Quantitative Methods II: Causal Inference and Mechanism

Instructer: Jiawei Fu

August 2024

Course Description

This is the second course in the Quantitative Methods sequence, focusing on design-based causal inference and mechanisms. We will explore both foundational concepts and the latest developments in the field. The course begins with causal inference using experimental designs, which provide ideal settings for understanding the fundamentals of causal inference compared to observational studies.

In the second part, we will apply these principles to non-experimental environments, where treatment assignment is not controlled by researchers, making causal inference significantly more challenging.

In the final sections, we will delve into causal mediation and other advanced topics. Throughout the course, we will rigorously derive and prove fundamental results. This rigorous approach is essential for students to fully comprehend the methods, critically engage with journal articles, and conduct future research in methodology.

Logistics

Lectures: Tuesday 10am-12pm
Recitation: Thursday 4pm-6pm
Office Hour: Friday 9pm-11pm at Room 312
Email: jf3739@nyu.edu

Requirements and Grading

Reading: The material is hard to understand if it is your first time to see. It is highly recommended to read and digest those required readings after the class.

Problem sets: The most effective way to encourage learning and assimilation of new knowledge is through hands-on assignments. Consequently, there will be one problem set each week. These assignments are designed to reinforce understanding rather than to challenge excessively. After engaging with the course materials, everyone should be able to complete them comfortably. Typically, the problem sets include two types of questions. (1) Algebra: These require the straightforward application of formulas discussed in class. (2) Proof: You will be asked to prove arguments similar to those we have covered in class, meaning that as long as you understand the lecture notes, you should be able to complete these tasks with ease.(50%)

Midterm and Final Exams: There will be two closed-book exams, one in the middle and the other at the end of the semester. (Each 25%)

1 Textbooks

1. Ding, P., 2024. A first course in causal inference. CRC Press.
2. Gerber, A.S. and Green, D.P., 2012. Field experiments: Design, analysis, and interpretation.
3. Imbens, G.W. and Rubin, D.B., 2015. Causal inference in statistics, social, and biomedical sciences. Cambridge university press.
4. Angrist, J.D. and Pischke, J.S., 2009. Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Week 1: Introduction

- Overview of the course
- Potential Outcome Framework
- Completely Randomization Experiment

Recommendations:

Ding, P., 2024. A first course in causal inference. CRC Press. Chapter 2

Imbens, G.W. and Rubin, D.B., 2015. Causal inference in statistics, social, and biomedical sciences. Cambridge university press. Chapter 1.

Part I: Causal Inference with Experiments

Week 2: Randomization Inference

- Finite-population Design-based Inference
- Foundations of Randomization Test
- Randomization Inference under Sharp Null
- Randomization Inference under Weak Null

Recommendations:

Ritzwoller, D.M., Romano, J.P. and Shaikh, A.M., 2024. Randomization Inference: Theory and Applications. arXiv preprint arXiv:2406.09521.

Romano, J.P. and Lehmann, E.L., 2005. Testing statistical hypotheses. Chapter 15

Imbens, G.W. and Rubin, D.B., 2015. Causal inference in statistics, social, and biomedical sciences. Cambridge university press. Chapter 5.

Ding, P., 2024. A first course in causal inference. CRC Press. Chapter 3.

Wu, J. and Ding, P., 2021. Randomization tests for weak null hypotheses in randomized experiments. Journal of the American Statistical Association, 116(536), pp.1898-1913.

Week 3: Neyman Inference

- Neyman Inference
- Super-population Framework

Recommendations:

Imbens, G.W. and Rubin, D.B., 2015. Causal inference in statistics, social, and biomedical sciences. Cambridge university press. Chapter 6.

Ding, P., 2024. A first course in causal inference. CRC Press. Chapter 4.

Week 4: Covariates Adjustment, Rerandomization, and Regression

- Semi-parametric regression and Lin's regression
- Covariate selection

Recommendations:

Li, X. and Ding, P., 2020. Rerandomization and regression adjustment. Journal of the Royal Statistical Society Series B: Statistical Methodology, 82(1), pp.241-268.

Tsiatis, A.A., Davidian, M., Zhang, M. and Lu, X., 2008. Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: a principled yet flexible approach. Statistics in medicine, 27(23), pp.4658-4677.

Week 5: Complex Design I: Stratified and Clustered Experiment

- Stratification and Post-Stratification
- Clustered design

Recommendations:

Middleton, J.A. and Aronow, P.M., 2015. Unbiased estimation of the average treatment effect in cluster-randomized experiments. Statistics, Politics and Policy, 6(1-2), pp.39-75.

Ding, P., 2024. A first course in causal inference. CRC Press. Chapter 5.

Abadie, A., Athey, S., Imbens, G.W. and Wooldridge, J.M., 2023. When should you adjust standard errors for clustering?. The Quarterly Journal of Economics, 138(1), pp.1-35.

Bugni, F., Canay, I., Shaikh, A. and Tabord-Meehan, M., 2022. Inference for cluster randomized experiments with non-ignorable cluster sizes. arXiv preprint ArXiv:2204.08356.

Week 6: Complex Design II: Matched-Pairs Design and Factorial Design

- Matched-Pairs Design
- Factorial Design

Recommendations:

Imai, K., 2008. Variance identification and efficiency analysis in randomized experiments under the matched-pair design. Statistics in medicine, 27(24), pp.4857-4873.

Bai, Y., 2022. Optimality of matched-pair designs in randomized controlled trials. American Economic Review, 112(12), pp.3911-3940.

Bai, Y., Romano, J.P. and Shaikh, A.M., 2022. Inference in experiments with matched pairs. Journal of the American Statistical Association, 117(540), pp.1726-1737.

Kahan, B.C., 2013. Bias in randomised factorial trials. Statistics in medicine, 32(26), pp.4540-4549.

Egami, N. and Imai, K., 2019. Causal interaction in factorial experiments: Application to conjoint analysis. Journal of the American Statistical Association.

Week 7: Complex Design III: Adaptive Design

- Multi-armed bandit problem
- Thompson sampling
- Adaptive Design

Recommendations:

Kasy, M. and Sautmann, A., 2021. Adaptive treatment assignment in experiments for policy choice. Econometrica, 89(1), pp.113-132.

Offer-Westort, M., Coppock, A. and Green, D.P., 2021. Adaptive experimental design: Prospects and applications in political science. American Journal of Political Science, 65(4), pp.826-844.

2 Part II: Causal Inference with Observational Studies

Week 8: Matching and Weighting

- Propensity Score
- Weighting
- Matching

Recommendations:

King, G. and Nielsen, R., 2019. Why propensity scores should not be used for matching. Political analysis, 27(4), pp.435-454.

King, G., Lucas, C. and Nielsen, R.A., 2017. The balance-sample size frontier in matching methods for causal inference. American journal of political science, 61(2), pp.473-489.

Week 9: Instrumental Variables

- IV under linear regression
- IV under Potential outcome framework; LATE
- Characterizing compliers

Recommendations:

Abadie, A., Gu, J. and Shen, S., 2024. Instrumental variable estimation with first-stage heterogeneity. Journal of Econometrics, 240(2), p.105425.

Angrist, J.D., Imbens, G.W. and Rubin, D.B., 1996. Identification of causal effects using instrumental variables. Journal of the American statistical Association, 91(434), pp.444-455.

Staiger, D.O. and Stock, J.H., 1994. Instrumental variables regression with weak instruments.

Aronow, P.M. and Carnegie, A., 2013. Beyond LATE: Estimation of the average treatment effect with an instrumental variable. Political Analysis, 21(4), pp.492-506.

Week 10: Difference in Differences

- Identification Assumptions
- DID and TWFE
- Synthetic Control

Recommendations:

Roth, J., Sant'Anna, P.H., Bilinski, A. and Poe, J., 2023. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. Journal of Econometrics, 235(2), pp.2218-2244.

Strezhnev, A., 2023. Decomposing triple-differences regression under staggered adoption. arXiv preprint arXiv:2307.02735.

Xu, Y., Zhao, A., and Peng.,D. 2024 Factorial Difference-in-Differences

Abadie, A., 2021. Using synthetic controls: Feasibility, data requirements, and methodological aspects. Journal of economic literature, 59(2), pp.391-425.

Week 11: Regression Discontinuity

- Continuity Based
- Unconfoundedness Based
- Synthetic Control
- Kink Design

Recommendations:

Cattaneo, M.D., Idrobo, N. and Titiunik, R., 2019. A practical introduction to regression discontinuity designs: Foundations. Cambridge University Press.

Card, D., Lee, D.S., Pei, Z. and Weber, A., 2015. Inference on causal effects in a generalized regression kink design. Econometrica, 83(6), pp.2453-2483.

3 Part III: Advanced Topics

Week 12: Causal Mediation Analysis

- Explicit Mediation Analysis

- Implicit Mediation Analysis (HTE, IOT)

Recommendations:

VanderWeele, T.J., 2015. Explanation in Causal Inference: Methods for Mediation and Interaction. Oxford University Press. Chapter 2

Imai, K., Keele, L. and Tingley, D., 2010. A general approach to causal mediation analysis. Psychological methods, 15(4), p.309.

Blackwell, M., Ma, R. and Opacic, A., 2024. Assumption Smuggling in Intermediate Outcome Tests of Causal Mechanisms. arXiv preprint arXiv:2407.07072.

Fu, J. and Slough, T., 2024. Heterogeneous Treatment Effects and Causal Mechanisms. arXiv preprint arXiv:2404.01566.

Week 13: Causal inference with Interference

- Partial Inference
- Two-stage Randomization
- Group Formation

Recommendations:

Hudgens, M.G. and Halloran, M.E., 2008. Toward causal inference with interference. Journal of the American Statistical Association, 103(482), pp.832-842.

Aronow, PM and Cyrus Samii. (2017). "Estimating Average Causal Effects Under General Interference." Annals of Applied Statistics 11(4):1912-1947

Week 14: Machine Learning with Causal Inference

- Double Machine Learning
- Covariate Selection and Adjustment
- Heterogeneous treatment effects

Recommendations:

Wager, S., Du, W., Taylor, J. and Tibshirani, R.J., 2016. High-dimensional regression adjustments in randomized experiments. Proceedings of the National Academy of Sciences, 113(45), pp.12673-12678.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. and Robins, J., 2018. Double/debiased machine learning for treatment and structural parameters.

Wager, S. and Athey, S., 2018. Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association, 113(523), pp.1228-1242.

List, J.A., Muir, I. and Sun, G., 2024. Using machine learning for efficient flexible regression adjustment in economic experiments. Econometric Reviews, pp.1-39.

Chernozhukov, V., Hansen, C., Kallus, N., Spindler, M. and Syrgkanis, V., 2024. Applied causal inference powered by ML and AI. arXiv preprint arXiv:2403.02467.