Introduction to Applied Statistics in Political Science: Description, Comparison, Estimation, Inference

Instructor: Jake Bowers
jwbowers@illinois.edu
http://jakebowers.org/
Methods Preceptor: Rahnuma Siddika
siddika2@illinois.edu

Fall 2026

Overview

When/Where We meet 3:30 to 5:50pm in 404 David Kinley Hall unless we need to switch to Zoom. Course materials will be mostly on Github. Some course materials and assignments will use Canvas.

Office Hours Please make an appointment on http://calendly.com/jakebowers if you want to come to Jake's office hours to ensure that we can meet and talk. I'm happy to schedule other times if those don't work for you.

Rahnuma's office hours are Fridays 10am to 12pm, on Zoom by default (or in person in 317B David Kinley Hall if you prefer). Please make an appointment on https://calendly.com/siddika2-illinois/fall2026rahnuma. If those times don't work for you, email Rahnuma to arrange another time.

The Methods Preceptor will wait before holding formal sections and will be open to input from the class about whether group sections will be useful or not.

Introduction This is the start of your practice with the skills and concepts of statistics as applied in political science. Political scientists use use statistics and data analysis skills political scientists to make arguments linking observation to theory. I say "start of a practice" because because the field constantly changes and because the subject is so deep and important that no one can ever truly master it: We are all always learning. So this course exists to help you start the learning that you will continue for your whole life working with data.

Statistics involves skills and concepts. If you don't have the skills, then the concepts are not concrete and are difficult to understand. In most statistics PhD programs the skills involve the mathematics of linear algebra and calculus and mathematical problem solving skills involved in deductive proofs and algebraic manipulation. In this course, we are going to use a flexible computer programming system in lieu of math in order to demonstrate to ourselves and make concrete the concepts that we must internalize and use in order to apply statistics to help us learn about the world and about our theories. Math, after all, is a language. R, the language we will be using, is another way to express abstract ideas. A by-product of using R to engage with statistical concepts is that you'll also practice how to use R to solve data problems. That is, you will start to practice some of the basics of "data science" or even "computational social science" on the way to practice some of the basics of statistics.

This course is also a graduate course in applied statistics or political methodology. This means that it exists within a series of continuous developments within multiple disciplines. The contents of this course will change over time because the disciplines change over time. What we thought worked well in the past may not be what we think works well today. What we teach in this course today may be seen by future scholars (we hope) as old fashioned. That is, this course is just like any other PhD level graduate class: we engage with the past in order to do things differently in the future. This class does not aim, therefore, to teach you to do political science as it was done in the 1950s, 1970s, 1990s,

or even last year. It is a tradition-based but future oriented class just like all of your other classes.

Goals and Expectations

This class aims to help you learn to think about what it means to do statistical inference for both descriptive and causal claims.

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 analyzing data.

I also hope that this course will help you continue to develop the acumen as a reader, writer, programmer and social scientist essential for your daily life as a social science researcher.

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

- Explain in their own words key concepts in statistics like "univariate description", "multivariate description", "statistical inference", "statistical adjustment" and describe how such concepts fit together in applied research.
- Articulate the major criteria for good data description and and make judgments about how to use those criteria to defend a strategy for description.
- Explain in their own words how statistical adjustment works using (1) the linear model and (2) stratification.
- Articulate the major criteria for good statistical adjustment and make judgments about how to use those criteria to defend a strategy for adjustment.
- Practice scientific computing using R and writing in R+markdown including showing competence in writing documents using R and markdown (and/or R and LaTeX) and Github for version control and collaboration.

To these ends I have designed a series of activities that should (1) give you opportunities to practice working with data and reasoning about statistics and (2) raise questions for discussion each week.

I do not lecture. Rather, each week we will meet to engage with the questions that you have.

Explorations Every week or so, I will ask you to complete a short assignment that encourages you to engage creatively with the topics of that section of the course. I anticipate that you will work on most of these assignments in groups and a few alone and that each of you will come to class prepared to discuss them. I don't think that the groups should have more than 3 people in them. However, I'm willing to have larger groups if you talk with me about it. The point of the explorations is for you to (1) practice learning on your own (making mistakes, confronting confusing error messages, finding help online and elsewhere) and in a group (this is how you will learn about statistics for the rest of your career, so these explorations are supposed to help you to practice it now), (2) engage with the topic of the week so that you are prepared to come to class with questions and ideas, (3) practice coding and confronting coding errors.

Daily R Five days per week, you will need to practice writing R code (you have 20 days out of every 28, no late work accepted, satisfactory if you practiced, fail if you did not, allowing 5 missing days. This is roughly 65 daily R assignments over the course of the term). This assignment requires that you make a GitHub Gist each day written in either R or R markdown or Quarto format in which you load a dataset from the web, or simulate some data, or load a dataset that comes with an R package, and learn something of interest to you about the units represented by those data. I'm imagining 2 to 10 lines of R code. In the ideal case, each gist should be written so that it runs from start to finish on any computer — not just yours. This might involve loading data from the web or using data bundled with R packages or generating some data within the code itself.

There is nothing to turn in each day. Make the gist **public** and put `dailyR` in its description, and I will find it. The counting is done by a script: each day it reads the public gists of everyone on the class roster and records who created one. A secret gist does not appear in the public list the script reads, so nothing can count it. Public gists can appear in GitHub's Discover page and search results, so do not put private or confidential data, including research data you cannot publish, in a Daily R gist. Without `dailyR` in the description, the script cannot tell your class practice from the gists you write for other reasons. Neither capitalization nor spacing matters: `DailyR`, `Daily R` and `daily-r` all work. Neither does the hour: a gist made at 11pm Monday counts for Monday, in Urbana-Champaign time.

What the script counts is *days* on which you created a gist, not the number of gists, so three gists written in one sitting is one day of practice. It reads the day a gist was created, so editing an old gist does not add a day and nothing can be backdated — that is what "no late work accepted" means here.

You will need a GitHub account, and you have already told me its username through the "Github username" assignment on Canvas. During the first week, make your first Daily R gist. It counts as a day of practice, and nothing else is ever reported. In return, each Monday morning you will get a note in a private repository that only you and I can see, telling you how many days you practiced in the week just finished and in the last 28. Watch the 28-day number: it is the one the rule above is written in, it falls when you have a rough patch, and it climbs back on its own when you start again. Full instructions are in `dailyR/student_instructions.md` in the github repository for this syllabus; the first-week Canvas announcement also links to them.

Replications Each of you will produce a short replication paper by the end of the term.[1] The idea is to practice using your new skills and concepts as applied to a topic of interest to you — but in a very targeted way. The idea is to understand what someone else did in the past, and perhaps to improve upon it now.

You should find a paper that uses statistical methods that you are willing to work to understand this term and where the data are available (ideally the data are easy to download, you can also contact the author of the paper after discussion with me). You can work on a paper that you yourself are writing, too.

You are allowed to do this work with a co-author or alone as you see fit.

I will be providing more detailed guidance about this assignment as the term goes on. You can see the current draft of the Final Paper Template (`final_paper_template.md`) if you want to see the kinds of details I articulated for the students last year.

[1] The idea comes from Gary King's assignment to his first year PhD students. I encourage you read that website and associated paper in which King explains his ideas. Our version will differ a little from his.

My Expectations

1. I assume you are eager to learn. Eagerness, curiosity and excitement will impel your energetic engagement with the class throughout the term. If you are bored, not curious, or unhappy about the class you should come and talk with me immediately. Energetic engagement manifests itself in meeting with your classmates outside of the class, in asking questions during the class, and in taking the assignments seriously.
2. I assume you are ready to work. Learning requires work. Making work about learning rather than grades, will make our work that much more difficult and time consuming. You will make errors. These errors are opportunities for you to learn — some of your learning will be about how to help yourself and some will be about statistics. If you have too much to do this term consider dropping the course. If you don't think you need this course consider dropping it. Graduate school is a place for you to develop and begin to pursue your own intellectual agendas: this course may be important for you this term, or it may not. That is up for you to decide.
3. I assume some previous engagement with high school mathematics.
4. You should ask questions when you don't understand things; chances are you're not alone. **This class is an opportunity to practice courage:** I expect you to make a guess when I ask a question (in writing or in person), I expect that you will ask a question when you have a problem.
5. **Do the work.** This does not mean divide the work up among your classmates so that you only do part of the work. Each person should engage with all of the work even if the people who writes it up changes from week to week.
6. All papers written in this class will assume familiarity with the principles of good writing in Becker (1986).
7. All final written work will be turned in as pdf files unless we have another specific arrangement.[2] I will not accept Microsoft, Apple, or any other proprietary format.
8. I assume you will use AI within the context of the course AI prompting setup (`ps530_ai_learning_guide.md`) (and that you will help me improve this setup and talk with me and others in the class about how you use AI to help you work harder and learn more rather than get work done quickly.)

Late Work I do not like evaluation for the sake of evaluation. Evaluation should provide opportunities for learning. So, if you'd prefer to spend more time using the paper assignment in this class to learn more, I am happy for you to take that time. I will not, however, entertain late submissions for any subsidiary paper assignments or other homeworks that are due throughout the term. If you think that you and/or the rest of the class have a compelling reason to change the due date on one of those assignments, let me know in advance and I will probably just change the due date for the whole class.

[2] For example, if you have some reason why pdf files make your life especially difficult, then of course I will work with you find another format.

Incompletes Incomplete grades at the end of the term are fine in theory but terrible in practice. I urge you to avoid an incomplete in this class. If you must take an incomplete, you must give me *at least* 2 months from the time of turning in an incomplete before you can expect a grade from me and it may well take me much longer. This means that if your fellowship, immigration status, or job depends on erasing an incomplete in this class, you should not leave this incomplete until the last minute.

Grades are Feedback Humans need feedback to close the gap between intention and action. They also need feedback to feel good about their progress and to motivate them. In this class I will use grades as feedback. All grades except for the final grade will be satisfactory, unsatisfactory (with the possibility of "outstanding"), and fail. These map roughly onto A+=outstanding, A=satisfactory, C=unsatisfactory, and F=fail (i.e. you didn't try).

I'll calculate your grade for the course this way: 30% daily R (you have 5 days out of every 7, no late work accepted, satisfactory if you practiced, fail if you did not, counted automatically as described above); 40% explorations (when you turn it in as a group everyone in the group receives the same grade, satisfactory if you are creative and thoughtful and diligent, unsatisfactory if you are not or if you don't seem to be getting the concepts, no late work accepted); 20% replication paper; 10% attendance (satisfactory if you show up, fail if not).

You can miss two classes without grade penalty.

I will drop your lowest exploration grade as well.

Because moments of evaluation are also moments of learning in this class, I do not curve. If you all perform at 100%, then I will give you all As.

You can redo any evaluation or the final paper in order to increase your grade on that assignment. If you want to resubmit something already graded, you need to let me know in advance so that I can make time to grade it again. If you want to resubmit work after the end of the term, that is also ok, but I may take many months to grade that work.

Computing We will be using R in class so those of you with laptops available should bring them to class. Of course, I will not tolerate the use of computers for anything other than class related work during active class time. Please install R on your computers before the first class session.

As you work on your papers, you will also learn to write about data analysis in a way that sounds and looks professional by using either R+markdown or Sweave (R+LaTeX). No paper will be accepted without a code appendix or reproduction Github repository made available to me. No paper will be accepted unless it is in Portable Document Format (pdf). No paper will be accepted with cut and pasted computer output in the place of well presented and replicable figures and tables. Although good empirical work requires that the analyst understand her tools, she must also think about how to communicate effectively: ability to reproduce past analyses and clean and clear presentations of data summaries are almost as important as clear writing in this regard.

Books

I'm am not requiring any particular books this term. The readings will be drawn from a variety of sources. I will try to make most of them available to you as we go if you can't find them easily online yourselves.

Recommended No book is perfect for all students. I suggest you ask around, look at other syllabi online, and just browse the shelves at the library and used bookstores to find books that make things clear to you. I will be adding some recommendations here. Let me know now if you have favorites.

If you discover any books or websites that are particularly useful to you, please alert me and the rest of the class about them. Thanks!

Academic Integrity According to the Student Code, 'It is the responsibility of each student to refrain from infractions of academic integrity, from conduct that may lead to suspicion of such infractions, and from conduct that aids others in such infractions.' Please know that it is my responsibility as an instructor to uphold the academic integrity policy of the University, which can be found here: http://studentcode.illinois.edu/article1_part4_1-401.html.

Disabilities To ensure that disability-related concerns are properly addressed from the beginning, students with disabilities who require assistance to participate in this class should see me as soon as possible. To obtain disability-related academic adjustments and/or auxiliary aids, students with disabilities must contact the course instructor and the Disability Resources and Educational Services (DRES) as soon as possible. To contact DRES you may visit 1207 S. Oak St., Champaign, call 333-4603 (V/TTY), or e-mail a message to mailto:disability@illinois.edu.

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.

Data: I'll be bringing in data that I have on hand. This means our units of analysis will often be individual people or perhaps political or geographic units, mostly in the United States. I'd love to use other data, so feel free to suggest and provide it to me — come to office hours and we can talk about how to use your favorite datasets in the class.

Theory: This class is about description, estimation, comparisons, and causal inference. Yet, statistics as a discipline exists to help us understand more than why the linear model works as it does. Thus, social science theory cannot be far from our minds as we think about what makes a given data analytic strategy meaningful. That is, while we spend a term thinking a lot about how to make meaningful statements about statistical inference, we must also keep substantive significance foremost in our minds.

I Description

August 25—Description in One Dimension

What makes a description useful or not useful? What is a good description? How would we know whether we have a good one or a bad one? Are there descriptions that are particularly robust to observations we'd like to ignore or mostly ignore because they are apt to mislead us?

Read: **Henceforth, "*" means "recommended" or "particularly useful" reading.**

*Rand R Wilcox (2012). *Introduction to robust estimation and hypothesis testing*. Academic Press, Chap 1–3

*Daniel Kaplan (2012). *Statistical Modeling A Fresh Approach*. Second. Macalester College, St. Paul, MN: Daniel Kaplan, Chap 2–3

September 1—Description in Two Dimensions I

Topics: Linear data models and fitting.

Questions: Why use straight lines to describe relationships? On what basis should we choose a straight line to "fit data" (why choose one straight line over others)? How to interpret slopes and intercepts (i.e. the descriptors of a line) given different ways of choosing lines? What is the relationship between the least squares method of line fitting and describing differences between two groups? When might we want to identify and perhaps diminish the influence of particular individual observations on an overall linear description? Cook's distance? Influence? Leverage?

Read: *Christopher H. Achen (1982). *Interpreting and Using Regression*. Newbury Park, CA: Sage, Chap 2

*Daniel Kaplan (2012). *Statistical Modeling A Fresh Approach*. Second. Macalester College, St. Paul, MN: Daniel Kaplan, Chap 4, 5–8

*Gareth James et al. (2013). *An Introduction to Statistical Learning*. Springer, Chap 2–3

*R.A. Berk (2008). *Statistical learning from a regression perspective*. Springer, Chap 1

*R. Berk (2010). "What you can and can't properly do with regression". In: *Journal of Quantitative Criminology*, pp. 1–7 (especially level 1 regression)

*Rand R Wilcox (2012). *Introduction to robust estimation and hypothesis testing*. Academic Press, Chap 5 and 10

September 8—Description in Two Dimensions II

Topics: Nonlinear data models and fitting

Questions: What other criteria for line choice might we have? Re-interpreting linear description as both smooth and comparison. How to argue that you have smoothed the data appropriately? How does the use of dummy variables or indicator variables and/or interaction terms allow us to engage with theoretical expectations that are not linear? What about transformations of predictors like simple polynomials? Or piecewise fits (smooth ones using splines, or not smooth ones using indicators and interactions)? Or locally smooth fits?

Read: Gareth James et al. (2013). *An Introduction to Statistical Learning*. Springer, Chap 3, 7[3]

William G Jacoby (2000). "Loess:: a nonparametric, graphical tool for depicting relationships between variables". In: *Electoral Studies* 19.4, pp. 577–613

The Fox and Weisberg Textbook Nonparametric Regression Appendix

[3] See https://www.statlearning.com/ for materials

Extra: We are not going to dive into the fitting of models that are nonlinear in parameters. But if you are interested in such things (for example the COVID SIR or SIER models or other strutural theoretical models as fit to data), here are a few resources.

- The Fox and Weisberg Textbook Nonlinear Least Squares Appendix
- https://datascienceplus.com/first-steps-with-non-linear-regression-in-r/
- https://www.sciencedirect.com/science/article/pii/S1201971220303039

September 15—Description in Two Dimensions III—Categorical Variables and Tables

Topics: Description when all the variables are categorical and/or binary.

Questions: What are strategies for learning from and summarizing relationships between two or three categorical variables? We have been smoothing scatterplots, but now must think about tables and other kinds of graphical devices. Also we'll continue to use indicator variables and interactions terms in linear models as short cuts to make certain summaries.

Read: Kosuke Imai (2018). *Quantitative social science: an introduction*. Princeton University Press, Chapter 2 using Tables to describe relationships.

See also:

- https://moderndive.com/2-viz.html
- https://medium.com/analytics-vidhya/visual-outputs-in-r-example-1-d71cb4be50eb
- https://bookdown.org/max/FES/visualizations-for-categorical-data-exploring-the-okcupid-data.html
- https://homepage.divms.uiowa.edu/~luke/classes/STAT4580/catone.html
- http://www.sthda.com/english/articles/32-r-graphics-essentials/129-visualizing-multivariate

II Statistical Inference: Estimation and Testing and the Unobserved

September 22—Estimation and Unobserved Causal Counterfactuals

Topics: Counterfactual approaches to formalizing causal statements; potential outcomes; the idea of **estimation** (work on the idea of "unbiased estimation" in the exploration.)

Question: What do we mean when we say "Z caused Y"? What is the counterfactual interpretation of this statement? What is the role of variables that we observe but which are not Z or Y? How might we use them to get more clear on the Z to Y relationship?

Read: *Henry E. Brady (2008). "Causation and explanation in social science". In: *Oxford handbook of political methodology*, pp. 217–270 for an excellent overview and then a discussion of Neyman's "average treatment effects" engagement with the fundamental problem of causal inference.

*Paul R. Rosenbaum (2017). *Observation and experiment : an introduction to causal inference*. Cambridge, MA: Harvard University Press, p. 374. ISBN: 9780674975576, Chap 1 and 2

*Paul R. Rosenbaum (May 2010). *Design of Observational Studies*. Springer. URL: http://www.springer.com/statistics/statistical+theory+and+methods/book/978-1-4419-1212-1, Chap 1 and 2

*Paul R. Rosenbaum (2002b). *Observational Studies*. Second. Springer-Verlag, Chap 1 and 2

*Guido W Imbens and Donald B Rubin (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press, Chap 1 and 2

*Thad Dunning (2012). *Natural experiments in the social sciences: a design-based approach*. Cambridge University Press, Chap 1 and 2

*A.S. Gerber and D.P. Green (2012). *Field experiments: Design, analysis, and interpretation*. WW Norton, Chap 1–3

*Estimands and Estimators Material and Slides

September 29–Estimation and Unobserved Populations

Topics: Estimating population quantities like means; estimating standard errors for those estimators (including analytic calculations but also the bootstrap).

Readings: *Sharon L Lohr (2010). *Sampling: design and analysis*. Chapman and Hall/CRC, Chap 2, esp. 2.8

*Peter M Aronow and Benjamin T Miller (2019). *Foundations of agnostic statistics*. Cambridge University Press, Chap 3 and 4, esp. 3.1–3.4, 4.1–4.2

October 6—Hypothesis tests

Topic: How does "statistically significant" relate to statistical inference about unobserved counterfactual quantities? The p-value as a measure of information against an argument summarizing the relationship between a **test statistic** and its **reference distribution** given a **null hypothesis**. Maybe we will talk about aspects of a "good test" like "low and controlled false positive rates" and "high power".

Reading: 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.

Paul R. Rosenbaum (May 2010). *Design of Observational Studies*. Springer. URL: http://www.springer.com/statistics/statistical+theory+and+methods/book/978-1-4419-1212-1, Chap 2

Daniel Kaplan (2009). *Statistical Modeling: A Fresh Approach*. http://www.macalester.edu/~kaplan/ism/. ISBN: 978-1448642397, Chap 15, 16.1, 16.6, 16.7, 17.5, 17.7, 17.8 discusses tests of hypotheses in the context of permutation distributions of linear model based test statistics. He wants to emphasize the F-statistic and R^2 and the ANOVA table, but his discussion of permutation based testing will apply to our concern with the effect of an experimental treatment on an outcome.

L. Gonick and W. Smith (1993). *The cartoon guide to statistics*. HarperPerennial New York, NY, Chap 8 explains the classical approach to hypothesis testing based on Normal and t-distributions.

G. Imbens and D. Rubin (2009). "Causal Inference in Statistics". Unpublished book manuscript. Forthcoming at Cambridge University Press., Chap 5 explains Fisher's approach to the sharp or strict null hypothesis test in the context of the potential outcomes framework for causal inference.

Richard Berk (2004). *Regression Analysis: A Constructive Critique*. Sage, Chap 4 provides an excellent and readable overview of the targets of inference and associated justifications often used by social scientists.

John Fox (2016). *Applied regression analysis and generalized linear models, Third Edition*. Sage, Chap 21.4 explains about bootstrap hypothesis tests (i.e. sampling model justified hypothesis tests).

Paul R. Rosenbaum (2002b). *Observational Studies*. Second. Springer-Verlag, Chap 2–2.4 explains and formalizes Fisher's randomization inference.

Paul R. Rosenbaum (2002a). "Covariance adjustment in randomized experiments and observational studies". In: *Statistical Science* 17.3, pp. 286–327 explains how one might use Fisher-style randomization inference with linear regression.

October 13—Confidence Intervals are Collections of Hypothesis Tests

Topics Given a reasonable data summary, what other guesses about said quantity are plausible? Continuing on statistical inference; Inverting hypothesis tests; null hypotheses and alternatives; a Central Limit Theorem. Maybe the bootstrap, but probably wait for population inference until next term.

Useful Reading Daniel Kaplan (2009). *Statistical Modeling: A Fresh Approach*. http://www.macalester.edu/~kaplan/ism/. ISBN: 978-1448642397, Chap 14

L. Gonick and W. Smith (1993). *The cartoon guide to statistics*. HarperPerennial New York, NY, Chap 7

Jerzy Neyman (1937). "Outline of a theory of statistical estimation based on the classical theory of probability". In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 236.767, pp. 333–380 invents the confidence interval.

John Fox (2016). *Applied regression analysis and generalized linear models, Third Edition*. Sage, Chap 21

G. Imbens and D. Rubin (2009). "Causal Inference in Statistics". Unpublished book manuscript. Forthcoming at Cambridge University Press., Chap 6 discusses and compares Fisher's approach to Neyman's approach. We will defer discussion about the the parts of the discussion regarding Normality until later in the course. Review their chapter 5.8 for discussion about inversion of the hypothesis test to create confidence intervals.

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.

S. Lohr (1999). *Sampling: Design and Analysis*. Brooks/Cole, Chap 2.7 a clear exposition of the random-sampling based approach.

R.A. Berk (2008). *Statistical learning from a regression perspective*. Springer and the problem of samples that are not random from populations that are imaginary.

October 20—No Class

October 27—Maybe Review Statistical Inference

Topics: Starting with "randomization-based" statistical inference to get the concepts clear: Estimating unobserved average causal effects (estimands), Testing hypotheses about unobserved average causal effects and/or unobserved individual causal effects, Creating confidence intervals by inverting hypothesis tests (and perhaps using standard errors).

III Adjustment: What does "controlling for" actually do?

November 3–Covariance Adjustment, Removing Linear additive relationships

Topics: Covariance adjustment, adjustment by linear model (interaction terms for stratification, or direct covariance adjustment aka "controlling for"), What does "controlling for" do? When does it make most sense? When does it make least sense?

Reading. *R.A. Berk (2008). *Statistical learning from a regression perspective*. Springer, Pages 1–8[4]

*Richard Berk (2004). *Regression Analysis: A Constructive Critique*. Sage, Chap 6–7 (skipping stuff on standardized coefs)

*Christopher H. Achen (2002). "Toward A New Political Methodology: Microfoundations and ART". in: *Annual Review of Political Science* 5 (1), pp. 423–450

*Christopher H Achen (2005). "Let's put garbage-can regressions and garbage-can probits where they belong". In: *Conflict Management and Peace Science* 22.4, pp. 327–339 (on the problem of kitchen sink regressions)

*John Fox (2008). *Applied regression analysis and generalized linear models*. Sage, Chap 11 on Overly Influential Points

A. Gelman and J. Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press Sections 9.0 – 9.2 (especially discussion of interpolation and extrapolation)

R. Berk (2010). "What you can and can't properly do with regression". In: *Journal of Quantitative Criminology*, pp. 1–7

November 10—Simple Stratification and Matching

Topics: Matching methods as a way to do the stratification-based approaches in a way that also allows for assessment of the question: "Did you adjust enough?"

Reading: *Paul R. Rosenbaum (2017). *Observation and experiment : an introduction to causal inference*. Cambridge, MA: Harvard University Press, p. 374. ISBN: 9780674975576, Chap 5 and 11

*Paul R. Rosenbaum (May 2010). *Design of Observational Studies*. Springer. URL: http://www.springer.com/statistics/statistical+theory+and+methods/book/978-1-4419-1212-1, Chap 1, 3, 7, 8, 9, 13 (available via SpringerLink from our library)

*A. Gelman and J. Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Chap 9.0–9.2 (on causal inference and especially interpolation and extrapolation)

*B.B. Hansen (Sept. 2004). "Full matching in an observational study of coaching for the SAT". in: *Journal of the American Statistical Association* 99.467, pp. 609–618 on full matching for adjustment

*B.B. Hansen and J. Bowers (2008). "Covariate Balance in Simple, Stratified and Clustered Comparative Studies". In: *Statistical Science* 23, p. 219. URL: doi:10.1214/08-STS254 on assessing balance.

[4] See here

November 17—Stratification Using Many Variables

Topics: Mahalanobis and propensity scores, calipers, penalties, exact matching. Maybe: fine balance.

Reading: See the readings from the last week.

November 24—No Class, Fall Break

December 1—Sensitivity Analysis for Estimation

Topics: All observational studies leave something out. How big must the influence of the unobserved variable be in order to overturn our substantive results?

Reading: Carrie A Hosman, Ben B Hansen, and Paul W Holland (2010). "The Sensitivity of Linear Regression Coefficients' Confidence Limits to the Omission of a Confounder". In: *The Annals of Applied Statistics* 4.2, pp. 849–870

Carlos Cinelli and Chad Hazlett (2020). "Making Sense of Sensitivity: Extending Omitted Variable Bias". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82.1, pp. 39–67

Guido W Imbens (2003). "Sensitivity to Exogeneity Assumptions in Program Evaluation". In: *The American Economic Review* 93.2, pp. 126–132

Emily Oster (2019). "Unobservable Selection and Coefficient Stability: Theory and Evidence". In: *Journal of Business & Economic Statistics* 37.2, pp. 187–204

Stephen Chaudoin, Jude Hays, and Raymond Hicks (2018). "Do We Really Know the WTO Cures Cancer?" In: *British Journal of Political Science* 48.4, pp. 903–928

December 8— Last Day of Class: Open Discussion

Topics: Open Discussion. Each person brings a question.

December 11— Final Project Due by 11:59pm CT

IV References

References

Achen, Christopher H (2005). "Let's put garbage-can regressions and garbage-can probits where they belong". In: *Conflict Management and Peace Science* 22.4, pp. 327–339.

– (1982). *Interpreting and Using Regression*. Newbury Park, CA: Sage.

– (2002). "Toward A New Political Methodology: Microfoundations and ART". In: *Annual Review of Political Science* 5 (1), pp. 423–450.

Aronow, Peter M and Benjamin T Miller (2019). *Foundations of agnostic statistics*. Cambridge University Press.

Becker, Howard S. (1986). *Writing for Social Scientists: How to Start and Finish Your Thesis, Book, or Article*. University of Chicago Press.

Berk, R. (2010). "What you can and can't properly do with regression". In: *Journal of Quantitative Criminology*, pp. 1–7.

Berk, R.A. (2008). *Statistical learning from a regression perspective*. Springer.

Berk, Richard (2004). *Regression Analysis: A Constructive Critique*. Sage.

Brady, Henry E. (2008). "Causation and explanation in social science". In: *Oxford handbook of political methodology*, pp. 217–270.

Chaudoin, Stephen, Jude Hays, and Raymond Hicks (2018). "Do We Really Know the WTO Cures Cancer?" In: *British Journal of Political Science* 48.4, pp. 903–928.

Cinelli, Carlos and Chad Hazlett (2020). "Making Sense of Sensitivity: Extending Omitted Variable Bias". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82.1, pp. 39–67.

Dunning, Thad (2012). *Natural experiments in the social sciences: a design-based approach*. Cambridge University Press.

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

Fox, John (2008). *Applied regression analysis and generalized linear models*. Sage.

– (2016). *Applied regression analysis and generalized linear models, Third Edition*. Sage.

Gelman, A. and J. Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gerber, A.S. and D.P. Green (2012). *Field experiments: Design, analysis, and interpretation*. WW Norton.

Gonick, L. and W. Smith (1993). *The cartoon guide to statistics*. HarperPerennial New York, NY.

Hansen, B.B. (Sept. 2004). "Full matching in an observational study of coaching for the SAT". In: *Journal of the American Statistical Association* 99.467, pp. 609–618.

Hansen, B.B. and J. Bowers (2008). "Covariate Balance in Simple, Stratified and Clustered Comparative Studies". In: *Statistical Science* 23, p. 219. URL: doi:10.1214/08-STS254.

Hosman, Carrie A, Ben B Hansen, and Paul W Holland (2010). "The Sensitivity of Linear Regression Coefficients' Confidence Limits to the Omission of a Confounder". In: *The Annals of Applied Statistics* 4.2, pp. 849–870.

Imai, Kosuke (2018). *Quantitative social science: an introduction*. Princeton University Press.

Imbens, G. and D. Rubin (2009). "Causal Inference in Statistics". Unpublished book manuscript. Forthcoming at Cambridge University Press.

Imbens, Guido W (2003). "Sensitivity to Exogeneity Assumptions in Program Evaluation". In: *The American Economic Review* 93.2, pp. 126–132.

Imbens, Guido W and Donald B Rubin (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

Jacoby, William G (2000). "Loess:: a nonparametric, graphical tool for depicting relationships between variables". In: *Electoral Studies* 19.4, pp. 577–613.

James, Gareth et al. (2013). *An Introduction to Statistical Learning*. Springer.

Kaplan, Daniel (2009). *Statistical Modeling: A Fresh Approach*. http://www.macalester.edu/~kaplan/ism/. ISBN: 978-1448642397.
– (2012). *Statistical Modeling A Fresh Approach*. Second. Macalester College, St. Paul, MN: Daniel Kaplan.
Lohr, S. (1999). *Sampling: Design and Analysis*. Brooks/Cole.
Lohr, Sharon L (2010). *Sampling: design and analysis*. Chapman and Hall/CRC.
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.
Neyman, Jerzy (1937). "Outline of a theory of statistical estimation based on the classical theory of probability". In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 236.767, pp. 333–380.
Oster, Emily (2019). "Unobservable Selection and Coefficient Stability: Theory and Evidence". In: *Journal of Business & Economic Statistics* 37.2, pp. 187–204.
Rosenbaum, Paul R. (2002a). "Covariance adjustment in randomized experiments and observational studies". In: *Statistical Science* 17.3, pp. 286–327.
– (2002b). *Observational Studies*. Second. Springer-Verlag.
– (May 2010). *Design of Observational Studies*. Springer. URL: http://www.springer.com/statistics/statistical+theory+and+methods/book/978-1-4419-1212-1.
– (2017). *Observation and experiment : an introduction to causal inference*. Cambridge, MA: Harvard University Press, p. 374. ISBN: 9780674975576.
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.
Wilcox, Rand R (2012). *Introduction to robust estimation and hypothesis testing*. Academic Press.