Quantitative Methods for Political Science I
Government 1000

Professor Jasjeet Singh Sekhon
Jacob Kline, TF

Class: 4-6pm Mondays
CBRSS Conference Room (34 Kirkland St.)

Assistant Professor Jasjeet Singh Sekhon
jasjeet.sekhon@harvard.edu
HTTP://jsekhon.fas.harvard.edu/
Phone: 496-2426, Fax: 496-5149
Office: CBRSS, 34 Kirkland St., Room #4
Office hours: 2–3pm Wednesdays, or by appointment, or by email anytime

Jacob Kline, Teaching Fellow
jkline@fas.harvard.edu
Phone: TBA
Office: TBA
Office hours: TBA
Sections: TBA

Catalog Description

Half Course. Introduction to major quantitative techniques used in political science. Covers exploratory data analysis, as well as descriptive and causal statistical inference of many types. The course emphasizes probability theory, least squares, robust estimation, matching and propensity scores. Note: Frequently taken by graduate students satisfying department requirements. Undergraduates looking for an introductory class should consider government 1001.

Discussion

We use an email list in this class quite frequently: mailto:gov1000-list@fas.harvard.edu If students have a question about the course material, they are advised to email the question to the entire list and ask for help. Letting everyone see each other’s questions and the answers improves everyone’s work. The primary focus of this course is to teach students (primarily undergraduates) enough statistical methods so they may conduct original quantitative empirical work on their own.

Prerequisites

High school algebra.
If you need a refresher you should try to get an introductory algebra book, such as from the Schaum’s Outline Series, and use it as a reference while doing homework assignments.

Evaluation

The only way to learn statistics is by doing. There is no substitute for answering questions. Therefore, there will be regular class assignments and two exams. Grades will be based on a weighted average of the midterm, final exam, regular class assignments and participation in both class and section. More information will be provided in class.

It is highly recommended that students form study groups in order to complete the homework assignments. Although it is recommended that people work together in order to complete the homework assignments, each individual student must hand in their own answers. Photocopies (or equivalent reproductions) of someone else’s answers are not acceptable.

Course Books and Software

Students may use any computer package to complete the assignments they wish. However, we recommend that students use one of the two main variants of the S statistical programming language: R or Splus. Both are installed on Harvard-MIT Data Center computers. R is available for a large variety of platforms for download: [http://www.r-project.org/](http://www.r-project.org/). R is open source software (released under the GNU public license) and is available at no charge. Splus is a commercial product and is available at no cost to Harvard students.

The five books listed below are required and available at various online bookstores. The first book is the main text book for the course. The second book is the secondary text and offers more intuition but less detail than the first. The third is a computer manual which does a nice job of linking theory with practice. The last two books supplement our discussion of least squares and logistic regression. The core course material will be communicated in lectures and associated notes and handouts. The textbooks are important reference guides which help communicate the concepts.

Course Plan and Tentative Outline

We will not be able to cover all of this material. But we will cover as much as we can. There is no serious time constraint because any material left over will be covered in Government 2000. We do not read many articles, but the articles we do read will be read carefully over several weeks.

1. The Nature of Statistics
   (a) Observational studies versus randomized experiments
   (b) Causal inference versus description
   (c) Methods of inductive inference

Readings:
- Ch. 1 of Wonnacott and Wonnacott
- Ch. 1, 2 of Freedman et al.

These readings will be examined closely over several weeks. They will get clearer as we learn more math.


2. Descriptive Statistics (in Section)
   (a) Frequency tables and graphs
   (b) Center of a distribution
   (c) Spread of a distribution
   (d) Linear Transformations
   (e) Calculations using relative frequencies

Readings:
- Ch. 2 of Wonnacott and Wonnacott
- Ch. 3, 4 of Freedman et al.
- Ch. 1 of Dalgaard

3. Probability
   (a) Probability models
   (b) Compound events
   (c) Conditional probability
   (d) Independence
   (e) Conditional Independence
(f) Bayes Theorem
(g) Other views of probability

Readings:

- Ch. 3 of Wonnacott and Wonnacott
- Ch. 13, 14, 15 of Freedman et al.

4. Probability Distributions

(a) Discrete random variables
(b) Mean and variance
(c) The binomial distribution
(d) Continuous distributions
(e) The normal distribution
(f) Functions of random variables

Readings:

- Ch. 4 of Wonnacott and Wonnacott
- Ch. 5, 6 Freedman et al.
- Ch. 2, 3 of Dalgaard

5. Two Random Variables

(a) Distributions
(b) A function of two random variables
(c) Covariance
(d) Correlation (see section 15-1 of Wonnacott and Wonnacott)
(e) Linear combination of two random variables

Readings:

- Ch. 5 of Wonnacott and Wonnacott
- Ch. 8, 9 Freedman et al.
- Ch. 4 of Dalgaard

6. Rules of Variances and Expectations (in Section)

See lecture notes.

7. Experimental Manipulation.

Readings: These readings will be examined closely over several weeks. They will get clearer as we learn more math.


8. Sampling

(a) Random Sampling
(b) Moments of the sample mean
(c) The shape of the sampling distribution
(d) Proportions (percentages)

Readings:

- Ch. 6 of Wonnacott and Wonnacott
- Ch. 19–21 Freedman et al.
- Ch. 7 of Dalgaard

9. Point Estimation

(a) Unbiasedness and Consistency
(b) Ordinary least squares (OLS)
(c) Robustness, least median of squares

Readings:

- Ch. 7 of Wonnacott and Wonnacott
- Ch. 7, 10 Freedman et al.
- Ch. 5 of Dalgaard

10. Confidence intervals (Chapter 8)

(a) A single mean
(b) Difference in two means, independent samples
(c) Proportions
(d) Confidence intervals and tests for $\beta$
(e) Small-sample $t$

11. Hypothesis Testing (Chapter 9)

(a) Hypothesis testing using confidence intervals
(b) $p$–values
(c) Classical hypothesis tests
(d) Bayesian confidence sets
12. Point Estimation and Simple Regression (Chapters 11–12)
   (a) The regression model
   (b) Sampling variability
   (c) Predicting $Y$ at a given level of $X$

Reading: Achen, pages 1-37

13. Multiple Regression and Regression and Correlation Compared (Chapters 13 and 15)
   (a) The regression model and its OLS fit
   (b) Regression coefficients as multiplication factors
   (c) Simple and multiple regression compared
   (d) Correlation and regression
   (e) Two regression lines
   (f) Correlation in multiple regression
   (g) Multicollinearity

Readings:
   • Achen, pages 37ff

14. Binary Regression

Readings:
   • Pampel, *Logistic Regression: A Primer*.

15. Propensity Scores

Readings:
   • We revisit Gerber and Green (2000) and Imai.

16. Regression Extensions (Chapter 14)
   (a) Non-spherical errors

17. Instrumental Variables (lecture notes).

18. Simulating Quantities of Interest: A Cautionary Tale
   Reading: King, Gary and Michael Tomz and Jason Wittenberg. *Making the Most of Statistical Analyses: Improving Interpretation and Presentation*.
Readings:


Supplementary Reading Material

The following books are certainly not required, but they may be of interest during the course. It is often very useful to read the same material covered by a variety of authors. Within each section, books are approximately ordered by increasing sophistication.

Computer Books


Mathematical Review


Probability and Statistics

Linear Regression and Applications


