Political Science 1600

INTRODUCTION TO QUANTITATIVE RESEARCH METHODS

Paul Testa
Assistant Professor
paul_testa@brown.edu

First Draft: August 29, 2017
This Draft: January 22, 2018

Overview

General Information

Canvas https://canvas.brown.edu/courses/1075234

Where/When We meet Tuesdays and Thursdays from 6:40–8:00 pm in Rockefeller Library 205.

Office Hours Tuesdays from **4:00-5:00 pm** in in **332 Blistein House**

Thursdays from **5:30 - 6:30 pm** in the lobby of Rockefeller and by appointment.

If you know in advance that you want to meet, please email me to reserve a 20-minute slot.

Course Summary

This class is an introduction to applied statistics as practiced in political science. It is computing intensive, and, as such, will enable students to execute basic quantitative analyses of social science data using the linear model with statistical inference arising from re-sampling and permutation based techniques as applied in the R statistical computing language (https://cran.r-project.org/) with RStudio (https://www.rstudio.com/.) By the end of the course, a successful student will be able to find social science data online, download it, analyze it, and write about how the analyses bear on focused social science or policy questions.

Course Goals

More than anything I assume a willingness to engage with mathematics, data analysis, computer programming, and the practice of social science thinking and

writing. I also assume you've taken at least one class in algebra at the level taught in most high schools in the United States and have used a personal computer to read and type email and other documents and have some experience with the Internet.

I also assume that you will read the syllabus and that you keep up to date on changes in the syllabus which will be announced in class. You should not expect a response to emails that ask a question already answered in the syllabus.

This is an experimental class so you should expect that the syllabus will change throughout the term. Make sure you have the syllabus with the latest date stamp. I will announce syllabus changes via the emails sent from the Canvas.

Community Standards

All students and the instructor must be respectful of others in the classroom. If you ever feel that the classroom environment is discouraging your participation or problematic in any way, please contact me.

Accessibility

Brown University is committed to full inclusion of all students. Please inform me if you have a disability or other condition that might require accommodations or modification of any of these course procedures. You may speak with me after class or during office hours. For more information contact Student and Employee Accessibility Services at 401-863-9588 or SEAS@brown.edu.

Academic Integrity

Neither the University nor I tolerate cheating or plagiarism. The Brown Writing Center defines plagiarism as "appropriating another person's ideas or words (spoken or written) without attributing those word or ideas to their true source." The consequences for plagiarism are often severe, and can include suspension or expulsion. This course will follow the guidelines in the Academic Code for determining what is and isn't plagiarism:

> In preparing assignments a student often needs or is required to employ outside sources of information or opinion. All such sources should be listed in the bibliography. Citations and footnote references are required for all specific facts that are not common knowledge and about which there is not general agreement. New discoveries or debatable opinions must be credited to the source, with specific references to edition and page even when the student restates the matter in his or her own words. Word-for-word inclusion of any part of someone else's written or oral sentence, even if only a phrase or sentence, requires citation in quotation marks and use of the appropriate conventions for attribution. Citations should normally include author, title, edition, and page. (Quotations longer than one sentence are generally indented from the text of the essay, without quotation marks, and identified by author, title, edition, and page.) Paraphrasing or summarizing the contents of another's work

> is not dishonest if the source or sources are clearly identified (author, title, edition, and page), but such paraphrasing does not constitute independent work and may be rejected by the instructor. Students who have questions about accurate and proper citation methods are expected to consult reference guides as well as course instructors.

We will discuss specific information about your written work in class in more detail, but if you are unsure of how to properly cite material, please ask for clarification. If you are having difficulty with writing or would like more information or assistance, consult the Writing Center, the Brown library and/or the Academic Code (https://www.brown.edu/Administration/Dean_of_the_College/curriculum/documents/academic-code.pdf) for more information.

Course Structure and Policies

Class

This course meets twice a week for 80 minutes on Tuesdays and Thursdays and places an emphasis on active learning. The first 20 to 30 minutes of class will be devoted to lecture, demonstration and review. The remainder of the class will focus on applications of these concepts through brief labs "labs" where you'll work with real data from a variety of sources. It is assumed that you will come to class having done each weeks assigned readings and reviewed material from the previous weeks lectures and labs. Slides and labs will be available on Canvas an hour before class.

Attendance

You may miss two classes without it having any effect on the attendance portion of your grade. After two absences, each additional absence (without a written note from the Univeristy) will reduce your final grade by 1 percent.

Readings

There is one required textbooks for the course (**Estimated cost:** ∼**$48.00**):

Imai, K. (2017). ***Quantitative Social Science: An Introduction*****. Princeton, NJ: Princeton University Press**

The primary textbook on which the course is structured. Most chapters are spread over multiple weeks. You should read this text with your laptop and R Studio open. Execute the code in the main text and ideally try to complete the assignments and exercises at the end of the chapter. **Approximate Cost: $48.00**[1]

Additional readings will be listed below and available to download on Canvas.

[1] Estimated from Amazon

Labs

The bulk of the work and learning you'll do in the course comes in the form of weekly labs in which you'll explore a given data set or paper using R. You'll be given an R Markdown document that will guide you through a set of exercises to teach concepts covered in the lectures and reading. You'll code in `R` and summaries of your findings in R Markdown. You will compile your document to produce an html document which you will **submit on Canvas by the end of each class.**

All work in this class **MUST BE SUBMITTED ONLINE VIA CANVAS.**

You are expected to work in collaboration with your peers. You may share code and discuss your results, but each of you must submit your own file.

Assignments

In addition to weekly labs, you will have periodic assignments the goal of which is to help you stay on track for writing your final paper. All assignments are due the Friday after the class with which they are associated.

The timeline of assignments for your final paper is as follows:

Week 3: Research Topics
Due Sunday, February 18, 2018 at 5:00 pm on Canvas

Week 5: Identifying Datasets
Due Sunday, March 4, 2018 at 5:00 pm on Canvas

Week 8: Outline of Research Proposal
Due Sunday, March 25, 2018 at 5:00 pm on Canvas

Week 12: Draft of Final Paper
Due Sunday, April 29, 2018 at 5:00 pm on Canvas

Week 15: Final Paper
Due Sunday, May 13, 2018 at 5:00 pm on Canvas

Assignments and labs must be submitted on time to Canvas. No late work will be accepted without prior approval of the instructor or a note from the University.

Grades

Your final grade for this course will be calculated as follows:

- **5% Class attendance**
- **10% Class involvement and participation**
- **40% Labs**
- **5% Assignments not including draft**
- **5% Draft of final Paper**
- **35% Final Papers**

Labs, assignments excluding the rough draft and final paper, will be graded graded out of 100 roughly on a ✓+ (100, completed on time, acceptable), ✓(88, completed on time, passable), ✓- (0 not submitted on time, unacceptable). If you are absent, you will receive a 0 on that day's lab, however, the lowest three lab grades will be dropped from your final lab grade. Your drafts and final paper will be graded on 100-point scales with rubrics provided beforehand.

Incomplete Work Assignments not turned in will be counted as zero in the calculation of the final grade.

Computers in class Please bring your laptops if you have them. We will install R and Rstudio together. If you do not own a laptop, you can still work in a group of other people who have laptops and will be able to complete the in-class worksheets without a problem. In fact, it is ideal if each group of 2-4 people works with one laptop and then shares the work among themselves. Of course, feel free to work on your own outside of class.

Time

This course meets 25 times over 14 weeks in the semester, including the first class and last class that will be held during reading period. Each class is 80 minutes long so you should expect to spend approximately 34 hours total in class; approximately 3 hours per week reading the textbook and reviewing material (42 hours total); approximately 15 hours on assignments for the final paper; approximately 20 hours researching, writing, and revising your final paper; and at least .5 hours meeting with me in person to discuss your work (Estimated Total Time: 111.5 hours)

Schedule

The general outline of topics for the course is as follows. The data and topics for some labs may change.

Week 0 — January 25, 2018— Introduction and Course Overview

Question: What am I getting myself into?

Read: Imai (2017) Chapter 1 (if you have the textbook)

Thursday **January 25, 2018—Lab:** Course overview

Reminder: Bring laptops if you have them. We'll take some time at the end of class to make sure everyone's setup with R, *R Studio* and LaTeX

Week 1. Data and Measurement

Read: Imai (2017) Chapter 3

Wickham and Grolemund (2017) http://r4ds.had.co.nz/tidy-data.html

Question: How do we describe the world around us?

Tuesday **January 30, 2018—Lab:** Levels of measurement; Measures of central tendency and dispersion; Loading and manipulating data in R

Thursday **February 1, 2018—Lab:** Measures of association; Aggregating and summarizing data; Exploratory data analysis and visualization

Week 2. Causation I

Read: Imai (2017) Chapter 2

Broockman and Kalla (2016) skim for understanding of basic question, data and design

Question: How do we know if X causes Y?

Tuesday **February 6, 2018—Lab** Potential outcomes and counterfactuals; The fundamental problem of causal inference; A "statistical" solution to that problem; The role of randomization in the "experimental ideal"

Thursday **February 8, 2018—Lab:** Estimating causal effects with experimental data

Week 3. Causation II / Prediction

Question: How do we know if X causes Y without randomly assigning X?

Read: Imai (2017) Chapter 2; Begin Chapter 4

Ferwerda and Miller (2014) skim for understanding of basic question, data and design

Tuesday **February 13, 2018—Lab** Tools for drawing causal inferences from observational data; Assessing claims of "as-if" random assignment. Using natural experiments to assess the effect of different types of foreign rule resistance? (Exploration of Ferwerda and Miller (2014))

Thursday **February 15, 2018—Lab:** Some basic principles of estimation. Simple OLS regression. Minimizing the sum of squared errors.

Assignment 1 **DUE: 5 pm on Sunday, February 18, 2018. Submit proposed research topics on Canvas**

Week 4. Prediction I

Read: Imai (2017) Chapter 4

Question: How do we make predictions?

Tuesday **February 20, 2018—No Class** Long weekend according to Brown's academic calendar?

Thursday **February 22, 2018—Lab:** Multiple regression. What it means to "control for X" or hold "Z constant". How to interpret multiple regression categorical and continuous predictions, and interactions.

Week 5. Prediction II

Read: Imai (2017) Chapter 4

Question: How do we make predictions adjusting for potentially confounding factors?

Tuesday **February 27, 2018—No Class** We'll discuss possible makeup dates

Thursday **March 1, 2018—Lab:** More practice modeling and interpreting linear regressions.

Assignment 2 **DUE: 5 pm on Sunday, March 4, 2018. Submit discussion of potential data sources to Canvas**

Week 6. Prediction III

Read: Imai (2017) Chapter 4

Question: How do we evaluate our models predictions?

Tuesday **March 6, 2018—Lab** Measures of model fit; Obtaining predicted values; Dangers of overfitting and underfitting

Thursday **March 8, 2018—Lab:** A brief introduction to Generalized Linear Models; Who supports the president? (Working with data from the NES and CCES)

Week 7. Probability I

Read: Imai (2017) Chapter 6

Levendusky (2009) Skim Chapter 3

Question: What is probability and how do we use it?

Tuesday **March 13, 2018—Lab**: Axioms of probability; Discrete and continuous probability distributions; Expectations, variance, and moments

Thursday **March 15, 2018—Lab:** Are citizens more polarized? (Replicating and extending the *The Partisan Sort*)

Week 8. Probability II

Read: Imai (2017) Chapter 6

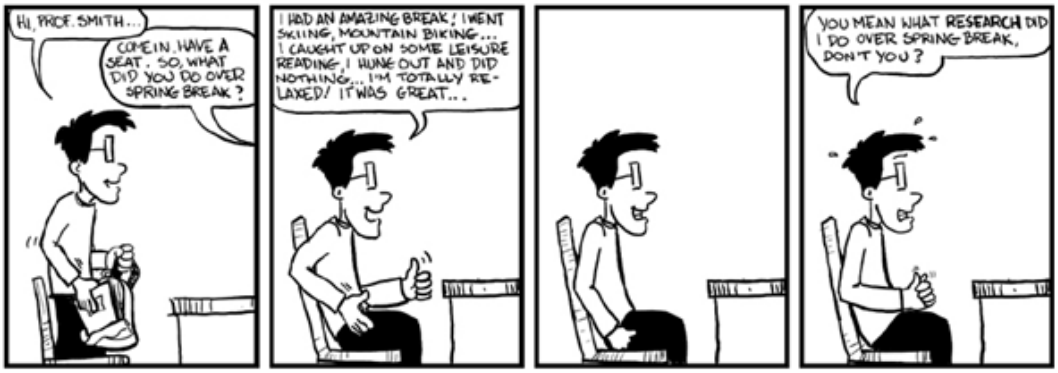
Question: What is probability and how do we use it?

Tuesday **March 20, 2018—Lab** Conditional probability; Bayes Rule; Likelihoods

Thursday **March 22, 2018—Lab:** The Law of Large Numbers and The Central Limit Theorem; More fun with partisanhip in the NES

Assignment 3 **DUE: 5 pm on Sunday, March 25, 2018. Submit outline of research paper on Canvas**

SPRING BREAK!!!



Week 9. Uncertainty I

Read: Imai (2017) Chapter 7

Question: How do we quantify uncertainty?

Tuesday **April 3, 2018—Lab** Asymptotic and simulation based approaches to sampling distributions, standard errors, and confidence intervals.

Thursday **April 5, 2018—Lab:** Exploration of Taubman Poll Data

Week 10. Uncertainty II

Read: Imai (2017) Chapter 7

Question: How do we quantify uncertainty?

Tuesday **April 10, 2018—Lab** Asymptotic and permutation based approaches to hypothesis test and p-values

Thursday **April 12, 2018—Lab:** Exploration of Taubman Poll Data

Week 11. Uncertainty III

Read: Imai (2017) Chapter 7

Question: How do we quantify uncertainty?

Tuesday **April 17, 2018—Lab** Challenges for valid inference. "Robust" estimation procedures. Taking time and space seriously. Testing multiple hypotheses.

Thursday **April 19, 2018—Lab:** Workshop on your drafts

Week 12. Explorations and Final Paper Workshops

Read: Imai (2017) Chapter 5

Question: What do you want to know?

Tuesday **April 24, 2018—Lab** Special topics and review

Thursday **April 26, 2018—Workshop:** Work on your final paper

Assignment 5 **DUE: 5 pm on Sunday, April 29, 2018. Submit draft of final paper to Canvas**

Week 13. Explorations and Final Paper Workshops

Read: Imai (2017) Chapter 5

Question: Can I get an extension?

Answer: No.[2]

Tuesday **May 1, 2018—Lab** Work on your final paper

Thursday **May 3, 2018—Last Class:** Some parting thoughts and last minute questions.

Final Papers due by 11:59 PM, Sunday, May 13, 2018 on Canvas

[2] I mean we can talk, but trust me you'll be much happier finishing on time.

References

Angrist, J. D. and Pischke, J.-S. (2013). Mastering Metrics. *Journal of Chemical Information and Modeling*, 53(9):1689–1699.

Broockman, D. and Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science (New York, N.Y.)*, 352(6282):220–4.

Ferwerda, J. and Miller, N. L. (2014). Political Devolution and Resistance to Foreign Rule: A Natural Experiment. *American Political Science Review*, 108(03):642–660.

Imai, K. (2017). *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press.

Levendusky, M. (2009). *The partisan sort: how liberals became Democrats and conservatives became Republicans*. University of Chicago Press.

Wickham, H. and Grolemund, G. (2017). R for Data Science.