Introduction to Program Evaluation

PBPL 28633 Spring 2020

Logistics

- Prerequisites: PBPL 26400
- Lecture: TuTh 11am-12:20pm in Keller 0023
- Instructor: Eric Potash (epotash@uchicago.edu)
- Office Hours: TuW 3:30-5:30pm in Keller 3105 or by appointment
- Grader: David McMillon (mcmillon@uchicago.edu)

Summary

This course will teach students how to answer public policy questions using regression analysis. We will discuss applications from the fields of education, health, job training, and others. Students will learn the statistical foundations of regression as well as its practical implementation using the R programming language. They will study the interpretation of regression results including causal inference through experimental as well as quasi-experimental designs. No previous programming experience is assumed. This course satisfies the Public Policy practicum METHODS requirement.

Resources

There is no required text. I will post slides to Canvas before each lecture. You may find these other books useful:

- *Mastering Metrics: The Path from Cause to Effect* by Angrist and Pischke
- *Impact Evaluation in Practice* by Gertler et al.
- *Data Analysis Using Regression and Multilevel/Hierarchical Models* by Gelman and Hill

Grading

Assignments (50%)

Assignments will be shared at the end of class on Thursday covering that week's material and will be due the following Thursday before class (12:30pm) via Canvas. Late assignments will be penalized by 10% per day, with 1 day starting at 12:31pm. Solutions will be posted on Mondays so no credit will be given for assignments submitted Monday. There will be 6 assignments total (none due in weeks 1, 6, or 10). Your lowest assignment score will be dropped.

Exams (50%)

There will be two take-home exams: a midterm (15%) in week 6 and a final (35%) after week 10.

Academic Dishonesty

Writing code is substantially different from writing essays: it is standard practice to google things that don't work, and copy a line or two from a manual or Stack Overflow. I encourage you to discuss general strategies for solving problems with your classmates and friends. Questions and answers on the discussion board will naturally include code. However, you should never ask to see a classmate's solutions nor copy code from your classmates. No one but you should type your code.

All students suspected of academic dishonesty will be reported to the Harris Dean of Students for investigation and adjudication. The disciplinary process can result in sanctions up to and including suspension or expulsion from the University. In addition to disciplinary sanctions, I will impose a grade penalty of 0 for students who have committed academic dishonesty. See the Harris and University of Chicago academic dishonesty policies.

ADA Accommodations

Any student who believes they may need assistance should inform the Harris Dean of Students office by the end of the first week of class. The Dean of Students office will coordinate any student accommodations with Harris instructors.

Software

In this course we will be using a programming language called R analyze and visualize data. You will need to install two pieces of software. First install the R programming language itself. Next install RStudio. This is what's called an *integrated development environment*. It makes using R programming easier and more productive. Assignments should be submitted using RMarkdown. In RStudio create a new RMarkdown PDF document (`File > New File > R Markdown` select output format `PDF`). Save the file and then click the "Knit" button (or `Ctrl+Shift+K`) to render it to PDF, installing the necessary software.

Syllabus

This syllabus is tentative and will be updated throughout the quarter. Lecture slides, readings, and assignments will be posted each week on Canvas.

1. Introduction to regression and causal inference

Why learn statistics? Why learn regression? Examples and challenges

2. Introduction to programming and R

What is programming? Basics including types, operators, functions, plotting

- Assignment 1 posted

3. Background in mathematics and statistics

Weighted averages; vectors, matrices, and linear algebra; lines; probability distributions

4. Inference and simulation

Estimates, standard errors, and confidence intervals; Sampling distributions and generative models

- Assignment 1 due, Assignment 2 posted

5. Introduction to Regression

Comparisons and effects; Historical origins of regression

6. Linear regression with a single predictor

Example: predicting presidential vote share from the economy

- Assignment 2 due, Assignment 3 posted

7. Fitting regression models in R

Checking model fitting using fake data simulation; Comparisons as regressions; Least squares

8. Linear regression with multiple predictors

Interpreting the coefficients, interactions, indicators

- Assignment 3 due, Assignment 4 posted

9. Cancelled

10. Modeling assumptions, diagnostics, and evaluation

Plotting the data and fitted model; Explained variance R^2

11. Transformations

Linear and logarithmic transformations; Example: wages and education

- Assignment 4 due

12. Midterm

13. Binary outcomes with linear and logistic regression

Linear probability model, odds ratios, interpreting logistic regression coefficients

14. Causal inference theory and randomized experiments

Basics; average causal effects; randomized experiments; properties, assumptions, and limitations of randomized experiments

15. Causal inference using regression on the treatment variable

Pre- and post-treatment covariates, treatments, and potential outcomes; Example: the effect of showing children educational television

- Assignment 5 posted

16. Difference-in-differences

- *Workers' Compensation and Injury Duration: Evidence from a Natural Experiment* by Meyer, Viscusi, and Durbin (1995)
- *Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania* by Card and Krueger (1993)

17. Regression discontinuity

- *Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement* by Angrist and Lavy (1999)
- *Does Welfare Inhibit Success? The Long-Term Effects of Removing Low-Income Youth from the Disability Rolls* by Deshpande (2016)
- Assignment 5 due, Assignment 6 posted

18. Experiments with non-compliance

Local average treatment effects, Wald estimator, introduction to instrumental variables.

- *The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment* by Chetty, Hendren, and Katz (2016)
- Assignment 6 due

19. Instrumental Variables

- Assignment 6 due

20. Instrumental Variables Continued