Introduction to the Design and Analysis of Randomized Experiments

Instructor: Jake Bowers
jwbowers@illinois.edu
http://jakebowers.org/

Escuela de Invierno de Métodos 2024

Overview

This course introduces the basic statistical concepts that guide the design and analysis of randomized experiments. The act of randomizing the assignment of an intervention offers special benefits to researchers interested in making counterfactual causal inferences, and the course begins by engaging with questions about why randomize (or why not randomize) and how randomized assignment is not the same as random sampling. It then introduces randomization based statistical inference; an approach to calculating standard errors for average treatment effect estimates and to calculating p-values for tests of hypotheses about causal effects. Discusses power analysis, and ends by engaging with some of the trickier issues in experimental design and analysis: what to do when experimental units drop out of the study or otherwise do not provide valid outcomes? how might we think about making causal inferences when the active treatment, the treatment of theoretical interest, cannot be directly randomized? what to do when we cannot randomly assign treatment directly to individual units but only to groups of them? Throughout we will be using the R programming language to demonstrate statistical concepts and also as tools for designing and analyzing randomized experiments. Most of the examples in the class will come from field experiments but given the focus on fundamentals the class will be useful for those working with randomized survey experiments and randomized lab experiments as well.

I will be open to adding and subtracting topics depending on student interest.

Goals and Expectations

This class aims to help you get started with the design and analysis of randomized experiments using randomization as the basis for statistical inference.

The point of the course is to position you to do the future learning that is at the core of your work as a researcher. That is, for most of your life you will not have classes or even textbooks available to help you learn how to work with data and statistics. Rather, you will have to learn on your own, with the help of the internet, friends, and friendly AI. So, this course aims to help you learn how to learn even as it helps you learn how to reason, decide and evaluate.

The **specific goals** of the course are that students:

- Explain in their own words key concepts in statistics like "causal inference","statistical inference", "hypothesis testing", "point estimation", "p-value", "confidence interval", "random assignment", and describe how such concepts fit together in applied research.
- Have practiced coding randomization, power calculations, and statistical analysis of experimental results
- Be familiar with standards and practices for many aspects of reproducible research
- Have practiced developing an experimental research design
- Have practiced created a pre-analysis plan for an experimental design

Books I will recommend that you read chapters from the following books:

> Alan S Gerber and Donald P Green (2012b). *Field experiments: Design, analysis, and interpretation*. WW Norton
>
> Jake Bowers, Maarten Voors, and Nahomi Ichino (Mar. 2021). *The Theory and Practice of Field Experiments: An Introduction from the EGAP Learning Days*. Berkeley, CA.: Evidence in Governance and Politics. URL: https://egap.github.io/theory_and_practice_of_field_experiments/ En español
>
> Graeme Blair, Alexander Coppock, and Macartan Humphreys (2023). *Research design in the social sciences: declaration, diagnosis, and redesign*. Princeton University Press. URL: https://book.declaredesign.org/

I also recommend some chapters from Glennerster and Takavarasha 2013, which presents material similar to Gerber and Green but less technically. Reading this volume alongside Gerber and Green might bring together intuition and statistics for some students.

Other readings will be assigned and distributed electronically.

Assumptions I assume some previous engagement with high school mathematics, probability and statistical computing in the R statistical programming language. If you haven't had experience with R but you love learning computing languages then you can still get a lot out of this course — you will learn a lot about R as kind of laboratory for learning about statistical theory and evaluating and analyzing experimental designs.

If you're new to R and/or statistics, I suggest using books like these or other resources for learning about R and basic statistics online.

> John Fox (2016). *Applied Regression Analysis and Generalized Linear Models*. 3rd. Los Angeles, CA: SAGE Publications
>
> Hadley Wickham and Garrett Grolemund (2017). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. First. Sebastopol, CA: O'Reilly Media

Credits Those who want to take this course for university credit should plan to create a proposal for an experimental design and/or a pre-analysis plan. I recommend meeting with me at least once during the week or even before the week of the course to discuss this. I will be presenting a template for research design during the class sessions.

Expectations **Ask questions when you don't understand things; chances are you're not alone.**

Schedule

Note: This schedule is preliminary and subject to change. If you miss a class make sure you contact me or one of your colleagues to find out about changes in the lesson plans or assignments.

1— Causal inference and the potential outcomes framework

What is causality, the potential outcomes framework, and how do experiments help causal identification?

Useful readings

- Alan S Gerber and Donald P Green (2012b). *Field experiments: Design, analysis, and interpretation*. WW Norton, Ch 1–3
- **Application:** A lab experiment: Lene Aarøe and Michael Bang Petersen (2013). "Hunger Games: Fluctuations in Blood Glucose Levels Influence Support for Social Welfare". In: *Psychological Science* 24.12, pp. 2550–2556 which uses a randomized experiment to follow the nicely designed observational study in Michael Bang Petersen et al. (2014). "Social Welfare and the Psychology of Food Sharing: Short-Term Hunger Increases Support for Social Welfare". In: *Political Psychology* 35.6, pp. 757–773
- **Application:** A field experiment: Alan S. Gerber, Donald P. Green, and Christopher W. Larimer (2008). "Social pressure and voter turnout: Evidence from a large-scale field experiment". In: *American Political Science Review* 102.1, pp. 33–48
- **Application:** An audit-experiment: Marianne Bertrand and Sendhil Mullainathan (2004). "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination". In: *American Economic Review* 94.4, pp. 991–1013
- **Recommended:** Jake Bowers, Maarten Voors, and Nahomi Ichino (Mar. 2021). *The Theory and Practice of Field Experiments: An Introduction from the EGAP Learning Days*. Berkeley, CA.: Evidence in Governance and Politics. URL: https://egap.github.io/theory_and_practice_of_field_experiments/, Module 1–3
- **Recommended:** Rachel Glennerster and Kudzai Takavarasha (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton: Princeton University Press, chs. 1–3
- **Recommended:** Abhijit Banerjee and Esther Duflo (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. New York: Public Affairs, ch. 1
- **Recommended:** Paul R. Rosenbaum (2020). *Design of Observational Studies*. 2nd. Springer Series in Statistics. Springer Sections 2.1–2.4
- **Recommended:** Paul R Rosenbaum (2017). *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press, Chap 2–3
- **Recommended:** Paul R Rosenbaum (2002). *Observational Studies*. Second Edition. New York, NY: Springer, Chap 2–2.4 explains and formalizes Fisher's randomization inference.
- **Recommended:** R.A. Fisher (1935). *The design of experiments. 1935*. Edinburgh: Oliver and Boyd, Chap 2 explains *the* invention of random-assignment based randomization inference in about 15 pages.
- **Recommended:** J. Neyman (1990). "On the application of probability theory to agricultural experiments. Essay on principles. Section 9 (1923)". In: *Statistical Science* 5. reprint. Transl. by Dabrowska and Speed, pp. 463–480; Donald B. Rubin (1990). "[On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9.] Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies". In: *Statistical Science* 5.4, pp. 472–480 *the* invention of random-sampling based randomization inference (estimators of the average causal effect and their standard errors).
- **Recommended:** Jake Bowers and Thomas Leavitt (2020). "Causality and Design-Based Inference". In: *The SAGE Handbook of Research Methods in Political Science and International Relations*. Ed. by Luigi Curini and Robert Franzese. Vol. 2. Thousand Oaks, CA: SAGE Publications. Chap. 41, pp. 769–804 Provides some statistical theory that connects Fisher's test and Neyman's estimator for causal inference.

2— Random assignment and identification under randomization

How does randomization help us learn about causal effects? Randomization strategies (simple, complete, blocked, clustered, etc.), random sampling versus random assignment.

Useful readings

- Alan S. Gerber and Donald P. Green (2012c). *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton & Co, ch. 2, remaining sections
- Dean Karlan and Jacob Appel (2016). *Failing in the Field: What We Can Learn When Field Research Goes Wrong*. Princeton: Princeton University Press ch. 2 (skim).
- **Recommended:** Rachel Glennerster and Kudzai Takavarasha (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton: Princeton University Press ch. 4.
- **Recommended:** Leonard Wantchekon (2003). "Clientelism and voting behavior: Evidence from a field experiment in Benin". In: *World Politics* 55.3, pp. 399–422
- **Recommended:** Paul Collier and Pedro C Vicente (2014). "Votes and violence: Evidence from a field experiment in Nigeria". In: *The Economic Journal* 124.574, F327–F355
- **Recommended:** Jake Bowers, Maarten Voors, and Nahomi Ichino (Mar. 2021). *The Theory and Practice of Field Experiments: An Introduction from the EGAP Learning Days*. Berkeley, CA.: Evidence in Governance and Politics. URL: https://egap.github.io/theory_and_practice_of_field_experiments/, Module 4 materials on randomization.

3— Statistical inference for counterfactual causal effects

More on the statistics of randomized experiments focusing on statistical power but reviewing testing and estimation. Sampling distributions, causal inference, hypothesis testing and statistical power.

Useful readings

- Alan S. Gerber and Donald P. Green (2012c). *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton & Co, ch. 3.
- **Recommended:** Jake Bowers, Maarten Voors, and Nahomi Ichino (Mar. 2021). *The Theory and Practice of Field Experiments: An Introduction from the EGAP Learning Days*. Berkeley, CA.: Evidence in Governance and Politics. URL: https://egap.github.io/theory_and_practice_of_field_experiments/, Module 5 & 6 materials on hypothesis testing and estimation.
- Karlan and Appel 2016 ch. 5.
- **Recommended:** Rachel Glennerster and Kudzai Takavarasha (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton: Princeton University Press ch. 6.
- **Recommended:** Bowers, Voors, and Nahomi Ichino 2021, Module 7 materials.

4— Clustered and blocked designs and power

How to think about designs where the researchers cannot directly administer a treatment to an individual unit of interest (cluster randomization)? And/or why we might construct an overall experiment from a set of mini-experiments (blocking or stratification)?

Useful readings

- Alan S. Gerber and Donald P. Green (2012c). *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton & Co, ch.8, Sections 8.1–8.3 only
- N. Ichino and M. Schündeln (Jan. 2012). "Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana". In: *Journal of Politics* 74.1, pp. 292–307
- Darin Christensen et al. (2021). "Building Resilient Health Systems: Experimental Evidence from Sierra Leone and The 2014 Ebola Outbreak". In: *The Quarterly Journal of Economics* 136.2, pp. 1145–1198

5— Maybe: Survey experiments

We could spend a session focusing on some of the challenges and opportunities posed by randomized experiments embedded within surveys.

Useful readings

- Gustavo Diaz, Christopher Grady, and James H Kuklinski (2020). "Survey Experiments and the Quest for Valid Interpretation". In: *The SAGE Handbook of Research Methods in Political Science and International Relations*. SAGE Publications
- Christopher Grady (2019). "10 Things to Know About Survey Experiments". In: *EGAP Methods Guides*. URL: https://egap.org/resource/10-things-to-know-about-survey-experiments/ (visited on 2019)
- Dawn Langan Teele, Joshua Kalla, and Frances Rosenbluth (2018). "The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics". In: *American Political Science Review* 112.3, pp. 525–541
- **Recommended:** Brian J Gaines, James H Kuklinski, and Paul J Quirk (2007). "The logic of the survey experiment reexamined". In: *Political Analysis* 15.1, pp. 1–20
- **Recommended:** Matthew S Winters and Rebecca Weitz-Shapiro (2013). "Lacking information or condoning corruption: When do voters support corrupt politicians?" In: *Comparative Politics* 45.4, pp. 418–436 and Taylor C Boas, F Daniel Hidalgo, and Marcus André Melo (2019). "Norms versus action: Why voters fail to sanction malfeasance in Brazil". In: *American Journal of Political Science* 63.2, pp. 385–400 and Trevor Incerti (2020). "Corruption information and vote share: A meta-analysis and lessons for experimental design". In: *American Political Science Review* 114.3, pp. 761–774.

6— Maybe: Causal effects of non-random compliance with randomized interventions

If an intervention has been randomized but the active dose has not, what can we do? It turns out that in this case we can estimate Complier Average Causal Effects (CACE) (also known as the Local Average Treatment Effect (LATE)) and we can also test hypotheses about the causal effects of non-randomly taking a dose of a randomized treatment.

Instrumental variables and the placebo controlled design and experiments built on previously executed randomizations (downstream experiments).

Useful readings

- Alan S. Gerber and Donald P. Green (2012c). *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton & Co, ch. 5
- Willa Friedman et al. (Jan. 2016). "Education as Liberation?" In: *Economica* 83.329, pp. 1–30 (Instrumental variables)
- Joshua L. Kalla and David E. Broockman (May 2020). "Reducing Exclusionary Attitudes through Interpersonal Conversation: Evidence from Three Field Experiments". In: *American Political Science Review* 114.2, pp. 410–425 (A placebo, control, treatment)
- Chapters 5 and 6 of Alan S Gerber and Donald P Green (2012a). *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: W.W. Norton.
- Chapter 5 and especially Section 5.3, "Instruments," of Paul R. Rosenbaum (2020). *Design of Observational Studies*. 2nd. Springer Series in Statistics. Springer
- Chapter 13, "Instruments" of Paul R Rosenbaum (2017). *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press
- Allison J. Sovey and Donald P. Green (2011). "Instrumental Variables Estimation in Political Science: A Readers Guide". In: *American Journal of Political Science* 55.1, pp. 188–200. ISSN: 1540-5907
- Paul R Rosenbaum (1996). "Identification of Causal Effects Using Instrumental Variables: Comment". In: *Journal of the American Statistical Association* 91.434, pp. 465–468
- Joshua D Angrist, Guido W Imbens, and Donald B Rubin (1996). "Identification of Causal Effects Using Instrumental Variables". In: *Journal of the American Statistical Association* 91.434, pp. 444–455
- Guido W Imbens and Paul R Rosenbaum (2005). "Robust, Accurate Confidence Intervals with a Weak Instrument: Quarter of Birth and Education". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 168.1, pp. 109–126 on weak instruments and the problem of 2SLS as an estimator
- Hyunseung Kang, Laura Peck, and Luke Keele (2018). "Inference for Instrumental Variables: A Randomization Inference Approach". In: *Journal of the Royal Statistical Society. Series A: Statistics in Society* 181.4, pp. 1231–1254

7— Maybe: Student presentations of experimental research designs and/or pre-analysis plans

References

Aarøe, Lene and Michael Bang Petersen (2013). "Hunger Games: Fluctuations in Blood Glucose Levels Influence Support for Social Welfare". In: *Psychological Science* 24.12, pp. 2550–2556.

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin (1996). "Identification of Causal Effects Using Instrumental Variables". In: *Journal of the American Statistical Association* 91.434, pp. 444–455.

Banerjee, Abhijit and Esther Duflo (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. New York: Public Affairs.

Bertrand, Marianne and Sendhil Mullainathan (2004). "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination". In: *American Economic Review* 94.4, pp. 991–1013.

Blair, Graeme, Alexander Coppock, and Macartan Humphreys (2023). *Research design in the social sciences: declaration, diagnosis, and redesign*. Princeton University Press. URL: https://book.declaredesign.org/.

Boas, Taylor C, F Daniel Hidalgo, and Marcus André Melo (2019). "Norms versus action: Why voters fail to sanction malfeasance in Brazil". In: *American Journal of Political Science* 63.2, pp. 385–400.

Bowers, Jake and Thomas Leavitt (2020). "Causality and Design-Based Inference". In: *The SAGE Handbook of Research Methods in Political Science and International Relations*. Ed. by Luigi Curini and Robert Franzese. Vol. 2. Thousand Oaks, CA: SAGE Publications. Chap. 41, pp. 769–804.

Bowers, Jake, Maarten Voors, and Nahomi Ichino (Mar. 2021). *The Theory and Practice of Field Experiments: An Introduction from the EGAP Learning Days*. Berkeley, CA.: Evidence in Governance and Politics. URL: https://egap.github.io/theory_and_practice_of_field_experiments/.

Christensen, Darin et al. (2021). "Building Resilient Health Systems: Experimental Evidence from Sierra Leone and The 2014 Ebola Outbreak". In: *The Quarterly Journal of Economics* 136.2, pp. 1145–1198.

Collier, Paul and Pedro C Vicente (2014). "Votes and violence: Evidence from a field experiment in Nigeria". In: *The Economic Journal* 124.574, F327–F355.

Diaz, Gustavo, Christopher Grady, and James H Kuklinski (2020). "Survey Experiments and the Quest for Valid Interpretation". In: *The SAGE Handbook of Research Methods in Political Science and International Relations*. SAGE Publications.

Fisher, R.A. (1935). *The design of experiments. 1935*. Edinburgh: Oliver and Boyd.

Fox, John (2016). *Applied Regression Analysis and Generalized Linear Models*. 3rd. Los Angeles, CA: SAGE Publications.

Friedman, Willa et al. (Jan. 2016). "Education as Liberation?" In: *Economica* 83.329, pp. 1–30.

Gaines, Brian J, James H Kuklinski, and Paul J Quirk (2007). "The logic of the survey experiment reexamined". In: *Political Analysis* 15.1, pp. 1–20.

Gerber, Alan S and Donald P Green (2012a). *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: W.W. Norton.

– (2012b). *Field experiments: Design, analysis, and interpretation*. WW Norton.

– (2012c). *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton & Co.

Gerber, Alan S., Donald P. Green, and Christopher W. Larimer (2008). "Social pressure and voter turnout: Evidence from a large-scale field experiment". In: *American Political Science Review* 102.1, pp. 33–48.

Glennerster, Rachel and Kudzai Takavarasha (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton: Princeton University Press.

Grady, Christopher (2019). "10 Things to Know About Survey Experiments". In: *EGAP Methods Guides*. URL: https://egap.org/resource/10-things-to-know-about-survey-experiments/ (visited on 2019).

Ichino, N. and M. Schündeln (Jan. 2012). "Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana". In: *Journal of Politics* 74.1, pp. 292–307.

Imbens, Guido W and Paul R Rosenbaum (2005). "Robust, Accurate Confidence Intervals with a Weak Instrument: Quarter of Birth and Education". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 168.1, pp. 109–126.

Incerti, Trevor (2020). "Corruption information and vote share: A meta-analysis and lessons for experimental design". In: *American Political Science Review* 114.3, pp. 761–774.

Kalla, Joshua L. and David E. Broockman (May 2020). "Reducing Exclusionary Attitudes through Interpersonal Conversation: Evidence from Three Field Experiments". In: *American Political Science Review* 114.2, pp. 410–425.

Kang, Hyunseung, Laura Peck, and Luke Keele (2018). "Inference for Instrumental Variables: A Randomization Inference Approach". In: *Journal of the Royal Statistical Society. Series A: Statistics in Society* 181.4, pp. 1231–1254.

Karlan, Dean and Jacob Appel (2016). *Failing in the Field: What We Can Learn When Field Research Goes Wrong*. Princeton: Princeton University Press.

Neyman, J. (1990). "On the application of probability theory to agricultural experiments. Essay on principles. Section 9 (1923)". In: *Statistical Science* 5. reprint. Transl. by Dabrowska and Speed, pp. 463–480.

Petersen, Michael Bang et al. (2014). "Social Welfare and the Psychology of Food Sharing: Short-Term Hunger Increases Support for Social Welfare". In: *Political Psychology* 35.6, pp. 757–773.

Rosenbaum, Paul R (1996). "Identification of Causal Effects Using Instrumental Variables: Comment". In: *Journal of the American Statistical Association* 91.434, pp. 465–468.

– (2002). *Observational Studies*. Second Edition. New York, NY: Springer.

– (2017). *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press.

– (2020). *Design of Observational Studies*. 2nd. Springer Series in Statistics. Springer.

Rubin, Donald B. (1990). "[On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9.] Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies". In: *Statistical Science* 5.4, pp. 472–480.

Sovey, Allison J. and Donald P. Green (2011). "Instrumental Variables Estimation in Political Science: A Readers Guide". In: *American Journal of Political Science* 55.1, pp. 188–200. ISSN: 1540-5907.

Teele, Dawn Langan, Joshua Kalla, and Frances Rosenbluth (2018). "The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics". In: *American Political Science Review* 112.3, pp. 525–541.

Wantchekon, Leonard (2003). "Clientelism and voting behavior: Evidence from a field experiment in Benin". In: *World Politics* 55.3, pp. 399–422.

Wickham, Hadley and Garrett Grolemund (2017). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. First. Sebastopol, CA: O'Reilly Media.

Winters, Matthew S and Rebecca Weitz-Shapiro (2013). "Lacking information or condoning corruption: When do voters support corrupt politicians?" In: *Comparative Politics* 45.4, pp. 418–436.