2.5 hours)
Matching	Regression discontinuity
Instrumental variables	Synthetic control
	• Machine learning in causal inference