Spring 2017

MULTIVARIATE TECHNIQUES 790-633
Wednesdays, 12:00 - 2:40, Hickman 313

Instructor: Professor Richard Lau (x2-9261)
Office Hours: Wednesdays, 3:00 - 4:00, and by appointment
Hickman 409

SYLLABUS

This course is designed to introduce students to various multivariate statistical techniques they will encounter in the course of conducting research and reading professional articles in political science. We will concentrate on ordinary least squares (OLS) regression, as this is the most general multivariate technique and the one most widely used in the field. Understanding OLS is crucial for learning even more advanced statistical techniques.

I will present common statistical formula and derive a few central statistical proofs in class. But the focus of the course will not be on mathematical statistics; rather, our focus will be on achieving an *intuitive understanding* of the various techniques and their limitations. More importantly, the goal of this class is to teach you that statistics, by themselves, tell us almost nothing. The data never speak for themselves. *Theory* is all important, and statistics are only useful to the extent they can help up *test, refine, or* (in some cases) *develop* theory. It is next to impossible to interpret any statistical result without some theory to help us understand what we are looking at. What you will learn in this class, if I am successful, is how to link your statistical results to your theory.

You cannot learn statistics by reading a book or sitting in a lecture, however; you must actively *use* the statistics you read about before you can really learn them. Hence a major focus of this class will be homework assignments aimed at applications of the techniques we are learning in class. Students may use their own data set for homework assignments, as long as it has appropriate data; or I will help you select an interesting one to use. I will be doing all of the homework assignments along with you, using the 2012 American National Election Study data. I have a very "clean" version of it, which you are welcomed to use, but I strongly recommend you find a dataset that is of interest – and research potential – to you.

Required Reading:

Our basic text for this class will be

McClendon, McKee J. 2002. *Multiple Regression and Causal Analysis.* Prospect Heights, IL: Waveland Press. You should buy this book, and keep it.

This book is available in the Barnes & Noble book store near the train station, and is on reserve at Alexander library. We will also do a little reading from

Knoke, David, George W. Bohrnstedt, and Alisa Potter Mee. 2002. *Statistics for Social Data Analysis* (4th ed. or later). Belmont, CA: Wadsworth/Thomson. ("KBM" in the detailed syllabus below). [They are up to the 5th edition now? Earlier or later editions of this book are fine.]

the same book you used last semester in 533.

These texts are supplemented by two journal articles:

Markus, Gregory B, and Philip E. Converse. 1979. "A Dynamic Simultaneous Equation Model of Electoral Choice." *American Political Science Review*, 73(December): 1055 - 1070.

King, Gary. 1986. "How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science*, 30(August): 606-627.

You can find them in the library, or download them from the library's electronic resources.

All of our homework assignments can be conducted on the computers in our computer labs, with the program *SPSS for Windows* (which is also on the computers in the various computer hubs around campus). You should all be familiar with this program from 533, and might have already bought a manual for that class. You can start using the program by pointing and clicking – but you will find that method to be very restrictive, and in the long run, inefficient. No serious data analyst points and clicks. In this class you will write your own syntax and conduct your statistical analyses in that manner.

I will illustrate all of the syntax you need for any of the procedures you will be running for this class, but they are simple. By far the most important, time consuming, and difficult task, is simply variable construction – that is, building the dataset you want to analyze. I takes practice to develop those skills. And as with all things technical, if all else fails, RTFM. Fortunately, SPSS has a wonderful built-in manual. Click on "Help" at the top of the screen, and then select "Command Syntax Reference" from the drop-down menu.

Grading:

Grades are based on class attendance and homework assignments, including a take-home final (basically a comprehensive homework assignment). *Homeworks are due at the start of class whenever they are due*, and (barring some massive computer failure affecting all the computers in the lab) *will not be accepted late*. Generally speaking I will grade HW assignments the afternoon they are due, so if you do not turn it in, I will not grade it. Your understanding in this class depends on your keeping up with the work – so do it!

You are each on you honor to do the homework assignments (the computer runs, and the interpretation there of) by yourself. However, this does not mean you cannot work with others on the homeworks. ***In fact, I encourage this, as it is a good way (probably the best way) to learn.*** In past years, subsets of students have had informal "doing the homework" meetings in the lab, and I encourage this (but try to schedule them well before the day before the homework is due!). But I do want everyone to run their own programs and to write up their own interpretations. Only on the final exam/homework assignment are you expected to work exclusively by yourself.

Tentative Weekly Schedule and Assignments

Date	*Topic*	*Reading*
Jan. 18	Introductory Class[1] What this class will attempt to do; finding an appropriate data set; and getting much of the math out of the way ... Begin bivariate regression	McClendon, Ch. 1
Jan 25	Finish Bivariate Regression HW#1: Setting up your data file in SPSS; Making a data codebook; Bivariate Regression Examples (due Feb. 10)	McClendon, Ch. 2
Feb. 1	Review Using SPSS [Keep working on HW# 1]	KBM, pp 405 - 414;
Feb. 8	Multivariate Regression Interpreting Regression Coefficients; The Geometry of OLS HW #2: Multivariate regression examples (due Feb. 24)	McClendon, Ch. 3
Feb. 15	Multivariate Regression Continued Interpreting R Square and the "Usefulness" of Variables in the Equation [Complete HW #2]	McClendon, Ch. 4, pp. 133 - 174
Feb. 22	Theoretical Importance, "Level" Importance, and Substantive Significance. Making Sure We Know What We've Learned So Far	Review McClendon pp. 80 - 81, 87 - 93; KBM, Ch. 8
Mar. 1	A Slight Tangent: One-Way ANOVA HW #3: ANOVA examples from your data (due Mar. 23)	KBM, Ch. 4
Mar. 8	Two-Way ANOVA and Interactions [Keep working on HW #3 ...]	McClendon, Ch. 5
Mar. 15	SPRING BREAK!	

[1] Appendix A of KBM (pp. 449 - 455 in my edition), and Appendix 3a in McClendon (pp 119 - 132) provide some valuable review of math we will use, that should also be covered in Math Camp.

Date	*Topic*	*Reading*
Mar. 22	Back to OLS: Interactions and Functional Form	McClendon, Chs. 6 & 7
	HW #4: Examining Some Nonlinear Models (due Apr. 6)	
Mar. 29	Another Tangent: Measurement Reliability and Validity	KBM, pp 405 - 414; Markus & Converse article
	[Finish HW # 4]	
Apr. 5	Pathologies of OLS	McClendon, 154-157 161-163, 174-197
	HW #5: How badly did we screw-up in our earlier assignments? (due Apr. 20)	
Apr. 12	Pathologies of OLS continued	King Article
	[Keep working on HW #5]	
Apr. 19	Finishing up OLS; Two-Stage Least Squares and Simple Causal Modeling	McClendon, Chs. 8 & 9
Apr. 26	Regression with "Limited" Dependent Variables	KBM, Ch. 9, pp 287 - 313
	Take-home Final, due Wednesday, May 3rd, by noon	

Interesting Statistics Articles in the Social (mostly Political) Science Literature

Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review*, 95(September): 529-546.

Bartels, Larry M. 1991. "Instrumental and 'Quasi-Instrumental' Variables." *American Journal of Political Science*, 35(July): 777-800.

Bartels, Larry M. 1996. "Pooling Disparate Observations." *American Journal of Political Science*, 40(August): 905-942.

Baron, Reuben M., & David A. Kenny. 1986. "The Moderator-mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology*, 51(6): 1173-1182.

Berry, William D., Matt Golder, and Daniel Milton. 2012. "Improving Tests of Theories Positing Interaction." *Journal of Politics*, 74(July): 653 - 671.

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis*, 14(Winter): 63-82.

Braumoeller, Bear F. 2004. "Hypothesis Testing and Multiplicative Interaction Terms." *International Organization,* 58(4): 807-820.

Braumoeller, Bear F., and Gary Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science*, 44(October): 844 - 858.

Buthe, Tim. 2002. "Taking Time Seriously: Modeling History and the Use of Narrative as Evidence." *American Political Science Review*, 96(September): 481 - 494.

Chatterjee, Sanjit, & Frederick Wiseman. 1983. "Use of Regression Diagnostics in Political Science Research." *American Journal of Political Science,* 27(August): 601-613.

Cleary, Paul D., & Ronald C. Kessler. 1982. "The Estimation and Interpretation of Modifier Effects." *Journal of Health and Social Behavior*, 23(2): 159-169.

Cohen, Jacob. 1990. "Things I Have Learned (So Far)." *American Psychologist*, 45(December): 1304-1312.

Crawford, Jarret T., Lee Jussim, and Jane M. Pilanski. 2014. "How (Not) to Interpret and Report Main Effects and Interactions in Multiple Regression: Why Crawford and Pilanski Did Not Actually Replicate Lindner and Nosek (2009)." *Political Psychology*, 35 (December): 857 - 862.

Daniels, Mark R., and R. Darcy. 1983. "Note on the Use and Interpretation of Discriminant Analysis." *American Journal of Political Science*, 27(May): 359 - 383.

Downs, George W., & David M. Ricke. 1979. "Interpreting Heteroskedasticity." *American Journal of Political Science,* 23(November): 816-828.

Friedrich, Robert J. 1982. "In Defense of Multiplicative Terms in Multiple Regression Equations." *American Journal of Political Science*, 26(3): 797-833.

Gerring, John. 2004. "What Is a Case Study and What Is It Good for?" *American Political Science Review*, 98(May): 341-354.

Glynn, Adam N. 2012. “The Product and Difference Fallacies for Indirect Effects.” *American Journal of Political Science*, 56(January): 257 - 269.

Huberty, Carl J. 1984. "Issues in the Use and Interpretation of Discriminant Analysis." *Psychological Bulletin*, 95(1): 156-171.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. “Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies.” *American Political Science Review*, 105(November): 765 - 789.

Jackman, Simon. 2000. “Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo.” *American Journal of Political Science*, 44(April): 375 - 404.

Jacoby, William G. 1999. "Levels of Measurement and Political Research: An Optimistic View." *American Journal of Political Science*, 43(January): 271 - 301.

Judd, Charles M. & Kenny, David A. 1981. “Process Analysis: Estimating Mediation in Treatment Evaluations.” *Evaluation Review*, 5(5): 602-619.

King, Gary. 1986. "How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science*, 30(May): 606-627.

King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. “Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation.” *American Political Science Review*, 95(March): 49 - 70.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. “Making the Most of Statistical Analyses: Improving Interpretation and Presentation.” *American Journal of Political Science*, 44(April): 347 - 361.

Kritzer, Herbert M. 1978. "An Introduction to Multivariate Contingency Table Analysis." *American Journal of Political Science*, 22(February): 187-226.

Lieberman, Evan S. 2005. “Nested Analysis as a Mixed-Method Strategy for Comparative Research.” *American Political Science Review*, 99(August): 435 - 452.

MacKinnon, David P., and Amanda J. Fairchild. 2009. “Current Directions in Mediation Analysis.” *Current Directions in Psychological Science*, 18(February): 16 - 20.

Miller, Jane E. 2005. *The Chicago guide to Writing about Multivariate Analysis*. Chicago: University of Chicago Press.

Montgomery, Jacob M. and Brendan Nyhan. 2010. “Bayesian Model Averaging: Theoretical Developments and Practical Applications.” *Political Analysis*, 18 (Spring): 245 - 270.

Mooney, Christopher Z. 1996. "Bootstrap Statistical Inference: Examples and Evaluations for Political Science." *American Journal of Political Science*, 40(May): 570-602.

Rosnow, Robert L., and Robert Rosenthal. 1989. “Definition and Interpretation of Interaction Effects.” *Psychological Bulletin*, 105: 143 - 146.

Southwood, Kenneth E. 1978. "Substantive Theory and Statistical Interaction: Five Models." *American Journal of Sociology*, 83: 1154-1203.

Sovey, Allison J., and Donald P Green. 2011. “Instrumental Variables Estimation in Political Science: A Readers’ Guide.” *American Journal of Political Science*, 55(January): 188 - 200.

Trier, Shawn, and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science*, 52(January): 201 - 217.

Westen, D, and Robert Rosenthal. 2003. "Quantifying Construct Validity: Two Simple Measures." *Journal of Personality and Social Psychology*, 84(March): 608-618.

Western, Bruce. 1995. "Concepts and Suggestions for Robust Regression Analysis." *American Journal of Political Science*, 39(August): 786-817.

Yzerbyt, Vincent Y., Dominique Muller, and Charles M. Judd. 2004. "Adjusting Researchers' Approach to Adjustment: One the Use of Covariates when Testing Interactions." *Journal of Experimental Social Psychology*, 40(May): 424-431.