Causal Inference Using Machine Learning

Master in Economics
Universidad Nacional de Tucumán

Spring 2024

Instructor: Andres Mena (asmena@face.unt.edu.ar)

Course link: zoom.us/andresmena

Course Description

This graduate-level course introduces the intersection of applied econometrics and machine learning techniques. It aims to equip students with essential tools for conducting causal inference in empirical research, public policy analysis, and business case studies. The course covers topics such as randomized experiments, regression discontinuity, instrumental variables, differences-in-differences, and synthetic control methods. A particular focus is on how machine learning and AI techniques can enhance these methods in high-dimensional settings by using statistical learners to estimate and infer low-dimensional causal effects.

Prerequisites

Students should have a background in econometrics, statistical inference, and machine learning. A graduate course in at least two of these three topics is expected. Students should also be familiar with programming in R or Python.

Textbooks

- **CIML**: Chernozhukov, Victor, et al. *Applied Causal Inference Powered by ML and AI.*
- **CIS**: Imbens, Guido, and Donald Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* Cambridge University Press.
- **MHE**: Angrist, Joshua, and Jorn-Steffen Pischke. *Mostly Harmless Econometrics.* Princeton University Press.
- **CIMix**: Cunningham, Scott. *Causal Inference: The Mixtape.* Yale University Press.

Course Schedule

Lecture 1 (10/23) - *Introduction to Causal Inference*: Potential Outcomes, Fundamental Problem of Causal Inference, Assignment Mechanisms, Selection Bias, Average Treatment Effect.

- **MHE**, Chapter 1, Chapter 2 pp 11-22
- **CIS**, Chapter 1

Additional Readings:

- **CIMix**, Chapter 1
- **CIS**, Chapter 2
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396), 945-960.
- Rambachan, A. (2018). Harvard Economics Math Camp 2018: Econometrics, Probability Review. *Lecture Notes*.
- Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2), 3-32.

Lecture 2 (10/30) - *Randomized Control Trials*: Bernoulli Trials, Completely Randomized Experiments, Inference with Two Sample Means, Fisher's Exact P-values.

- **CIS**, Chapters 5-6

Additional Readings:

- Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2020). Sampling-Based Versus Design-Based Uncertainty in Regression Analysis. *Econometrica*.

Lecture 3 (11/06) - *Randomized Control Trials II*: Covariates in RCT, Heterogeneous Treatment Effects.

- **CIML**, Chapter 2, pp 43-57
- Chernozhukov, V., Demirer, M., Duflo, E., & Fernández-Val, I. (2018). Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India (Working Paper No. 24678). National Bureau of Economic Research.

- **CIS**, Chapters 23, pp. 513-529

Additional Readings:

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

Lecture 4 (11/13) - *Selection on Observables* : Conditional Expectation Function, OLS, Causal Regression, Inference in OLS, Semiparametric estimators, Doubly robust methods.

- **MHE**, Chapter 3 pp 27-64
- **CIML**, Chapter 1 pp 13-26

Additional Readings:

- **CIMix**, Chapter 2, pp 39-95

- Bruce E. Hansen. Econometrics. Princeton University Press. pp 14-57

Lecture 5 (11/20) - *High-Dimensional Regression*: LASSO, Elastic-Net, Inference under High Dimensionality

- **CIML**, Chapter 3
- Hastie, Tibshirani, & Friedman. The Elements of Statistical Learning. Springer. pp 61-73

Additional Readings:

- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.

- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), 301-320.

- Belloni, A., Chernozhukov, V., & Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2), 29-50.

- Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2), 608-650.

11/25 - *Midterm Presentations*

Lecture 6 (11/27) - *Double Machine Learning*: Frisch-Waugh-Lovell theorem, Double Lasso, Partially Linear regression, Neyman Orthogonality, Influence Functions, Inference using DML.

- **CIML**, Chapter 4 pp 105-114
- **CIML**, Chapter 10 pp 251-265
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2017). Double machine learning for treatment and causal parameters. *Econometrica*, 85(1), 53-80.
- Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685-725.

Additional Readings:

- Hines, O., Dukes, O., Diaz-Ordaz, K., & Vansteelandt, S. (2022). Demystifying Statistical Learning Based on Efficient Influence Functions. *The American Statistician*, 76(3), 292–304.

- Fisher, A., & Kennedy, E. H. (2020). Visually Communicating and Teaching Intuition for Influence Functions. *The American Statistician*, 75(2), 162–172.

- Chernozhukov, V., Escanciano, J. C., Ichimura, H., Newey, W. K., & Robins, J. M. (2022). Locally robust semiparametric estimation. *Econometrica*, 90(4), 1501-1535.

Lecture 7 (12/04) - *Regression Discontinuity*: RDD Framework, Estimation in RDD, RDD with Covariates, High-Dimensional Covariates in RDD, Empirical Example of RDD

- **CIML**, Chapter 17
- Abdulkadiroglu, A., Joshua Angrist, and Parag Pathak (2014). "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," Econometrica, 81(1): 137-196.

Additional Readings:

- **CIML**, Chapter 6

- **CIMix**, pp 241-282

- Cattaneo, Matias, Nicholas Idrobo, and Rocio Titiunik (2019). "A Practical Introduction to Regression Discontinuity Designs," Cambridge University Press.

Lecture 8 (12/10) - *Instrumental Variables I*: Selection on Unobservables, Homogeneous vs Heterogeneous TE, Identification LATE, 2SLS, Inference 2SLS.

- **MHE**, Chapter 4, pp 113-146
- Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2), 467-475.

Additional Readings:

- Angrist, J. D. (1990). Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records. The American Economic Review, 80(3), 313–336.

- Angrist & Imbens (1995): ”Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity” JASA

- Angrist, Imbens, & Rubin (1996): ”Identification of Causal Effects Using Instrumental Variables,” JASA

- Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. Journal of Economic Perspectives, 15(4), 69-85.

Lecture 9 (12/10) - *Instrumental Variables II*: Use of Covariates in IV Models, Instrument Validity, Weak Instruments and Many Instruments, Semiparametric IV, RCT with imperfect compliance.

- Blandhol, C., Bonney, J., Mogstad, M., & Torgovitsky, A. (2022). When is TSLS actually LATE? National Bureau of Economic Research Working Paper No. 29709.
- Słoczyński, T. (2024). When should we (not) interpret linear IV estimands as LATE? Working paper.
- Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. Journal of Econometrics, 113(2), 231-263.

Additional Readings:

- Angrist, J., & Kolesár, M. (2021). One instrument to rule them all: The bias and coverage of just-ID IV. Review of Economics and Statistics, 103(3), 476-490.

- Kitagawa, T. (2015). A test for instrument validity. Econometrica, 83(5), 2043-2063.

- Hong, H., & Nekipelov, D. (2010). Semiparametric efficiency in nonlinear LATE models. *Quantitative Economics*, 1(2), 279–304.

- Balke & Pearl (1997) Bounds on Treatment Effects from Studies with Imperfect Compliance, JASA

Lecture 10 (12/12) - *Instrumental Variables ML:* Partially Linear IV Models, DML Inference on LATE, DML Inference in Interactive IV Regression Models, DML Inference with Weak Instruments, Robust DML Inference under Weak Identification.

- **CIML**, Chapter 4. p. 354-369

Additional Readings:

- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2017). Double machine learning for treatment and causal parameters. *Econometrica*, 85(1), 53-80.

Lecture 11 (12/12) - *Differences-in-Differences:* 2x2 DiD design, Parallel trends assumption, Anticipatory effects, Inference on DID.

- **CIMix**, pp 411-433
- Garthwaite, Craig, Tal Gross, and Matthew J. Notowidigdo (2014). "Public Health Insurance, Labor Supply, and Employment Lock," Quarterly Journal of Economics, 129(2): 653-696

Additional Readings:

- Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? The Quarterly journal of economics, 119(1):249–275.

- Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. American Economic Review: Papers and Proceedings, 93(2), 133-138.

- Rambachan, Ashesh, and Jonathan Roth (2022). "An Honest Approach to Parallel Trends," Forth-coming, Review of Economic Studies

- Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. American Economic Review: Insights, 4(3):305–22.

- Malani, A., & Reif, J. (2015). Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform. Journal of Public Economics, 124, 1-17.

Lecture 12 (12/16) - *Differences-in-Differences under Staggered Adoption*: TWFE estimation, Negative Weights, Forbidden Comparison, Diagnostics and Solutions.

- Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation.

- Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200-230.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254-277.

Additional Readings:

- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175-199.

- De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review, 110(9):2964–2996.

- Jakiela, P. (2021). Simple diagnostics for two-way fixed effects. *arXiv:2103.13229.*

- Gardner, J., Thakral, N., Tô, L. T., & Yap, L. (2024). Two-Stage Differences in Differences.

Lecture 13 (12/16) - *Differences-in-Differences and Machine Learning* Semiparametric DiD. DML for 2x2 Design. DML with staggered adoption. DML for Fuzzy DiD.

- **CIML**, Chapter 16
- Chang, N.-C. (2019). Double/debiased machine learning for difference-in-differences models. Department of Economics, University of California, Los Angeles.
- Callaway, B., Drukker, D., Liu, D., & Sant'Anna, P. H. C. (2023). Double/Debiased Machine-learning estimator for Difference-in-Difference with Multiple Periods.

Additional Readings:

- Abadie, A. (2005). Semiparametric difference-in-differences estimators. Review of Economic Studies, 72(1), 1–19. 00321

- Mena, Andres (2024) Double Debiased Machine learning for Fuzzy Difference-in-Differences. Working Paper

Lecture 14 (12/18) - *Synthetic Control Methods*: Context requirements, Convex hull condition, Sparcity, Inference on SCM.

- **CIMix**, pp 511-540
- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *American Economic Review*, 105(3), 391-425.

Additional Readings:

- Doudchenko, N., & Imbens, G. W. (2016). Balancing, regression, difference-in-differences and synthetic control methods: A synthesis. *NBER Working Paper.*

- Card, D. (1990). The impact of the Mariel boatlift on the Miami labor market. *ILR Review*, 43(2), 245–257.

- Firpo, S., & Possebom, V. (2018). Synthetic control method: Inference, sensitivity analysis, and confidence sets. *Journal of Causal Inference*, 6(2), 20160026.

- Chen, J. (2023). Synthetic control as online linear regression. *Econometrica*, 91(2):465–491.

Lecture 15 (12/18) - *Synthetic Control Methods and Extensions*: Synthetic DID. Constrained SCM. Challenges on Inference. Bayesian SCM.

- Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 112(12), 4088-4118.
- Chernozhukov, V., Wuthrich, K., and Zhu, Y. (2024). A t-test for synthetic controls. *arXiv:1812.10820v7.*
- Martinez, I., & Vives-i-Bastida, J. (2024). Bayesian and Frequentist Inference for Synthetic Controls. *arXiv:2206.01779v3*

Additional Readings:

- Li, K. T. (2020). Statistical inference for average treatment effects estimated by synthetic control methods. *Journal of the American Statistical Association*, 115(532):2068–2083.

- Andrews, D. W. K. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica*, 68(2):399–405.

- Brand, J. & Mena, A. (2024) Demand estimation using constrained Synthetic Control. Working Paper.

12/19 - *Final Presentations*

Grading

The final grade will be based on a replication exercise (20%), a midterm presentation (10%), a final presentation (20%), and a research proposal (50%).