Introduction to Statistics in Political Science

Political Science 812 Fall 2016

Lecture Location Van Vleck 8337
Lecture Time Monday and Wednesday, 1:20–2:35 p.m.
Section Location Van Hise 274 Computer Lab
Section Times Friday, 9:55–10:45 a.m. and 11:00–11:50 a.m.

Instructor Dave Weimer
Email weimer@lafollette.wisc.edu
Office 215 North Hall
Office Hours Mondays and Wednesdays, 10–11:30 a.m. and 2:45–4 p.m.

Teaching Assistant Dmitrii Kofanov
Email kofanov@wisc.edu
Office Hours Mondays and Wednesdays, 11:30 a.m.–12:30 p.m., North Hall 121

Overview

Political scientists employ increasingly sophisticated statistical methods. Understanding these methods—and new ones that will undoubtedly become available—requires a firm foundation in mathematical statistics. This course is intended to provide this foundation so that students can continue their methods training with subsequent courses in the department (PS 813 and PS 818) as well as other advanced courses and, most importantly, through independent learning. It also provides some applications that illustrate concepts and introduce students to empirical political science research.

I seek to help every student achieve basic competence in the material. Please use my office hours to get help on the material covered in lecture. Do NOT allow yourself to fall behind. If necessary, I am willing to slow the pace to keep everyone moving forward. However, I cannot make adjustments unless you communicate to me any problems you are having with the material.

Textbook

The primary textbook for this course is


Sections

Weekly sections will focus primarily on statistical computing, including instruction in using statistical software and practical computer exercises. Time will also be set aside to go over problem sets.
Statistical computing

The course will give you experience with two computational resources. One resource is the widely used statistical package, Stata, which has capabilities for implementing stochastic simulations. The other resource is R, an implementation of the S statistical programming language. It can be downloaded for free from http://www.r-project.org/.

Grading

Grading will be divided between problem sets (15 percent), a midterm exam (25 percent), a final exam (50 percent), and a data analysis report (10 percent).

Problem sets

Short problem sets will be handed out in class, typically due the following week. The problem sets will cover both theory and application. You are welcome to discuss the problem sets with each other and run programs together, but the final write-ups should be your own. Also, note that simply copying Stata or R output without reformatting is not appropriate. In addition to the problem sets, skill in R will be developed through translation of Stata warm-up exercises.

Midterm exam

There will be an in-class midterm on October 26. In addition to counting towards your final grade, the exam should serve as an indicator of your progress in the course.

Final exam

There will be a cumulative final exam held during exam week. The date will be scheduled early in the semester.

Data analysis report

Students will complete a report employing basic methods to answer an empirical question of their own choosing. Data will typically come from a common political science data set (American National Election Study, Correlates of War, etc.). A literature review is unnecessary. Papers should be roughly five pages with appendices as needed. They should be submitted at the last class (December 14).

Topics and readings

The dates are tentative and will be adjusted to reflect our progress in learning. I suggest you read through the material before class and again after it is discussed in class. Even a quick skim of the material beforehand is very beneficial.
Introduction and Overview (Sept. 7)
Overview of estimation, inference, and presentation in political science; frequentist and subjectivist interpretations

Introductory case: The Butterfly Ballot


L&M, Chapter 1

Probability Foundations (Sept. 12, 14, 19, and 21)
Laws of probability
Bayes’ theorem
Decision analysis

L&M, Chapter 2


Calculus Review (Sept. 26, 28, and Oct. 3)
Differential calculus
Integral calculus

Random Variables (Oct. 5, 10, and 12)
Probability mass and density functions
Cumulative distribution functions
Central and Non-central Moments
Moment Generating Functions

L&M, pp. 118–162

Common Univariate Distributions (Oct. 17 and 19)
Bernoulli, binomial, and hypergeometric distributions
Poisson and negative binomial distributions
Exponential, gamma and beta distributions
Normal distribution


Midterm Exam (Oct. 26)
*Multivariate Distributions* (Oct. 24 and 31)
Bivariate and multivariate distributions
Marginal distributions
Conditional distributions
Mathematical covariance and correlation
Bivariate normal distribution
Functions of random variables

L&M, pp. 162–193

*Stochastic Simulation* (Nov. 2)
Monte Carlo Simulation
Agent-based models

L&M, pp. 274–278


*Limits and Asymptotic Distributions* (Nov. 7 and 9)
Probability limits
Law of large numbers
Central limit theorem
Normal approximation to the binomial distribution

L&M, pp. 246–258

*Desirable Properties of Estimators* (Nov. 14 and 16)
Bias
Efficiency
Mean squared error
Consistency

L&M, pp. 312–322, 330–333

*Maximum Likelihood* (Nov. 21 and 23)
Maximum likelihood
Method of moments
Properties of maximum likelihood estimators

L&M, pp. 281–287, 290–297
Classical Inference and Hypothesis Testing (Nov. 28, 30, and Dec. 5)
Introduction to hypothesis testing
Neyman-Pearson lemma
Tests of hypothesis about parameters of normal populations

L&M, Chapter 6, pp. 385–418, 457–480

Analysis of Categorical Data (Dec. 7)
Contingency tables
Chi-square test
Fisher exact test

L&M, pp. 519–528


Introduction to Ordinary Least Squares (Dec. 12 and 14)
Linear statistical models
Bivariate ordinary least squares

L&M, Chapter 11