

ED 711 – Intermediate Quantitative Research Methods (version 2)

Tuesdays, 4:10-6:55, 105 Withers Hall

Instructor: Dr. Stephen Porter, Dept. of Leadership, Policy, and Adult and Higher Education
Email: srporter@ncsu.edu (*best way to contact me*)
Skype: stephen.porter50010
Office hours: Poe 616A, by appointment

The focus of this course is on multivariate statistical analysis using multiple regression and logistic regression, with applications in educational research. Unlike ED710, we will be using multiple independent variables simultaneously rather than one at a time, with both continuous (e.g., test scores) and discrete (e.g., dropout) dependent variables.

By the end of the course you should be able to:

- Understand how to estimate and interpret regression results
- Perform model diagnostics to detect violations of regression assumptions
- Analyze categorical dependent variables, such as yes/no outcomes
- Creating meaningful tables of results from statistical output
- Understand articles in education journals that use these methods

We will be covering these topics as a way of introducing you to some of the primary statistical techniques used in education. To use any of these techniques in your dissertation may require additional study on your part.

Previous coursework

To succeed in this course you must have taken either ED 710 or ST 507. If you have not taken either course [here at NCSU](#), please see me immediately.

Because this course assumes prior work in graduate-level statistics, I assume that you have a foundation of high school algebra and introductory statistics. I will thus assume that everyone is very familiar with the material covered in ED 710. If a) you are still uncertain about some of the concepts from your introductory statistics courses and/or b) cannot do high-school algebra (e.g., variables, equations, exponents, solving linear equations), then you should get up to speed in these areas before you take this course.

Readings

Required text:

Wooldridge, J.M. (2009). *Introductory Econometrics*. 4th edition, South-Western/Cengage.

In addition to the required text listed above, you will read additional book chapters and articles. All additional readings are noted with an * and will be available via Moodle.

Evaluation

Your grade will be determined as follows:

1. *Take-home exams (60%)* – You will be given 3-4 assignments during the semester that cover various topics in the course; you will have one week to complete each assignment. These assignments will generally involve analyzing datasets and reporting results from your analyses, as well as interpreting results reported in journal articles. You may use your notes and assigned readings to do the assignments, but you may not consult with any other people or sources (e.g., the Internet). I refer to these short papers as exams to stress that you should complete these assignments completely on your own, without any computer or statistical assistance from anyone else (other than me, of course). Late assignments will not be accepted, because I review the assignment answers during class on the day the assignment is due; you will receive a 0 for that assignment.

I will grade these assignments progressively harder throughout the semester, as you become more familiar with Stata and the material.

2. *Take-home final (40%)* – This will be similar to the take-home exams, only longer and covering the entire semester of material.

I will use the standard grading scale at NCSU:

97 ≤	A+	≤	100	77 ≤	C+	<	80
93 ≤	A	<	97	73 ≤	C	<	77
90 ≤	A-	<	93	70 ≤	C-	<	73
87 ≤	B+	<	90	67 ≤	D+	<	70
83 ≤	B	<	87	63 ≤	D	<	67
80 ≤	B-	<	83	60 ≤	D-	<	63
				0 ≤	F	<	60

Academic misconduct

Academic misconduct in any form is in violation of North Carolina State University policy and will not be tolerated. This includes, but is not limited to: copying or sharing answers on assignments, plagiarism, and having someone else do or assist you with your academic work. See <http://policies.ncsu.edu/policy/pol-11-35-01> for more information.

Submitting an assignment to me electronically is the equivalent to signing the university Honor Pledge: "I have neither given nor received unauthorized aid on this test or assignment."

Disabilities

Reasonable accommodations will be made for students with verifiable disabilities. In order to take advantage of available accommodations, students must register with Disability Services for Students at 1900 Student Health Center, Campus Box 7509, 515-7653. For more information on NC State's policy on working with students with disabilities, please see the Academic Accommodations for Students with Disabilities Regulation (REG02.20.01).

Software

We will be using Stata IC (version 12) for the course. Stata is similar to SPSS, but is much more powerful in terms of what it can do. You have three options for using Stata:

1. Virtual Computing Lab (VCL) - With the VCL, you make an advance reservation to have Stata run virtually on your home computer, during class and whenever you need it to work on an assignment. Information on using the VCL can be found here:
 - <https://vcl.ncsu.edu/help>
 - <https://sites.google.com/a/ncsu.edu/cedstata>
2. Rent – you can rent Stata for either six-months (\$65) or one year (\$98).
3. Purchase – you can purchase a copy of Stata for \$179; this is a perpetual license, meaning that you will only pay this fee once.

You can rent or purchase Stata here: <http://www.stata.com/order/new/edu/gradplans/direct-ship-pricing>.

Deciding between the VCL and renting/purchasing is generally a tradeoff between cost and convenience. The VCL is free, but can sometimes take up to 30 minutes to load onto your computer. Some students also find using software via the VCL difficult; for example, if you are not careful about where you save your files, you can lose them after the session ends. Conversely, renting or purchasing Stata involves a cost, but under the Stata licensing terms you can install Stata on three different computers (e.g., home desktop, laptop and office computer) as long as only one copy is being used at a time. In general, if you know you will be doing a quantitative dissertation, you should consider buying a copy, particularly if you will be working with categorical dependent variables (Stata makes it very easy compared to SPSS and SAS).

The set of reference manuals comes with the software as a pdf file, and I will be providing you with syntax, so there is absolutely no need to purchase any documentation. If you have not used Stata before, and wish to have a supplemental text, I recommend purchasing this text if you can find a cheap, used copy online: Kohler, Ulrich, & Kreuter, Frauke. (2005). *Data Analysis Using Stata*. College Station: Stata Press. Otherwise, I would not bother with a text on Stata, as I will be showing you how to use the software in class. UCLA also has an excellent website linking to various Stata resources: <http://statcomp.ats.ucla.edu/stata/default.htm>.

You should be prepared to take detailed notes throughout the course, both on the statistical material and also on how to use Stata. I will post copies of the relevant Stata output on Moodle for you to notate. I will also post the Stata command files that I use in class each week; you should download this file before every class. For the most part, you will be able to cut and paste these commands and make changes in them for your assignments.

Besides using the Moodle course website, you will also be using Dropbox to submit your assignments. I will send out invitations for Dropbox in a few weeks.

Course schedule

I would rather cover 75% of the syllabus very well, than cover 100% and have students leaving the course scratching their heads, so I will change the schedule as needed. Any changes to the course schedule will be discussed in class. Depending on the pace of the class, we may not cover all topics listed on the syllabus. I may also change topics and add readings as the semester progresses.

You will note that there are no dates for when take-homes will be handed out (other than the final). I leave this open, because I pace the class according to how well students seem to be assimilating the material. I may occasionally spend two or three weeks on a topic, rather than one. This allows me to hand out assignments when I am confident that the majority of class will do well, rather than according to a preset schedule. I will tell you at least a week in advance when a take-home will be handed out.

Important dates:

- October 23rd – no class
- November 27th – last class
- December 5th – final exam opens at noon
- December 12th – final exam due at noon

Topics

Introductions and expectations; review of mathematics and statistics; using Stata

1. Wooldridge, chapter 1 and Appendix A.
2. Review of algebra (see Appendix to this syllabus below)
3. ***If you have never used Stata before:** open the documentation that comes with Stata, and read the following chapters in *Getting Started With Stata*: 1-5, 9, 10, 13, 16, and 17 (I have also posted GSWS on Moodle).

Introduction to linear regression; using Stata continued

1. Wooldridge, chapter 2.
2. *Wainer, H. (1997). Improving tabular displays, with NAEP tables as examples and inspirations. *Journal of Educational and Behavioral Statistics*, 22(1), 1-30.

Estimation

1. Wooldridge, chapter 3.

Hypothesis testing

1. Wooldridge, chapter 4.
2. *Horowitz and Spector (2005). Is there a difference between private and public education on college performance? *Economics of Education Review*, 24(2), 189–195.

Properties, standardization and functional form

1. Wooldridge, chapter 5 and chapter 6 (pp. 184-199).
2. *Article to be determined.

Goodness of fit and predictions

1. Wooldridge, rest of chapter 6.

Dummy variables

1. Wooldridge, chapter 7.
2. Review Horowitz and Spector.
3. *Article to be determined.

Heteroskedasticity

1. Wooldridge, chapter 8.

Model specification

1. Wooldridge, chapter 9.
2. *Babyak, M.A. (2004). What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine*, 66, 411-421.

Logistic regression with binary dependent variables

1. Wooldridge, chapter 17.
2. *DesJardins, S. (2001). Assessing the effects of changing institutional aid policy. *Research in Higher Education*, 42(6), 653-678.

Logistic regression with nominal dependent variables

1. *Article to be determined.

Appendix – Getting Up to Speed with Algebra

Understanding basic high school algebra is crucial for success in this course. Please review this website:

http://www.wtamu.edu/academic/anns/mps/math/mathlab/beg_algebra/

It will make the rest of the semester much easier for you, I promise.

If you think your algebra skills are pretty good, just do tutorials 10, 19, 25 and 31. These are practice tests for the four sections relevant for this course. If you don't do well on some problems, go back and review the tutorials for that particular kind of problem.

Basically, my expectation is that you should be able to do the practice tests and answer all of the problems correctly, with the following exceptions:

- Tutorial 10: ignore problems 4a-4f and 5a
- Tutorial 19: ignore problems 4b-4f
- Tutorial 25: ignore problems 7a-7b and 8a
- Tutorial 31: ignore problems 1c-1f, 4a-4b and 5a-5c

If you are looking for a thorough review, I suggest the following tutorials:

- 3,4,8 and 9 (the others in this section are pretty basic and I assume you don't need to review them)
- 11-18
- 20-24
- 26-30

All of these sections are important, but the most important are 20-24, because they cover graphing and the slope of a line. We will be estimating slopes all semester, so this is a vital concept for you to understand.