PS231A - QUANTITATIVE ANALYSIS IN POLITICAL RESEARCH

Department of Political Science • UC Berkeley

FALL 2019

Class Time:	`Th 9:00-11:59AM`
Location:	`791 Barrows`
Instructor:	`Joel A. Middleton`
Email:	`joel.middleton@berkeley.edu`
Phone:	`202-431-1412`
Office:	`728 Barrows`
Office hours:	`M 10:30-11:59AM or by appointment`
Section time:	`Th 5:00-6:59PM`
Location:	`791 Barrows`
GSI:	`Samuel Trachtman`
Email:	`sam.trachtman@berkeley.edu`

OVERVIEW

This course will be an introduction to statistical methods for political researchers. Throughout the semester I will strive to provide a rigorous introduction to statistics while eschewing strong parametric assumptions. The goal is to lay the foundations for clear thinking about statistical inference that will aide you in future classes such as PS 231b and PS 236a/b.

I will assume that students have passed the math camp exam and that they have a strong grasp of calculus and matrix algebra. I will also assume some exposure to linear regression and the ability to work in `R` for the problem sets.

Not to worry, I generally find that students who do not exactly meet these requirements can be successful. Each one of you is capable of mastering the material, so I generally welcome anyone willing to give it their all. However, we wont spend time reviewing foundational concepts.

I use `R` in this class. In past iterations I have allowed Stata, however, sticking with `R` will avoid any translation issues and also help prepare you for future courses in this department.

I expect you to come to class having read the assigned reading for that day. Class period will be an opportunity to clarify the material and discuss your questions. As much as possible, I am going to ask your peers to answer your questions as I think it is useful to have a chance to try to explain concepts.

BOOKS

The required book for this course will be *Theory of Agnostic Statistics* by Aronow and Miller (AM). It will be distributed via the course website. You may want to have a copy of this printed and bound. Copy Central on Bancroft has the book on file and will print these on demand. It should cost approximately $25.

I will also suggest related readings from four other books. You may choose to purchase these books or not. For a more traditional political science perspective on statistics, I recommend *Statistical Modeling and Inference for the Social Sciences* by our own Sean Gailmard (G). A good book for a first course in probability is *Introduction to Mathematical Statistics*, 7th edition by Hogg, McKean and Craig (HMC). For a perspective on regression that is similar to Aronow and Miller, but which goes deeper into econometrics, I recommend the out of print (but available used) *A Course in Econometrics* by Goldberger (GB). For a book on causality, I recommend *Causal Inference for Statistics, Social and Medical Sciences* by Imbens and Rubin (IR).

I wont be distributing my notes for several reasons. First, I'll be writing on the board most of the time because I think this is the right speed for this material. As such I wont have slides to distribute. Second, I think the act of taking notes in class helps learn the material. Third, the book has the material written out with a great degree of care and the supplementary books are another good resource should you find both the class notes and the book difficult.

GRADES

The course grade will be determined as follows: 60% problem sets; Midterm 15%; Final 20%; Participation 5%.

PROBLEM SETS

There will be one problem set for each of the six sections of the course. Dates in the Course Schedule with an asterisk indicate that a problem set will be due.

Students should plan to work on problem sets together. You will, however, submit your problem sets individually and the goal of working together should be that you become able to solve similar problems on your own.

Working together should only be a problem if I believe you are simply copying the work of another.

COURSE SCHEDULE

I will make every effort to adhere to the schedule posted below. Adjustments may be necessary. We may cover some topics in less detail if we fall behind or add a topic if we we advance more quickly than expected.

August 29, September 5, 12, 19 - Probability Theory and Random Variables

Reading: AM Chapters 1 & 2
Further reading: G Chapters 1-4, 5; HMC Chapters 1, 2; GB Chapters 2-6

September 26, October 3- Learning from Random Samples

Reading: AM Chapter 3
Further reading: G Chapters 7-9; HMC Chapters 4, 5; GB Chapters 8-11

October 10, 17 - Regression

Reading: AM Chapter 4
Further reading: GB Chapters 13-17

October 24 - Parametric Models

Reading: AM Chapter 5
Further reading: G Chapter 6; HMC Chapter 3

October 31 - Midterm Exam

November 7 - Parametric Models (continued)

Reading: AM Chapter 5 (continued)
Further reading: G Chapter 6; HMC Chapter 3

November 14, 21- Missing Data

Reading: AM Chapter 6

December 5 - Causal Inference

Reading: AM Chapter 7
Further reading: IR Chapters 1-7

Final Exam

Assigned: December 16, 9AM
Due: December 17, 6PM