PUBP 602 Quantitative Methods I

Fall 2015, CRN 10181

Prof. Paul Manna   http://pmanna.people.wm.edu/
Department of Government   pmanna@wm.edu
Thomas Jefferson Program in Public Policy   221-3024
30 Morton Hall

Course meetings: 201 Morton Hall, MW 9:30-10:50am
Office hours: T 12n-1:30pm, F 2-3:30pm, and by appointment
Teaching Assistant: Rob Marty; ramarty@email.wm.edu; office hours: M 12:30-2pm, Th 3:30-5pm, and F 9:30-11am

1. Course overview

Quantitative evidence has become increasingly important for informing solutions to challenging policy problems. However, statistics do not "speak for themselves." If poorly generated they mean nothing. If poorly presented or explained they can misinform, mislead, and, on occasion, even do great harm. Thus, possessing the ability to create, interpret, and present quantitative evidence is essential for anyone beginning a career in public policy. Even if you never personally generate statistical results after you leave William & Mary you will inevitably have to evaluate the quantitative findings of others.

With that preface let me welcome you to PUBP 602. This class serves four important purposes. First, it provides an introduction to applied statistical methods and statistical computing. Second, it prepares students for more advanced statistics courses. Third, it emphasizes communicating statistical results to readers who may lack statistical training. Finally, it will help you develop the habits of mind that will make you careful critics and users of data. To serve those ends, in this class you will develop quantitative skills by actually practicing them.

Some of you may possess little or no quantitative background. That's okay. The work is challenging, but not impossible. If you already have statistical experience, please talk with me immediately to determine whether you should take this course.

At the outset, I also would like to emphasize one thing: In this class I will not ask you to do things that waste your time. In fact, *everything I will teach or require of you is based on the best practices of professional policy analysts and academic researchers*. In my own work I try to live up to the standards I will be teaching you.

2. Course materials

2.1 Readings

We will use the following required text in PUBP 602 and PUBP 603.
- Jeffrey M. Wooldridge. 2013. *Introductory Econometrics, A Modern Approach, 5th edition.* ISBN 9781111531041.

These great supplemental resources are free.
- Online Statistics: An Interactive Multimedia Course of Study. http://onlinestatbook.com/2/index.html.
- Stata's YouTube channel: https://www.youtube.com/user/statacorp. Stata is the statistical computing package we will use in class.
- UCLA Institute for Digital Research and Education (IDRE). Resources to help you learn and use Stata. http://www.ats.ucla.edu/stat/stata/.

2.2 Software

As I mentioned during orientation, we will use Stata in PUBP 602, PUBP 603, and in your other courses such as the Policy Research Seminar. I *strongly* urge you to buy the program. (By the way: Neither the Public Policy Program nor I earn any money as a result of that recommendation!) Fortunately here's how to buy it at a discounted rate via the Stata GradPlan:
- Surf to http://www.stata.com/order/new/edu/gradplans/student-pricing/.
- On that page, I would recommend purchasing Stata/IC 14 with perpetual license or Stata/SE 14 with perpetual license.
- Do NOT buy Small Stata 14. It has very limited capabilities.
- Once you make your transaction, you will receive a few emails from the Stata staff and they will provide you with instructions for how to download the program.

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
30	Exam 1
30	Exam 2
20	Final paper

In general, I will base grades on the following percentage 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.

In a graduate class such as this, any grade below a "B" on any assignment, exam, or paper suggests that a student is having trouble grasping basic course ideas, which are essential building blocks for future courses and the work world. Please talk with me or Rob (or both) if you find yourself having difficulty.

Finally, because errors sometimes creep into grade calculations (and on rare occasions assignments become lost after they have been handed in) please keep a copy of all work submitted for this course until final grades have been processed.

3.2 Homework assignments

These assignments will focus on basic concepts and real life policy questions. It is crucial that you complete these assignments on time. 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 and the writing component, is polished and professionally done.

3.3 Exams

We will have two take-home exams. Both will ask you to perform calculations and will emphasize interpreting results. The exams will be open-book and open-note.

3.4 Final Paper

The course's capstone paper will provide you an opportunity to use your quantitative skills in an area that you choose. I will make some data sets available but you may also use data from another source. More details about the paper will be forthcoming. In previous years, some students have used this assignment as a writing sample for internship and job applications.

4. Other important notes

4.1 Daily class operation

You will develop professional habits of mind and get the most out of class by doing these things.

The night before class:
- Do the readings. Do not expect to understand the material after only one read, but reading things over before class will help to prime your thinking for the day's topics.
- Check Blackboard for files to download for class. Download handouts, data sets, and Stata .do files and save to your personal H:\ drive. *I strongly encourage you to print out the .pdf handouts that I post to Blackboard, or have the ability to annotate them electronically*.
- Charge your laptop battery. Unfortunately, outlet power is not always conveniently located in our classroom, so don't rely on plugging in your machine during class. During class we will frequently break from lecture and discussion to have you try some procedures and activities in Stata.

At the beginning of each class before lecture begins:
- Arrive on time and quietly take your seat if you are late. Please sit at a table with 2 or 3 people. Do not sit alone. Our class activities work best when you are in teams.
- Have Stata running on your laptop computer or be sitting next to someone who does.
- Close your email and Internet entertainment and disable all other electronic distractions.

During class:
- Ask questions when you do not understand something.
- Do not attend to email, the Internet, or other electronic distractions, including phones. If you are awaiting an important call for a job interview or something similar, please alert me to that fact ahead of class and excuse yourself quietly to take the call if it comes in.

4.2 Appropriate use of technology in class

As section 4.1 suggests, your laptops will be powerful educational tools for this class because they will allow us to do applied exercises together to help you learn course concepts. However, do not let your laptops or other electronic devices distract you, your fellow classmates, or me from our in-class work. Students who use electronic devices in class inappropriately suggest that they possess neither the intellectual focus nor the respect for others needed to do real professional work. Those students end up developing reputations that make it difficult for faculty members to give them strong recommendations to other professors and future employers.

Further, a mounting body of research has shown that students learn less when they rely on laptops and other electronic devices for note-taking in class. I would strongly encourage you to consider taking hand-written notes and reserve your laptop for the statistical computing exercises we will do in class.

4.3 Academic misconduct

I begin each semester by assuming that academic misconduct will not be an issue, but as a policy matter I mention this on every syllabus. For any questions about policies regarding cheating, plagiarism, or other types of misconduct, please refer to the web site of the William & Mary Honor Council and the relevant pages about the Honor Code from the Student Handbook. If I discover a student violating the Honor Code I will initiate an Honor Council proceeding and, at a minimum, recommend to the Council that the student receive an "F" for the course. I take this issue extremely seriously, and hope you do, too.

5. Schedule

We will adjust this schedule as needed. Any changes to assignment or exam due dates will provide you with more time, not less time, to complete the work. You'll notice that the reading assignments repeat for some days. That is intentional because re-reading certain pages in a new context will help to deepen your understanding of prior concepts while establishing new ones.

5.1 Preliminaries

W 8/26	Course overview
	Measurement
	• OnlineStatBook, Ch. 1 Introduction, Levels of Measurement. Read the short sections called "Types of Scales" and "Consequences of Levels of Measurement." http://onlinestatbook.com/2/introduction/levels_of_measurement.html
	• Andrew Gelman. 2015. What's the most important thing in statistics that's not in the textbooks? April 28. http://andrewgelman.com/2015/04/28/whats-important-thing-statistics-thats-not-textbooks/. Be sure to at least skim the comments, too; some are quite interesting.
	• Paul Krugman. 2013. There is No "True" Unemployment Rate, July 14. http://krugman.blogs.nytimes.com/2013/07/14/there-is-no-true-unemployment-rate/
M 8/31	Probability basics
	• Wooldridge, B.1 Random Variables and Their Probability Distributions (pp. 722-727) and B.2 Joint Distributions, Conditional Distributions, and Independence (pp. 727-730).
	• Online Statistics, Ch. 5. Probability, "Basic Concepts." http://onlinestatbook.com/2/probability/basic.html.

W 9/2	Stata nuts and bolts *Advice for today's material: I would suggest that you skim over the material in the Stata GS manual and then watch the intro video. Then go back and read more carefully parts of the GS manual for topics where you would like some additional clarification.* • Stata Getting Started [GS] manual (either for Mac or Windows). Access this by opening Stata and then going to the "Help" menu. Select "PDF Documentation." If you do not own Stata you can do this reading using a machine in one of the computer labs. Read Part 2. The Stata user interface, Part 4. Getting help, and Part 13. Using the Do-file Editor—automating Stata. • Stata intro video from the Research Support Center for the Family, Home, and Social Sciences College at BYU. Watch from the 0:00 to 54:00 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
M 9/7	Probability distributions and random variables • Wooldridge, B.1 Random Variables and Their Probability Distributions (pp. 722-727).
W 9/9	Application: The normal distribution • Wooldridge, B.5 The Normal and Related Distributions (pp. 745-752).

Friday, September 11. HOMEWORK #1 DUE TO ROB BY 5PM

5.2 Data description

M 9/14	Describing one variable: Numerical summaries of central tendency and spread • Wooldridge, A.1. The Summation Operator and Descriptive Statistics (pp. 703-705); B.3. Features of Probability Distributions (pp. 730-737). • Stata YouTube channel, Descriptive statistics in Stata. https://www.youtube.com/watch?v=kKFbnEWwa2s • UCLA IDRE site. (1) "Descriptive information and statistics," http://www.ats.ucla.edu/stat/stata/modules/descript.htm, (2) "Labeling data," http://www.ats.ucla.edu/stat/stata/modules/labels.htm, and (3) "Creating and recoding variables," http://www.ats.ucla.edu/stat/stata/modules/vars.htm
W 9/16	Describing relationships between two discrete variables: Crosstabulation tables • Wooldridge, B.2 Joint Distributions, Conditional Distributions, and Independence (pp. 727-730). • Stata YouTube channel, Tables and cross tabulations in Stata (relevant content starts at around the 4:00 mark). https://www.youtube.com/watch?v=3WpMRtTNZsw
M 9/21	Describing relationships between two continuous variables: Scatterplots and correlation • Wooldridge, B.4 Features of Joint and Conditional Distributions (pp. 737-745). • Stata YouTube channel, Basic scatterplots in Stata. https://www.youtube.com/watch?v=GhVGpe3lb3E • Stata YouTube channel, Pearson's correlation coefficient in Stata. https://www.youtube.com/watch?v=o7ko844ff-g
W 9/23	Describing relationships between two variables: Scatterplots and correlation (cont.)

Friday, September 25. HOMEWORK #2 DUE TO ROB BY 5PM

M 9/28	Data description in visualizations and presentations • Edward Tufte, Feynman-Tufte principle, http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001kE. • Andrew Gelman and Antony Unwin. 2013. Infovis and Statistical Graphics: Different Goals, Different Looks. January 20. http://www.stat.columbia.edu/~gelman/research/published/vis14.pdf. Read pp. 7-11 very carefully (section on "Some goals involving the visual display of quantitative information"); skim the rest.
W 9/30	Data description in visualizations and presentations • Stata Graphics [G] manual. G-1 Introduction and Overview. • UCLA IDRE site, "Introduction to Graphs in Stata," http://www.ats.ucla.edu/stat/stata/modules/graph8/intro/graph8.htm.
M 10/5	Data description in visualizations and presentations exercise • Bring your write-up for Homework #2 to class.
W 10/7	No class—Washington program trip
M 10/12	No class—fall break
W 10/14	Describing relationships between two variables: Regression • Wooldridge, A.2 Properties of Linear Functions (pp. 705-709). Then read the first two paragraphs (p. 710) of A.3 Some Special Functions and Their Properties and skim the pictures in section A.3. How do the pictures in A.2 differ from the pictures in A.3? • Wooldridge, B.4 Features of Joint and Conditional Distributions, re-read part on "Conditional Expectation" (pp. 741-742). • Wooldridge, Chapter 2. The Simple Regression Model, sections 2.1, 2.2 (study the figures; you can skim the text), and 2.3. • Stata YouTube channel, Simple linear regression in Stata. You can skip the part about inference. https://www.youtube.com/watch?v=HafqFSB9x70

Friday, October 16. TAKE-HOME EXAM #1 DUE TO MANNA BY 5PM.

M 10/19	Describing relationships between two variables: Regression (cont.) • Re-read the assignment from 10/14. Be sure you can interpret the discussion in example 2.1 (p. 23), example 2.2 (p. 24), example 2.3 (p. 32), example 2.4 (pp. 33-34), and example 2.5 (p. 34). • Woodridge, A.3 Proportions and Percentages (pp. 707-709). Be sure you understand the discussion in example A.3 (p. 709). If you ever misuse or conflate the concepts "percentage change" and "percentage point change," later you will be placed in the pillory in Colonial Williamsburg to pay for your statistical crimes. See: http://www.flickr.com/photos/36652543@N06/6046817425/.

5.3 *Statistical inference*

W 10/21	Population, sample, and sampling distributions • Wooldridge, C.1 Populations, Parameters, and Random Sampling (pp. 755-756) and C.2 Finite Sample Properties of Estimators (pp. 756-763). Try hard to understand the message conveyed in Figure C.1 (p. 759) and Figure C.2 (p. 761).
M 10/26	Population, sample, and sampling distributions (cont.) • Wooldridge, C.3 Asymptotic or Larger Sample Properties of Estimators (pp. 763-767).

W 10/28	Confidence intervals and hypothesis testing • Wooldridge, C.5 Interval Estimation and Confidence Intervals (pp. 770-777). The paragraph on p. 771 that contains the phrase "A confidence interval is often interpreted as follows…" is very important, as is the "rule of thumb" at the bottom of p. 775.

Friday, October 30. HOMEWORK #3 DUE TO ROB BY 5PM

M 11/2	Confidence intervals and hypothesis testing (cont.) • Wooldridge, C.6 Hypothesis Testing (pp. 777-789). The section on "Practical Versus Statistical Significance" (p. 788) is very important.
W 11/4	Confidence intervals and hypothesis testing (cont.) • Online Statistics, Ch. 11 Logic of Hypothesis Testing, "Significance Testing and Confidence Intervals" http://onlinestatbook.com/2/logic_of_hypothesis_testing/sign_conf.html.
M 11/9	Inference for means and differences of means • Online Statistics, Ch. 12 Tests of Means, (1) "Single mean," http://onlinestatbook.com/2/tests_of_means/single_mean.html and (2) "Difference between two means," http://onlinestatbook.com/2/tests_of_means/difference_means.html • Stata YouTube channel: (1) One-sample t-test in Stata. https://www.youtube.com/watch?v=HwzCyqW-0dc and (2) t-test for two independent samples in Stata. https://www.youtube.com/watch?v=by4c3h3WXQc
W 11/11	Inference for means and differences of means (cont.)

Friday, November 13. HOMEWORK #4 DUE TO ROB BY 5PM

M 11/16	Inference for regression • Online Statistics, Ch. 14 Regression, "Inferential statistics for b and r," http://onlinestatbook.com/2/regression/inferential.html. You can focus on the discussion of the slope rather than the correlation coefficient. • Stata YouTube channel, Simple linear regression in Stata. https://www.youtube.com/watch?v=HafqFSB9x70.
W 11/18	Inference for regression (cont.)
M 11/23	Inference for cross-tabulation (contingency) tables • Online Statistics, Ch. 17 Chi Square, "Contingency tables," http://onlinestatbook.com/2/chi_square/contingency.html • Stata YouTube channel, Pearson's chi-squared and Fisher's exact test in Stata. https://www.youtube.com/watch?v=DBsMPZqJj-o
W 11/25	No class—Thanksgiving break
M 11/30	Error in hypothesis testing • Online Statistics, Ch. 11 Logic of Hypothesis Testing, "Type I and Type II Errors" http://onlinestatbook.com/2/logic_of_hypothesis_testing/errors.html
W 12/2	Multiple regression warmup • Wooldridge, re-read Chapter 2. The Simple Regression Model, sections 2.1, 2.2, 2.3. • Wooldridge, Chapter 3. Multiple Regression Analysis: Estimation, section 3.1.

Wednesday, December 9. TAKE-HOME EXAM #2 DUE TO MANNA BY 5PM.
Wednesday, December 16. FINAL PAPER DUE (VIA EMAIL) TO MANNA BY 5PM