

UNIVERSITY OF MICHIGAN

Sociology 610
Statistical Methods II
Mark Mizruchi
Winter 2014

Office: 4234 LSA, 764-7444
Office Hours: by appointment
Email address: mizruchi@umich.edu
Homepage: <http://www-personal.umich.edu/~mizruchi/>

GSI: Kelly Russell (klruss@umich.edu), office and hours TBA

This course is the second portion of the two-semester sequence required of all sociology department graduate students. It consists of two weekly class sessions plus a lab-discussion. In the first semester we covered basic concepts of probability, sampling distributions, confidence intervals, and statistical inference. This term our focus is on multivariate techniques, especially multiple linear regression. The lab sessions will be used to discuss problems encountered in the lectures and written assignments and to refine students' statistical computing skills. The course assumes knowledge of the material covered in Sociology 510. It assumes no mathematical knowledge beyond high school algebra, but students will have an opportunity to develop elementary skills in more advanced mathematical techniques.

Requirements

Requirements for this course include homework (problem-solving) assignments, a term project (to be described in class), and midterm and final examinations. The primary text is *Basic Econometrics* (fifth edition), by Gujarati and Porter. Although this book is designed for students in economics, it is actually less mathematical than the sociologically-oriented text that I had used prior to adopting this book, and I hope that you will find it useful. As with all textbooks in this field, this book is extremely, even outrageously, expensive and you are not required to purchase it. I recommend that you find a way to secure a copy, however. It is best viewed as an investment, one that you may find yourself using as a reference long after you have completed this course. I have also included some assigned readings from another text, *Applied Regression Analysis, Linear Models, and Related Methods* by John Fox. Fox's book provides more detailed treatment of several important topics than that provided by Gujarati and Porter. For those of you who would like a simple, more narrative, introduction to multiple regression, I have ordered copies of *Multiple Regression: A Primer*, by Paul D. Allison. This book is not a substitute for Gujarati and Foster, but it may serve as a useful supplement. All

readings, including those in the textbooks and those listed as recommended, are optional. If you can absorb the material without doing the readings, that is fine. But you are responsible for the material in the readings. I will also post, on the class Ctools site shortly after the end of each class, the notes from that class. This will hopefully provide a useful review. I would not use this as a substitute for your own note-taking, however, since it will be in my words, not yours.

COURSE OUTLINE¹

January 9: Introduction to the course

January 14: Review- levels of measurement, sampling theory, hypothesis testing, analysis of variance

Reading: reread your 510 materials

January 16: Simple regression

Reading: Gujarti and Porter, pp. 15-21, 27-28, 34-48 (skim), 55-61, 61-69 (skim; we will return to these issues in detail later in the term); Recommended: Allison, pp. 97-115

January 21: Statistical inference in regression

Reading: Gujarti and Porter, pp. 69-83, 107-126

January 23: An introduction to multiple regression: causality and partial correlation

Reading: Gujarti and Porter, pp. 213-215, 188-192; Recommended: Allison, pp. 1-46

January 28: The mathematics of multiple regression I: the normal equations

Reading: Gujarti and Porter, pp. 192-207, 234-246, 249

¹ This outline provides an approximate description of topics and dates. Historically I have tended to deviate from it, in ways that cannot be predicted at the beginning of the term. I can therefore not promise that the described topics and dates will match.

January 30: An introduction to matrix algebra

Reading: Gujarti and Porter, pp. 838-848

February 4: Multiple regression in matrix form

Reading: Gujarti and Porter, pp. 849-861

February 6: Analysis of variance as regression: dummy variables

Reading: Gujarti and Porter, pp. 277-282; Fox, pp. 155-161, 135-144

February 11: Two-way analysis of variance

Reading: Gujarti and Porter, p. 283, 288-290; Fox, pp. 328-335

February 13: Analysis of covariance

Reading: Gujarti and Porter, pp. 283-284; Fox, pp. 192-195; Recommended: Allison, pp. 166-170

February 18: Regression diagnostics I: Outliers and influential observations

Reading: Gujarti and Porter, pp. 496-498

February 20: Regression diagnostics II: Multicollinearity

Reading: Gujarti and Porter, pp. 320-347; Recommended: Allison, pp. 137-150

February 25: Review; MIDTERM EXAMINATION DISTRIBUTED

February 27: Exam in progress; no class

March 4, 6: Vacation

March 11: Multiple regression assumptions

Reading: Gujarati and Porter, pp. 61-69; Fox, pp. 295-307; Recommended: Allison, pp. 119-134

March 13: Heteroskedasticity, weighted least squares, and robust standard errors

Reading: Gujarati and Porter, pp. 365-378, 385-387, 389-395

March 18, 20: Generalized least squares

Reading: Gujarati and Porter, pp. 412-429, 434-437, 440-447

March 25: Polynomial regression

Reading: Gujarati and Porter, pp. 210-213; Recommended: Allison, pp. 153-166

March 27, April 1, 3: Maximum likelihood estimation and logistic regression

Reading: Gujarati and Porter, pp. 541-547, 552-567, 580; Recommended: Fox, pp. 438-472

April 8: Poisson and negative binomial regression

Reading: Gujarati and Porter, pp. 574-579

April 10: An introduction to event-history analysis

Reading: Paul D. Allison, *Event History Analysis*, pp. 9-42

April 15: Panel regression: fixed and random effects models

Reading: Gujarati and Porter, pp. 591-608

April 17: Review; FINAL EXAMINATION DISTRIBUTED

April 22: Final examination due