

University of California, Los Angeles  
Department of Political Science

Westwood  
Copy Center  
Gayley and  
Braxton

## Statistical Methods

Political Science 200a  
Fall Quarter 2012  
Office Hours: Tu 1:15 – 2:45, or by appointment.  
Course Assistant: Bronwyn Lewis

Professor James DeNardo  
3345 Bunche Hall  
e-mail: [jdenardo@ucla.edu](mailto:jdenardo@ucla.edu)  
[bronwyn.ashleigh@gmail.com](mailto:bronwyn.ashleigh@gmail.com)  
Lab SSCL login info

This course explores the uses and abuses of statistical reasoning in social and political studies. Students learn the whys and hows of data analysis (including statistical computing) and the logic of statistical inquiry. The course strives to develop critical judgment about quantitative studies of political, economic, and social life and to cultivate the skills necessary to undertake such research on one's own.

### Homework, Graded Assignments, and Exams

Becoming adept at data analysis requires regular practice. Learning by doing is the way to go and that will be our philosophy. To complete the course successfully, each of the following elements must be completed on or before the scheduled dates. Extensions require prior permission from Professor DeNardo.

- due on the Tue next week  
to both (digital CA & DeNardo adequate) submission will be adequate
- 1). Weekly Homework and Summaries (10%). Each week students prepare written summaries of the substantive articles on the syllabus (about 200 words per article). A good summary should describe the central argument or finding of the article, its data analytical strategy (or salient statistical lessons one can draw from it), and provide some reaction to the author's position (by considering implications, extensions, connections, limitations, inspirations...). Students also complete regular problem sets, often computational, that develop ideas from lab or lecture. Homeworks and summaries are required but ungraded elements, scored 1 or 0 for complete or not. *not textbook*

- physical & pdf/docx by email attachment
- 2). Data analytical research project (50%). A data analytical study of roughly 15-20 pages (excluding graphical displays), based on a substantial political data set provided by the instructor. Papers should emulate the style and purpose of professional journal articles. *grading will be done by DeNardo*

- 3). Midterm project (40%). An in-class exam or take-home project, TBA, designed to consolidate ideas and prepare for the research project.

Written work should be typed and double-spaced, using 12 point fonts, one

inch margins, and numbered pages. Papers should be submitted digitally in doc(x) or pdf formats, and in hardcopy, to each instructor.

### **Required Texts**

David S. Moore and George P. McCabe, *Introduction to the Practice of Statistics*, W. H. Freeman, 4<sup>th</sup> or 5<sup>th</sup> edition (IPS). Available for purchase at ASUCLA bookstore. Survey carefully, reading sections relevant to our discussions in detail.

Edward R. Tufte, *Data Analysis for Politics and Policy*, Prentice Hall. (DAPP)  
Available for purchase at ASUCLA bookstore in reprint.

Xeroxed handout set. The handout set contains lecture slides, datasets, and examples. It will be used regularly in lecture and workshop, and should be brought to each of our meetings. Available at Westwood Copy Center in Westwood Village.

Articles on the syllabus can be downloaded from our secure website or public sources like JSTOR.

### **Highly Recommended**

The following books use statistical reasoning and evidence to challenge deeply held and widely shared beliefs about politics and social affairs. They exemplify central ideas from the course, while providing a concrete sense of the possibilities and challenges that confront a statistical approach to understanding the world:

Stephen J. Gould, *Full House: The Spread of Excellence from Plato to Darwin*, (Random House, paper). ISBN: 0609801406. Reprint edition (October 1997).

Bjorn Lomborg, *The Skeptical Environmentalist: Measuring the Real State of the World* (Cambridge University Press, paper). ISBN: 0521010683; reprint edition September 2001.

Available from Amazon.

### **Helpful Supplementary Texts**

W. Wooton and I. Drooyan, *Intermediate Algebra*, or some similar text for mathematical refreshment.

M. O. Finkelstein and B. Levin, *Statistics for Lawyers*, Springer Verlag.

Mosteller, Fienberg, and Rourke, *Beginning Statistics with Data Analysis*, Addison-Wesley.

## Weekly Reading Assignments

### *I. Numeracy, Measurement, and Quantification*

How numerate people think about quantities large and small, and why numeracy is required to understand the world. Strategies of measurement in complex, noisy, opaque settings.

1. Edward Tufte, DAPP, Chapter 1.
2. D. R. Hofstadter, "Number numbness, or why innumeracy may be just as dangerous as illiteracy," *Scientific American*, May 1982. Interesting ideas about measuring and conceptualizing numerical evidence.
3. P. Heidenrich, "How many Iraqis were killed?," *Foreign Policy* Spring 1993. 2,000,000? 1,000,000? 500,000? 100,000? A numerate interpretation of casualty data from the first war in Iraq. Compare methods of reasoning to Hofstadter's.
4. Tim Groseclose and Jeffrey Milyo, "A Measure of Media Bias," *The Quarterly Journal of Economics*, Vol. 120, No. 4 (Nov., 2005), pp. 1191-1237. Sophisticated scaling techniques for measuring "latent" variables.
5. "The Size of the Cosmos," in Alan Lightman, *The Discoveries: Great Breakthroughs in 20<sup>th</sup> Century Science* (Pantheon, 2005). One of the greatest data analytical studies of all time, with an interesting political subtext. Notice how an ingenious strategy of *control* produces a momentous discovery.
6. Ansolabehere, Rodden, and Snyder, "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting," *American Political Science Review*, Vol. 102, No. 2 (May 2008).

### *II. (Statistical) Pattern Recognition and Explanation*

How explanations arise from the interplay between evidence and beliefs. Why is effective social research difficult? Theory building and testing.

1. Daniel Scott Smith, "Child-Naming Practices, Kinship Ties, and Change in Family Attitudes in Hingham, Massachusetts, 1641 to 1880," *Journal of Social History*, Vol. 18, No. 4 (Summer, 1985), pp. 541-566.
2. Stephen J. Gould, "American Polygeny and Craniometry before Darwin," in *The Mismeasure of Man* (Norton, 1996). History of early investigations about the vexed relationship between race and IQ, with attention to measurement, selection, and statistical control..
3. David Barash, "The Stubborn Pull of Dogma," *New York Times*, 9/27/05. A special nemesis in the social sciences.
4. John Alford, et. al., "Are Political Orientations Genetically Transmitted?," *American Political Science Review*, Vol. 99, No. 2, (May 2005). Are political dogmas hard wired, and how could one tell?
5. Larry M. Bartels, "Partisanship and Voting Behavior, 1952-1996" *American Journal of Political Science*, Vol 44, Issue 1 (Jan 2000), 35-50. The relationship between partisanship and voting is a fundamental problem in electoral studies. Is the relationship stable over time?

### III. *Shapes of Distributions: Statistical Description and Comparison*

Pay careful attention to authors' reasoning about some fundamental distributions, including their shapes, their evolution over time, and their social, political, and economic causes and consequences.

- ✓ 1. M & M, IPS. Shapes of distributions and how to visualize, summarize, and compare them.
2. David R. Mayhew, "The Case of the Vanishing Marginals," *Polity* 7 1974. A seminal article on party competition, grounded in a careful comparison of vote split distributions.
3. S.J. Olshansky, B.A. Carnes, and C.K. Cassel, "The Aging of the Human Species," *Scientific American* 268, No. 4, April 1993. How is the distribution of age evolving among humans? A profound issue that will shape everyone's future.
4. Gary C. Jacobson, "The Marginals Never Vanished: Incumbency and Competition in Elections to the U.S. House of Representatives, 1952-82," *AJPS*, February 1987. Interesting response to Mayhew, using an argument that considers *spread* and *level* in the Mayhew distributions.
5. William Grimes, "Winnowing the Field of America to One Representative,

"NY Times book review of *The Average American*. Who is the "average" American? An interesting question... or not?

6. John W. Tukey, *Exploratory Data Analysis*, Chapters 1-3. The original source for many of the techniques described in M & M. Seek to understand the basic ideas and approach without getting stalled on the numerous spins of each technique.

Highly Recommended (available from Amazon or the like)

7. S. J. Gould, *Full House* (Random House). A stimulating consideration of humans' place in the distribution of life. Does evolution favor complexity?

#### IV. Origins of Distributions: *The Distribution of Wealth and Power*

- 1\*. Martina Morris and Bruce Western, "Inequality in Earnings at the Close of the Twentieth Century," *Annual Review of Sociology*, 1999, 25: 623-57. Polarization of the income distribution in the United States.
- 2\*. Daron Acemoglu, Simon Johnson, and James Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* (2001). Large scale patterns of wealth and poverty in the world, and their origins in colonial institutions.
- 3\*. Michael Ross, "Is Democracy Good for the Poor?" *American Journal of Political Science*, Vol. 50, No. 4 (2006), pp. 860-874. A careful study of regime design and human welfare, with instructive discussions about missing data, statistical control, and social measurement systems.
4. Michael Mitzenmacher, "A Brief History of Generative Models for Power Law and Lognormal Distributions," *Internet Mathematics*, Vol. I, No. 2: 226-251. Mathematical theories about the origins of skew distributions and "black swans."
5. Lars-Erik Cederman, "Generating State-Size Distributions: A Geopolitical Model," Center for Comparative and International Studies (CIS) working paper (September, 2003).

#### V. Causal Relationships and Regression Techniques

1. Edward Tufte, DAPP, Chapter 3.

2. M & M, IPS, Chapter 2. The material in 1 and 2 is very important and requires careful study. Mastery will come only with several readings and experience.
3. Lisa Blaydes and Drew Linzer, "Elite Competition, Religiosity, and Anti-Americanism in the Islamic World," *American Political Science Review* Vol. 106, No. 2 May 2012. Political origins of anti-Americanism.
4. Michael Ross, "Does Oil Hinder Democracy?" *World Politics*, Vol. 53 (April 2001), pp. 325-61. Does the "curse of oil" impede democratic development? How can we isolate the effect of oil on regime formation in a densely multivariate causal environment?
5. James DeNardo, "Turnout and the Vote: The Joke's on the Democrats," *American Political Science Review*, Vol. 74, No. 2 (June 1980), 406-420. Regression analysis informed by mathematical modeling.
6. F. J. Anscombe, "Graphs in Statistical Analysis," *The American Statistician* 27 (1973). An early precursor of central developments in modern statistics. How does Anscombe demonstrate the pivotal role of graphical display?
7. Kenneth Scheve and David Stasavage, "Democracy, War, and Wealth Lessons from Two Centuries of Inheritance Taxation " working paper (October 2011).

#### Supplementary Reading

8. W.S. Cleveland and R. McGill, "The Many Faces of a Scatterplot," *Journal of the American Statistical Association* 79 (1984), p. 807-22. Read the first half, and skim the second. Introduces Cleveland's *lowess* scatterplot smoother and exemplifies Anscombe's program.

#### VI. Regression Methods For Observational Data: Estimating Causal Relationships, Confounding Effects, Model Dependence, and Statistical Control

1. DAPP, Chapter 4.
2. F. Mosteller and J. Tukey, "Woes of Regression Coefficients," Chapter 13 in *Data Analysis and Regression*. Focus on the fundamental (and recurring) theme—the model you fit affects the answer you get (observational inference is *model-dependent*, making *statistical control* and *model specification* paramount issues.)
3. Gary King, "How Not to Lie with Statistics" *AJPS*, August 1986.

Interpreting regression coefficients and understanding their limitations.

4. John J. Donohue III and Steven Levitt, "The Impact of Legalized Abortion on Crime," *The Quarterly Journal of Economics* (May, 2001). Interesting but controversial work about the effects of abortion on crime. Notice how many confounding effects enter the story and the elaborate strategy of statistical control.
5. John J. Donohue and Justin Wolfers, "Uses and Abuses of Empirical Evidence in the Death Penalty Debate," *Stanford Law Review*, Vol. 58, No. 791 (2006). *Complexity* is a fundamental feature of social life and a central challenge to robust causal inference. Model uncertainty, sensitivity analysis, and the limits of observational evidence.
6. James D. Fearon, "Counterfactuals and Hypothesis Testing in Political Science," *World Politics* Vol 43, No. 2, January 1991. An interesting discussion of counterfactual reasoning, a central construct in theories of causal inference.
7. Andrew Gelman, "Causality and Statistical Learning," *American Journal of Sociology* Volume 117, Number 3 (November 2011): 955–966.

#### VII. Model Specification in Complex Multivariate Settings with Messy Data and Limited Theory

1. Daniel Treisman, "The Causes of Corruption: A Cross-National Study," *Journal of Public Economics* 76 (2000), pp. 399-457. Thoughtful application of regression modeling in a highly multivariate, cross-national study, using messy data.
2. Arthur S. Goldberger and Charles F. Manski, "Review Article: The Bell Curve by Herrnstein and Murray," *Journal of Economic Literature* Vol. 33 (June 1995). Pay careful attention to the discussions about statistical control and model selection in a complicated regression problem.
3. Edward E. Leamer, "Let's Take the Con Out of Econometrics," *American Economic Review*, Vol. 73, No. 1 (March 1983), pp. 31-43.
4. R. Klitgaard et. al., "Regression Without A Model," *Policy Sciences* 13 (1981). Practical introduction to model specification and the fundamental idea of *cross-validation*.

5. Lawrence H. Summers, "The Scientific Illusion in Empirical Macroeconomics," *The Scandinavian Journal of Economics*, Vol. 93, No. 2, (Jun., 1991), pp. 129-148.
6. Xavier Sala-i-Martin, "I Just Ran Four Million Regressions," NBER Working Paper No. 6252, November 1997.

#### VIII. *Strategies of Statistical Control in Observational Studies*

1. James Fowler, et. al., "Genetic Variation in Political Participation," *American Political Science Review*, Vol. 102, No. 2, (May 2008). Using nature to hold things constant.
2. Donald B. Rubin, "For Objective Causal Inference, Design Trumps Analysis," *The Annals of Applied Statistics*, Vol.2, No.3 (Sept 2008).
3. Clayton Nall, "The Road to Division: Interstate Highways and Geographic Polarization," working paper (February, 2012).
4. Daniel E. Ho, "Why Affirmative Action Does Not Cause Black Students to Fail the Bar," *The Yale Law Journal* (2005). Critique of Richard Sander, "A Systematic Analysis of Affirmative Action in American Law Schools," *Stanford Law Review*, 367 (2004). Careful causal reasoning is indispensable when using statistical adjustments to achieve control in observational data.
5. Robert M. Feinberg and Gregory N. Price, "THE FUNDING OF ECONOMICS RESEARCH: DOES SOCIAL CAPITAL MATTER FOR SUCCESS AT THE NATIONAL SCIENCE FOUNDATION?" *The Review of Economics and Statistics*, Vol. 86, No. 1 (Feb., 2004), pp. 245-252.

#### IX. *Experiments, Natural and Designed, for Causal Inferences*

1. J.B. McConahay, "Experimental Research in Political Psychology." A readable summary of basic principles in experimental design.
2. M & M, 2.7, 3.1-3.3.
3. Shanto Iyengar, Mark Peters, and Donald Kinder, "Experimental Demonstrations of the "Not-So-Minimal" Consequences of Television News Programs," *American Political Science Review* (Vol. 76, pp. 848-858).
4. Alan S. Gerber and Donald P. Green, "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment,"



*American Political Science Review* 94: 653-663.

5. Kosuke Imai, "Do Get-Out-The-Vote Calls Reduce Turnout? The Importance of Statistical Methods for Field Experiments," forthcoming *American Political Science Review*. When it comes to randomization, the devil is in the details.
6. Richard Berk, "Randomized Experiments as the Bronze Standard," working paper, UCLA Department of Statistics (May, 2005).
7. Robert LaLonde, "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review*, Vol. 76, No. 4 (Sep. 1986).
8. Roberto Agodini and Mark Dynarski, "ARE EXPERIMENTS THE ONLY OPTION? A LOOK AT DROPOUT PREVENTION PROGRAMS," *The Review of Economics and Statistics*, February 2004, 86(1): 180-194.

#### Supplementary Readings

9. Donald T. Campbell and Julian C. Stanley, *Experimental and Quasi-Experimental Designs for Research* (Rand McNally, 1963).

#### X. Sampling, Polling, Extrapolation, Generalization

1. M & M, IPS, 3.4, 4.1, 4.2, 4.5.
2. Stephen Fienberg, "Randomization and Social Affairs: The 1970 Draft Lottery" *Science* 171, 22 January 1971 . A famously non-random drawing.
3. Julian L. Simon, "Resources, Population, Environment: An Oversupply of False Bad News," *Science* 208, 27 June 1980. Many sources of information are not based on systematic sampling.
4. Howard Wainer, "A Selection of Selection Anomalies," *Chance* (1988).
5. Michael O. Finkelstein and Bruce Levin, *Statistics for Lawyers* Springer-Verlag, Chapter 9, "Sampling Issues."
6. Douglas Rivers, "Sample Matching: Representative Sampling From Internet Panels," YouGovPolimetrix. Web based polling.
7. Frederick Mosteller, "Errors: Non-sampling Errors."

8. Andrew Gelman and Gary King, "Why are American Presidential Campaign Polls So Variable When Votes Are So Predictable?," *British Journal of Political Science*, 23 (1993), pp. 409-451. (XRX)