

Advanced Methods of Political Analysis

What? POLI706, University of South Carolina, Spring 2014

Who? Tobias Heinrich, heinriect@mailbox.sc.edu, GAMB333. No official office hour; just chat me up or email me.

Who else? Stephanie Davis, sramseydavis@gmail.com, Mondays and Wednesday, 9-10am.

When and where? Mondays, 2.20-5.05pm in GAMB106.

Outline

Unless you are a political philosopher or 100% game theorist, skills of statistical analysis are the *sine qua non* for a modern social scientist. This course is one of the early and many formal and informal stepping stones on your way to learning these skills.

The demands on statistical sophistication in order to be taken seriously and to get published are rising rapidly. Whereas in 1970s, analyzing correlations was fine, by the 1980s simple regressions were the norm. Within the 1990s, likelihood approaches rose. Since the 2000s, structural estimation, Bayesian statistical inference, and causal inference became more prominent albeit not necessarily required. In recent years, machine-learning approaches crept up.

This course helps you along to learning the tools to carry out empirical social science research. The focus will be on the modern, practical aspects of this. To this end, the course features four main building blocks, namely the rigorous introduction of the basics of the linear regression model; classical and modern social scientific inference; carrying out data analyses; and a guide to the wider world of statistical tools.

If you participate actively and take this course *very* seriously, you will learn the following:

- understand the basics of the linear regression model;
- read and comprehend social science research that uses linear regressions as a means for inference;
- learn how to perform inference in a traditional as well as modern way;
- acquire the basic competency for exploring greater bodies of statistical literature to find tools for particular future research needs.

Required material

The only material that requires a purchase is Jeffrey M. Wooldridge's *Introductory Econometrics: A Modern Approach*, 5th edition. All other readings and code examples (in R) will be provided.

The following books are likely to come in handy and are great resources.

- Andrew Gelman, Jennifer Hill, *Data Analysis Using Regression and Multilevel/ Hierarchical Models*.
- Hadley Wickham, *ggplot2: Elegant Graphics for Data Analysis*.

The entire course will revolve around R and LaTeX as the pieces of software to carry out analyses and write up output, respectively. RStudio is the recommended GUI for R.

Expectations

Lest this is not perfectly clear: Unless you participate actively and take every *very* seriously, you will fail to learn the material, and thus by the end will not have acquired the skills to conduct the inferential part of modern social science research. It is the expectation that, if you take this course, that you will immerse yourself in it.

“Immerse yourself”? It means that you carefully read the assigned chapters, immediately raise points that are not clear, devote numerous hours to homework assignments, midterm, and final, and consult with instructor and TA. It ought to go without saying that you should in class on time and every time.

The instructor and the TA are here to help you learn the material—make use of us. That said, for many, many, many problems you will encounter (in particular with R and LaTeX), you will trivially find the answer through a simple search on the internet. Do this, please, before you consult us. After all, learning efficient self-help on these matters will save many headaches in the future.

There are also specific expectations about everything that will be handed in the course. It is this instructor's firm belief that anything you produce and that you deem to be worthy to be shown to someone else should be packaged in a way that it concisely and clearly conveys the content. This means, most often, the use of graphs, formatted tables, properly labeled output, etc.

And: If in doubt about anything, talk to the TA or me at any point.

Grading, Rules, etc.

There are three sets of works that will determine your grade: (roughly) weekly exercises, midterm, and a final paper. Their weighting is 27.5, 27.5, and 45, respectively.

Roughly every week, you have the chance to work through exercises that connect to the lecture (and previous lectures). Each contributes equally. You are highly encouraged to collaborate on this and consult material from outside of the course. What you turn in, however, has to be your own. What matters is that you understand the material; it does not matter how exactly you arrived at the comprehension. Consequently, for any exercise, you have the option(s) to fix your errors. That is, you may redo the parts of the homework that you got wrong and show it to the

TA. Your grade then gets adjusted. That means that, if by the end of the semester you understand everything even though you struggled early in the semester, there is no reason for you not get a 100 on each exercise. The final for getting a grade adjusted is a week after the last day of class (ie. *Cinco de Mayo* 2014).

The midterm (roughly after Week 8) will feature data analysis with interpretation, graphs, etc. It will be a partial paper. The same re-grading rule applies. Grading will be done by the instructor.

The final paper should be a standalone novel piece of research from you that, in principle, could be submitted to a social science journal or be presented at a conference. It has to come with all the bells, whistles, and formalities of a research paper (pitch, theory, literature review, data description, analyses, and graphs as well as the provision of replication material). More details on the paper will be provided as the semester progresses. All issues related to the paper will be handled by the instructor. Over the course of producing the paper, you will be asked to supply various partial steps, each of which needs to get at least an “okay” by the instructor for you to proceed. In this course, you have to submit a one-page sketch of your research idea (by Week 5), and the frontend of the paper (introduction, literature review, and a sketch of the research design) (by Week 10). The final paper is due a week after the last day of class.

In line with the goal to help you learn modern statistical inference, the tools used to carry out analyses should reflect the goal. You are strongly urged to adopt LaTeX and R as the tools to carry out and summarize your work throughout the course. If you opt to use for some other software, the instructor will be unlikely to be able to offer help.

I will use the following grading scale to map between percentages and letter grades. A 92-100, B+ 87-91, B 80-86, C+ 77-79, C 70-76, D+ 67-69, D 60-66, and F 0-59. The standard rules of rounding are applied. The University of South Carolina Honor Code applies.

Reasonable accommodations are available for students with a documented disability. If you have a disability and may need accommodations to fully participate in this class, contact the Office of Student Disability Services: 777-6142, TDD 777-6744, email sasds@mailbox.sc.edu, or stop by LeConte College Room 112A. All accommodations must be approved through the Office of Student Disability Services.

Course schedule

This is a rough schedule of the sequence of what will be covered in class. The instructor will most probably amend things quite a bit. So, always check what will be next.

WEEK 1

Why are we doing this? Why statistical analysis? What do we want? Also: How not to be caught with FUQs.¹

- Wooldridge, Chapter 1.

¹ Fundamentally Unidentified Questions.

WEEKS 2 AND 3

The fundamentals of linear regression.

- Wooldridge, Chapters 2 and 3.

WEEK 4

Basics of programming, scripting and actually doing regressions in R.

WEEK 5

Including qualitative variables and interactions in your regression. Also: One-page pitch of paper is due.

- Wooldridge, Chapter 7.
- Thomas Brambor, William Clark, and Matt Golder, 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14(1).
URL: <http://pan.oxfordjournals.org/content/14/1/63.short>

WEEK 6

Statistical inference, done classically.

- Wooldridge, Chapter 4.

WEEKS 7 AND 8

Statistical inference, done in a modern way.

- Gary King, Michael Tomz, and Jason Wittenberg, 2000. "Making the Most of Statistical Analyses." *American Journal of Political Science* 44(2).
URL: <http://www.jstor.org/discover/10.2307/2669316>
- <http://ggplot2.org>

WEEK 9

Reading and interpreting research articles. In groups, you will carefully read articles and present them, discuss the quality of statistical analysis, and replicate core findings.

WEEKS 10 AND 11

More technical aspects of linear regressions: asymptotics, misspecifications, omitted variables, and model fit. Also: Frontend of paper is due (Week 10).

- Wooldridge, Chapters 5, 6, 8, and 9.

WEEK 12

More about statistical issues: selection effects, serial correlation, etc.

WEEK 13

Looking for more appropriate techniques. In groups, you will carefully read very, very advanced statistics and econometrics articles and then discuss what their goal is and how it could help your research.

WEEK 14

What else is there in stats? What ought to be next for you? Maximum likelihood, machine learning, Thomas Bayes, missing data, ...

- Philip A. Schrodtt, 2013. "Seven Deadly Sins of Contemporary Quantitative Political Analysis." *Journal of Peace Research*.
URL: <http://jpr.sagepub.com/content/early/2013/10/08/0022343313499597.full>
- Hal Varian, 2013. "Big Data: New Tricks for Econometrics."
URL: <http://www.aeaweb.org/aea/2014conference/program/retrieve.php?pdfid=495>.