GESIS Summer School in Survey Methodology 2022

Syllabus for course:
“Using Directed Acyclic Graphs for Causal & Statistical Inference”

Lecturer: Julian Schuessler
E-mail: julians@ps.au.dk
Homepage: www.julianschuessler.net

Date: 03-05 August 2022
Time: 09:00-11:00 | 13:00-15:00
Venue: Online via Zoom

About the Lecturer:
Julian Schuessler is a post-doc at the Centre for the Experimental-Philosophical Study of Discrimination, Aarhus University, Denmark. He defended his PhD on the use of causal graphs for causal and statistical inference problems in the social sciences at the University of Konstanz in 2020. In 2019, he received the American Statistical Association’s “Causality in Statistics Education Award”. Beyond causal inference and statistical methods, his research interests are in the empirical study of discrimination, political behaviour, and public opinion, using observational and experimental data.

Selected Publications:

Course Description:
This course uses causal graphs (or “directed acyclic graphs”, DAGs) as a remarkably simple, yet general and powerful framework to describe and discuss a large set of problems that empirical social scientists need to tackle. Is my question of interest descriptive or causal? How can I communicate my assumptions effectively to others, and can I test them? How can I tell correlation from causation? How do I choose control variables for my regression models? After discussing how DAGs can be used to answer these foundational questions, the course also covers basics of causal interaction and effect heterogeneity, causal mediation, nonresponse/selection bias (and adjustments for it) and, if time permits, instrumental variables and panel data analysis from a DAG perspective.

Keywords:
causal graphs; causal inference; mediation; instrumental variables; nonresponse

Course Prerequisites:
Participants should be willing to learn and use formal reasoning and must have at least Bachelor-level knowledge of statistics. Basic knowledge of R is helpful.

Target Group:
Participants will find the course useful if:
- They are interested in causal questions and want to understand the assumptions associated with regression control, mediation analysis and instrumental variables better
They are interested in non-causal question, want to use data suffering from nonresponse, and want to understand how to use causal assumptions in this case

Course and Learning Objectives:
By the end of the course participants will:
- Know how to use causal graphs to visualize causal assumptions, define quantities of interest, and to determine testability of assumptions via d-separation
- Know how to graphically determine identification of causal and descriptive quantities like average causal effects, causal interaction, effect heterogeneity, natural direct and indirect effects, and population distributions from data with nonresponse
- Know under what graphical assumptions instrumental variable and panel data analysis typically operate
- Will have some basic knowledge about how all of this relates to implementation in standard statistical software

Organizational Structure of the Course:
This short course throughout will change between short lecture-style inputs and individual or small-group hands-on exercises supervised by the lecturer and a teaching assistant (4hrs/day). Participants are encouraged to bring their own research ideas to develop them further using the material from the class. The lecturer will also be available for individual consultation in the afternoons (1hr/day).

Software and Hardware Requirements:
We will use briefly discuss some elements of the R packages “dagitty”, “sensemakr”, “mediation”, “AER”, “estimatr”. Most of the course will not depend on using R. For those who have never used R, installation instructions and a short introductory video are linked to below under “Preparatory Reading”.

Day-to-day Schedule and Literature:

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic(s)</th>
</tr>
</thead>
</table>
| 1   | Introductions, interests and preferences of participants, self-assessments  
Statistical control and Simpson’s paradox  
Descriptive and causal questions  
Graph basics and d-separation  
Interventions and causality  
Back-door criterion and unobserved confounding  
Post-treatment bias |
|     | Compulsory reading (have to be read before class):  
  Only pp. 245—252 needed!  
Play around with DAGitty: [http://dagitty.net/dags.html](http://dagitty.net/dags.html)  
Suggested reading (suggested, yet do not have to be read before class):  
| 2   | Sensitivity analysis for unobserved confounding  
Causal interaction and effect heterogeneity  
Causal mediation: Controlled and natural effects, post-treatment confounding, sensitivity analyses |
### Compulsory reading:
None.

### Suggested reading:

### 3

**Non-response/missing data and its consequences**

Adjustment for non-response from a graphical perspective

Instrumental variables in linear & nonparametric models

Choosing control variables in IV models

Panel data and fixed effects

Wrap-up

### Compulsory reading:

### Suggested reading:

### Preparatory Reading:
- For those who have never used R before (18 mins): [https://www.youtube.com/watch?v=DuQ5Q0a6Sw](https://www.youtube.com/watch?v=DuQ5Q0a6Sw)

### Additional Recommended Literature:

**Textbooks:** *(Aronow/Miller and Imai do not discuss graphs)*


Other books and papers:
- Pearl, Judea. "Indirect Confounding and Causal Calculus (On three papers by Cox and Wermuth)." (2015).