Quantitative Research Methods

Boston University
Political Science 841
Fall 2018

Last revised: September 4, 2018

Professor: Taylor C. Boas
Email: tboas@bu.edu
Office location: 232 Bay State Rd., rm. 311B
Office phone: 617-353-4214
Office hours: Tuesday/Thursday 11–12:30

Lecture location: CAS 204B
Lecture time: Tuesday/Thursday 2–3:15

Teaching Fellow: Luisa Godinez Puig, luisagp@bu.edu. Office hours (rm. 311D): M 12:30–2:30, Th 3:30–4:30
Lab location: PLS 312B
Lab time: Monday 10:10–11:55

Course Description

Quantitative research methods are important tools that political scientists and others use to test empirical claims about the world around us. This course offers an introduction to probability, descriptive statistics, hypothesis testing, and regression analysis, the foundations upon which nearly all quantitative analysis in social science builds. We will place emphasis both on theory, i.e., the concepts, logic, and mathematics underlying statistics, and applications, using software to analyze real data and implement the techniques we have learned. We will also learn how to assess work done by other scholars that uses quantitative methods.

I do not assume that students have any prior mathematical background beyond (a) high school algebra, and (b) the material covered in the BU social science math boot camp. For more advanced topics from the boot camp that students may not have covered prior to BU, such as matrix algebra and calculus, we will review the relevant techniques in class before they are covered on homework and/or exams.

As this is a required course for political science graduate students, it necessarily serves different constituencies. For all students, this course will give you a basic understanding of the tools used for empirical research in much of the social sciences, allowing you to better comprehend articles published in top journals. It will prepare you for the next course in the political science sequence—PO 843: Maximum Likelihood Estimation—as well as other advanced courses that you might take in other departments. For those who are skeptical of quantitative research, this course will help inform your skepticism, showing explicitly the assumptions required for hypothesis testing and regression analysis, why these assumptions are often not met in practice, and what problems arise as a result. For those who find themselves inclined toward statistical analysis, this course will help teach you how to do good quantitative research rather than simply crunching numbers for numbers' sake. Finally, for those who are agnostic about quantitative methods, I hope to convince you that they should be part of your research toolkit!

Software

The required computer software for the course is R (http://www.r-project.org/), an open-source statistical analysis package that is increasingly becoming the standard for quantitative analysis in political science as well as other disciplines. We will also be making use of RStudio (https://www.rstudio.com), an interface that makes R more user-friendly. Both are available for free; you should download and install them on your computers. In addition to being free and open-source, R is powerful, flexible, and has a large user community with plenty of free advice and expertise available on the Internet. It is also the software used in PO 843 and more advanced courses at ICPSR. The Teaching Fellow will be covering R during weekly sections, using labs that are designed to complement lectures and readings.

I will be distributing various files related to the course via Dropbox, a file sharing system. This requires you to install a free application on your computer. On the first day of class I will explain how Dropbox will work for our purposes.

Readings

The required textbooks for the course are:

> Richard J. Larsen and Morris L. Marx, *An Introduction to Mathematical Statistics and Its Applications*, 5th ed. (Prentice Hall, 2012).
>
> Will H. Moore and David A. Siegel, *A Mathematics Course for Political & Social Research* (Princeton, 2013).

A new edition of the Larsen and Marx textbook has recently been published, but not much changes in statistics textbooks from year to year, so I've stuck with the 5th edition, which can be found cheaper online. Moore and Siegel is the textbook used in the BU social science math boot camp, so you should already have it, but if not, please buy it online.

Most statistics textbooks cover a pretty standard set of topics, yet with different levels of mathematical sophistication. Larsen and Marx is a fairly math-oriented textbook. Depending on your background and learning style, you might want to supplement it by also reading on your own from a different book. Here are some options I am familiar with. There are also countless others.

> David S. Moore, George P. McCabe, and Bruce Craig, *Introduction to the Practice of Statistics* (W.H. Freeman). Middle-of-the-road level of mathematical sophistication. I've used this textbook in the past, and if you buy the 7th edition, I can give you an old syllabus with reading assignments from this book that correspond to our topics.

David Freedman, Robert Pisani, and Roger Purves, *Statistics* (W. W. Norton). Very little math; explains almost everything in words.

In addition to the assigned readings from the textbook, for most weeks of the course I have listed a political science journal article that uses one or more of the techniques being covered that week. Most of the time we will discuss these articles in class. The political science journal articles are all available electronically via JSTOR. When you read these, feel free to skim the literature review and theoretical set-up, but you should pay careful attention to the statistical analysis—more so than you would normally—and be prepared to talk about it.

Instruction in statistical computing in R will be based on customized labs that parallel the topics in the course, rather than on a textbook per se. However, you should download the following free book as a very useful reference:

James E. Monogan III, *Political Analysis Using R* (Springer, 2015). http://link.springer.com/book/10.1007/978-3-319-23446-5.

Sections

The Teaching Fellow will teach a weekly 2-hour section. The section will involve instruction in R and review of exams, homework assignments, and theoretical material from lecture. There is no attendance or participation grade related to the section. However, this will be the only place that R, which is required for homework, will be taught. If you don't learn it in section, you will have to learn on your own. In prior years, students also found the review material covered in section to be very valuable.

Problem sets

Problem sets (20% of the final grade), consisting of both math and computer exercises, will be assigned nearly every week and will be due the following week. For math problems, it is fine to turn in handwritten work, but you are encouraged to learn LaTeX to type up the problem sets. LaTeX is required in some advanced courses and is becoming the standard document preparation platform for quantitative work in political science.

Problem sets are graded on a check, check-minus, check-plus basis. It is impossible to learn statistics without practice, so these problem sets are very important. Working in groups is allowed and is often the best way to learn; however, each student must turn in his or her own write-up. If working in a group, make sure you understand all of the answers, because you will be on your own come exam time!

The final homework assignment, due on the last day of class, will consist of writing a short paper (approximately 5 pages) analyzing a data set of your choice using multiple regression. The assignment will be distributed approximately one month before it is due, as there are intermediate deadlines. This assignment counts twice as much as the other problem sets.

Exams

Two midterm exams (20% each) will be held in class, approximately one-third and two-thirds of the way through the course. The final exam (40%) will be held during the scheduled exam period, as noted below. The exams are closed-book, but you will be allowed to bring in a single page (front and back) of notes. The exams are not cumulative—you will not have the same types of problems from midterm 1 on the subsequent exams—but statistics itself is cumulative, so you cannot forget what you learned in the first half of the course and still do well on the final.

You will be allowed to use R on your laptops during the exams, but no other uses of your computer are permitted. The sheet of notes may not contain any worked-out problems from previous years' exams. Any violations of these rules will be considered cheating.

The week after each midterm, the homework assignment will consist of correcting your mistakes and turning in a revised version. To facilitate this, I will return exams with errors highlighted or underlined, but I will not distribute a solution sheet until after the revisions are turned in. The revised exam will be graded the same as other homework and will not affect your midterm exam grade.

Schedule and Readings

Sept. 4, 6: Introduction to Probability

Larsen and Marx, 2.1–2.5.

Moore and Siegel, 9

Christopher H. Achen, "Advice for Students Taking a First Political Science Graduate Course in Statistical Methods," *The Political Methodologist* 10, 2 (Spring 2002): 10–12.

Sept. 11, 13: Random Variables and Expectations

Larsen and Marx, 3.1, 3.3, 3.4, 3.5 (skip Example 3.5.2 and Theorem 3.5.2), 3.6 (stop at "Higher Moments").

Moore and Siegel, 10.1–10.2, 10.4–10.5, 10.7 (skip or skim 10.71–10.72)

Sept. 18, 20: Discrete Probability Distributions

Larsen and Marx, 3.2 (Binomial only), 4.1, 4.2 (stop at "Intervals Between Events").

Moore and Siegel, 10.6 (skim multinomial and negative binomial distributions)

Robert Weissberg, "Collective vs. Dyadic Representation in Congress," *American Political Science Review* 72, 2 (June 1978): 535–547.

Sept. 25, 27: Continuous Probability Distributions and the Central Limit Theorem

Larsen and Marx, 4.3.

Moore and Siegel, 11.1–11.1.3, 11.2–11.2.2, 11.3 (Normal Distribution only)

Donald E. Stokes, "Party Loyalty and the Likelihood of Deviating Elections," *Journal of Politics* 24, 4 (Nov. 1962): 689–702.

Oct. 2: Midterm Exam

Oct. 4, 11: Confidence and Significance

Larsen and Marx, 5.1, 5.3, 5.4 (sections on "Unbiasedness," 313–316, and "Efficiency," 317–319), 6.1–6.3.

Michael Coppedge, "The Dynamic Diversity of Latin American Party Systems," *Party Politics* 4, 4 (1998): 547–568.

Oct. 16: Power and Error

Larsen and Marx, 6.4 (skip the "Power Curves" section and everything beginning with "Decision Rules for Nonnormal Data").

Zaller, John. 2002. "The statistical power of election studies to detect media exposure effects in political campaigns." *Electoral Studies* 21: 297–329.

Oct. 18, 23: Inferences for Means

Larsen and Marx, 7.1–7.2, 7.4 (stop at "When the Normality Assumption Is Not Met"), 9.2.

Thad Dunning and Lauren Harrison, "Cross-Cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali," *American Political Science Review* 104, 1 (Feb. 2010): 21–39.

Oct. 25: Inferences for Tabular Data

Larsen and Marx, 7.5, 10.5 (skip "Testing for Independence: the General Case" and continue with Case Study 10.5.1).

Edward D. Mansfield and Jack Snyder, "Democratization and the Danger of War," *International Security* 20, 1 (Summer 1995): 5–38.

Oct. 30: Midterm Exam

Nov. 1, 6, 8, 13: Simple Linear Regression

Larsen and Marx, 11.1, 11.2 (stop at "Nonlinear Models"), 11.3 (stop at "Drawing Inferences about E(Y|x)").

John Neter, et al., *Applied Linear Regression Models* (Richard D. Irwin, Inc., 1983). Chapter 6: "Matrix approach to simple regression analysis" (skip material on weighted least squares).

Moore and Siegel, 12 (review—covered in math boot camp)

Panagopoulos, Costas. 2013. "Extrinsic Rewards, Intrinsic Motivation and Voting." *Journal of Politics* 75, 1: 266–280.

Nov. 15, 20: Multiple Regression

Eric A. Hanushek and John E. Jackson, *Statistical Methods for Social Scientists* (Academic Press, 1977). Chapter 5: "Multivariate Estimation in Matrix Form."

Douglas L. Kriner and Francis X. Shen, "Iraq Casualties and the 2006 Senate Elections." *Legislative Studies Quarterly* 32, 4 (November 2007): 507–530.

Nov. 27, 29: Violating Regression Assumptions

Peter Kennedy, *A Guide to Econometrics* (MIT Press, 1998). Chapters 3, 6, 7, 8, 9, 10, 11.

Dec. 4: Interaction Terms

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14: 63–82.

Christenson, Dino P., and David M. Glick. 2014. "Chief Justice Roberts's Health Care Decision Disrobed: The Microfoundations of the Supreme Court's Legitimacy." *American Journal of Political Science* 59, 2: 403–418.

Dec. 6: Extra day, TBD

Dec. 11: Conclusion

Dec. 18, 3–5 p.m.: Final exam