This version: April 2, 2018

Political Science 271
Advanced Statistical Applications

Spring Quarter 2017
SSB 104, Wednesday 3-5:50PM

Molly Roberts
SSB 399
meroberts@ucsd.edu
Office Hours: Monday, 1–3 PM

Prerequisites

Political Science 270 (or equivalent)

Overview

This course is the third course in the quantitative research methods sequence at the UCSD Political Science department. Building on Political Science 270 and Political Science 273, this course teaches advanced statistical tools for empirical political science. The course is focused around building models for various research tasks. First, we start out with models for causal inference, reviewing some of the material in Political Science 273. However, unlike 273, we will spend more time focusing on how these models are estimated. Next, we will build models for prediction and will discuss how models for prediction differ from those aimed at estimating causal effects. Last, we will briefly cover models for measurement and models for discovery, and discuss how these models can be used within the research process.

Assessment

There are no written exams in the class, and your grade will be based on a combination of:

- **Homeworks (50%):** Five problem sets will be given throughout the quarter, skewed heavily toward the beginning of the quarter. Problem sets will contain analytical, computational, and data analysis questions. Each problem set will be counted equally toward the calculation of the final grade. The following instructions will apply to all problem sets unless otherwise noted.

- Late submission will not be accepted unless you ask for special permission from the instructor in advance. Problem set write-ups should be turned in in hard copy, a separate copy of the problem set write-up and code will be turned in electronically.
- Working in groups is encouraged for conceptual and sometimes technical discussion, but each student must submit their own writeup of the solutions that shows their independent work on the assignment. In particular, you should not copy someone else's answers or computer code. We also ask you to write down the names of the other students with whom you solved the problems together on the first sheet of your solutions. At times, the instructor will specify that for particular problems or problem sets that students should not work with others.
- For analytical questions, you should include your intermediate steps, as well as comments on those steps when appropriate. For data analysis questions, include annotated code as part of your answers. All results should be presented so that they can be easily understood.

- **Final project (40%)**: The final project will be a poster and short research memo which typically applies a method learned in this course to an empirical problem of your substantive interest. The memo should outline a research paper that could potentially be written after the class has been completed.
  - I encourage you to work with another student on your poster and memo. By co-authoring you will (1) learn how to effectively collaborate with someone else on your research, which is very important in political science where most cutting-edge research is collaborative and (2) more likely have a good, potentially publishable paper (multiple brains are usually better than one).
  - Unless you already have a concrete research project suitable for this course (e.g., from your dissertation project), we recommend that you start with replicating the results in a published article and then improve the original analysis using the methods learned in this course (or elsewhere). Oftentimes, the most time-consuming part of a research project is data collection (which is not the focus of this course) and using data someone has already archived for their publication and made publicly available gets around this problem.

- Students are expected to adhere to the following deadlines:
  - April 25: Turn in a brief description of your proposed project. By this date you need to have found your coauthor, acquired the data you plan to use, and completed a descriptive analysis of the data (e.g. simple summary statistics, crosstabs and plots). Meet with the instructor to discuss your proposal during her office hours. You may be asked to revise and resubmit the proposal in two weeks from the meeting.
  - June 6: Poster session: Class time will be spent in a poster session on this day where students present the results of their paper and comment on one another's work.
  - June 6: Memo due. Please turn in one printed copy of your memo by the end of the day, and email electronic copies to the instructor.

- **Participation and presentation (10%):** Students are strongly encouraged to ask questions and actively participate in discussions during lectures and recitation sessions.

Academic Honesty and Plagiarism

All of your graded work must be done by you. If you are unfamiliar with the University's policy on academic integrity, please see http://students.ucsd.edu/academics/academic-integrity/policy.html.

Syllabus and Plan

The syllabus will be updated periodically throughout the course, so that we can keep with the cadence of the class. I will post to Piazza when such updates are made.

Reading and Textbooks

We will read chapters from these books throughout the course. We recommend that you purchase the King book. The others we will only read a few chapters from, and will be available on electronic reserve.

- James, Gareth, Daniela Witten, Trever Hastie and Robert Tibshirani. *An Introduction to Statistical Learning with Applications to R.* Springer-Verlag, 2013. (available online at: `http://www-bcf.usc.edu/~gareth/ISL/ISLR\%20First\%20Printing.pdf`.)
- King, Gary. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference.* University of Michigan Press, 1998.
- Ward, Michael and Ahlquist, John. *Maximum Likelihood Strategies for Social Science.* Cambridge University Press, 2017.

Piazza

We will be using Piazza for general discussion and questions and answers throughout the class. Piazza allows students to see other students' questions and learn from them as well as answer them. Your respectful and thoughtful participation in the discussion forum will count toward your participation grade. Please do not e-mail the instructor with questions (post them on Piazza!) unless they are personal in nature. I will check the Piazza forum daily to provide my own answers and contributions.

Software

We will be using `R` an open-source statistical package. You can download it from the web here:

`http://cran.r-project.org/`

COURSE SCHEDULE

1 April 3: Course Introduction and Review of Causal Inference

2 April 11: Linear Regression Reframed and Basic Probability

Reading

Chapter 1, Chapter 3 King

3 April 18: Maximum Likelihood

Reading

Chapter 4.1-4.3, King

Chapter 1, Ward and Ahlquist

4 April 25: Optimization, Uncertainty, and Properties of MLE

Reading

Chapter 4.3-4.6, King

Chapter 2, Ward and Ahlquist

5 May 2: Binary Dependent Variables and Event Count Models

Reading

Chapter 5, King

Chapter 3, 8, and 10, Ward and Ahlquist

6 May 9: Models for Prediction

Reading

Shmueli, Galit. 2010. "To Explain or to Predict?" *Statistical Science*, Vol. 25, No. 3.

Sections 2.1.1 and 2.1.2, Chapter 5 and 6, James et al.

7 May 16: Models for Prediction

Reading

Chapter 5 and 8, James et al.

8 May 23: Models for Measurement

Reading

Chapter 10.1-10.2, James et al.

Poole, Keith T., and Howard Rosenthal. "A spatial model for legislative roll call analysis." *American Journal of Political Science* (1985): 357-384.

9 May 30: Models for Discovery

Reading

Chapter 10.3, James et al.

Imai, Kosuke, and Dustin Tingley. "A statistical method for empirical testing of competing theories." *American Journal of Political Science* 56.1 (2012): 218-236.

10 June 6: Poster Session