GOVT 302 Quantitative Methods

Fall 2016, CRN 14149

Professor Manna
Department of Government
318A Tyler Hall

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

Course Meetings: MWF 10:00-10:50am, 123 Tyler Hall
Office Hours: M 11:00am-12:00n, Th 1:15-3:15pm, 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 using maximum likelihood estimation, 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. *Important note: Both of these books have been dramatically changed from earlier editions so you need to be sure to get the editions noted here.*

- Colin Lewis-Beck and Michael Lewis-Beck. 2015. *Applied Regression: An Introduction, 2nd edition*. Newberry Park, CA: SAGE. ISBN: 9781483381473. https://us.sagepub.com/en-us/nam/applied-regression/book244616.
- 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.

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.
- Just in case: https://www.etsy.com/shop/NausicaaDistribution/items.

Other required readings and materials, including handouts for each class day, are on Blackboard.

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 purchasing the software and I can show you how.

3. Assignments and grading

3.1 Grading

I will calculate course grades based on the following items.

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

I will base final course grades on the following scale, with partial-percents typically rounded to the nearest full percent: A 100-93, A– 92-90, B+ 89-88, B 87-83, B– 82-80, C+ 79-78, C 77-73, C– 72-70, D+ 69-68, D 67-63, D– 62-60, 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 your work 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 understand general concepts, perform calculations, and interpret results. You will have a choice between taking a final take-home 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 final 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 = Lab session; class will meet in Swem Library, Dulin Learning Center (first floor).

4.1 Preliminaries

W 8/24	Class introduction
F 8/26	Stata nuts and bolts • Long and Freese, Ch. 2 Introduction to Stata. Skim entire chapter, then go back and read carefully pp. 23-41; the discussion on pp. 40-41 (i.e., section 2.10) is incredibly important.
M 8/29	Probability basics *Note: The Moore and Siegel book is available as an e-book from Swem, too. • 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
W 8/31	Discrete and continuous 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
F 9/2	Probability applications • Re-read the material from 8/29 and 8/31 while reviewing your notes from those classes.
M 9/5, L	Lab exercise #1: Stata hands-on practice. Meet in Dulin Learning Center (first floor of Swem) • Long and Freese, Ch. 2 Introduction to Stata, pp. 46-59 • Watch Manna's video demo on Blackboard.
W 9/7	Estimation, probability, and sampling distributions • Peter Kennedy, *A Guide to Econometrics (4e)*, pp. 313-316. BB • Jeffrey M. Wooldridge, *Introductory Econometrics (5e)*, pp. 755-763; try hard to understand the message conveyed in Figure C.1 (p. 759) and Figure C.2 (p. 761). BB
F 9/9	Estimation, probability, and sampling distributions (cont.) • Re-read the material from 9/7 while reviewing your notes from that day. ****HOMEWORK #1 DUE BY 5:00PM****

4.2 Method of ordinary least squares (OLS) regression

M 9/12	Drawing a regression line • Lewis-Beck & Lewis-Beck, pp. 1-8.
W 9/14	Drawing a regression line (cont.) • No new reading.
F 9/16	Interpreting bivariate OLS • Lewis-Beck & Lewis-Beck, pp. 8-21 and 38-39. • UCLA IDRE. Read Section 1.0 Introduction, 1.2 Examining Data, and 1.3 Simple Linear Regression: http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter1/statareg1.htm. Here's an incredibly helpful annotated version of the output from section 1.3: http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter1/statareg_annotated1.htm.
M 9/19	Interpreting bivariate OLS (cont.) / Hypothesis testing and confidence intervals • Lewis-Beck & Lewis-Beck, pp. 29-38. • Ezra Hauer. 2004. The harm done by tests of significance. *Accident Analysis and Prevention* 36: 495-500. http://andrewgelman.com/wp-content/uploads/2014/12/1154-The-Harm-done-by-tests-of-significance.pdf.

W 9/21	Hypothesis testing and confidence intervals (cont.) • Lewis-Beck & Lewis-Beck, re-read pp. pp. 36-38. • 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, pp. 1-3 (read), 4-13 (skim) and 13-15 (read) http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf. • American Statistical Association. 2016. Press release: American Statistical Association releases statement on statistical significance and p-values. March 7. https://www.amstat.org/newsroom/pressreleases/P-ValueStatement.pdf. • FiveThirtyEight. n.d. Hack your way to scientific glory. http://projects.fivethirtyeight.com/p-hacking/.
F 9/23	Interpreting multivariate OLS • Lewis-Beck & Lewis-Beck, pp. 55-64. • 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. • Rodolfo Espino and Michael M. Franz. 2002. Latino phenotypic discrimination revisited: The impact of skin color on occupational status. *Social Science Quarterly* 83(2): 612-623. http://onlinelibrary.wiley.com/doi/10.1111/1540-6237.00104/epdf.
M 9/26, L	Lab exercise #2: Interpreting OLS results
W 9/28	Using dummy independent variables and interaction terms • Lewis-Beck & Lewis-Beck, pp. 64-69.
F 9/30	Using dummy independent variables and interaction terms (cont.) • Lewis-Beck & Lewis-Beck, pp. 69-74. • Thomas Brambor, William Roberts Clark, and Matt Golder. 2006. Understanding interaction models: Improving empirical analyses. *Political Analysis* 14: 63-82. http://pan.oxfordjournals.org/content/14/1/63.full.pdf+html. 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. ****HOMEWORK #2 DUE BY 5:00PM****
M 10/3	Multicollinearity • Lewis-Beck & Lewis-Beck, pp. 75-83. • UCLA IDRE site. Scroll to Section 2.4 Checking for Multicollinearity at http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm.

4.3 OLS assumptions and diagnostics

W 10/5	The data generating process • 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
F 10/7	Recovering and working with regression residuals • Lewis-Beck & Lewis-Beck, re-read pp. 1-20, focusing on the visuals on those pages. Specifically, consider what it means for the plotted data points to deviate from the regression line. How would you describe the distance between each point and the line?
M 10/10	No class—fall break

W 10/12	Model specification / Measurement • Lewis-Beck & Lewis-Beck, pp. 23-25. • Nate Silver. 2010. Study claiming link between stimulus funding and partisanship is manifestly flawed. FiveThirtyEight. April 1. http://fivethirtyeight.com/features/study-claiming-link-between-stimulus/. • Paul Krugman. 2013. There is no "true" unemployment rate. New York Times. July 14. http://krugman.blogs.nytimes.com/2013/07/14/there-is-no-true-unemployment-rate/.
F 10/14	Measurement (cont.) / Stochastic component • Lewis-Beck & Lewis-Beck, pp. 25-28 and 39-49.
M 10/17	Stochastic component (cont.) • Kenneth J. Meier and Laurence J. O'Toole, Jr. 2002. Public management and organizational performance. The effect of managerial quality. *Journal of Policy Analysis and Management* 21(4): 629-643. http://onlinelibrary.wiley.com/doi/10.1002/pam.10078/epdf. ****TAKE-HOME EXAM #1 DUE BY 5:00PM****
W 10/19	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 10/21, L	Lab exercise #3: OLS diagnostics practice

4.4 Models for limited dependent variables

M 10/24	Method of maximum likelihood estimation (MLE) • Peter Kennedy, *A Guide to Econometrics (4e)*, pp. 21-22. BB • Long and Freese, pp. 83-86.
W 10/26	Method of maximum likelihood estimation (MLE) (cont.) • Re-read assignment and review your notes from 10/24.
F 10/28	OLS meets MLE: Heteroskedastic regression example • Vincent L. Hutchings et al.. 2004. Congressional representation of black interests: Recognizing the importance of stability. *Journal of Politics* 66(2): 450-468. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-2508.2004.00159.x/epdf. ****HOMEWORK #3 DUE BY 5:00PM****
M 10/31	Binary dependent variables – Theory and Estimation • Long and Freese, Ch 5. Models for Binary Outcomes: Estimation, Testing, and Fit, pp. 187-197.
W 11/2	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. Focus on `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). • Paul Manna and Laura L. Ryan. 2011. Competitive grants and educational federalism: President Obama's Race to the Top Program in theory and practice. *Publius: The Journal of Federalism* 41(3): 522-546. Focus on the results in Table 1. http://pmanna.people.wm.edu/research/Manna&Ryan2011Publius.pdf.
F 11/4, L	Lab exercise #4: Binary dependent variables practice
M 11/7	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 11/9	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. Focus on `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. • Travis N. Ridout and Glen R. Smith. 2008. Free advertising: How the media amplify campaign messages. *Political Research Quarterly* 61(4): 598-608. http://prq.sagepub.com/content/early/2008/04/08/1065912908314202.full.pdf+html.
F 11/11, L	Lab exercise #5: Count dependent variables practice
M 11/14	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 11/16	Ordered dependent variables – Interpretation • Long and Freese, Ch 5. Models for Ordinal Outcomes. Focus on `predict` pp. 339-341; `mchange` and `mchangeplot` pp. 341-351 & 364-370; `mtable` pp. 351-359; `mgen` pp. 359-370; `listcoef`: pp. 335-338. • Paul Manna. 2002. The signals parents send when they choose their children's schools. *Educational Policy* 16(3): 425-447. http://pmanna.people.wm.edu/research/Manna2002_EducationalPolicy.pdf.
F 11/18, L	Lab exercise #6: Ordered dependent variables practice
M 11/21	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 11/23	No class—Thanksgiving break ****HOMEWORK #4 DUE BEFORE YOU LEAVE FOR BREAK****
F 11/25	No class—Thanksgiving break
M 11/28	Nominal dependent variables – Interpretation • Long and Freese, Ch 8. Models for Nominal Outcomes. Focus on `predict` pp. 412-415; `change` and `mchangeplot`: pp. 415-419 & 424-425; `mtable` pp. 423-424; `mgen` pp. 432-434; `listcoef` and `mlogitplot` pp. 435-444. • Courtenay Ryals Conrad and Will H. Moore. 2010. What stops the torture? *American Journal of Political Science* 54(2): 459-476. http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5907.2010.00441.x/epdf.
W 11/30, L	Lab exercise #7: Nominal dependent variables practice
F 12/2	Class wrap up • Re-read David A. Freedman. 1991. Statistical Models and Shoe Leather. *Sociological Methodology*, vol. 21: 291-313. BB

****Wednesday, December 14. FINAL PAPER or FINAL EXAM DUE BY 5PM****

Recommended winter break reading:

- Kevin A. Clarke and David M. Primo. 2012. *A Model Discipline: Political Science and the Logic of Representations*. New York: Oxford University Press.