POL2504H1F Statistics for Political Scientists

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

Christopher Cochrane
Assistant Professor
Department of Political Science
University of Toronto

office: SS 3124
office hours: F 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, three written analytical assignments, and an in-class final exam.

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 - 60%

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 & 4 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.

2.3 Exam - 30%

The exam in this class is designed to test your understanding of the logic, application, interpretation, and limitations of key statistical methods in real-world research scenarios.

2.3.1 Assignment Schedule and Weighting

Assignment#1. Probability Theory	Begins Sept.14	Due Sept.28 (15%)
Assignment#2. Data Summary	Begins Oct.5	Due Oct.19 (15%)
Assignment#3. Final Report	Begins Oct.19	Due Nov.23 (30%)

2.3.2 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 extention, provided that they notify me, prior to the assignemnt due date, that their work will be late. In the interests of fairness, students who require more than 7 days worth of extention 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 class unless they have submitted the assignment.

2.3.3 Exam

The exam for this course will be held during the final class.

2.3.4 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 Universitys 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. 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 in-stock at major book suppliers (amazon, indigo, etc), or, alternatively, are available online via the class page on Coursepeer. 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.[1]

[1]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

R and RStudio are open-source and freely available. There are a number of excellent introductions to R that are available, for free, on-line.

You will no doubt have many questions about Stata and R as you work your way through statistics and programming. All of the questions that you have, however, are likely to have been asked and answered on one of the many discussion boards and communities devoted to Stata (e.g., Statalist) and R (e.g., The R section on Stackoverlow). Thus, when you have questions about these programs, the first stage is to look for an answer on the web; if that does not work, the second stage is to post that question on our class webpage; and, if that does not work, the third stage is to send your question to me. Please do not post questions to online forums without first going through these three other stages. Also, if you find an answer to your question, I would appreciate if you post your questions and a link to the answer on the class webpage.

3.1 Required Readings

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

Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences, 4th Edition. Upper Saddle River, NJ: Pearson Education Inc.

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

3.2 Optional

Mitchell, Michael N. 2012. A Visual Guide to Stata Graphics, 3rd Edition. College Station, TX: Stata Press.

3.3 Required Software

Stata I/C, Release 11 or 12 [software]. $179 (US) for single-user perpetual license [via GradPlan]; $65 (US) for single-user 6-month lease [via GradPlan].[2]
R: A Language and Environment for Statistical Computing. Free open source software that you can download from http://www.r-project.org/.

3.4 Recommended Software

Stat/Transfer, Release 10 or 11 [software]. $69 (US) for single-user perpetual license [via GradPlan].[3]

RStudio. Free open source software that you can download from http://rstudio.org/.
Any code editor for R (e.g., Tinn-R for PC or any Mac OS equivalent).

Licensed Software Office, Information Commons, Robarts Library, 1st Floor (130 St. George Street) Phone: 416-978-4990, lic.software@utoronto.ca, between 9:30am and 5:00pm, Monday - Friday. You may need to present your UTORid to staff at the Information Commons.

[2]A single-user license may be installed on more than one computer. However, the end-user license agreement stipulates that although 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.

[3]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 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 14: Introduction

4.1.1 Lab

None

4.1.2 Readings

- POL2504 Course Syllabus.

4.1.3 Assignments

Assignment 1 begins (15%)

4.2 September 21: NO CLASS

4.3 September 28: Probability and the Scientific Method

4.3.1 Lab

Getting Started with Stata

4.3.2 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.
- Lakatos, Imre. 1970. "Falsification and the Methodology of Scientific Research Programmes." In Imre Lakatos and Alan Musgrave (eds.), Criticism and the Growth of Knowledge. New York, NY: Cambridge University Press, 91-195.
- Popper, Karl. [1935] 2002. The Logic of Scientific Discovery. New York, NY: Routledge, 37-56.

- Mlodinow, Leonard. 2008. The Drunkard's Walk: How Randomness Rules our Lives. New York, NY: Vintage Books, 222 pages.
- Acock, Alan C. 2010. A Gentle Introduction to Stata. College Station, TX: Stata Press. Chapter 1.

4.3.3 Assignments

Assignment 1 ends - Due at the beginning of class

4.4 October 5: Measurement and Data

4.4.1 Lab

Summarizing Data

4.4.2 Readings

- Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 1-71.
- Acock, Alan C. 2010. A Gentle Introduction to Stata. College Station, TX: Stata Press, 21-86; 317-348.

R Blaydes, Lisa, and Drew A. Linzer. 2012. "Elite Competition, Religiosity, and Anti-Americanism in the Islamic World." American Political Science Review 106(2): 225-243.

4.4.3 Assignments

Assignment 2 begins

4.5 October 12: Descriptive Inference

4.5.1 Lab

Estimation

4.5.2 Readings

- Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 73-182.
- Acock, Alan C. 2010. A Gentle Introduction to Stata. College Station, TX: Stata Press, 87-115.

4.6 October 19: Associational Inference

4.6.1 Lab

Associational Inference

4.6.2 Readings

- Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 182-254.
- Acock, Alan C. 2010. A Gentle Introduction to Stata. College Station, TX: Stata Press, 117-178.

R Holland, Paul W. 1986. Statistics and Causal Inference. Journal of the American Statistical Association 81(December): 945-960.

R Charney, Evan. 2008. Genes and Ideologies. Perspectives on Politics 6(2): 299-319.

R 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.

R Charney, Evan. 2008. Politics, Genetics, and 'Greedy Reductionism.' Perspectives on Politics 6(2): 337-343.

4.6.3 Assignments

Assignment 2 ends - Due at the beginning of class
Assignment 3 begins

4.7 October 26: Correlation and Regression

4.7.1 Lab

Correlation and Regression

4.7.2 Readings

- Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 255-320.
- Acock, Alan C. 2010. A Gentle Introduction to Stata. College Station, TX: Stata Press, 179-241.
- 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 November 2: Multivariate Regression

4.8.1 Lab

Multivariate Regression

4.8.2 Readings

- Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 321-411.
- Achen, Christopher. 2002. Toward a New Political Methodology: Microfoundations and ART. Annual Review of Political Science 5: 423-50.

4.9 November 9: Regression Diagnostics and Postestimation

4.9.1 Lab

Post-Estimation

4.9.2 Readings

- Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 412-441.
- Acock, Alan C. 2010. A Gentle Introduction to Stata. College Station, TX: Stata Press, 243-288.
- Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. Understanding Interaction Models: Improving Empirical Analyses. Political Analysis 14(1): 63-82.
- 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.

4.10 November 16: Logistic Regression

4.10.1 Lab

Logistic Regression

4.10.2 Readings

- Acock, Alan C. 2010. A Gentle Introduction to Stata. College Station, TX: Stata Press, 289-315.
- Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 483-518.
- Long, J. Scott, and Jeremy Freese. 2006. Regression Models for Categorical Dependent Variables Using Stata. College Station, TX: Stata Press, 131-181.

4.11 November 23: Time Series, Multilevel Models, and Missing Data

4.11.1 Lab

Time Series, Multilevel Models, and Missing Data: Nuisance and Opportunity

4.11.2 Readings

- Acock, Alan C. 2010. A Gentle Introduction to Stata. College Station, TX: Stata Press, 349-370.
- Agresti, Alan, and Barbara Finlay. 2009. Statistical Methods for the Social Sciences. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 519-544.
- Little, Roderick J.A. 1992. Regression with Missing Xs: A Review. Journal of the American Statistical Association 87(420): 1227-1237.
- -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- Stimson, James A. 1985. Regression in Space and Time: A Statistical Essay. American Journal of Political Science 29(4): 914-947.

4.11.3 Assignments

Assignment 3 ends - Due at the beginning of class

4.12 November 30: Statistical Computing

4.12.1 Lab

Getting Started with R

4.12.2 Readings

- Zuur, Alain F., Elena N. Ieno, and Erik H.W.G. Meesters. 2009. A Beginner's Guide to R. New York, NY: Springer.

4.13 Exam Review: TBA

4.14 Optional Computing Primers: TBA

4.15 December 7: Final Exam