LPO 8800: Statistical Methods in Education Research

Sean P. Corcoran

Fall 2020

E-mail: sean.corcoran@vanderbilt.edu
Office: Payne Hall 205C
Office Hours: Wednesdays 10am-12:30pm (by appt.)
https://calendly.com/sean-p-corcoran

Web: **seanpcorcoran.org**
Classroom: 6 Magnolia Circle Room 203
Class Hours: Thursdays 9:35am-12:25pm
Phone: (615) 322-8021

Course Description

This course is an introduction to the theory, methods, and practice of statistics. It is intended as a foundational prerequisite for graduate students who intend to complete the core quantitative methods sequence in the Leadership & Policy Studies program (Regression I/II, Causal Inference, and Research Practicum). Topics will include probability theory, descriptive statistics, population distributions, hypothesis testing and confidence interval estimation, correlation, and regression. While concepts will be introduced with some mathematical rigor, the primary emphasis is the practical application and conceptual understanding of statistics. The course will be taught using the Stata statistical software package and large-sample datasets commonly used in education policy and social science research.

Prerequisites

A prior undergraduate or graduate course in statistics is recommended, though not required. Note that the class begins at a very introductory level but advances quickly. If you have concerns about your prior preparation for this class, please see me.

Books

The following textbook is required:

- *Statistical Methods for the Social Sciences, 5th edition* by Alan Agresti, Pearson, 2018.

If you have no experience with Stata, you may benefit from the following simple introductory text (not required). I will provide other resources for learning Stata.

- *A Gentle Introduction to Stata, Revised 6th edition* by Alan C. Acock, Stata Press, 2018.

Other readings listed in the course schedule below will be made available via Brightspace.

Course Structure

The class will meet weekly, in person. Due to the current public health situation, my intent is to limit our in-person interactions by putting as much material as possible online. See the "COVID-19" section below for more information.

This is a graduate course designed for students at the doctoral and master's level. I expect that students enrolled in the course are motivated by a desire to learn the course material. Please come to class having carefully read any relevant textbook chapters and supplemental materials. Attendance in class and participation in in-class lab exercises is extremely important.

Stata

Stata is the statistical software used in this course. I recommend the most recent release (Version 16), but other recent vintages will suffice (e.g., 13-15). I presently use Stata 15. Be aware that small differences exist between versions, and that files created in recent versions of Stata may not open in older versions (e.g., 12 or before). Purchasing options are available via the following website: https://www.stata.com/order/new/edu/gradplans/student-pricing/. A 6-month license for Stata/IC can be purchased for $48, but if you intend to use Stata in your dissertation work I recommend purchasing a perpetual license of Stata/SE or the more powerful multi-processor (MP) version. Under normal conditions, Stata is freely available to you in the Wyatt 132 computer lab and elsewhere on campus. During the pandemic, Vanderbilt is making Stata available to you virtually through VMWare. See Brightspace for details.

There are many great resources for learning Stata, including the Acock text noted above. UCLA has some nice online resources for learning Stata (https://stats.idre.ucla.edu/stata/) and the Stata YouTube site is also quite helpful. There are some handy Stata "cheat sheets" posted here: https://geocenter.github.io/StataTraining/portfolio/01_resource/. I will also upload some Stata references to Brightspace.

Course Requirements

Your grade for the course will be based on **ten** problem sets (40%), a statistical project (20%), a midterm (20%), and final exam (20%). The problem sets will vary in length and points possible, but each will be weighted equally when calculating your final grade, using the percent correct on each. Since 11 problem sets are assigned, you will be permitted to drop your lowest score.

The (tentative) schedule of problem set assignments is shown in the course schedule below. These are subject to change based on the pace of the course. Please submit your problem set solutions to me via email at sean.corcoran@vanderbilt.edu. Include your last name and problem set number in the filename (e.g., *Corcoran_PS1.pdf*). Late assignments will not be accepted, particularly after problem set solutions have been provided or discussed in class.

Unless otherwise indicated, the file you submit to me should be a log of your Stata session, saved as a text file (with the .txt extension) or—better yet—converted to a PDF. Begin by copying the problem set instructions into the Stata do-file editor. Comment out the questions. Insert after

each question the commands you used to respond to that question. The resulting log file will include the instructions (in the form of comments), your commands, and the output. Edit this file as appropriate, for example by adding interpretations of your output and any other commentary that might be asked for. Graphical output can be submitted separately, preferably as a PDF file. You are encouraged to work together on the problem sets, but all work submitted must be that of the individual student. Duplicate assignments will not be accepted.

Other Important Information

1. **Brightspace**: All materials pertaining to this course, including lecture notes, problem sets, and datasets, will be available on Vanderbilt's Brightspace platform (https://brightspace.vanderbilt.edu). Check in frequently for new materials and announcements. I will aim to have all lecture notes and related materials posted in advance of class. Occasional delays are to be expected, however.

2. **Classroom etiquette**: Please bring your laptop to class. To help promote a productive learning environment, please devote your time and attention to the class itself. Please do not use Facebook, instant messaging, email, or other digital distractions while in class. Please silence your cell phone as well.

3. **Academic integrity**: All academic work at Vanderbilt is done under the Honor System. Students are expected to conform to the highest standards of academic integrity in this course. Any attempt to pass off someone else's work as your own is a violation of this standard, and there are many ways this can happen beyond blatant cheating. Full details of the Vanderbilt Honor System may be found here: http://www.vanderbilt.edu/student_handbook/the-honor-system/ If you have any doubts about how the Honor Code applies to your work in this class, please ask me—not another student—for clarification. Uncertainty about application of the Honor Code does not excuse a violation.

4. **Accommodations**: Students who need an academic accommodation based on the impact of a disability must initiate the request with the Vanderbilt Access Center. Professional staff will evaluate the request with required documentation, recommend reasonable accommodations, and prepare an official letter for faculty dated in the current semester in which the request is made. Students should contact the Access Center as soon as possible, as timely notice is needed to coordinate accommodations. They are located in the Baker Building, Suite 108; https://www.vanderbilt.edu/student-access/disability/.

5. **Mandatory reporter obligation**: All university faculty and administrators are mandatory reports. What this means is that all faculty, including me, must report allegations of sexual misconduct and intimate partner violence to the Title IX Coordinator. In addition, all faculty are obligated to report any allegations of discrimination to the Title IX Coordinator (615-343-9004).

   I am willing to discuss such incidents with you, but I can only do so in the context of us both understanding my reporting obligations. If you want to talk with someone in confidence, officials in the Student Health Center, the University Counseling Center, and officials in the Office of the Chaplain and Religious Life (when acting as clergy) can all

maintain confidentiality. In addition, officials in the Project Safe Center (Crisis Hotline: 615-322-7233) have limited confidentiality, in that they have to report the incidents they are told of, but can do so without providing identifying information about the victim(s).

6. **Mental health and wellness**: If you are experiencing undue personal and/or academic stress during the semester that may be interfering with your ability to perform academically, Vanderbilt's Student Care Network offers a range of services to assist and support you. I am available to speak with you about stresses related to your work in this course, and I can assist you in connecting with the Student Care Network. The Office of Student Care Coordination (OSCC) is the central and first point of contact to help students navigate and connect to appropriate resources on and off-campus, develop a plan of action, and provide ongoing support. You can schedule an appointment with the OSCC at https://www.vanderbilt.edu/carecoordination/ or call 615-343-WELL.

   The Student Care Network also offers drop-in services on campus on a regular basis. You can find a calendar of services at https://www.vanderbilt.edu/studentcarenetwork/satellite-services/

   If you or someone you know needs to speak with a professional counselor immediately, the University Counseling Center offers Crisis Care Counseling during the summer and academic year. Students may come directly to the UCC and be seen by the clinician on call, or may call the UCC at (615) 322-2571 to speak with a clinician. You can find additional information at https://www.vanderbilt.edu//ucc/

Learning During a Time of COVID-19

This class is being taught during an unprecedented time for higher education. It is almost certain we will encounter unexpected challenges, schedule changes, technical problems, and other inconveniences. I ask for your patience and optimism as we traverse this semester together.

1. Please be familiar with Vanderbilt University's Return to Campus Plan, and keep current on the latest developments. If you are experiencing symptoms of COVID-19 or have been recently exposed to someone who has tested positive, **do not come to class!** Remain in quarantine for 14 days and continue the course online.
2. We will meet once per week, in person, unless otherwise stated. Please wear your mask and maintain social distancing and I will do the same. The classroom has been set up in a way to ensure everyone can safely remain six feet apart at all times. Please maintain proper distancing when entering and exiting the classroom.
3. Several students have been approved for remote-only learning. I intend to live-stream class meetings and—I hope—make some lecture materials available asynchronously. A recurring Zoom link is available on Brightspace, and all classes will be recorded for later viewing. Some technical issues are to be expected, so please be patient!
4. For those joining the class via Zoom:
    - Please turn your video on. I would like the class to be as interactive and engaged as possible, and believe this will work best if we can see you.
    - Please join on time, and make sure your audio and video are working.
    - Please keep your audio on mute if you are not speaking.
    - Please **do ask questions**! In this small class it is fine to just unmute yourself and start speaking. If I am not at the computer, I may not be able to see chat entries or virtual hand-raising.
5. My office hours will be virtual (on Zoom) this semester. Please visit my **Calendly** site to schedule a 15- or 30-minute appointment. Select "15-min office hours meeting" or "30-min office hours meeting" to find an available time. If you are unavailable during my office hours, there is an option to propose another 15- or 30-minute block in my schedule.
6. Graduate school is already a stressful and anxious time, and the current public health crisis has only served to add to this. Please make your own mental health and well-being a priority this semester. Take advantage of the many resources Vanderbilt has to offer (see the previous page) and do not hesitate to reach out to me anytime you feel things are getting difficult.

Statistical Project Instructions

Overview

The statistical project is an opportunity for you to apply your knowledge of statistics to real-world data, and answer a specific research question. You will analyze your data using Stata, interpret your results, and communicate your findings in writing.

Requirements

For the statistical project, you may choose one of the following options: (1) provide a written report on a topic to be determined by me, or (2) provide a written analysis of a research question of your choice, using appropriate data. I will provide data for the former, and guidance on the latter. PhD students are recommended to choose option 2.

Your written analysis should range between 12—15 pages, including tables and graphs. You may wish to include some graphs and tables in the main body of the paper, and relegate others to an appendix (appendices do not count toward the page limit).

Deadlines

You must decide which option you are pursuing (and if 2, a topic) by **October 15**. The project itself will be due on or before the last day of class, **December 3**. Please submit your completed project as a PDF document via email, and use your last name and the words "Statistical Project" as the filename.

Guidelines

Think of your statistical project as communicating a story through data analysis. Like any good story, it should have a beginning, middle, and end, and a "storyline" that is clear, logical, and compelling. Be selective about tables, graphs, and statistics you choose to include. Overwhelming your reader with output and statistical results is a surefire way to lose and bore them. It also makes you look like an amateur. Be sure to reference all figures and tables in the text itself.

You may use any technique learned in class in your analysis, unless otherwise specified (relevant in option 1). One of the objectives of this project is a demonstration that you know which techniques to apply in specific situations. In many cases there is more than one technique that can be used. Choose the one that most clearly communicates the result you are trying to convey to the reader. Show some range in your use of methods, and avoid overt repetition.

If you choose option 1, you must fully address all of the questions I assign. These will be available later in the semester. If you choose option 2, you may craft your own research question as you see fit—but the quantity and rigor of the analysis should match or exceed that required under option 1.

A sample project and "model answer" will be posted on Brightspace, as an example of past work. Additional details on the project will be provided later in the semester.

Class schedule

The schedule is tentative and subject to change.

Lecture 1 (Aug 27): Introduction to concepts of probability and statistics

★ Agresti chapters 1-2

- *Problem set 1 assigned Aug 27, due Sept 3*

Lecture 2 (Sept 3): Describing univariate distributions (I)

★ Agresti chapter 3

- Loeb, S., Dynarski, S., McFarland, D., Morris, P., Reardon, S., & Reber, S. (2017). *Descriptive analysis in education: A guide for researchers*. (NCEE 2017-4023). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Available at: https://ies.ed.gov/ncee/pubs/20174023/pdf/20174023.pdf
- *Problem set 2 assigned Sept 3, due Sept 10*

Lecture 3 (Sept 10): Describing univariate distributions (II)

★ Agresti chapter 3

- Loeb et al. (2017)—see Lecture 2
- *Problem set 3 assigned Sept 10, due Sept 17*

Lecture 4 (Sept 17): Probability and sampling distributions (I)

★ Agresti chapter 4

- *Problem set 4 assigned Sept 17, due Sept 24.*

Lecture 5 (Sept 24): Probability and sampling distributions (II)

★ Agresti chapter 4

- *Problem set 5 assigned Sept 24, due Oct 1.*

Lecture 6 (Oct 1): Statistical inference—estimation

★ Agresti chapter 5

- Romer, D. (2020). "In Praise of Confidence Intervals." *AEA Papers and Proceedings* 110: 55-60. Available at: https://pubs.aeaweb.org/doi/pdfplus/10.1257/pandp.20201059
- Review for midterm exam

MIDTERM (Oct 8)

Lecture 7 (Oct 15): Statistical inference—significance tests

★ Agresti chapter 6

- Greenland, S., S. J. Senn, K. J. Rothman, J. B. Carlin, C. Poole, S. N. Goodman and D. G. Altman (2016). "Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations." *European Journal of Epidemiology* 31(4): 337—350. Available at: https://link.springer.com/article/10.1007/s10654-016-0149-3
- *Problem set 6 assigned Oct 15, due Oct 22.*
- Please notify Prof Corcoran about your intended statistical project topic.

Lecture 8 (Oct 22): Power and effect size

★ Agresti chapter 6

★ Kraft, M. A. (2020). "Interpreting Effect Sizes of Education Interventions." *Educational Researcher*. Available at: https://journals.sagepub.com/doi/10.3102/0013189X20912798

- Hill, C. J., H. S. Bloom, A. R. Black & M. W. Lipsey (2008). "Empirical Benchmarks for Interpreting Effect Sizes in Research." *Child Development Perspectives* 2(3): 172—177. Available at: https://srcd.onlinelibrary.wiley.com/doi/abs/10.1111/j.1750-8606.2008.00061.x
- *Problem set 7 assigned Oct 22, due Oct 29*

Lecture 9 (Oct 29): Hypothesis testing—two groups

★ Agresti chapter 7

- *Problem set 8 assigned Oct 29, due Nov 5.*

Lecture 10 (Nov 5): Bivariate covariance and correlation

★ Agresti chapters 8-9

- *Problem set 9 assigned Nov 5, due Nov 12.*

Lecture 11 (Nov 12): Simple linear regression (I)

★ Agresti chapter 9

- *Problem set 10 assigned Nov 12, due Nov 19.*

Lecture 12 (Nov 19): Simple linear regression (II)

★ Agresti chapter 9

- *Problem set 11 assigned Nov 19, due Dec 3 (after Thanksgiving).*

THANKSGIVING BREAK (Nov 26)

- *No class—enjoy!*

REVIEW FOR FINAL EXAM (Dec 3)

- Note: this class will meet virtually, via Zoom.

FINAL EXAM (Dec 7-12)

- Final exam will be a take-home exam.

Schedule at a glance

Aug 27	Lecture 1: Introduction to concepts of probability and statistics	PS 1 assigned
Sep 3	Lecture 2: Describing univariate distributions (I)	PS2 assigned
Sep 10	Lecture 3: Describing univariate distributions (I)	PS3 assigned
Sep 17	Lecture 4: Probability and sampling distributions (I)	PS4 assigned
Sep 24	Lecture 5: Probability and sampling distributions (II)	PS5 assigned
Oct 1	Lecture 6: Statistical inference—estimation	
Oct 8	**Midterm**	
Oct 15	Lecture 7: Statistical inference—significance tests	PS6 assigned
Oct 22	Lecture 8: Power and effect size	PS7 assigned
Oct 29	Lecture 9: Hypothesis testing—two groups	PS8 assigned
Nov 5	Lecture 10: Bivariate covariance and correlation	PS9 assigned
Nov 12	Lecture 11: Simple linear regression (I)	PS10 assigned
Nov 19	Lecture 12: Simple linear regression (II)	PS11 assigned
Nov 26	**NO CLASS - Thanksgiving**	
Dec 3	**Review for final exam**	
Dec 7-12	**Final**	