POLI 783: Introductory Statistics

University of North Carolina at Chapel Hill
Fall 2016

Instructor

Stephen Gent
Office: Hamilton 352
Email: gent@unc.edu
Office Hours: Tuesdays 11:00am-12:00pm; Thursdays 1:00–3:00pm

TA/Lab Instructor

Dan Gustafson
Office: Hamilton 303
Email: dgustaf@live.unc.edu
Office Hours: Mondays & Wednesdays 9:00–10:30am

Course Information

Class: Tuesdays & Thursdays 9:30–10:45am, Hamilton 351
Lab: Thursdays 11:00–11:50am, Hamilton 351
Course Website (Sakai): https://sakai.unc.edu/portal/site/poli783f16

Course Description

Over the last half century, political scientists have come to rely increasingly on quantitative methods for their research. This trend has occurred in virtually all subfields and is ongoing, as one can see in the publication of ever more complex quantitative analysis in the top-tier journals. Given this development, statistical literacy has become indispensable for political scientists. Whether you just want to keep up with the literature in your field or want to be engaged in your own quantitative research program, a solid knowledge of statistical methods has become just as important as knowledge of the theories in your field. Learning statistics may not be what you had in mind when you decided to attend graduate school in political science, but the truth is that much of the substantive literature in the discipline will remain inaccessible without a thorough understanding of the language of statistics. POLI 783 teaches you that language, so that you can pursue more fruitfully your substantive interests.

The main emphasis in the course is on those concepts and procedures that lie at the heart of all statistical analysis. Thus, we discuss the key principles of statistical inference along with statistical measures that are widely used in political analysis. A thorough understanding of this material is essential if you want to become a competent producer and/or consumer of quantitative research in political science.

Course Objectives

Provided that you do the work that is required, you can expect to learn the following skills from POLI 783:

- You will learn fundamental concepts that underlie statistical analysis.
- You will learn the logic of statistical inference and discover how this logic can be used to draw inferences about the political world.
- You will learn how to translate substantive research questions into statistical terms, so that the language and procedures of statistics can inform your own research.
- You will obtain some hands-on experience in data analysis, including an introduction to Stata and R.
- You will be prepared for more advanced courses in statistics and data analysis.
- You will learn the mathematical skills you need to know in order to succeed in this and other graduate methods classes in the department.

Prerequisites

There are no prerequisites for POLI 783, except that basic knowledge of algebra is assumed. Prior course work in statistics is helpful, although it should be noted that POLI 783 goes considerably beyond the material covered in most undergraduate statistics courses.

Knowledge of basic calculus and familiarity with mathematical notation will be necessary for POLI 783. The lab will cover all of the mathematics that you will need in this course. Below I list a couple of reference books that might be useful to review the mathematics commonly used in political science research.

Grading

Students will receive one combined final grade for both the class and the lab.

Homework (50%)

You are expected to complete a series of homework assignments for the class during the semester. Some of these assignments involve data analysis using Stata, while others require you to solve analytical statistical problems. You are expected to submit each class homework assignment by 9:30 am on the day it is due. Late submission of these assignments will result in a grade reduction.

There will also be some homework assignments for the lab over the course of the semester. The homework assignments will be rather small and are intended to help you keep up a gradual study of these mathematical concepts rather than resorting to cramming sessions.

Exams: Midterm (15%) and Final (35%)

Two closed-book, in-class exams will be given during the semester. A mid-term exam will take place in class on Tuesday, October 25, while a cumulative final exam will take place on Tuesday, December 13, in Hamilton 351 at 8:00am (as scheduled by the university). The exams will only cover the material from the statistics coursework, not the lab. However, the exam will require you to use the some of the math skills learned in the lab.

Collaboration and UNC Honor Code

You can learn a great deal from your classmates, and I encourage you to discuss the course materials, including the homework assignments. However, there is a fine line between accepting help and having someone else do the work for you, or between giving help and finishing someone else's work. For your own benefit you should avoid crossing this line. Plagiarism is a UNC Honor Code violation and will be prosecuted accordingly.

Course Materials

The following books are required for POLI 783:

- Pollock, Philip H. 2016. *The Essentials of Political Analysis*, 5th ed. CQ Press.
- Pollock, Philip H. 2015. *A Stata Companion to Political Analysis*, 3rd ed. CQ Press.
- Westfall, Peter H., & Kevin S. S. Hennnig. 2013. *Understanding Advanced Statistical Methods*. CRC Press.

If you are interested in an additional resources on math for political science:

- Jonathan Kropko. 2016. *Mathematics for Social Scientists*. Sage Publications.
- Will H. Moore & David A. Siegel. 2013. *A Mathematics Course for Political and Social Research*. Princeton University Press.

Course Philosophy

POLI 783 forms a foundation for all of your future quantitative methods courses in the PhD program. My primary goal as the instructor of this course is to make sure that every first-year student gains a solid understanding of key concepts in statistical methods. I will aim to pace the course accordingly. If this requires us to slow down and spend extra time on some topics, that is not a problem. I would rather you gain a firm grasp of key, central concepts than a shallow conception of a broad range of topics.

Below I provide a list of topics to cover in this course, but not a specific schedule. We may cover some topics in less than one class period; other topics will likely require multiple days. We may not be able to get to the last set of topics on the list. Don't worry about this. We will focus on the most critical concepts first. You will learn about the topics that we don't get to in future courses.

Words of Advice

For some of you, this may be your first course in statistics. I recognize this and have designed the course with this in mind. My main piece of advice is to do all that you can to keep up. Much of the material covered in this course is cumulative, so it can be difficult if you fall behind. Make sure that you keep up with the reading and the homework problems. Don't be afraid to ask questions in class. If you need more explanation about something, you can be sure someone else in the class is thinking the same thing. If you need additional help, don't hesitate to come to see me or Dan in our office hours.

On the other hand, for some of you, the first part of the course may seem relatively easy, especially if you've already taken a course in research methods as an undergraduate or in a masters' program. I would advise you not to be lulled in a sense of complacency, or you may quickly find yourself falling behind once we hit more advanced topics. In my experience, there is always something to learn about even the most basic statistical concepts, and a firm understanding of statistical methods generally requires you to revisit these concepts multiple times during your career as a graduate student and beyond.

Course Organization

Part I: Introductory Statistics

The first part of the course is a brief introduction to the use of statistics in political science. This will take up the first few weeks of the semester. Starting with substantive research questions, we ask what statistical analysis can do to provide insights into the political world. You will also learn how to use the statistical software package Stata. This part of the class is pitched at the level of a undergraduate social statistics class: it provides a "cookbook" approach to statistics that requires little mathematics and no prior experience with statistics. The main text for Part I is Pollock's *Essentials of Political Analysis*. The *Stata Companion* will be used for homework assignments in this part of the course.

Part II: Mathematical Statistics

The second part is the heart of the course. It gets behind the "cookbook" statistics, so to speak. It delves into the mathematical statistical theory that underlies the techniques discussed in Part I, focusing in particular on the logic of statistical inference. To this effect, Part II addresses probability theory, sampling distributions, the theory of hypothesis testing, and estimation theory. This sketch of the statistical model of inference goes well beyond a typical undergraduate social statistics class; it is much more abstract and mathematically demanding. This part of the course lays the essential foundation for more advanced courses in applied data analysis. The main text for Part II is Westfall and Henning (WH).

Part I: Introductory Statistics

Descriptive Statistics

- Levels of measurement, measures of central tendency and dispersion
- Pollock: Ch. 2
- *Recommended: Read or skim Pollock: Intro, Ch. 1, 3–5.*

Inferential Statistics

- Sampling, confidence intervals, significance tests
- Pollock: Ch. 6–7 (pp. 123–170)

Correlation and Linear Regression

- Correlation, OLS regression
- Pollock: Ch. 8

Logistic Regression and MLE

- Logistic regression, maximum likelihood estimation
- Pollock: Ch. 9

Part II: Mathematical Statistics

Discrete Random Variables

- Discrete probability distribution functions, Bernoulli, binomial, Poisson
- WH: Ch. 2.1–2.3; Ch. 3.1–3.2
- *Recommended: Read WH Ch. 1*

Continuous Random Variables

- Continuous probability distribution functions, cumulative distribution functions, uniform, normal
- WH: Ch. 2.4, 2.6; Ch. 3.2

Multivariate Distributions

- Conditional distributions, marginal distributions, joint distributions, independence, Bayes' Theorem
- WH: Ch. 5.1–5.2, 5.4, 5.6; Ch. 6.1–6.2, 6.4–6.6

Expected Value, Variance, and Functions of Random Variables

- Expected values, distributions and expectations of functions of random variables, Jensen's inequality, variance, Chebyschev's inequality, covariance, correlation
- WH: Ch. 8.1–8.3; Ch. 9.1–9.9; Ch. 10.1–10.2

Moments and Moment Generating Functions

- Moments, moment generating functions
- See Sakai for reference materials

Sampling Distributions

- Sampling, law of large numbers, central limit theorem, normal family (Chi-squared, Student's t, and F distributions)
- WH: Ch. 7.1; Ch. 8.4, 8.6; Ch. 10.4; Ch. 16.1–16.2, 16.4, 16.6, 16.9

Estimation

- Estimators, bias, consistency, efficiency, confidence intervals
- WH: Ch. 11; 14.1–14.2; 16.6–16.7

Hypothesis Testing

- Hypothesis tests, types of error, power
- WH: Ch. 15; 16.8, 16.10; 18.3

Preview of Future Courses (if time)

- Least squares estimation
- Likelihood functions, maximum likelihood estimation
- Bayesian statistics

Lab Topics

The lab will cover material that complements and enriches the content covered in the course lectures. Lab sessions early in the semester will review all of the mathematics that will be needed in the course. The lab will also provide an introduction to the R statistical programming package, which you will use in future courses in the methods sequence. Examples of topics that will be covered in the lab include:

- Basic number theory, set theory & functions
- Probability theory
- Differential calculus
- Integral calculus
- Distributions and central limit theorem
- Programming and data management (in R and Stata)
- Applying statistical principles to research design

Important Dates

No Class:

- Thursday, September 1 (APSA)
- Thursday, October 20 (Fall Break)
- Thursday, November 24 (Thanksgiving)

Exams

- Midterm exam: October 25 (in class)
- Final exam: Tuesday, December 13 at 8:00am in Hamilton 351