GOVT 302 Quantitative Methods

Fall 2015, CRN 14149

Professor Manna
Department of Government
30 Morton Hall

http://pmanna.people.wm.edu/
pmanna@wm.edu
221-3024

Course Meetings: MWF 1:00-1:50pm, 1 Morton Hall
Office Hours: T 12:00n-1:30pm, F 2:00-3:30pm, and by appointment

1. Course description

This course introduces students to basic statistical theory and advanced quantitative methods with a focus on statistical inference using multiple regression analysis. We will address ordinary least squares regression and limited dependent variable models, including models for binary, categorical, ordered, and count dependent variables. Examples will come from subfields across political science. Students will develop a firm grounding in quantitative methods, have opportunities to explore statistical models using real data, and become proficient users of the statistical software package Stata.

Prerequisites are GOVT 301 or GOVT 307 or instructor permission, provided there is evidence that the student has had some introduction to research design, statistical analysis, and statistical computing.

2. Course materials

2.1 Readings

These required books are available for purchase on-line and at the college bookstore:

- Larry D. Schroeder, David L. Sjoquist, and Paula E. Stephan. 1986. *Understanding Regression Analysis: An Introductory Guide.* Newberry Park, CA: SAGE. ISBN: 9780803927582. http://www.sagepub.com/books/Book1909.
- J. Scott Long and Jeremy Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata, 3rd Edition.* College Station, TX: Stata Press. ISBN 9781597181112. http://www.stata.com/bookstore/regmodcdvs.html. *Important note: This book has been dramatically changed from earlier editions so you need to be sure to get the 3rd edition.*

These resources are free:

- OnlineStatBook. http://onlinestatbook.com/2/index.html. Great for a review of basic concepts as needed.
- UCLA Institute for Digital Research and Education (IDRE). Resources to help you learn and use Stata. http://www.ats.ucla.edu/stat/stata/.
- Stata's YouTube channel: https://www.youtube.com/user/statacorp.

Other required readings are posted to the course Blackboard site or are available on-line.

2.2 Software

We will use Stata in this course. The program is available in the William & Mary computer labs, and at the public machines in Swem library. If you are interested in purchasing Stata, fortunately it is available at a discounted rate through William & Mary, which is a member of the Stata GradPlan. Let me know if you are interested in this option and I can show you how to purchase it.

3. Assignments and grading

3.1 Grading

I will calculate course grades based on the following items. You need to complete all items to receive course credit. Students not completing all items will receive an Incomplete.

Percent	Item
20	Homework assignments
25	Exam 1
25	Exam 2
30	Exam 3 or final paper

I will base final course grades on the following scale, with partial-percents typically rounded to the nearest full percent: A 93-100, A– 90-92, B+ 88-89, B 83-87, B– 80-82, C+ 78-79, C 73-77, C– 70-72, D+ 68-69, D 63-67, D– 60-62, F <60.

Finally, because errors sometimes creep into grade calculations—and on rare occasions papers become lost as they are handed in—please keep a copy of all work you submit until I have processed final grades.

3.2 Homework assignments

A key way to learn statistics is to do statistics. In order to put that principle into practice you will complete four homework assignments. The assignments will provide you with opportunities to experiment with a variety of methods using real data. Grading will stress two things: (1) the degree to which you have made a strong effort to complete all parts of each assignment; and (2) the extent to which your work, especially the statistical computing component, is polished and professionally done.

3.3 Exams

Everyone will take two take-home exams. Both will ask you to perform calculations and will emphasize interpreting results. You will have a choice between taking a third exam or completing a final paper. All three exams will be open-book and open-note.

3.4 Final Paper

This assignment is for students who do not take the third exam. The course's capstone paper will provide you an opportunity to use your quantitative skills on a topic that you choose. I will make some data sets available but you may also use data from another source. More details about the paper's requirements will be forthcoming.

3.5 Academic misconduct

I begin by assuming academic misconduct will not become an issue in this class. If it does, for any questions about policies regarding cheating, plagiarism, or other types of academic dishonesty please see the William & Mary Honor Council's web site and the discussion of the Honor Code in the Student Handbook. If I discover a student cheating I will initiate an Honor Council proceeding and, at a bare minimum, recommend that the student receive an F for the course.

4. Schedule of topics, readings, and assignments

We will adjust this schedule as needed.
BB = Reading is on Blackboard; L = Class meets in Swem Library, Dulin Learning Center (first floor).

Other important dates: The drop-add deadline is Sept. 4; the withdraw deadline is Oct. 23.

4.1 Preliminaries

W 8/26	Class introduction • David A. Freedman. 1991. Statistical Models and Shoe Leather. *Sociological Methodology*. vol. 21: 291-313. BB
F 8/28	Probability basics • Will H. Moore and David A. Siegel. 2013. *A Mathematics Course for Political and Social Research*, Ch. 9. An Introduction to Probability, pp. 175-182 (section 9.1). BB
M 8/31	Probability distributions • Moore and Siegel, Ch. 10 An Introduction to (Discrete) Distributions, pp. 209-218 (section 10.4 and 10.5) and pp. 218-229 (section 10.6). BB • Moore and Siegel, Ch. 11 Continuous Distributions, read pp. 242-247 (section 11.1); skim pp. 258-271 (section 11.3). BB
W 9/2	Stata nuts and bolts • Long and Freese, Ch. 2 Introduction to Stata, pp. 23-41. The discussion on pp. 40-41 (i.e., section 2.10) is incredibly important. • Stata intro video from the Research Support Center for the Family, Home, and Social Sciences College at BYU. Watch from the 0:00 to 31:00 mark. You don't need to download the files mentioned at the start of the presentation. https://www.youtube.com/watch?v=QaI_a_l2jqo
F 9/4	Creating and managing variables in Stata • Long and Freese, Ch. 2 Introduction to Stata, pp. 41-59. • Continuation of Stata intro video from the Research Support Center for the Family, Home, and Social Sciences College at BYU. Watch from the 31:00 to 54:00 mark. https://www.youtube.com/watch?v=QaI_a_l2jqo
M 9/7	Stata questions Sampling distributions • Online Statistics, Introduction to Sampling Distributions, http://onlinestatbook.com/2/sampling_distributions/intro_samp_dist.html. • Peter Kennedy, *A Guide to Econometrics (4e)*, pp. 313-316. BB
W 9/9	Sampling distributions (cont.) • Re-read Peter Kennedy, *A Guide to Econometrics (4e)*, pp. 313-316. BB

4.2 Method of ordinary least squares (OLS) regression

F 9/11	Drawing a regression line • Schroeder, Sjoquist, and Stephan, pp. 11-22
M 9/14, L	Lab exercise #1: Preliminaries. Meet in Swem Dulin Learning Center (first floor of Swem)
W 9/16	Drawing a regression line (cont.) • No new reading

F 9/18	Interpreting bivariate OLS • Schroeder, Sjoquist, and Stephan, pp. 22-29. • UCLA IDRE site. Scroll to Section 1.3 Simple Linear Regression at http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter1/statareg1.htm. Here's an annotated version of the output: http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter1/statareg_annotated1.htm. ****HOMEWORK #1 DUE BY 5:00PM****
M 9/21	Interpreting bivariate OLS (cont.) / Hypothesis testing and confidence intervals • Schroeder, Sjoquist, and Stephan, pp. 36-47 (stop before Left-Tail Tests).
W 9/23	Hypothesis testing and confidence intervals (cont.) • Schroeder, Sjoquist, and Stephan, p. 53. • Andrew Gelman and Eric Loken. 2013. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time, Nov. 14, pp. 1-3. http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf. • Andrew Gelman. 2015. Good, mediocre, and bad p-values, April 13. http://andrewgelman.com/2015/04/30/good-mediocre-bad-p-values/.
F 9/25	Interpreting multivariate OLS • Schroeder, Sjoquist, and Stephan, pp. 29-36, 51-53, and 62-65. • UCLA IDRE site. Scroll to Section 1.4 Multiple Regression at http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter1/statareg1.htm. Here's an annotated version of the output: http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter1/statareg_annotated2.htm.
M 9/28, L	Lab exercise #2: Interpreting bivariate and multivariate OLS results
W 9/30	Using dummy independent variables and interaction terms • Schroeder, Sjoquist, and Stephan, pp. 56-59. • Thomas Brambor, William Roberts Clark, and Matt Golder. 2006. Understanding interaction models: Improving empirical analyses. *Political Analysis* 14: 63-82. Read pp. 63 thru top of 66; try skimming the rest. An updated link for the website mentioned in the article is: http://mattgolder.com/interactions. BB
F 10/2	Using dummy independent variables and interaction terms (cont.) • No new reading

4.3 Unpacking OLS assumptions

M 10/5	Recovering and working with regression residuals • UCLA IDRE site. At this link look at Section 2.0 Regression Diagnostics and see Section 2.8 Summary for an overview of some of the techniques we will learn: http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm. ****HOMEWORK #2 DUE BY 5:00PM****
W 10/7	Model specification • Schroeder, Sjoquist, and Stephan, pp. 65-69 (stop before Stepwise Regression).
F 10/9	Model specification (cont.) / Measurement • Schroeder, Sjoquist, and Stephan, pp. 70-71 on Proxy Variables and Measurement Error. • UCLA IDRE site. Scroll to Section 2.1 Unusual and Influential Data at http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm.
M 10/12	No class—fall break

W 10/14	Measurement (cont.) / Error term • Schroeder, Sjoquist, and Stephan, pp. 72-75 (sections on Autocorrelation and Heteroskedasticity). • UCLA IDRE site. Scroll to Section 2.2 Checking Normality of Residuals and Section 2.3 Checking Homoskedasticity at http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm.
F 10/16	Error term (cont.) • Long and Freese, pp. 103-105 (section 3.1.9).
M 10/19	Multicollinearity • Schroeder, Sjoquist, and Stephan, pp. 71-72. • UCLA IDRE site. Scroll to Section 2.4 Checking for Multicollinearity at http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm. ****TAKE-HOME EXAM #1 DUE BY 5:00PM****
W 10/21	OLS diagnostics demonstration • It would be useful to re-read what we've discussed from the UCLA IDRE site at http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm.
F 10/23, L	Lab exercise #3: OLS diagnostics practice

4.4 Models for limited dependent variables

M 10/26	Method of maximum likelihood estimation (MLE) • Peter Kennedy, *A Guide to Econometrics (4e)*, pp. 21-22. BB • Long and Freese, pp. 83-86.
W 10/28	Method of maximum likelihood estimation (MLE) (cont.)
F 10/30	OLS meets MLE • Gary King. 1998. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. Ch. 1 Introduction, pp. 3-13. As you read, think hard about equations 1.1 and 1.2 (p. 8) and King's discussion about their parts. BB ****HOMEWORK #3 DUE BY 5:00PM****
M 11/2	Binary dependent variables – Theory and Estimation • Long and Freese, Ch 5. Models for Binary Outcomes: Estimation, Testing, and Fit, pp. 187-197.
W 11/4	Binary dependent variables – Interpretation • Long and Freese, Ch 5. Models for Binary Outcomes: Estimation, Testing, and Fit and Ch 6. Models for Binary Outcomes: Interpretation `predict`: pp. 206-209 `mchange`: pp. 246-264 `mtable`: pp. 270-284 (can skip section 6.3.3) `mgen`: pp. 286-300 `listcoef`: pp. 228-235 (especially, pp. 232-233)
F 11/6, L	Lab exercise #4: Binary dependent variables practice
M 11/9	Count dependent variables – Theory and Estimation • Long and Freese, Ch 9. Models for Count Outcomes, pp. 481-489 (Poisson regression), 507-510 (negative binomial regression), and 511-512 (testing for overdispersion).

W 11/11	Count dependent variables – Interpretation • Long and Freese, Ch 9. Models for Count Outcomes. Note: The examples of the commands below mainly focus on the Poisson model, but they could be used for either Poisson or negative binomial specifications. `predict`: pp. 501 `mchange`: pp. 493-498 `mtable`: pp. 496-499; 516-518 `mgen`: pp. 500-504; 516-518 `listcoef`: pp. 490-493; 514-515
F 11/13, L	Lab exercise #5: Count dependent variables practice
M 11/16	Ordered dependent variables – Theory and Estimation • Long and Freese, Ch 7. Models for Ordinal Outcomes, pp. 309-320. ****TAKE-HOME EXAM #2 DUE BY 5:00PM****
W 11/18	Ordered dependent variables – Interpretation • Long and Freese, Ch 5. Models for Ordinal Outcomes `predict`: pp. 339-341 `mchange` and `mchangeplot`: pp. 341-351; 364-370 `mtable`: pp. 351-359 `mgen`: pp. 359-370 `listcoef`: pp. 335-338
F 11/20, L	Lab exercise #6: Ordered dependent variables practice
M 11/23	Nominal dependent variables – Theory and Estimation • Long and Freese, Ch 8. Models for Nominal Outcomes, pp. 385-398 (multinomial logit), 407-410 (testing the IIA assumption), and 465-469 (multinomial probit).
W 11/25	No class—Thanksgiving break ****HOMEWORK #4 DUE BEFORE YOU LEAVE FOR THANKSGIVING BREAK****
F 11/27	No class—Thanksgiving break
M 11/30	Nominal dependent variables – Interpretation • Long and Freese, Ch 8. Models for Nominal Outcomes `predict`: pp. 412-415 `mchange` and `mchangeplot`: pp. 415-419; 424-425 `mtable`: pp. 423-424 `mgen`: pp. 432-434 `listcoef` and `mlogitplot`: pp. 435-444
W 12/2, L	Lab exercise #7: Nominal dependent variables practice
F 12/4	Class wrap up • Re-read David A. Freedman. 1991. Statistical Models and Shoe Leather. *Sociological Methodology*, vol. 21: 291-313. BB

****Wednesday, December 16. FINAL PAPER or TAKE-HOME EXAM #3 DUE BY 5PM****