

# POL2504H1F Statistics for Political Scientists

Wednesday 12-2pm, Bloor Street West-371 (FE), Rm.36

Christopher Cochrane  
Assistant Professor  
Department of Political Science  
University of Toronto

office: SS 3110  
hours: W 2-4pm  
email: christopher.cochrane@utoronto.ca  
phone: (416) 208-5121

## 1 Course Description

This course examines the theory and application of statistics in political science. At the end of this course, students will be able to confidently and properly interpret, assess and employ the most commonly used statistical methods in political science research. Coverage includes: probability theory and scientific epistemology; concept development and measurement; project workflow and data management; descriptive, associational, and causal inference; estimation and statistical significance; and linear and logistic regression. The course assumes no prior training in mathematics or statistics.

## 2 Assessment

Assessment is based on course participation and six written analytical assignments.

### 2.1 Participation - 10%

The free exchange of information is a cornerstone of scientific research. There are a number of discussion forums (e.g., Statalist), websites (e.g., UCLA Statalab), and programs (e.g. CLARIFY) through which quantitative researchers share their expertise with those who need help. The participation component of your grade in this course is designed to foster and reward this ethic. Your grade is based on the quality and quantity of your contributions to in-class discussion, as well as your contributions to questions, answers and discussions on the class web page.

### 2.2 Assignments - 90%

The assignments in this course are designed to improve your understanding of statistical methods and concepts, and to prepare you to write a major quantitative research paper. Assignments#3-6 should be accompanied by Stata syntax files and original unaltered versions, in Stata format, of the data file(s) that you have used in these assignments. I should be able to replicate entirely your analysis by saving your data files to my computer and then running your Stata syntax. Assessment is based in part on the ease and accuracy of the replication.

#### 2.2.1 Assignment Schedule and Weighting

Assignment#1. Probability Theory	Begins Sept.15	Due Sept.29 (10%)
Assignment#2. Hypotheses	Begins Sept.29	Due Oct.6 (15%)
Assignment#3. Data Summary	Begins Oct.6	Due Oct.13 (10%)
Assignment#4. Association	Begins Oct.27	Due Nov.3 (15%)
Assignment#5. Linear Regression	Begins Nov.10	Due Nov.24 (25%)
Assignment#6. Logistic Regression	Begins Nov.24	Due Dec.1 (15%)

### 2.2.2 Late Penalties and Extensions

The late penalty for all coursework is 5% for each day, or part thereof, that the work is late (including weekends). However, students who are ill, or who have dependents who are ill, are entitled to a combined maximum of 7 days of undocumented extension, provided that they notify me, prior to the assignment due date, that their work will be late. In the interests of fairness, students who require more than 7 days worth of extension over the course of the semester will have to provide documentation, as per University policy, to justify the extra time. Note, however, that in the case of Assignment#1, students are not permitted to attend the third class unless they have submitted the assignment.

### 2.2.3 Statement on Academic Integrity

From The Handbook on Academic Integrity:

Honesty and fairness are considered fundamental values shared by students, staff and faculty at the University of Toronto. The University's policies and procedures that deal with cases of cheating, plagiarism and other forms of academic misconduct, are designed to protect the integrity of the institution and to maintain a community where competition is fair. As a result, U of T treats cases of academic misconduct very seriously. If it has been alleged that you committed an academic offense, you will find that the allegation is dealt with formally and seriously, and that the penalties can be severe if it is determined that you did cheat. All of the policies and procedures surrounding academic offenses are dealt within one policy: The Code of Behaviour on Academic Matters (the 'Code'). This booklet on Academic Integrity is intended to supplement the Code, but not to take its place as the official document on these matters. Nor does this booklet take the place of legal counsel. The full text of the Code of Behaviour on Academic Matters can be found in your Faculty Calendar or online at [www.utoronto.ca/govcncl/pap/policies/behaveac.html](http://www.utoronto.ca/govcncl/pap/policies/behaveac.html). The purpose of this booklet is to: (1) Outline clearly and simply what academic offenses are, to help you avoid committing one unwittingly; (2) Give you a sense of what to expect should you be suspected of committing an academic offense; (3) Inform you of your rights and responsibilities with respect to the procedures under the Code. As a student, you are responsible for ensuring the integrity of your work and for understanding what constitutes an academic offense. If you are not sure if your actions or methods are acceptable, always ask your instructor. Your instructor can explain, for example, the nuances of plagiarism and how to use secondary sources appropriately; he or she will also tell you what kinds of aids calculators, dictionaries, etc. are permitted in a test or exam. Ignorance of the rules does not excuse cheating or plagiarism.

## 3 Texts and Materials

Required and recommended books for this course are available at the University of Toronto Bookstore. The Department of Political Science has purchased Stata for the computers in FE36, as well as for the computers in the Political Science Graduate Students' Computer Lounge in Sid Smith. Additionally, you may purchase Stata for your own computers via the Stata GradPlan at the University of Toronto.<sup>1</sup>

---

<sup>1</sup>Students at the University of Toronto are able to purchase software from Stata at greatly reduced prices. In order to benefit from these prices, order online via the 'GradPlan' section of the Stata website <http://www.stata.com/order/new/edu/gradplans/cgpcampus-order.html>. After 24 hours of receiving an email confirmation of your purchase from Stata, you may pick up your software by taking a copy of the confirmation to the Licensed Software Office, Information Commons, Robarts Library, 1st Floor (130 St. George Street) Phone: 416-978-4990, [lic.software@utoronto.ca](mailto:lic.software@utoronto.ca), between 9:30am and 5:00pm, Monday - Friday. You will be asked to pay an additional \$3 handling fee (as of July 25, 2010) when picking up your software. You may need to present your UTORid to staff at the Information Commons.

### 3.1 Required

Mlodinow, Leonard. 2008. *The Drunkard's Walk: How Randomness Rules Our Lives*. Toronto, ON: Random House of Canada.

Pollock III, Philip H. 2008. *The Essentials of Political Analysis*, 3rd Edition, (including *A Stata Companion to Political Analysis*). Washington, DC: CQ Press.

### 3.2 Recommended

Acock, Alan C. 2010. *A Gentle Introduction to Stata*, 3rd Edition. College Station, TX: State Press.

Levitt, Steven, and Stephen J. Dubner. 2005. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York, NY: Harper Collins.

Stata I/C, Release 11 [software]. \$179 (US) for single-user perpetual license [via GradPlan]; \$65 (US) for single-user 6-month lease [via GradPlan].<sup>2</sup>

Stat/Transfer, Release 10 [software]. \$69 (US) for single-user perpetual license [via GradPlan].<sup>3</sup>

A high quality (e.g., Patriot Xporter©) USB Key [hardware], or some other device to back-up your data and documents.

## 4 Class Schedule

**Class Attendance:** Due to the lab component, class attendance is especially important. Students who miss a class will be behind. Moreover, students who begin assignments outside of class are likely to require assistance from others, as they will not have received the lab instruction to help with the assignment. In the event that you have to miss a class, please use the discussion board on the class website to catchup. I encourage you to use this board whenever you have a question, and to monitor this board regularly for an opportunity to answer the questions of others. I will monitor the discussion and interject when absolutely necessary, though, ideally, I would encourage you to answer each other's questions to the best of your abilities.

**Readings:** Many of the readings for this class are available electronically. You may link to these readings in this document by clicking on the title of the article that you wish to download. Readings that are not available electronically will be made available via other means. An -R- indicates that the reading is recommended, but not required.

### 4.1 September 15: Introduction

#### 4.1.1 Discussion

What role do statistics play in the different subfields of political science? What role could they play?

#### 4.1.2 Lab

Getting Started with Data Management and Stata

---

<sup>2</sup>A single-user license may be installed on more than one computer. However, the end-user license agreement stipulates that you may not have the software *installed* simultaneously on more than three computers, and that you may not *use* the software simultaneously on more than one computer.

<sup>3</sup>Stat/Transfer allows you to easily transfer data files from one format (e.g., SPSS) to another (e.g., Stata). This program will save you a lot of time if you are planning to use statistical software in the course of your dissertation research. Unlike Stata, however, Stat/Transfer is not the kind of program that one would normally use on a regular basis. Thus, the availability of Stat/Transfer on the computers in the Political Science Graduate Computer Lounge will likely suffice for most students.

### 4.1.3 Readings

- POL2504 Course Syllabus.
- King, Gary. 1991. On Political Methodology. *Political Analysis* 2(1): 1-30.

### 4.1.4 Assignments

Assignment 1 begins (10%)

## 4.2 September 22: Scientific Theories and Hypotheses

### 4.2.1 Discussion

What are the defining characteristics of a scientific theory? How do we adjudicate between rival scientific explanations of the same phenomenon?

### 4.2.2 Lab

From Concepts to Variables

### 4.2.3 Readings

- Hempel, Carl G. 1966. *Philosophy of Natural Science*. Englewood Cliffs, N.J.: Prentice Hall, Inc., 3-32.
- Friedman, Milton. [1953] 1969. *Essays in Positive Economics*. Chicago, IL: The University of Chicago Press, 3-43.
- Kuhn, Thomas S. [1962] 1970. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press, 1-42.
- Pollock III, Philip H. 2008. *The Essentials of Political Analysis*, 3rd Edition. Washington, DC: CQ Press, ch.1 & ch.3.
- Popper, Karl. [1935] 2002. *The Logic of Scientific Discovery*. New York, NY: Routledge, 37-56.
- -R- Hempel, Carl G. 1966. *Philosophy of Natural Science*. Englewood Cliffs, N.J.: Prentice Hall, Inc., 110 pages.
- -R- Przeworki, Adam and Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York, NY: John Wiley and Sons, 17-30.

## 4.3 September 29: Introduction to Probability

### 4.3.1 Discussion

What are the implications of Mlodinow's arguments for empirical research in the social sciences?

### 4.3.2 Lab

Constructing Hypotheses

### 4.3.3 Readings

Mlodinow, Leonard. 2008. *The Drunkard's Walk: How Randomness Rules our Lives*. New York, NY: Vintage Books, 222 pages.

#### **4.3.4 Assignments**

Assignment 1 ends - Due at the beginning of class

Assignment 2 begins (15%)

### **4.4 October 6: Descriptive Inference I / Summary Statistics**

#### **4.4.1 Class**

What is the purpose of a graph? What are the characteristics of effective (and ineffective) graphs?

#### **4.4.2 Lab**

Summarizing Data with Tables, Figures and Statistics

#### **4.4.3 Readings**

- Agresti, Alan, and Barbara Finlay. 1997. *Statistical Methods for the Social Sciences*. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 35-79.
- Kastellec, Jonathan P., and Eduardo L. Leoni. 2007. Using Graphs Instead of Tables in Political Science. *Perspectives on Politics* 5(4): 755-771.
- Pollock III, Philip H. 2008. *The Essentials of Political Analysis*, 3rd Edition. Washington, DC: CQ Press, ch.2.
- Mitchell, Michael N. 2008. *A Visual Guide to Stata Graphics*, 2nd Edition. College Station, TX: Stata Press, 1-32.

#### **4.4.4 Assignments**

Assignment#2 ends - Due at the beginning of class.

Assignment#3 begins (10%)

### **4.5 October 13: Descriptive Inference II / Probability and Sampling**

#### **4.5.1 Discussion**

What does probability theory imply about case selection?

#### **4.5.2 Lab**

Estimation

#### **4.5.3 Readings**

- Pollock III, Philip H. 2008. *The Essentials of Political Analysis*, 3rd Edition. Washington, DC: CQ Press, ch.5 & ch.6.
- Ritchie, Ferris J. 1999. *The Statistical Imagination*. New York, NY: McGraw Hill, 154-220.

#### **4.5.4 Assignments**

Assignment#3 ends - Due at the beginning of class.

## 4.6 October 20: Associational Inference I / Measures of Association

### 4.6.1 Discussion

How are descriptive and associational inferences related? How are they different?

### 4.6.2 Lab

Measures of Association

### 4.6.3 Readings

- Pollock III, Philip H. 2008. *The Essentials of Political Analysis*, 3rd Edition. Washington, DC: CQ Press, ch.7.
- Rosato, Sebastian. 2003. The Flawed Logic of Democratic Peace Theory. *American Political Science Review* 97(4): 585-602.
- Slantchev, Branislav L., Anna Alexandrova, and Erik Gartzke. 2005. Probabilistic Causality, Selection Bias, and the Logic of the Democratic Peace. *American Political Science Review* 99(3): 459-462.

## 4.7 October 27: Associational Inference II / Hypothesis Testing

### 4.7.1 Discussion

How are associational and causal inferences related? How are they different?

### 4.7.2 Lab

Hypothesis Testing

### 4.7.3 Readings

- Holland, Paul W. 1986. Statistics and Causal Inference. *Journal of the American Statistical Association* 81(December): 945-960.
- Charney, Evan. 2008. Genes and Ideologies. *Perspectives on Politics* 6(2): 299-319.
- Alford, John R., Carolyn L. Funk, and John R. Hibbing. 2008. Beyond Liberals and Conservatives to Political Genotypes and Phenotypes. *Perspectives on Politics* (6)2: 321-328.
- Charney, Evan. 2008. Politics, Genetics, and 'Greedy Reductionism.' *Perspectives on Politics* 6(2): 337-343.
- -R- Levitt, Steven, and Stephen J. Dubner. 2005. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York, NY: Harper Collins.

### 4.7.4 Assignments

Assignment#4 begins (15%)

## 4.8 November 3: Linear Regression I / Correlation and Linear Regression

### 4.8.1 Discussion

What are the strengths and limitations of regression as a tool in social science research? How might Achen and King respond to McGregor?

### 4.8.2 Lab

Interpreting and Using Regression

### 4.8.3 Readings

- Pollock III, Philip H. 2008. *The Essentials of Political Analysis*, 3rd Edition. Washington, DC: CQ Press, ch.8.
- Achen, Christopher H. 1982. *Interpreting and Using Regression*. Newbury Park, CA: Sage Publications. [78 pages]
- McGregor, James P. 1993. Procustes and the Regression Model: On the Misuse of the Regression Model. *PS: Political Science and Politics* 26(4): 801-804.
- King, Gary. 1991. 'Truth' is Stranger than Prediction, More Questionable than Causal inference. *American Journal of Political Science* 35(4): 1047-1053.

### 4.8.4 Assignments

Assignment#4 ends - Due at the beginning of class.

## 4.9 November 10: Linear Regression II / Model Assumptions and Diagnostics

### 4.9.1 Discussion

OLS models are built on a number of assumptions. Which of these assumptions are likely to be frequently violated in social science research? What are the implications of these violations? What are some possible solutions or workarounds?

### 4.9.2 Lab

Regression Diagnostics

### 4.9.3 Readings

- Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis* 14(1): 63-82.
- Chatterjee, Samprit, and Ali S. Hadi. 2006. *Regression Analysis by Example*, 4th Edition. Hoboken, NJ: John Wiley and Sons, 85-105.
- King, Gary. 1986. How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science. *American Journal of Political Science* 30(3): 666-687.
- King, Gary, Michael Tomz, and Jazon Wittenberg. 2000. Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science* 44(2): 341-355.

#### **4.9.4 Assignments**

Assignment#5 begins (25%)

### **4.10 November 17: Logistic Regression I / Regression for Binary Dependent Variables**

#### **4.10.1 Discussion**

What are the potential benefits and drawbacks of working with binary as opposed to linear dependent variables?

#### **4.10.2 Lab**

Regression for Binary Dependent Variables

#### **4.10.3 Readings**

- Pampel, Fred C. 2000. *Logistic Regression: A Primer*. Thousand Oaks, CA: Sage, 1-39.
- Pollock III, Philip H. 2008. *The Essentials of Political Analysis*, 3rd Edition. Washington, DC: CQ Press, ch.9.

### **4.11 November 24: Logistic Regression II / Postestimation and Diagnostics**

#### **4.11.1 Lab**

Post-estimation and Presentation

#### **4.11.2 Readings**

- Long, J. Scott, and Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata*. College Station, TX: Stata Press, 131-181.
- King, Gary. 1995. Replication, Replication. *PS: Political Science & Politics* 28(3): 443-449.

#### **4.11.3 Assignments**

Assignment#5 ends - Due at the beginning of class

Assignment#6 begins (15%)

### **4.12 December 1: Heads Up! Cross-Sectional, Time-Series, and Multi-Level Analyses**

#### **4.12.1 Discussion**

What are the benefits and challenges of cross-time, cross-level and cross-sectional data?



#### 4.12.2 Readings

- McNabb, David E. 2004. Research Methods for Political Science: Quantitative and Qualitative Methods, 287-301.
- Snijders, Tom and Roel Bosker. 1999. Multilevel Analysis: An Introduction to Advanced Multilevel Modeling. Thousand Oaks, CA: Sage Publications, 6-16.
- -R- Beck, Nathaniel, and Jonathan N. Katz. 1995. What do do (and not to do) with Time-Series Cross-Section Data. American Political Science Review 89(3): 634-647.
- -R- Achen, Christopher H., and W. Phillips Shively. Cross-Level Inference. Chicago, IL: University of Chicago Press, 1-28.
- -R- Stimson, James A. 1985. Regression in Space and Time: A Statistical Essay. American Journal of Political Science 29(4): 914-947.

#### 4.12.3 Lab

Replication

#### 4.12.4 Assignments

Assignment#6 (15%) ends - Due at the beginning of class.