

PLAD 7100
Political Research with Quantitative Methods
Fall 2020

Statistics component

Wednesday

10:00am–12:30pm

<https://virginia.zoom.us/j/92365795863?pwd=VSszVWxGejBsTm5Ea3lBl3c4dEhRUT09>

Nicholas Winter

nwinter@virginia.edu

S385 Gibson Hall

Office hours: Wednesday 2–3:30

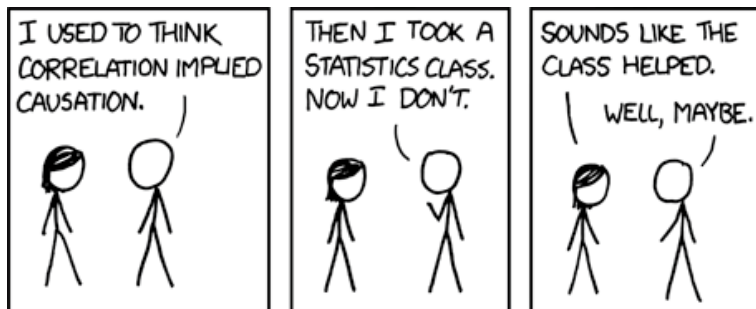
☞ by appointment

Personal Zoom Meeting Room:

<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 to explore and evaluate political science theories, 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

The course will meet synchronously each week during the scheduled time, on Zoom. Our synchronous meetings will include a good bit of interactive lecture, plus some time spent individually and in small groups (in breakout rooms) working problems. In addition, there will be some asynchronous video material for you to watch.

The synchronous lectures will be recorded. In fact, I have done this for years, in order to enable students to review material as they are doing homework or preparing for exams. Thus, the online format actually changes this class relatively little, except of course for the lack of human contact.

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 consist 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 (relatively 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 through the “UVa Hive,” a virtual computer lab that you can connect to remotely; see https://virginia.service-now.com/its/?id=itsweb_kb_article&sys_id=a2b62242dbe3c7c04f32fb671d96196a for information on the Hive. (My understanding is that Macs and non-English Windows systems may not be able to connect to the Hive.)

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”: IC, SE, or MP, with increasing capabilities (and price). The least expensive (IC) 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 IC) or a year (\$94 for IC), 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. Stata is also available on the computers in the Politics Department computer lab in Gibson Hall; this may be of use at some point in the future. Final note: the current version of Stata is 16. The Hive has Stata 15; that will be fine 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/or the HathiTrust emergency 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 Collab.

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

Acock, Alan C. 2018. *A Gentle Introduction to Stata*. Rev. 6th ed. College Station, Texas: Stata Press. ISBN 978-1-59718-269-0. (Note that the 5th edition will serve your needs.)

Wonnacott, Thomas H. and Ronald J. Wonnacott. 1990. *Introductory Statistics for Business and Economics*. Fourth Edition. New York: John Wiley & Sons. ISBN 978-0471615170.

https://www.amazon.com/gp/offer-listing/047161517X/ref=dp_olp_used?ie=UTF8&condition=used

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

A homework assignment will be posted in the “Resources” section of the Collab site weekly, due in class the following week.

Course Schedule and Outline

August 25: Welcome & Introduction to the Course

September 2: Basic Probability & Descriptive Statistics

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

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

September 9: Probability Distributions and Random Variables

Wonnacott & Wonnacott, chapters 4-5

September 16 & 23: Sampling and Point Estimation & Introduction to Statistical Computing using Stata

Wonnacott & Wonnacott, chapters 6-7

Abelson, chapter 1

Acock, chapters 1-3

Optional: Stata Documentation; in particular *Getting Started with Stata* and selections from the *Stata User's Manual*

September 30: Interval Estimation and Hypothesis Testing

Wonnacott & Wonnacott, chapters 8-9

Abelson, chapter 2

Acock, chapters 4, 7

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

October 7: Differences of means; catch up; review and Q&A for 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 8 and October 13.

October 14: Data management and analysis using Stata

Wonnacott & Wonnacott, section 2-7 & chapter 17

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

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

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 21: Bivariate Regression I

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

Achen, pages 1-37

Kelejian & Oates, pages 1-9, 25-33, and 43-86

Acock, chapter 8

October 28: Bivariate Regression II: Topics

Wonnacott & Wonnacott, 12-3 through 12-5

Achen, pages 37-51

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

Abelson, chapter 5

Optional, on ANOVA (useful especially if you ever read work in social psychology): Acock, chapter 9

November 4: Multiple Regression I

Wonnacott & Wonnacott, chapter 13

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

Acock, chapter 10

November 11: Multiple regression II

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 18: Regression Topics: indicator variables, interaction terms, colinearity/micronumerosity

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 Collab; also from <http://links.jstor.org/sici?sici=0092-5853%28198608%2930%3A3%3C666%3AHNTLWS%3E2.0.CO%3B2-3>.

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

Abelson, chapter 7

November 25: No class—Happy Thanksgiving

December 2: Regression Topics: non-linear relationships, outliers, model specification; OLS vs. matching; Ketchup and review; etc.

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://sites.google.com/site/mkmtwo/Miller-Matching.pdf>

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=dmo031>

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=proo10>

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=groo34>

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=groo23>

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=groo04>

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>