PLAD 7100
Political Research with Quantitative Methods
Fall 2023

Statistics component
Monday 12:30-1:45
Wednesday 11:15–12:45
296 Gibson Hall

Nicholas Winter
nwinter@virginia.edu
S385 Gibson Hall
Office hours: Thursday 11–1 *&* by appointment
Office hours both in-person and virtual:
https://virginia.zoom.us/my/nwinter)

"Inductive logic is unlike deductive or symbolic logic. In deductive reasoning, when you have true premises and a valid argument, the conclusion must be true too. Valid deductive arguments do not take risks. Inductive logic takes risks. You can have true premises, a good argument, but a false conclusion. Inductive logic uses probability to analyse that kind of risky argument."
　　　　　— Ian Hacking. 2001. *An Introduction to Probability and Inductive Logic*, xi.

http://xkcd.com/552/
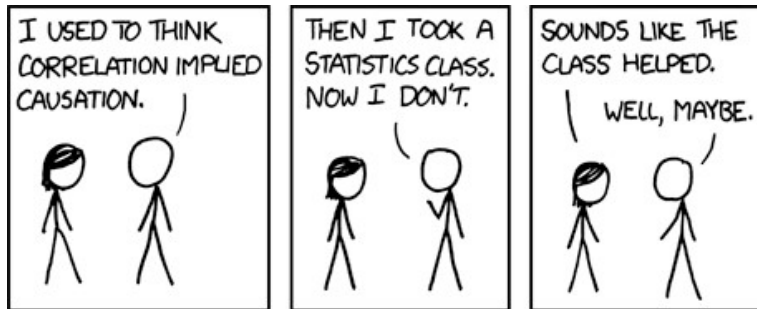
PLAD7100 includes two semi-autonomous components: statistics and math. The statistics component of this course will introduce you to some basic theory and methods for conducting quantitative analyses in political science, with a focus on statistics and data analysis. The mathematics component will introduce you to the mathematical background that will be used (to some extent) in this course and that will be assumed in future quantitative methodology courses. The primary goals are to give you tools to begin applying quantitative methods in your own work, to prepare you for further coursework, and to situate statistics within the broader endeavor of data analysis (and that within the broader endeavor of social science research).

Thus, in the statistics component we will begin at the beginning—with basic probability theory—then move through single-variable statistical analysis, and conclude with regression analysis. The lectures and problem sets will include a moderate amount of statistical theory, because I believe strongly that familiarity with the underlying theory is critical to the smart application of statistical techniques, but will emphasize the relevance of that theory to data analysis. The later problem sets will shift the emphasis toward application and data analysis.

There are no prerequisites for the course. Both components are required; that is, you may not take one without the other.

This document describes the statistics component; see the separate math component syllabus for information on that component.

Course Format and Requirements

Our meetings will include a good bit of interactive lecture, plus some time spent individually and in small groups working problems. In addition, there may be some asynchronous video material for you to watch. The lectures will be recorded and made available online. I have done this for years, in order to enable students to review material as they are doing homework or preparing for exams.

Additional requirements for the course include reading, more-or-less weekly homework assignments, a midterm exam, and a final exam.

Understanding statistics requires learning by doing, and for that reason there are a number of homework assignments— about one a week. The early assignments will consists of "paper-and-pencil" exercises to help solidify your understanding of basic concepts and procedures; the later assignments will ask you to conduct analysis of real political science data that I provide. I encourage group work on homework assignments, although each student should write up and turn in their own set of answers.

The exams will be "take-home" with a (generous) time limit. They are closed-book, except that you are allowed one page of notes (8½ by 11, front only) for the midterm, and two pages for the final.

Because the material in this course is cumulative, and because the in-class exercises will be critical to your learning, attending class consistently and staying current on the reading and homework is absolutely vital for your success. *Therefore, no incompletes will be given in this course*, barring extraordinary circumstances. In addition, I do not recommend auditing, though I don't forbid it. If you audit the course, I strongly encourage you to attend consistently and to complete the reading and homework assignments.

Software

The statistical software we will use is Stata (http://www.stata.com). Stata is available on the Politics lab computers in Gibson 296 and 196, and also through the "Remote Apps" system, a virtual computer lab that you can connect to remotely; see https://in.virginia.edu/remote-apps.

Though you do not need to purchase your own copy of Stata, I strongly encourage you to do so, *especially* if you anticipate further empirical work beyond this course. It is available for Windows, Mac, and Unix/Linux platforms with a student discount; see https://www.stata.com/order/new/edu/gradplans/student-pricing/. Stata comes in three "flavors": BE (Basic Edition), SE (Standard Edition), or MP (Multi-processor), with increasing capabilities (and price). The least expensive (BE) will be sufficient to analyze the data sets for this class and for all or most work you will do in graduate school. As you will see on the site, there are options to rent Stata for 6 months ($48 for BE) or a year ($94 for BE), or purchase a copy that is valid forever ($225). If you plan to purchase a perpetual license, there are discounts for group orders, so you may want to coordinate with each other and reach out to sales@stata.com. Final note: the current version of Stata is 18, although version 17 (or even 16 or 15) should meet your needs for this course (and most else you might need).

Readings

There are several books for this course. All but the Acock book are available electronically through the UVa Library, and used copies are available online. I recommend getting a hard copies—especially of the Wonnacott and Wonnacott book—if you can. That said, I will make available a PDF of Wonnacott & Wonnacott through Canvas.

Required

Abelson, Robert P. 1995. *Statistics as Principled Argument*. Hillsdale, NJ: L. Erlbaum Associates. ISBN 0805805281

Achen, Christopher H. 1982. *Interpreting and Using Regression*. Thousand Oaks, CA: Sage. ISBN 0803919158
    (Available electronically through the UVA library.)

Wonnacott, Thomas H. and Ronald J. Wonnacott. 1990. *Introductory Statistics for Business and Economics*. Fourth Edition. New York: John Wiley & Sons. ISBN 978-0471615170. Very inexpensive used copies are available; e.g.: https://www.amazon.com/gp/offer-listing/047161517X/ref=dp_olp_used?ie=UTF8&condition=used

Optional

Acock, Alan C. 2018. *A Gentle Introduction to Stata*. Rev. 6th ed. College Station, Texas: Stata Press. ISBN 978-1-59718-269-0.

In addition, we will read substantial portions of the following book, which is unfortunately out of print. I will make available copies of the relevant sections.

Kelejian, Harry H. and Wallace E. Oates. 1989. *Introduction to Econometrics: Principles and Applications.* Third Edition. New York: Harper and Row.

Copies of assigned articles and other readings will be available on-line.

Note: At the end of the syllabus I include a set of references to *Stata Journal* articles, most written by Nicholas Cox, that explain some tips, tricks, and useful techniques for managing data, creating analysis graphs, and presenting model results. These are all supplementary (and optional) but they contain a wealth of useful Stata lore.

Homework schedule

Homework assignments will generally be posted to Canvas on Wednesday or Thursday after class, and will be due on Friday at noon a week later (that is, eight or nine days after posting). You will submit homework through Canvas as a PDF file, so please be sure you have access to a scanner or a phone app that can create a single PDF from images of documents, as you will likely prefer to do at least some of the assignments by hand.

Course Schedule and Outline

August 23: Welcome & Introduction to the Course

August 28: Descriptive Statistics

Wonnacott & Wonnacott, skim chapter 1, read sections 2-1, 2-2, 2-3, 2-5 & 2-6

Optional: Kranzler, John H. 2017. "Overcoming Math Anxiety." In *Statistics for the Terrified.* 6th ed. New York: Rowman & Littlefield, chapter 1

August 30: Probability I

Wonnacott & Wonnacott, chapter 3

September 4: No class (Labor Day)

September 6: Probability II

September 11 & 13: Probability Distributions: Discrete

Wonnacott & Wonnacott, sections 4-1 through 4-3

September 18 & 20: Probability Distributions: Continuous, Covariance

Wonnacott & Wonnacott, sections 4-4 through 4-6; chapter 5

September 25 & 27: Sampling, Point Estimation, and Statistical Inference

Wonnacott & Wonnacott, chapters 6-7

Abelson, chapter 1

Acock, chapters 1-3

October 2: No class (UVa Fall Break)

October 4: Confidence Intervals

Wonnacott & Wonnacott, sections 8-1 and 8-2

October 9: Significance Testing

Wonnacott & Wonnacott, chapter 9

Abelson, chapter 2

Acock, chapters 4, 7

Cohen, Jacob. 1994. "The Earth is Round (p<.05)" *American Psychologist* 49(12):997-1003. Available from Canvas.

October 11: Proportions & Differences of means

Wonnacott & Wonnacott, sections 8-3 through 8-5

October 16: Catch up, review, Q&A for midterm

October 18: Midterm

Midterm: You will take the midterm during a 4-hour period of your choice within a 72-hour window to be determined between October 16 and October 19.

October 23: Univariate Data Analysis; Graphical Analysis and Display

Wonnacott *&* Wonnacott, section 2-7

Abelson, chapters 3, 4

Kastellec, Jonathan P., and Eduardo L. Leoni. 2007. "Using Graphs Instead of Tables in Political Science." *Perspectives on Politics* 5(4), **pages 755-763**.

Schwabish, Jonathan A. 2014. "An Economist's Guide to Visualizing Data." *Journal of Economic Perspectives* 28 (1):209-34.

Acock, chapters 5, 6

Optional supplementary readings

Epstein, Lee, Andrew D. Martin, and Matthew M. Schneider. 2006. "On the Effective Communication of the Results of Empirical Studies, Part I." *Vanderbilt Law Review* 59(6):1811-71.

Epstein, Lee, Andrew D. Martin, and Christina L. Boyd. 2007. "On the Effective Communication of the Results of Empirical Studies, Part II." *Vanderbilt Law Review* 60(3), **pages 801-31**.

Healey, Christopher G., and J. T. Enns. 2012. "Attention and Visual Memory in Visualization and Computer Graphics." *Visualization and Computer Graphics, IEEE Transactions on* 18 (7):1170-88.

October 25: Bivariate Data Analysis, Chi-square, Multivariate Data Analysis

Wonnacott *&* Wonnacott, chapter 17

Murnane, Richard J. and John B. Willett. "Reducing Observed Bias by the Method of Stratification." In *Methods Matter: Improving Causal Inference in Educational and Social Science Research,* 286-304

October 30 & November 1: Introduction to Bivariate Regression

Wonnacott *&* Wonnacott, chapter 11 and 12-1, 12-2

Achen, pages 1-37

Kelejian *&* Oates, pages 1-9

Acock, chapter 8

November 6 & 8: Bivariate Regression Estimation and Inference

Wonnacott *&* Wonnacott, 12-3 through 12-5

Achen, pages 37-51

Kelejian *&* Oates, pages 25-33, and 43-86

Optional, on ANOVA (useful especially if you ever read work in social psychology):

Wonnacott *&* Wonnacott, section 10-1

Acock, chapter 9

November 13: Bivariate Regression Inference II

Kelejian *&* Oates, pages 89-104 *&* 123-131

Abelson, chapter 5

November 15 & 20: Multiple regression I

Wonnacott *&* Wonnacott, chapter 13

Kelejian *&* Oates, chapter 4 (134-161) and 200-202

Acock, chapter 10

Kastellec, Jonathan P., and Eduardo L. Leoni. 2007. "Using Graphs Instead of Tables in Political Science." *Perspectives on Politics* 5(4), **pages 763-768**.

Epstein, Lee, Andrew D. Martin, and Christina L. Boyd. 2007. "On the Effective Communication of the Results of Empirical Studies, Part II." *Vanderbilt Law Review* 60(3), **pages 831-46**.

Abelson, chapter 6

November 22: No class (Thanksgiving break)

November 27 & 29: Regression Topics: indicator variables, interaction terms, colinearity/micronumerosity, etc.

Wonnacott *&* Wonnacott, sections 14-1 & 14-2

Achen, pages 51-79

King, Gary. 1986. "How Not to Lie With Statistics: Avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science* 30(3):666-87.
Available from Canvas; also from https://www.jstor.org/stable/2111095

Kelejian *&* Oates, pages 205-211 and 178-186

Abelson, chapter 7

December 4: Regression Topics: Non-Linear Relationships, Outliers, Model Specification; OLS vs. Matching; Ketchup and Review; What's Next?

Wonnacott *&* Wonnacott, chapters 14

Bartels, Larry M. 1990. "Five Approaches to Model Specification." *Political Methodologist* 3(2):2-6.

Kelejian *&* Oates, pages 251-256

Abelson, chapter 8, 9

Miller, Michael K. 2013. "The Uses and Abuses of Matching." Working Paper, George Washington University. https://drive.google.com/file/d/1o2vQZTNmg7b_R5geEhopGGgwT9bkOt3n/view

Justin Esarey's blog comments on Miller, and the ensuing discussion, at http://politicalmethodology.wordpress.com/2013/07/25/matching-madness-causal-inference-in-political-methodology/

Final exam will be a closed-book, timed, take-home exam taken within a 72-hour window after classes end. Precise timing TBA.

***Stata Journal* Articles of Interest**

Tables of estimation results

Gallup, John Luke. 2012. "A New System for Formatting Estimation Tables." *Stata Journal* 12(1):3-28.

Cox, Nicholas J. 2012. "Speaking Stata: Output to Order." *Stata Journal* 12(1):147-158.

Newson, Roger B. 2012. "From Resultssets to Resultstables in Stata." *Stata Journal* 12(2):191-213.

In addition, see my handout, "Producing Presentation-Quality Tables and Figures for Regression Models," which describes my own command, mktab, as well as a few other options.

Data management and tabulation

Cox, Nicholas J. 2011. "Speaking Stata: Compared With ..." *Stata Journal* 11(2):305-314.

Baum, Christopher F. and Nicholas J. Cox. 2007. "Stata Tip 45: Getting Those Data Into Shape." *Stata Journal* 7(2):268-271. http://www.stata-journal.com/sjpdf.html?articlenum=dm0031

Cox, Nicholas J. 2003. "Speaking Stata: Problems with tables, Part I" *Stata Journal* 3(3):309-324. http://www.stata-journal.com/sjpdf.html?articlenum=pr0010

Graphing

Cox, Nicholas J. 2010. "Speaking Stata: Graphing Subsets" *Stata Journal* 10(4):670-681.

Cox, Nicholas J. 2010. "Speaking Stata: The statsby Strategy" *Stata Journal* 10(1):143-151.

Buis, Maarten L. and Martin Weiss. 2009. "Stata Tip 81: A Table of Graphs." *Stata Journal* 9(4):643-647.

Cox, Nicholas J. 2009. "Speaking Stata: Paired, Parallel, or Profile Plots for Changes, Correlations, and Other Comparisons." *Stata Journal* 9(4):621-639.

Cox, Nicholas J. 2008. "Speaking Stata: Between Tables and Graphs." *Stata Journal* 8(2):269-289. http://www.stata-journal.com/sjpdf.html?articlenum=gr0034

Cox, Nicholas J. 2008. "Stata tip 27: Classifying data points on scatter plots." *Stata Journal* 5(4):604-606. http://www.stata-journal.com/sjpdf.html?articlenum=gr0023

Cox, Nicholas J. 2008. "Speaking Stata: Graphing Categorical and Compositional Data." *Stata Journal* 4(2):190-215. http://www.stata-journal.com/sjpdf.html?articlenum=gr0004

Model Interpretation

Williams, Richard. 2012. "Using the margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects." *Stata Journal* 12(2):308-331.

Jann, Ben. 2013. Plotting regression coefficients and other estimates in Stata. University of Bern Social Sciences Working Papers Nr. 1. Available from http://ideas.repec.org/p/bss/wpaper/1.html