Introductory Statistics for Education Research (EDUC 667-003)
Fall 2022

Instructor: Dr. Wendy Chan
Email: wechan@upenn.edu
Phone: 215-573-2038
Office: GSE Room 342
Office Hours: By appointment (Virtual)

Course Logistics: This class will meet once per week on Thursdays from 10:15 AM - 12:45 PM in 3440 Market St Room 501.

Course Description: This course is designed to provide students with a basic foundation in the use of statistical methods for quantitative research. Throughout this course, you will learn concepts in descriptive and inferential statistics in order to read, comprehend and communicate results from methodological and applied studies in education research. In addition to developing a strong foundation in statistics, you will also learn methods for data analysis using statistical software. To this end, you will have opportunities to organize, analyze, and summarize parts of actual datasets from education. This is the first course in statistics for students preparing to become researchers in the social sciences and education, and will provide preparation for more advanced coursework in statistical methods and quantitative research (should you choose to do so). While there are no formal prerequisites for EDUC 667, students should have a working knowledge of basic algebra.

Course Objectives:

- Understand the role of statistical methods in educational research

- Learn how statistical methods and results are used and communicated in research literature in education

- Understand the use of statistics in testing research hypotheses and in analyzing the results from research studies

- Develop a basic foundation in gathering, coding, preparing and analyzing data using statistical software

- Acquire preliminary technical literacy to comprehend and critique research methods used in journals such as *Educational Evaluation and Policy Analysis* and *Evaluation Review*

Course and Teaching Philosophy: This course is designed to provide you with opportunities to improve your understanding of quantitative methods and your ability to conduct statistical analysis. You should finish this course with an enhanced ability to think critically about how data is generated and how researchers and practitioners apply inferential methods to reach conclusions. In statistics, you will often be asked to explain and communicate concepts and research findings in a way that is easily interpretable by others.

As with any course, *you* are responsible for making sure that you understand the material. If you feel that you need help, please do not hesitate to ask me. Do not let yourself fall behind because concepts tend to build upon one another. This class is an open learning environment so please ask questions if there is any confusion. Your questions may help others who are confused and will also help me gauge student understanding. There will be a few times where I will decide that we need to move on and

ask that you follow up outside of class, but it is never inappropriate to ask a question. If you take responsibility for making sure that you understand the material and seek help when you don't, I will take responsibility for being there to help you.

Textbooks, Software, and Other Course Material:

- Moore, David S., McCabe, George P. & Craig, Bruce. Introduction to the Practice of Statistics; 8th edition (Please note that the 6th edition of the textbook is available as an e-book and a copy of it has been provided on the Canvas website for this course. It is perfectly acceptable to use this earlier version. You do not need the CD-ROM).

- We will use R and R Studio in this class, which are free and available for download at `https://www.r-project.org/` and `https://www.rstudio.com/`, respectively. You are welcome to use another software program, but please be aware that I will not be able to troubleshoot issues related to any other software program other than R. However, many software programs, including R, have useful help forums online. If you run into computing issues, I encourage you to try searching for solutions online first as another person may have already solved it.

- Throughout the course, I will post articles related to the topics we cover in class. You will be asked to submit responses on *Canvas* to these articles as part of your Participation grade. All of the readings must be completed *prior* to the start of each corresponding class meeting.

Course Grading:

- Problem Sets   30%

- Midterm Exam   30%

- Final Exam   30%

- Participation   10%

Problem Sets: There will be **5** problem sets. The problem sets are designed to reinforce the material from lectures and readings as well as provide opportunities to apply the concepts using data. The quality of writing and explanation is a crucial component to the assignments. Please ensure that you allocate enough time for each problem set so that you can try the problems and return to them at a later point, if needed. For problems involving R, be sure to summarize and interpret the findings. You do not need to provide the R code. Please do not upload an R script file as your completed assignment.

Problem sets will be posted, collected, graded, and returned via *Canvas*. Each assignment must be uploaded **before** the start of class on the date it is due.

Collaboration Policy: You are free to work together on the problem sets. In fact, you are *encouraged* to work in groups. Working in groups is an excellent way to learn, motivate each other, and reinforce understanding of key concepts. However, the completed problem sets that are turned in must be individual work and must be written *in your own words*. Frequently, students (and professors!) think that they understand something, but when they sit down to try it themselves, they realize they do not. Writing out your assignment in your own words is how you will know that you really understand.

Late Assignment Policy:

Late assignments will **not** be accepted. If you have any questions, concerns, or comments about a problem set, please inform me *in writing* (by e-mail) and *prior* to any due date.

Midterm/Final Exam: There will be a take home midterm and final exam. For any questions that involve R, you must provide the code in an Appendix with your completed exam. You have a **week** to complete each exam. Unlike the problem sets, each exam is *individual* work so there should not be collaboration of any kind. In the event of truly extenuating circumstances, such as illness, illness in the family, an important and immovable event, you may request an extension. Please contact me as soon as possible if you need an extension.

Participation: Participation is based upon your responses to the articles posted on *Canvas*. Additionally, attendance and participation in class discussions, both through answering questions I pose and through questions you ask, will also count towards your Participation grade. Attendance is *incredibly* important, especially in a statistics class where the concepts build upon one another. Exams often reveal that students have done well on some material, for which they attended the relevant classes and poorly on other material, for which they missed class. Please try to arrive on time for class to begin at our scheduled time. You are responsible for learning the material that you may have missed.

Other Information:

Academic Honesty: Please consult the GSE Student Handbook on the following webpage for details on expected student conduct: `http://www.gse.upenn.edu/policies/academicintegrity`. Please be sure to read the material in this document. Plagiarism or cheating of any kind will be dealt with according to University policy, which can be found at: `http://www.upenn.edu/academicintegrity`.

Use of cell phones and laptops: Please do not use your cell phone in class. Of course, you may set your phone to vibrate and/or your video camera off to answer emergency calls or send emergency texts. Please do not use your laptop for anything other than taking notes in class and using R.

Communicating with me: Email is the best way to reach me. If I do not respond within 48 hours, feel free to send a follow up email.

Topic Schedule (Fall 2022)

Week	Date	Readings and Topics
1	09/01	**Readings** MMC $1.1 - 1.2$ **Topics** • Course introduction • Displaying and describing distributions • Introduction to R
2	09/08	**Readings** MMC $2.1 - 2.2$ **Topics** • Scatterplots and correlation • R tutorial continued
3	09/15	**Readings** MMC $2.3 - 2.6$ **Topics** • Least squares regression • Two-way tables • R tutorial on regression • **Problem Set 1 due**
4	09/29	**Readings** MMC $3.1 - 3.3$ **Topics** • Sampling design • Design of experiments
5	10/06	**Readings** MMC $4.1 - 4.2; 4.5$ **Topics** • Probability models • General probability rules • **Problem Set 2 due**

Week	Date	Readings and Topics
6	10/13	**Readings** None **Topics** • Continuation of probability
7	10/20	**Readings** MMC $4.3 - 4.4$ **Topics** • Continuation of probability • Random variables • **Midterm Exam Due**
8	10/27	**Readings** MMC 1.3 **Topics** • Density curves • Normal distribution
9	11/03	**Readings** MMC $5.2; 6.1$ **Topics** • Sampling distribution of sampling mean • Confidence intervals • **Problem Set 3 due**
10	11/10	**Readings** MMC $6.2 - 6.3$ **Topics** • Significance tests • Use and abuse of statistics and statistical tests
11	11/17	**Readings** MMC $7.1 - 7.2$ **Topics** • The t distribution • Comparing two means • **Problem Set 4 due**

Week	Date	Readings and Topics
12	11/22	**Readings** MMC $8.1 - 8.2$ **Topics** • Inference for a single proportion • Inference for two proportions
13	12/01	**Readings** MMC $10.1; 11.1$ **Topics** • Simple linear regression • Multiple regression • **Problem Set 5 due**
14	12/08	**Readings** MMC 12.1 **Topics** • One-way ANOVA
15	12/15	**Readings** MMC $13.1 - 13.2$ **Topics** • Two-way ANOVA • **Final Exam Due**