

Statistics and Data Analysis II: The Basics of Regression

ICPSR Summer Program, 2018

Instructor

Shane P. Singh (University of Georgia)
Office: 213 Newberry Hall
Office Hours: 5:15-6:45PM
Email: singh@uga.edu

Teaching Assistants

Chase Meyer (University of Florida)
Office: 324 Newberry Hall
Office Hours: 10:00AM-NOON
Email: chasebmeyer05@gmail.com

Bryce Williams-Tuggle (Vanderbilt University)
Office: 323 Newberry Hall
Office Hours: 12:45-2:45PM
Email: bryce.g.williams-tuggle@vanderbilt.edu

Class Meeting Time: 3:00PM-5:00PM

Class Location: G127 Angell Hall

Description and Goal of the Course: This course is an introduction to ordinary least squares (OLS) bivariate and multiple regression, and it assumes familiarity with introductory statistics, probability, and basic algebra. OLS, while useful in and of itself, also serves as the foundation of many more advanced data analytical techniques. The primary goal of the course is to develop a practical, applied, and intuitive (rather than strictly theoretical or mathematical) understanding of OLS regression. Participants should leave with a solid working foundation of the application of OLS, the ability to use it in their own research, and the skills to build on it in their further study of more advanced statistical topics.

Required Readings:

Agresti, Alan. 2018. *Statistical Methods for the Social Sciences*. 5th ed. Boston: Pearson.

Lewis-Beck, Colin, and Michael Lewis-Beck. 2016. *Applied Regression: An Introduction*. 2nd ed. Thousand Oaks, CA: SAGE Publications.

The books are available for purchase at Ulrich's Bookstore (1200 S. University Ave.) and online.

Course Content:

Readings: Be sure to complete the readings prior to each class period. The content of the readings will not always be discussed in class due to the amount of material we will cover in just four weeks. However, it is important that you comprehend the content, as it is fundamental to your ability to meaningfully understand and use regression analysis.

Daily Class Meetings: Lectures will complement the readings. In some cases, material from the readings will not be covered. And, in many cases, we will cover material not in the readings. We will also frequently analyze real data.

Weekly Lab Sessions: The weekly lab sessions will provide you the opportunity to get “hands-on” experience with course content and to receive guidance on the practice assignments and homework assignments. These sessions will also serve as an opportunity to become familiar with statistical software. Information about the time and location of the lab sessions will be provided.

Course Homework Assignments: In addition to some short practice assignments, you will have four more substantial assignments, which will be graded. The assignments will require statistical software, which is freely available for use from the Summer Program. All assignments will be available electronically on CTools.

The following is required for each homework assignment:

- Type and staple your homework (hand calculations can be written out).
- Present exercises in numerical order, and label your answers clearly.
- Embed any figures or tables within the homework; do not attach them at the end.

Grade Breakdown:

- Assignment 1: 20%; **Due on Day 8**
- Assignment 2: 30%; **Due on Day 13**
- Assignment 3: 30%; **Due on Day 16**
- Assignment 4: 20%; **Due on Day 19**

A Note on Software: While the in-class data analysis will be conducted with Stata, students are welcome to complete their assignments using Stata, R, or SPSS. (Please get our permission if you wish to use another software package.) The instructor and/or teaching assistants will be available to assist students with the software during the lab sessions.

Readings and Course Schedule: *The course syllabus is a general plan for the course; deviations announced to the class by the instructor may be necessary.*

WEEK 1

Day 1: Introduction and Review of Statistical Inference

Day 2: Association between Two Variables

- Agresti: Chapter 8

Day 3: Constructing and Interpreting a Bivariate OLS Regression

- Agresti: Chapter 9, Sections 9.1-9.4
- Lewis-Beck and Lewis-Beck: Chapter 1

Day 4: Statistical Inference with Bivariate OLS Regression

- Agresti: Chapter 9, Section 9.5
- Lewis-Beck and Lewis-Beck: Chapter 2, pgs. 29-39

Day 5: The Assumptions of OLS Regression

- Agresti: Chapter 9, pgs. 266-267 and Section 9.6
- Lewis-Beck and Lewis-Beck: Chapter 2, pgs. 23-28

WEEK 2

Day 6: Applying What We've Learned to This Point

Day 7: Confounding Factors, Causality, and Control

- Agresti: Chapter 10

Day 8: Multiple Regression

- Agresti: Chapter 11, pgs. 307-319; Chapter 13, Section 13.3
- Lewis-Beck and Lewis-Beck: Chapter 3, pgs. 55-64 and 72-73

Day 9: Collinearity and Comparing the Effects of Variables in Multiple Regression

- Agresti: Chapter 11, pgs. 319-324 and Section 11.7
- Lewis-Beck and Lewis-Beck: Chapter 4, pgs. 75-78 and 83-86

Day 10: Categorical Independent Variables

- Agresti: Chapter 12, Section 12.1; Chapter 13, pgs. 387-392
- Lewis-Beck and Lewis-Beck: Chapter 3, pgs. 64-69

WEEK 3

Day 11: Interactions between Variables 1

- Lewis-Beck and Lewis-Beck: Chapter 3, pgs. 69-71

Day 12: Interactions between Variables 2

- Agresti: Chapter 11, Section 11.4; Chapter 13, pgs. 392-397

Day 13: Applying What We've Learned to This Point

Day 14: Model Selection

- Agresti: Chapter 11, Section 11.5; Chapter 14, Section 14.1

Day 15: Regression Diagnostics 1

- Agresti: Chapter 14, pgs. 426-429, Section 14.3, and pgs. 448-449
- Lewis-Beck and Lewis-Beck: Chapter 2, pgs. 44-49; Chapter 4, pgs. 79-83

WEEK 4

Day 16: Regression Diagnostics 2

- Agresti: Chapter 14, pgs. 430-433, Sections 14.5-14.6, and pgs. 449-450
- Lewis-Beck and Lewis-Beck: Chapter 2, pgs. 39-44 and 49-52; Chapter 4, pgs. 86-88

Day 17: Categorical Dependent Variables

- Agresti: Chapter 5, Section 5.5; Chapter 15, Sections 15.1-15.5

Day 18: Applying What We've Learned to This Point

Day 19: Catch-Up (if Necessary), Wrap-Up, and What's Next?

- Agresti: Chapter 13, Section 13.5; Chapter 16, Sections 16.2 and 16.8
- Lewis-Beck and Lewis-Beck: Chapter 4, pgs. 93-95