POLS 206: Political Methodology II

Winter 2020

Course meetings:

T R 2:30 - 3:55pm, Ellison 3814 (Pritchett Room)

Instructor:

Paasha Mahdavi, paasha@ucsb.edu, Ellison 3807.

Office Hours: R 1:30 - 2:30pm, or by appointment.

Course description and objectives

This class is the second course in a three-course quantitative methods sequence for political science graduate students, picking up where POLS 205 left off. Our broader goal is to understand how to use statistical analysis to make evidence-based claims on political science issues. Our focus is on ordinary least squares (OLS) multivariate regression, with an introduction to non-linear models with binary outcomes. Topics we will cover include research design, sampling, hypothesis testing, omitted variable bias, measurement error, diagnostics, and other issues related to model specification. Throughout the quarter we will apply these topics and tools to real political science problems both in class and in take-home assignments and projects.

By the end of the quarter you will be able to ...

1. Understand how to apply data-based arguments to questions in political science;
2. Interpret results from multivariate regression analysis;
3. Diagnose problems related to data integrity and model specification;
4. Identify potential challenges in drawing conclusions from statistical findings;
5. Use `R` to conduct basic and intermediate statistical analysis.

Prerequisites

This course is intended for doctoral students in political science, but exceptions can be made for graduate students from other departments at UCSB. POLS 205 (Political Methods I) is the formal prerequisite for this course. Familiarity with substantive concepts in political science, policy, economics, and political economy is highly desirable but not required. I also assume that you have a working knowledge of `R`, such as that obtained in POLS 205.

Grading and course requirements

1. Problem sets: 50%
2. Take-home final exam: 50%

Required Books

- Michael A. Bailey. 2015. *Real Stats: Using Econometrics for Political Science and Public Policy*. Oxford University Press. 1st edition.
- Garrett Grolemund and Hadley Wickham. 2017. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media.
    - Available at `https://r4ds.had.co.nz/`
- Hadley Wickham. 2019. *Advanced R*. Chapman & Hall/CRC Press. 2nd edition.
    - Available at `https://adv-r.hadley.nz/`

Suggested Books

- Damodar N. Gujarati. 2005. *Essentials of Econometrics*. 3rd Edition. Boston: Irwin/McGrawHill.
- Andrew Gelman and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Jeffrey M. Wooldridge. *Introduction to Econometrics: A Modern Approach*. Cengage Learning. 5th edition.
- Gary King, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton University Press.
- Stephen L. Morgan and Christopher Winship. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press. 8th edition.
- Neil Weiss. *Introductory Statistics*. Pearson. 9th edition.
- Joshua D. Angrist and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Kosuke Imai. 2017. *A First Course in Quantitative Social Science*. Princeton University Press.
- Jane Miller. *The Chicago Guide to Writing about Multivariate Analysis*. 2nd edition.
- Peter Aronow and Benjamin Miller. *Theory of Agnostic Statistics* (September 6, 2015, edition)

Course Webpage

Additional readings for the course are papers and book chapters which can be downloaded from the course webpage on GauchoSpace. I will post lecture handouts on the course webpage prior to each course topic. In-class applications will be posted online before each class as well.

Problem Sets

Problem sets will be posted on the course webpage at least one week before they are due. I will also post any data for problem sets on the course webpage. There will be five problem sets throughout the quarter. They will be graded as complete / partial / incomplete. To earn full credit, you must show compelling evidence of effort to solve each problem thoroughly.

These assignments are perhaps the most important part of the course – statistics is a subject that absolutely requires learning by doing! Start working on problem sets as soon as they are posted. You can work in groups but each person must submit their problem sets individually. Note that group study can be an excellent way to learn statistics but you must put in proper effort to fully grasp the material; simply showing up to a group study session without having reviewed and worked through the problem set is a recipe for failure. After problem sets are turned in, we will review selected answers to problem sets during class.

Weekly empirical summaries

In addition to problem sets, you will post to the discussion forum on GauchoSpace a short empirical summary of a published article on Monday evenings every week, starting in Week 2. Preferably, the article you choose should come from your weekly reading assignments in the other courses you are taking this quarter. The summary should be a bullet-point list of the article's (1) causal relationship of interest, (2) research design, (3) mode of statistical inference, and (4) shortcomings related to any of these three concepts. Although these paragraphs are ungraded, I will be reading them each week and, at times, discussing them (and the corresponding articles) in class.

Take-home Final Exam

The final exam will involve replication and extension of an existing study. You will be asked to replicate the study's main findings and then use their data to extend the analysis using tools and techniques learned in the course. The end product will be a paper that resembles the "methods and results" section of a published journal article. You can use either the dataset I provide you, or (preferably) a dataset of your choosing that is published in a known repository online (e.g., DataVerse, ICPSR, etc.). It will be due during Finals week.

Course Plan

*Tentative; subject to change as the quarter develops. Most of the required readings are from Real Stats (**RS**), and all pages within each chapter are to be read unless indicated otherwise.*

Week #	Topic	Readings	Deliverables
Introduction			
1	Inference & probability review	RS 1–2; HMP 7.5–7.9	**Project plan**
Regression basics			
2	Multivariate OLS review	RS 3–5.1; Gujarati 8.1–8.3	**Problem set 1**
3	Dummy variables, transformations	RS 6–7.3; Benoit (2011)	**Problem set 2**
Biased estimates and misspecification			
4	Omitted variables	Angrist & Pischke 8; RS 5.2 & 14	**Problem set 3**
5	Measurement error, outliers, heteroskedasticity	RS 5.3–5.5; Wooldridge 8.1–8.3, 9.4; Gujarati 13	
6a	F-tests, R^2, diagnostics	RS 7.4; Gujarati 8.4, 8.8	**Replication**
Heterogeneity			
6b	Interactions	RS 10; Wooldridge 7.4; Brambor et al (2011); Hainmueller et al (2018)	**Problem set 4**
Distributions and Limited Dependent Variables			
7	Sampling theory	Wooldridge Appx C; KKV 208–230	
8	LPM, Logit, Probit	RS 12; Gelman and Hill 5	**Extension**
9	Generalized linear models	Angrist & Pischke 3.4.2; Gelman and Hill 6	**Problem set 5**
Concluding concerns			
10	Research design revisited	RS 16; Morgan and Winship (2014)	
11			**Final exam**