EDUC 3418: Causal Moderation and Mediation Analysis
School of Education
University of Pittsburgh
Spring, 2024

Thursday 12:00 pm – 2:40 pm
5404 Wesley W Posvar Hall

Instructor: Xu Qin (xuqin@pitt.edu)
Office Hour: By appointment

Teaching Assistant: Tetsuya Yamada (tey15@pitt.edu)
Office Hour: By appointment

Course Overview:

This course is designed for graduate students who are interested in applying advanced quantitative methods to their research. Research questions regarding how, for whom, and where a treatment achieves its effect on an outcome are often key to the advancement of scientific knowledge and have become increasingly valued in various fields. Such questions can be answered by causal moderation and mediation analysis, which assesses the heterogeneity of the treatment effect across individual and contextual characteristics and uncovers the mediation mechanism underlying the treatment effect.

This course introduces the theoretical concepts of moderated effects and mediated effects under the potential outcomes framework, cutting-edge methodological approaches, and how to implement the methods with user-friendly `R` packages. Specifically, we will begin with a general introduction to mediation and moderation analyses, followed by an introduction to causal inference under the potential outcomes framework, which lays the conceptual foundation for the whole course. Next, we will introduce definition, identification, estimation, and inference of causal moderated effects and causal mediated effects separately. The causal analyses rely on the assumption of no unmeasured confounding, which is usually violated in real applications. We will introduce sensitivity analysis methods for assessing the potential influence of unmeasured confounding. We will also introduce several extended topics, including (1) multilevel causal mediation analysis, which accounts for multilevel data structures, (2) causal moderated mediation analysis, which assesses the heterogeneity of causal mediation mechanisms across subpopulations, and (3) power analysis, which determines the sample size required to detect the causal effects with sufficient power. We will provide tutorials on how to use easy-to-implement `R` packages to conduct the causal analyses throughout the course.

The course blends statistical theories and data applications -- avoiding the extremes of presenting unneeded theory in isolation, or of giving application tools without the foundation needed for practical understanding. Examples will be drawn from education, psychology, and other social sciences disciplines.

Prerequisites:

Prerequisites of the course are EFOP 2018 (Introduction to Statistics) and EFOP 2410 (Applied Regression Analysis), or their equivalents. It would be better if you have taken EFOP 3472 (Causal Inference in Educational Research)/ BIOST 2068 (Introduction to Causal Inference, or their equivalents. Students are expected to have some basic knowledge about R.

Required Books (All requested to be available online):

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.
https://pitt.primo.exlibrisgroup.com/permalink/01PITT_INST/e8h8hp/alma99102031525806236

Hong, G. (2015). *Causality in a social world: Moderation, mediation and spill-over*. John Wiley & Sons.
https://pitt.primo.exlibrisgroup.com/permalink/01PITT_INST/1j5jsjp/alma9998501659906236

VanderWeele, T. (2015). *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press.
https://pitt.primo.exlibrisgroup.com/permalink/01PITT_INST/1j5jsjp/alma9998523037106236

Recommended Reading*:*

Pearl, J., & Mackenzie, D. (2018). *The book of why: the new science of cause and effect*. Basic books.

A Turing Award-winning computer scientist and statistician shows how understanding causality has revolutionized science and will revolutionize artificial intelligence. The book explores the subject of causality and causal inference from statistical and philosophical points of view for a general audience.

Leite, W. (2016). *Practical propensity score methods using R*. Sage Publications.

The book is available on reserve at Hillman. You will be able to check it out for 2 hours at a time and make whichever material scans are necessary.

Software:

We will use R, a free software environment for statistical computing and graphics. It can be downloaded here: https://cran.r-project.org/. Students are expected to have some basic knowledge about R. If you would like to learn more about R, you may refer to the following books:

Introductory Statistics with R
http://www.academia.dk/BiologiskAntropologi/Epidemiologi/PDF/Introductory_Statistics_with_R__2nd_ed.pdf

R for Beginners
http://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf

R Graphics
http://www.e-reading.biz/bookreader.php/137370/C486x_C06.pdf

A book list for learning R programming:
https://www.r-bloggers.com/learning-r-programming-by-reading-books-a-book-list-rstats/

Course Format:

Before class: **Within the week before the lecture date listed on the last page, slides and other materials for the lecture will be posted on Canvas**. To get prepared for the lecture, please read the associated book chapters listed at the end of the slides before the lecture date. Going over slides beforehand is not mandatory.

In class: **On the lecture date, we meet in the classroom**. **If you are unable to attend due to a valid emergency, please let the instructor know beforehand.** In each lecture, the instructor will introduce the concepts and show how to use statistical methods to address substantive research questions by conducting analyses and interpreting the results. You are strongly encouraged to ask questions in class.

Collaboration and Study Groups:

Collaborative learning is central to this course and serves as a way to foster deep learning. 4 - 5 students from different disciplines are assigned to one study group. Discussions about the course materials within study groups are strongly encouraged. **In class, you will be given a chance to discuss in groups. After class, you may choose different ways to discuss with your group members.** Your group can discuss by posting your questions and comments on the Canvas group discussion board in your own group's view (click your group number on the right of your Canvas homepage). Your group can also meet in person or online. **Should there be any questions that your group cannot address, please post them on the discussion board that is visible to the whole class (see main menu on the left of your homepage). We will respond periodically.**

Homework Assignments:

There will be three assignments based on data provided by the instructor. Assignments will provide hands-on experience with real data applications. **Each assignment assignment will be posted on Canvas three weeks before the due date and must be submitted by 11:59 pm on the due date listed on the last page.**

- Please submit your assignment following the instruction here: https://community.canvaslms.com/t5/Student-Guide/How-do-I-upload-a-file-as-an-assignment-submission-in-Canvas/ta-p/274
- Collaborations within assigned study groups are encouraged. Please understand that each student must turn in individual homework assignments, not group work. Your text should reflect your own understanding of the material. **Students who submit group homework assignments will be given zeros.** To properly acknowledge the contribution of your collaborators, please indicate on the cover page of each assignment the names of the people with whom you worked.

- Because solutions to assignments will be posted on Canvas right after the due time, **late assignment will not be accepted**. In a valid emergency, appropriate accommodations will be made. It is best, if possible, to contact the **instructor** prior to the due date.
- If you have questions/concerns about your grades, please directly email the instructor and TA rather than leaving a comment on the Canvas grading page, because we will not get automatic notifications.

Final Presentation:
Presentation of re-analyses of existing data. Students may analyze their own data or may choose from among several data sets made available by the instructor. The presentation should contain the following elements.

1. A clearly stated research question along with a brief explanation of why this question is important for theory or policy. A detailed literature review is not required.

2. A description of the sample and data, including a clear definition of the variables.

3. A brief justification for choosing a causal inference method covered in the course; and an introduction to the basic idea of the method.

4. A clarification of the causal estimands that would supply an answer to your research question(s), the identification of the causal estimands, and the underlying assumptions.

5. A description of the analysis procedure.

6. A concise description of your analysis results. If possible, it is recommended to visualize the results.

7. A discussion of your findings, including the limitations of the analysis.

On 03/29/2024 and 04/05/2024, each student has 15 minutes to meet with the instructor and TA online to discuss the final presentation proposal. A sign-up sheet will be provided as the dates approach. For your better preparation of the meeting, a plan form will be provided. Please submit it before the individual meeting. We will go over your plan and discuss your questions during the meeting.

You will present your work during the regular class time on 04/18/2022 or 04/25/2022. No matter on which day you choose to present, you are required to attend both sessions because learning about others' work may inspire your own work. Each of you will have around 15 minutes to present and receive feedback. A sign-up sheet will be provided as the dates approach. Writing papers are not required. If you would like to write a paper, the instructor is happy to read your draft and provide comments.

Evaluation:

You will be evaluated on the basis of your assignments (75%) and final presentation (25%). Letter grades will be based on actual points earned as follows:

Point	Letter		Point	Letter
Above 93	A		77-80	C+
90-93	A-		74-77	C
87-90	B+		70-74	C-
84-87	B		67-70	D+
80-84	B-		64-67	D
			60-64	D-
			Below 60	F

Academic Integrity:

Please make sure you read the university guidelines on Academic integrity (http://www.pitt.edu/~provost/ai1.html). Attention to this policy is particularly important in a course like EDUC 3418, in which collaboration with other students is encouraged. If, for instance, you work closely with other students during the planning, execution, or interpretation of your data analyses – a process that we encourage and fully support – you should make sure that the other students' contributions are recognized explicitly in your written account. If you have any questions about what constitutes appropriate collaboration, or how to define what constitutes your own work, please meet with me.

Special Accommodation:

If you have a disability for which you are or may be requesting an accommodation, you are encouraged to contact both your instructor and Disability Resources and Services, 140 William Pitt Union, (412) 648-7890/(412)383-7355 (TTY), as early as possible in the term. DRS will verify your disability and determine reasonable accommodations for this course.

Tentative Course Outline

Week	Date	Topic	Assignment
1	Jan 11	Introduction to moderation and mediation analysis	Assignment 1 Posted
2	Jan 18	Causal inference I	
3	Jan 25	Causal inference II	
4	Feb 1	Causal inference III	
5	Feb 8	Causal moderation analysis I	
6	Feb 15	Causal moderation analysis II	Assignment 1 Due Assignment 2 Posted
7	Feb 22	Causal mediation analysis I	
8	Feb 29	Causal mediation analysis II	
9	Mar 7	Causal moderation and mediation analysis in multisite randomized trials	
10	Mar 14	No class	Assignment 2 Due Assignment 3 Posted
11	Mar 21	Causal moderated mediation analysis I	
12	Mar 28	Causal moderated mediation analysis II	
13	Apr 4	Power analysis	
14	Apr 11	Review	Assignment 3 Due
15	Apr 18	Final presentation	
16	Apr 25	Final presentation	