GOVT 391-02 Quantitative Methods

Fall 2013, CRN 15282

Professor Manna
Department of Government
30 Morton Hall

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

Course Meetings: MWF 1:00-1:50am, 39 Morton Hall
Office Hours: T 3:30-5:00pm, F 2:00-3:30pm, and by appointment
Teaching Assistant: Lizzy Pelletier, efpelletier@email.wm.edu

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 those 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. 2006. *Regression Models for Categorical Dependent Variables Using Stata, 2nd Edition.* College Station, TX: Stata Press. ISBN 9781597180115. http://www.stata.com/bookstore/regmodcdvs.html.

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 conducting independent research or taking additional statistics courses beyond this class I would encourage you to purchase 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

The only 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 computer lab, Morton 244.

4.1 Preliminaries

W 8/28	Class introduction • David A. Freedman. 1991. "Statistical Models and Shoe Leather." *Sociological Methodology*. vol. 21: 291-313. BB
F 8/30	Working with different variable types in Stata • Long and Freese, Ch 2. "Introduction to Stata," pp. 16-18 (section 2.1), 22 (section 2.5), and 31-51 (sections 2.10-2.14). • Try to memorize the list that Long and Freese provide in section 2.10 (pp. 31-33).
M 9/2	Probability basics in Stata • Long and Freese, Ch 2. "Introduction to Stata," re-read pp. 37-41.
W 9/4	Probability distributions and data description in Stata • UCLA IDRE site, "Introduction to Graphs in Stata," at http://www.ats.ucla.edu/stat/stata/modules/graph8/intro/graph8.htm. Study the first six graphs shown (histograms and box-plots). How would you describe them? • National Institutes of Standards and Technology, "Gallery of Distributions" at http://www.itl.nist.gov/div898/handbook/eda/section3/eda366.htm. Do two things here: (1) Look at the visuals (and try to puzzle through the formulas) for these distributions: Normal, t, F, Chi-Square, and Poisson. How would you describe them? (2) Try to understand the distinction between these terms used to describe each distribution: "Probability Density Function" and "Cumulative Distribution Function."
F 9/6, L	Lab exercise #1: Stata fundamentals, variables, and probability – Meet in Morton 244
M 9/9	Sampling distributions • Online Statistics, "Introduction to Sampling Distributions," http://onlinestatbook.com/2/sampling_distributions/intro_samp_dist.html. • Online Statistics, "Sampling Distribution of the Mean," http://onlinestatbook.com/2/sampling_distributions/samp_dist_mean.html. • Peter Kennedy, *A Guide to Econometrics (4e)*, pp. 313-316. BB
W 9/11	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 9/13	Scatterplots and correlation • Schroeder, Sjoquist, and Stephan, "The Linear Correlation Coefficient," pp. 23-26. • UCLA IDRE site, "Introduction to Graphs in Stata." Study the graphs appearing after the pie chart http://www.ats.ucla.edu/stat/stata/modules/graph8/intro/graph8.htm. ****HOMEWORK #1 DUE BY 5PM**** **(8/28-9/6: Stata and probability fundamentals)**
M 9/16	Drawing a regression line • Schroeder, Sjoquist, and Stephan, pp. 11-22.
W 9/18	Drawing a regression line (cont.). • No new reading.

F 9/20	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.
M 9/23	Interpreting bivariate OLS (cont.) / Hypothesis testing and confidence intervals • Schroeder, Sjoquist, and Stephan, pp. 36-47 (stop before "Left-Tail Tests").
W 9/25	Hypothesis testing and confidence intervals (cont.) • Schroeder, Sjoquist, and Stephan, pp. 36-47.
F 9/27	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 9/30, L	Lab exercise #2: Interpreting bivariate and multivariate OLS results – Meet in Morton 244
W 10/2	Using dummy independent variables and interaction terms • Schroeder, Sjoquist, and Stephan, pp. 56-59. • Optional: Matt Golder – Multiplicative Interaction Terms website: https://files.nyu.edu/mrg217/public/interaction.html. Especially useful resources on this site are: (1) Thomas Brambor, William Roberts Clark, & Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14: 63-82; and (2) the section of the web page called "Standard Errors."
F 10/4	Using dummy independent variables and interaction terms (cont.) ****HOMEWORK #2 DUE BY 5PM**** **(9/9-9/30: Sampling distributions, Method of OLS and interpretation)**

4.3 Unpacking OLS assumptions

M 10/7	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.
W 10/9	Model specification • Schroeder, Sjoquist, and Stephan, pp. 65-69.
F 10/11	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 10/14	No class—fall break
W 10/16	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 10/18	Error term (cont.) ****TAKE HOME EXAM #1 DUE BY 5PM**** **(8/28-10/4: Preliminaries and Method of OLS)**
M 10/21	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.
W 10/23	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 10/25, L	Lab exercise #3: OLS diagnostics practice – Meet in Morton 244

4.4 Models for limited dependent variables

M 10/28	Method of maximum likelihood estimation (MLE) • Peter Kennedy, *A Guide to Econometrics (4e)*, pp. 21-22. BB • Long and Freese, pp. 75-78.
W 10/30	Method of maximum likelihood estimation (MLE) (cont.)
F 11/1	Binary dependent variables – Theory • Long and Freese, Ch 4. "Models for Binary Outcomes," pp. 131-135. ****HOMEWORK #3 DUE BY 5PM**** **(10/7-10/25: OLS diagnostics)**
M 11/4	No class—Manna conference travel
W 11/6	Binary dependent variables – Interpretation • Long and Freese, Ch 4. "Models for Binary Outcomes," pp. 136-140 and 157-177.
F 11/8, L	Lab exercise #4: Binary dependent variables practice – Meet in Morton 244
M 11/11	Count dependent variables – Theory • Long and Freese, Ch 8. "Models for Count Outcomes," pp. 349-350 and 356-357 (on Poisson regression) and 372-373 (on negative binomial regression).
W 11/13	Count dependent variables – Interpretation • Long and Freese, Ch 8. "Models for Count Outcomes," pp. 357-372 (on Poisson regression) and 374-381 (on negative binomial regression). NOTE: Pay close attention to section 8.3.3 on testing for overdispersion (pp. 376-377).
F 11/15, L	Lab exercise #5: Count dependent variables practice – Meet in Morton 244 ****TAKE HOME EXAM #2 DUE BY 5PM**** **(10/7 – 11/6: Unpacking OLS assumptions, MLE, Binary DVs)**
M 11/18	Ordered dependent variables – Theory • Long and Freese, Ch 5. "Models for Ordinal Outcomes," pp. 183-187.
W 11/20	Ordered dependent variables – Interpretation • Long and Freese, Ch 5. "Models for Ordinal Outcomes," pp. 188-193 and 202-217.
F 11/22, L	Lab exercise #6: Ordered dependent variables practice – Meet in Morton 244

M 11/25	Nominal dependent variables – Theory • Long and Freese, Ch 6. "Models for Nominal Outcomes with Case-Specific Data," pp. 223-228 (on multinomial logit), 243-246 (on testing the IIA assumption), and 272-276 (on multinomial probit). ****HOMEWORK #4 DUE BEFORE YOU LEAVE FOR THANKSGIVING BREAK**** **(11/11-11/22: Count DVs, Ordered DVs)**
W 11/27	No class—Thanksgiving break
F 11/29	No class—Thanksgiving break
M 12/2	Nominal dependent variables – Interpretation • Long and Freese, Ch 6. "Models for Nominal Outcomes with Case-Specific Data," pp. 228-235, re-read 243-246 (on testing the IIA assumption), and 246-260.
W 12/4, L	Lab exercise #7: Nominal dependent variables practice – Meet in Morton 244
F 12/6	Class wrap up • Re-read David A. Freedman. 1991. "Statistical Models and Shoe Leather." *Sociological Methodology*, vol. 21: 291-313. BB

****Wednesday, December 18. FINAL PAPER or TAKE HOME EXAM #3 DUE BY 5PM****

Recommended winter break reading

- Steven Johnson. 2006. *The Ghost Map: The Story of London's Most Terrifying Epidemic—and How it Changed Science, Cities, and the Modern World.* New York: Riverhead Books. A multimedia website about the book is here: http://www.theghostmap.com/.
- Kevin A. Clarke and David M. Primo. 2012. *A Model Discipline: Political Science and the Logic of Representations*. New York: Oxford University Press.