Advanced Regression Models and Analysis

06-08 July (12 hours)

Robert C. Luskin

University of Texas at Austin

Course description

This course is about "regression models," roughly and broadly defined as statistical models to explain some single dependent variable. This definition includes such nonlinear models as logit and probit, as well as the more familiar linear ones, although we shall focus mainly on the latter. Despite skipping proofs and derivations, I hope to convey something more than a cookbook sense of these models, their accompanying assumptions, and the analyses they sustain. The general questions we shall address include: What sorts of models imply and reflect what sorts of relationships between independent and dependent variables? What assumptions must we make, and what do they mean? How likely are the assumptions to be violated, and with what consequences? How can we tell when violations occur? What alternative assumptions might we make? What estimators provide statistically desirable estimates? Where several different estimators might serve, what are their advantages and disadvantages? What do the estimates tell us, and how certainly?

Software

The lab sessions will use R, although students more familiar with another appropriate software . package are welcome to use it instead. Easy-to-follow R do-files will be provided, as will a practice dataset, although students are also welcome to use their own data.

Prerequisites

Students should be comfortable with basic mathematical and statistical concepts and notation. We shall devote some time at the beginning to a bird's-eye view of the statistical background.

Schedule

Monday, 06 July

- Statistical concepts: Random variables, expectations, properties of estimators
- Why regression models? What do they say?
- The classical assumptions. What do they mean?
- Estimation by OLS: Rationale and properties
- Numerical examples and interpretation

Tuesday, 07 July

- Review of 06 July

- Lab
- Fit (including R^2) and inference (hypothesis testing)
- Violations of assumptions: how they may arise, how likely they are, and what to do about them
- Alternative estimators, including GLS, MLE, and 2SLS

Wednesday, 08 July

- Review of 07 July
- Lab
- Nonlinear models
- Models with discrete explanatory variables
- Models for discrete dependent Variables
- Recap

Background Reading

Kmenta, Jan. 1997. Elements of Econometrics (2nd ed.). Ann Arbor, MI: University of Michigan Press. Chapter 1.

and

Gujarati, Damodar N. and Dawn C. Porter. 2019. *Basic Econometrics* (5th ed.). McGraw-Hill Eduication: New York, NY.

or

Wooldridge, Jeffrey M. 2019. Introductory Econometrics: A Modern Approach (7th ed.). Cengage Learning: Boston, MA.

Short biography of the instructor

Robert C. Luskin is Associate Professor of Government at the University of Texas at Austin, Visiting Affiliated Professor at CEVIPOF, *Sciences-Po*, Paris, and Research Advisor at the Center for Deliberative Democracy at Stanford University. He has been a Fellow at the Center for Advanced Study in the Behavioral Sciences at Stanford and has served on Editorial Boards of *Political Analysis* and the *American Political Science Review*. His general interests include public opinion, voting behavior, political psychology, and statistical methods, and he has long been particularly interested in the effects of political information on the texture and outcomes of representative democracy. He has published papers on these and other topics in the *American Political Science Review*, the *American Journal of Political Science*, the *Journal of Politics*, the *British Journal of Political Science*, *Political Analysis*, and other scholarly journals.