

Statistical Analysis in Sociology

Sociology 312

Winter 2018

Instructor: Patrick Greiner

Office: PLC 623

Email: pgreiner@uoregon.edu

Instructor Office Hours: Friday 12pm - 2pm

Lecture Location: Straub 145

GTF: Natasha Erickson

Office: PLC 708

Lab: MCK 442

Email: nericks3@uoregon.edu

Lab Hours: Tuesday 10am – 12pm

Course Description: The aim of this course is to develop your understanding of the fundamental statistical methodologies used in sociological research. Ultimately, the class is organized in order to help you develop a literacy in statistical research that will aid you in understanding and communicating contemporary research findings. With this goal in mind, the course will focus on building understandings of basic statistical concepts, such as the spread and shape of distributions, central aspects of probability, inference from samples to populations, and the fundamentals of model construction and interpretation. The course will also be organized in order to facilitate your development of a working understanding of the statistical software, R, which will provide you with the tools to develop a skill that is increasingly seen as important and practical in many aspects of modern industry and analysis.

Learning Objectives:

- Students will develop an understanding of the distribution of data, and will learn to distinguish between the various forms of categorical and quantitative variables
- Students will develop the skills to describe and understand the relationship between variables using graphics, tables and regression models.
- Students will learn to understand the concepts of statistical inference, and apply such concepts in determining the validity of statistical findings.
- Students will learn the basics of operating the statistical software package, RStudio, and will be proficient in performing basic analyses in it.
- Students will be proficient in reading and interpreting the findings of statistical research

Graded Assignments

- Participation Exercises (10%)- Attendance is required if one hopes to gain a working understanding of the class concepts, as a result, there will be regular in-class participation exercises.
- Weekly homework (50%)- There will be weekly assignments in order to help students gauge their proficiency in the topics and to prevent them from falling behind. Homework will be submitted online via Canvas.
- Final Exam (40%)- A take home final exam will be administered at the end of the term. Students will be evaluated on their ability to read and correctly interpret and explain the findings presented within a recent academic paper.
 - The Final will be made available online the morning of Wednesday, March 21st, and will be due by 11:59pm on March 22nd. I will be available during our scheduled final hours (in Straub 145 at 12:30pm) to answer questions.

Lab Sections:

Lab section attendance is not required, but it is highly recommended. These sections will provide a space to receive guidance as you complete your homework assignments, as well as to get clarification on any concepts that you are finding difficult or elusive.

Course Schedule:

Week 1- Familiarizing yourself with R, and Introductions to Data

Tuesday, 1/9: Introduction to course and book

Homework 0: Download and install RStudio and necessary packages (qss and swirl)

Thursday, 1/11: Using RStudio, and Playing With Data?

Required Reading: Imai, Ch. 1, pp 10-28

Sunday, 1/14: **Homework 1-**

Responses to Imai exercises 1.5.1 due on Canvas at 11:59pm. Also upload R script file of the code you used to get your responses.

Week 2- Understanding, exploring, and describing data

Tuesday, 1/16: Approaches to identifying causality in the social sciences

Required Reading: Imai, Ch. 2, pp 32-63

Thursday, 1/18: Univariate descriptive statistics and an introduction to distributions

Required Reading: Imai, Ch. 2, pp 63-69

Sunday, 1/21: **Homework 2-**

Responses to Imai exercises 2.8.1 due on Canvas at 11:59pm. Also upload R script file of the code you used to get your responses.

Week 3- Distributions, Spread and Measures of Center

Tuesday, 1/23: Understanding and visualizing distributions

Required Reading: Imai, Ch. 3, pp 75-88

Thursday, 1/25: Understanding and visualizing distributions cont.

Required Reading: finish Imai, Ch. 3, pp 75-88

Sunday, 1/28: **Homework 3-**

Responses to Imai exercises 3.9.2 due on Canvas at 11:59pm. For question 2, rather than making a quantile-quantile plot change the histogram bin sizes a few times. Include a discussion of what happens when you change the histogram bin sizes. Also upload R script file of the code you used to get your responses.

Week 4- Measuring Association Between Categorical Variables

Tuesday, 1/30: Two-way Tables

Thursday, 2/1: Mean Differences

Wednesday, 2/7: **Homework 4-**

Turn in responses to questions 1-3 and 5 in Imai 2.8.3 on Canvas by 11:59 pm. In your response to question 5 include a discussion of how you might apply the natural experiment technique to another question that interests you. Also upload R script file of the code you used to get your responses.

Week 5- Measuring Association Between Quantitative Variables

Tuesday, 2/6: Correlation coefficients

Required Reading: Imai, Ch. 4, pp 96-107; 139-143

Thursday, 2/8: Correlations and introductions to regression lines

Required Reading: Imai, Ch. 4, pp 143-155

Wednesday, 2/14: **Homework 5-**

Recreate the analysis of presidential vote share on Imai, pp. 149-155. Turn in your R script and output on Canvas. Include an explanation of the coefficients from your regression analysis.

Week 6- Measuring Association- The Basics of Regression Analysis

Tuesday, 2/13: Understanding the math behind the regression analysis

Required Reading: Imai, Ch 4, pp. 156-161

Thursday, 2/15: Assumptions of regressions, and ways we go wrong

Required Reading: Nau, "Testing assumptions of linear regression",
<http://people.duke.edu/~rnau/testing.htm>

Wednesday, 2/21: **Homework 6-**

Read abstract, & pp 163-180 in Fruedenburg and Fisher's "Post-industrialization and Environmental Quality: An empirical analysis of the environmental state".

Skim Rosa and York's "Societal Processes and Carbon Dioxide (CO₂) Emissions: Comment on "Post Industrialization and Environmental Quality: An Empirical Analysis of the Environmental State"".

Discuss whether or not Freudenberg and Fisher violated any assumptions of regression analyses.

Week 7- Inference, and Understanding Sample Distributions

Tuesday, 2/20: Understanding probability and the Bayesian v. frequentist debate

Required Reading: Imai, Ch 6, pp. 242- 277

Thursday, 2/22: Distributions and laws of large numbers

Required Reading: Imai, Ch 6, pp. 277-278; 286-307 (focusing on continuous variables and normal distributions)

Wednesday, 2/28 (Note- Sunday 2/25 is the final day to drop or change grade structure):

Homework 7-

Responses to Imai questions 1-3 in exercise 6.6.1 (this assignment is extra credit!) due on Canvas at 11:59pm. Also upload R script file of the code you used to get your responses.

Week 8- Inference, and Understanding Sample Distributions cont.

Tuesday, 2/27: Introductions to inference

Required Reading: Imai, Ch 7, pp. 314- 342 (stop at "Hypothesis testing")

Thursday, 3/1: Distributions and laws of large numbers

Required Reading: Imai, Ch 7, 342-363; 378-390

Wednesday, 3/7: ***Homework 8-***

Responses to questions 1 and 2 of Imai exercise 7.5.3 due on Canvas at 11:59pm. Also upload R script file of the code you used to get your responses.

Week 9- Using Control Variables and Categorical Predictors

Tuesday, 3/6: Controls and multivariate regression analysis

Required Reading: Imai, Ch 4, pp. 161- 170

Thursday, 3/8: Categorical variables and interactions

Required Reading: Imai, Ch 4, 170- 176; 181- 182

Wednesday, 3/14: ***Homework 9-***

Responses to questions 1 and 2 of Imai exercises 4.52 due on Canvas at 11:59pm. Also upload R script file of the code you used to get your responses.

Week 10- Interaction Effects (cont.) and Log Transformed Modeling, and review

Tuesday, 3/13: Log transformations

Required Reading: Nau, “The logarithm transformation”:
<https://people.duke.edu/~rnau/411log.htm>

Thursday, 3/15: Review

Wednesday, 3/21: *No Homework- Prep for final*

Final Week-

Wednesday, 3/21– Thursday, 3/22: *The Final will be made available online the morning of Wednesday, March 21st, and will be due by 11:59pm on March 22nd. I will be available during our scheduled final hours (in Straub 145 at 12:30pm) to answer questions.*