Introduction to Statistics in Political Science

Political Science 812 Fall 2014

Lecture Location Ogg Room, 422 North Hall
Lecture Time Tuesday & Thursday 9:30am–10:45pm
Section Location SSCC Computer Lab, 3218 Social Sciences
Section Times Friday 11–11:50am

Instructor Alexander Tahk
Email atahk@wisc.edu
Office 212 North Hall
Office Hours Monday 1:30–3:30pm

Teaching Assistant Sarah Bouchat
Email bouchat@wisc.edu
Office Hours Monday 3–5pm (Peet's Coffee, Memorial Union) & Wednesday 6–8pm (Fair Trade Coffee)

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 self-learning. It will also provide some applications that illustrate concepts and introduce students to empirical political science research.

Textbooks

The primary textbook for this course is:


However, you do *not* need the most recent edition of this textbook. An earlier edition should cover the same material and may be less expensive.

There are also two free textbooks that may be useful. These are not as complete as DeGroot & Schervish. The first book is very terse but includes most of the material we will cover and integrates R. The second provides better explanations but skips many topics we will cover.


Many other treatments of this material are also available. Two good possibilities are:


Either of these should be an adequate substitute for DeGroot and Schervish if desired.

For a bit more or less rigorous treatment, respectively, consider:


For a mathematical review, I also recommend the following book:


**Sections**

Weekly sections will be held in the SSCC Computer Lab. These sections will focus primarily on statistical computing, including instruction in using statistical software and practical computer exercises.

**Statistical computing**

Computational components of the problem sets will make use of R, an implementation of the S statistical programming language. It can be downloaded for free from [http://www.r-project.org/](http://www.r-project.org/). We will also use Stata at the end of the course when we cover linear regression.

**Grading**

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

**Problem sets**

There will be short problem sets handed out in class, typically one every Tuesday and due on the following Tuesday (unless otherwise noted on the problem set) to the mailbox of your TA, Sarah Bouchat. These will be graded on a check-plus/check/check-minus/zero basis. Late assignments are strongly discouraged. A pattern of late assignments will result in a grade penalty. Assignments more than one week late will not be accepted.

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 R or Stata output without reformatting is not appropriate.
Midterm exam

There will be an in-class midterm on October 23. 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 during the first week of classes.

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 and are due on the last class (December 11).

Prerequisites

This course has no formal prerequisites. However, you are assumed to have been exposed to differential calculus and basic integral calculus. No background in linear algebra is needed.

Topics and readings

The syllabus is organized around topics rather than by day. We will typically spend around one week per topic, but may spend more or less. I suggest you read through the material before class and again after its discussed in class. Even a quick skim of the material beforehand is very beneficial.

Introduction and overview

Overview of estimation, inference, and presentation in political science
Frequentist and subjectivist interpretations
Introductory case: the butterfly ballot

Reading: IPSUR, Chapters 1 & 2
OIS, Chapters 1.1–1.5
DS, Chapters 1.1–1.4

Probability foundations

Laws of probability
Random variables
Bayes’ theorem

Reading: IPSUR, Chapter 4
OIS, Chapters 2.1–2.2
DS, Chapters 1.5–1.11, 2.1–2.3
Probability distributions

Probability mass functions
Probability density functions
Cumulative distribution functions
Common discrete density functions

Reading: IPSUR, Chapters 5.1.1, 5.7, 6.1.1 & 6.4
DS, Chapters 3.1–3.3

Common univariate distributions

Bernoulli, binomial, and hypergeometric distributions
Poisson and negative binomial distributions
Normal distribution
Exponential, gamma and beta distributions

Reading: IPSUR, Chapters 5.2, 5.3, 5.6, 6.2, 6.3, & 6.5
OIS, Chapter 3
DS, Chapters 5.1–5.8

Bivariate distributions

Bivariate and multivariate distributions
Marginal distributions
Conditional distributions
Multinomial distribution
Bivariate normal distribution

Reading: IPSUR, Chapter 7.1, 7.3, 7.4 & 7.6
DS, Chapters 3.4–3.7, 5.9–5.10

Basics of mathematical statistics

Mean and variance
Covariance and correlation
Non-central and central moments
Functions of a random variable

Reading: IPSUR, Chapters 5.1, 5.4, 6.1 & 7.2
OIS, Chapter 2.4
DS, Chapters 3.8–3.9, 4.1–4.7

Limits and asymptotic distributions

Probability limits
Law of large numbers
Central limit theorem
Normal approximation to the binomial distribution

Reading: IPSUR, Chapter 8
OIS, Chapter 4.4
DS, Chapter 6.1–6.3

**Point estimation**

Bias
Consistency
Mean squared error

*Reading:* IPSUR, Chapter 9.1
OIS, Chapters 4.1
DS, Chapter 7.1, 8.7

**Maximum likelihood**

Maximum likelihood
Method of moments
Properties of maximum likelihood estimators

*Reading:* DS, Chapter 7.5–7.6

**Inference and hypothesis testing**

Introduction to hypothesis testing
Neyman-Pearson lemma
Difference of means

*Reading:* IPSUR, Chapter 10.1–10.4
OIS, Chapters 4.3 & 5
DS, Chapters 9.1–9.2, 9.4–9.6


**Interval estimation**

Introduction to confidence intervals
Large-sample confidence intervals

*Reading:* IPSUR, Chapter 9.2–9.5
OIS, Chapters 4.2
DS, Chapter 8.5

**Analysis of categorical data**

Contingency tables
Chi-square test
Fisher exact test

*Reading:* IPSUR, Chapter 14
OIS, Chapter 6
DS, Chapter 10.1–10.5

**Analysis of variance**

One-way analysis of variance

*Reading:* IPSUR, Chapter 10.6  
DS, Chapter 11.6  

**Introduction to ordinary least squares**

Linear statistical models  
Bivariate ordinary least squares

*Reading:* IPSUR, Chapter 11  
OIS, Chapter 7  
DS, Chapters 11.2–11.3  