

Sociology 613
nominal title: "Seminar in Multiple Regression and Path Analysis"
Advanced Graduate Statistics
Fall 2021

Pat Hastings (pat.hastings@colostate.edu). Office hours are Monday/Wednesday 2-3:15pm and by appointment in Clark B248. Anytime my office door is open, you are welcome to come in.

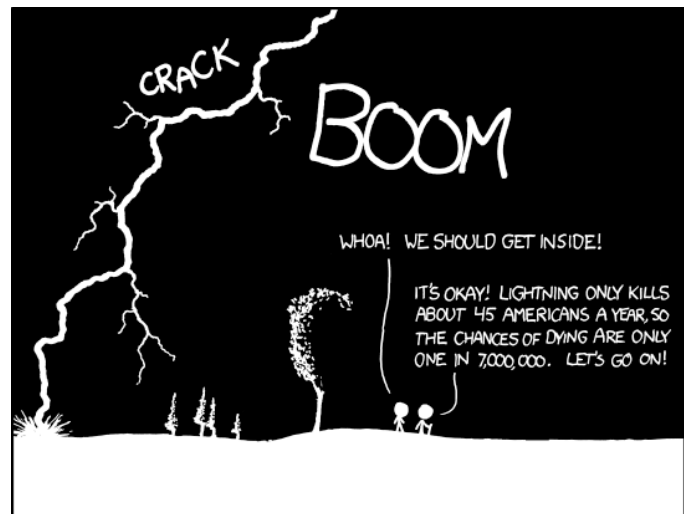
Lecture: Monday/Wednesday 3:30-4:45pm in Clark B252

Course Website: <https://colostate.instructure.com/courses/126735>

Due dates and other details of the syllabus will most likely change over the semester. Changes will be announced in class and by email.

Overview

This course is designed to help graduate students advance their ability to conduct, read, evaluate, and publish sociological research using quantitative methods. We will discuss the most commonly used statistical methods and most commonly encountered statistical issues by sociologists. Specific topics will include: regression, causal inference, mediation, moderation, model fit and selection, logistic regression, missing data techniques, methods for panel and multi-level data analysis, replication and transparency, and an introduction to machine learning. The main focus will be on using the methods, and less focus will be on the mathematical details. Throughout the course, students will use statistical software to put their skills into practice, and they will write an empirical research paper to gain competence and confidence with quantitative methods.



THE ANNUAL DEATH RATE AMONG PEOPLE WHO KNOW THAT STATISTIC IS ONE IN SIX.

Learning Goals

- Identify the appropriate statistical methods and models given a specific type of data and research question.
- Interpret the results of various statistical models and discuss their relevance for testing hypotheses, answering research questions, and evaluating competing theories.

- Gain expertise with a statistical software package to conduct statistical analyses.
 - Know how to effectively read, review, and write a quantitative paper.
 - Build a strong statistical foundation in order to be able to learn new advanced statistical methods as needed.
-

Prerequisites, or, What You Should Already Know

A prerequisite for this course is a graduate-level statistics course, such as Soc 511. You should be familiar with the basics of probability, statistical inference, significant tests, correlation, and regression, as well as the use of a statistical software package. I recognize students' previous training (and retention of it) may vary widely, and we will adjust the course as necessary. If your background in statistics and methods is not very strong, you can still do very well in this course, but you should expect to put in some extra work at the beginning. If you are unsure about your preparedness for this course, talk to me.

Evaluation

Your grade will be assigned using the standard letter +/- system based on:

- Course assignments (60%)
- A research paper (40%)

Course assignments (60%)

Most of the assignments will involve both paper-based and software-based components. I will drop your lowest score and your grade will be the average of the rest. I strongly encourage you to work with other students in the class, however, you must each turn in your own assignment. Unless stated otherwise, assignments will be due in Canvas by midnight on the Sunday after they are assigned.

Research paper (40%)

Each student will write a final paper that uses statistical methods from this course to answer a theoretically motivated research question of your choosing. The analysis you conduct should include a regression analysis (unless you discuss with me and we agree otherwise) using data that can effectively address your research question. You will write about your findings in the form of a typical published sociology journal article, and my hope is that many of you will ultimately be able to publish the paper that you start in this class. There will be five parts to this paper to be turned in:

- Abstract (< 300 words)
- Extended abstract with research questions and theoretical motivation, basic outline of data and methods
- Data and methods section (no results)
- A sh!tty first draft (with results)
- Final paper

Textbook

There is one required book:

Agresti, Alan. 2018. *Statistical Methods for the Social Sciences* (5th edition). Pearson (ISBN: 9780134507101)

We will discuss options for accessing the book at the beginning of the course. Any other readings for the class will be posted on the course website.

Statistical Computing

I will teach the course using Stata (www.stata.com). Stata is flexible, powerful, relatively user-friendly, and commonly used by social scientists. It has a large and diverse user community with many user-written commands that keep Stata continually up to date with new developments. You will need to use Stata for most classes, the assignments, and your final paper. We will discuss options during the beginning of the course.

If you prefer to complete the course using R or Python, you can discuss this option with me. I also use these tools. Note, however, that I will be providing class examples and assignment solutions in Stata. Relative to these other programs, Stata has the lower bar to entry and it's not too hard to move from one program to another once you really understand what you are trying to do!

How can I do well?

Since this is a statistics course, it will be different from the typical sociology course. Here is some advice:

- Most of the material is cumulative, so it is absolutely essential that you keep up with the course material. If you find yourself falling behind, ask for help!
- Being good at statistics requires thinking through how to solve problems. Statistics cannot be learned simply by reading a book or listening to a lecture. You should not expect to really understand the material until after you have completed the relevant assignments.
- Learning to do statistics and use statistical software is in many ways like learning a language. It gets easier the more you use it. Start the assignments early so you have time to work on them in multiple sessions and so can talk with others and get help if you need it.
- Please ask questions if you do not understand something. If it is unclear to you, it is probably unclear to other students as well. If it doesn't make sense in lecture, it will probably not make sense later when you are staring at your notes and trying to do the course assignments.

- Finally, I am aware that many of you are dealing with challenging situations. My hope is that this class is enjoyable and useful to you, without being a tremendous burden. If there are external factors that are detrimentally affecting your performance in this class, please consider talking with me so we can adjust our expectations appropriately.

COURSE OUTLINE

Please note, this is a tentative schedule. If (and when) there are changes to this schedule, you will receive adequate notice. All substantive topics below will be simultaneously covered in Stata. The numbers in parentheses roughly correspond to the relevant chapter of the Agresti textbook. Other readings will be assigned throughout the semester.

#	Date	Substantive topics	Due
1	8/23, 8/25	<ul style="list-style-type: none"> • Here we go! What's the course about? • Statistics review: sampling, distributions, descriptive and inferential statistics, hypothesis testing, statistical significance (1-8) 	Ungraded assessment (due 8/25)
2	8/30, 9/1	<ul style="list-style-type: none"> • Stata review: do files, basic commands, survey weights • Linear and multiple regression basics, correlation (9-10, 11.1) • Estimation and interpretation of Regression Models (11, 13) 	Assignment 1
3	9/8	<p><i>Note: 9/6 is a holiday. No class!</i></p> <ul style="list-style-type: none"> • Causality, DAGs • Mediation, path analysis, SEM (10) 	Assignment 2
4	9/13, 9/15	<ul style="list-style-type: none"> • Moderation, interaction terms (10.3, 11.4) • Writing a quantitative paper. 	Assignment 3
5	9/20, 9/22	<ul style="list-style-type: none"> • More with variables in models: standardized coefficients, logged variables, polynomials (11.7, 14.5-6) • Comparing regression models, model selection, and goodness of fit statistics (11.5, 14.1) • Making good tables and when/why figures are better 	Assignment 4
6	9/27, 9/29	<ul style="list-style-type: none"> • Regression assumptions and diagnostics (14.2-3) • General Linear Model, Poisson and Negative Binomial Regression (14.4) 	Assignment 5 <i>Paper: Abstract (< 250 words)</i>
7	10/4, 10/6	<ul style="list-style-type: none"> • Logistic regression (15) • Why not just do LPM? Good question... 	Assignment 6

Syllabus – Sociology 613: Advanced Graduate Statistics

8	10/11, 10/13	<ul style="list-style-type: none"> Ordered logit, Multinomial logit 	<i>Paper: Extended abstract with research questions and theoretical motivation, basic outline of data and methods</i>
9	10/18, 10/20	<ul style="list-style-type: none"> Missing data issues and multiple imputation (16.1) 	Assignment 7
10	10/25, 10/27	<ul style="list-style-type: none"> Panel data and fixed effects models 	Assignment 8
11	11/1, 11/3	<ul style="list-style-type: none"> Multi-level data, random effects, HLM (16.2) 	Assignment 9
12	11/8, 11/10	<ul style="list-style-type: none"> The dark side of statistics: p-hacking, the garden of forking paths Transparency and replication 	<i>Paper: Data and methods section (no results)</i>
13	11/15, 11/17	<p><i>Note: the following week is Fall Break. No class!</i></p> <ul style="list-style-type: none"> Reviewing a quantitative paper 	Assignment 10
14	11/29, 12/1	<ul style="list-style-type: none"> Introduction to machine learning. New methods or a new vocabulary? Supervised and unsupervised methods 	<i>Paper: A shitty first draft (with results)</i>
15	12/6, 12/8	<ul style="list-style-type: none"> TBD. Potential topics depending on time and interest: More machine learning, clustering, latent variables, factor analysis, more causal inference 	
		<ul style="list-style-type: none"> Final Project (exact due date TBA) 	<i>Paper: Final paper</i>

Other Important Matters

In-Class Decorum: Please be prepared to give the class your full concentration. Avoid checking email/social media/etc during class time. Colorado State University has stated five Principles of the Community: inclusion, integrity, respect, service and social justice (<http://diversity.colostate.edu/principles-of-community/>). Your conduct in this class should adhere to these to help us generate an open, tolerant, and respectful learning environment that we can all flourish in.

COVID-19: Important information for students: All students are expected and required to report any COVID-19 symptoms to the university immediately, as well as exposures or positive tests from a non-CSU testing location. If you suspect you have symptoms, or if you know you have been exposed to a positive person or have tested positive for COVID, you are required to fill out the COVID Reporter (<https://covid.colostate.edu/reporter/>). If you know or believe you have been exposed, including living with someone known to be COVID positive, or are symptomatic, it is important for the health of yourself and others that you complete the online COVID Reporter. Do not ask your instructor to report for you. If you do not have internet access to fill out the online COVID-19 Reporter, please call (970) 491-4600. You may also report concerns in your academic or living spaces regarding COVID exposures through the COVID Reporter. You will not be penalized in any way for reporting. When you complete the COVID Reporter for any reason, the CSU Public Health office is notified. Once notified, that office will contact you and, depending upon each situation, will conduct contact tracing, initiate any necessary public health requirements and notify you if you need to take any steps. For the latest information about the University's COVID resources and information, please visit the CSU COVID-19 site: <https://covid.colostate.edu/>.

Academic Integrity: This course will adhere to the CSU Academic Integrity Policy as found on the Student' Responsibilities page of the CSU General Catalog (<https://catalog.colostate.edu/general-catalog/policies/students-responsibilities/#academic-integrity>) and in the Student Conduct Code (<https://resolutioncenter.colostate.edu/wp-content/uploads/sites/32/2018/08/Student-Conduct-Code-v2018.pdf>) Do your own work. Don't cheat. If you are unsure what is permissible, please speak with the instructor. Violations will result in a grading penalty and be addressed through the appropriate University mechanisms.

Accommodations: If you have a diagnosed learning or physical disability, which may require special accommodations, please talk to me at the beginning of the course. The Student Disability Center (<https://disabilitycenter.colostate.edu>) can also help facilitate your individual needs. I will work with you and to make sure any individual needs are appropriately accommodated.