

PoliSci 4782 Political Analysis II

Introduction

Seamus Wagner

The Ohio State University

Welcome to PoliSci4782

A course focusing on **statistical models**:

- How do statistical models work in a general sense?
- What models do we want to use in different empirical settings?
 - *Linear regression is only one option, and sometimes not a good option.*
- How should we test model performance and make comparison across models?
 - *a question that becomes increasingly important in the age of big data and machine learning*
- How do we improve our analysis before and after running models?

Why Focusing on Statistical Models

Statistical models help:

- understand relationships between multiple variables
- empirically evaluate our theory (hypothesis) about causality
 - *how our quantities of interest (outcome variables) are causally affected by other variables*
- make predictions about outcome variables
 - *once we have the statistical insights about how different outcomes are generated (data generating process)*

Statistical Models

What does a “statistical model” mean in this course?

- It is an equation that shows how one variable (outcome) is *probabilistically* determined by a bunch of other variables (explanatory variables/predictors) in a linear or nonlinear way.
- It represents our theory (hypothesis) about the *data generating process* of the outcome.

How Do We Do with Models?

We do the following steps: (1) specify a model (with the help of our theory), (2) estimate the model (with a method, e.g. OLS), and (3) interpret results, in particular two things:

- estimates of effect parameters (i.e. “coefficients”)
 - they tell us how things are related
- uncertainty of those estimates (i.e. “standard errors”)
 - they tell us whether the corresponding estimates are statistically reliable

We often need to try different model specifications, evaluate their performance, and tweak both data and our models during analysis (throw away some observations/impute missing values/adjust standard errors ...)

Learning Objectives

- understand the basics of most statistical models commonly used in social scientific research and industries
- evaluate the quality of quantitative analyses in media and academic journals critically
- use R to process and analyze quantitative data comfortably and proficiently
- employ statistical models to address a variety of questions in their professional and personal life

How Will We Achieve the Objectives?

- Learn conceptual and theoretical materials in lectures and complete comprehensive checks afterwards
- Learn practical skills in lab tutorials and complete problem sets afterwards
- Both lectures and tutorials are pre-recorded videos available on Carmen (no later than each Monday)
- Both comprehensive checks and problem sets are posted in the same time with videos and due at the end of each Sunday (Carmen submission)
- Contact me via **buckeye email** anytime (respond within 24 hours each weekday)
- Chat with me via Zoom meetings (in the office hour mode) on Friday, 1:00-3:00 PM

Keep up with the lectures and lab tutorials on a weekly basis and submit weekly assignments on time!

How Your Score Is Calculated?

- 15 comprehensive checks (one for each lecture, starting from Week 2): 30%
- 8 problem sets (one for each lab session, starting from Week 2): 30%
- final exam (online, open-book, open-note): 20%
- participation (combination of participation, questions, and office hours): 20%

If you foresee a delay of assignment submission, let me know as early as possible (5 days ahead of the deadlines)

What Comes Next?

- Strongly recommend you to read through our syllabus by yourself (available on the syllabus page or in course files)
- First lab session: introduction to R
 - interface of RStudio
 - basic syntax of R language
 - loading and manipulating datasets
- Assignments (comprehensive checks and problem sets) will start from Week 2

If you have any questions regarding class policies, logistics, or schedule, sent them to wagner.1470@osu.edu.

Thank you!