Regression Analysis

Course Objectives:

The main body of this course focuses on "intermediate" level single-equation regression techniques. Though we begin with a review of bivariate regression, this course presumes that students have mastered basic statistics and regression.

Student Responsibilities:

Four major problem sets will be distributed during the course, and these constitute the bulk of student responsibilities. Occasional, less challenging work sheets may also be handed out. These will help illustrate particular estimation procedures, and should help students gain hands-on experience with key ideas as well as with SPSS for Windows, our class estimation package. Each assignment may involve computation by hand, or use of computer packages, or both. Information on available data sets will be provided.

Course Sequence:

I. Review of Bivariate Regression and Correlation.
II. Scalar Introduction to Multiple Regression.
III. A Matrix Introduction to Multiple Regression. The General Linear Model.
V. Topics in Regression (As time and class interest allow): Introduction to Maximum Likelihood Estimation, Discrete Dependent Variables.

Readings:

Main Course Texts:

Gujarati’s NOTE: 3rd and 4th editions are perfectly acceptable, and NO data disk is needed.
John Fox (1992) *Regression Diagnostics*. A good Sage Series monograph that treats many key estimation problems. Good diagnosis is essential for proper treatment, and Fox describes several diagnostic techniques with admirable clarity.


*Regression Analysis ‘Blue Bible’*. This copier package, available through ICPSR, will contain all course assignments and many helpful hand-outs. A bargain at twice the price.

**At Your Discretion:**

To keep course material within reach of all students, several Sage monographs have been ordered, and these will serve as more intuitive introductions to all major topics treated in this course. These monographs are listed as recommended for purchase, and students are free to select those which cover topics of greatest interest, or in which they feel they need the most help (see descriptions below). Supplemental monographs to consider for purchase (listed in order of their use in this course):

1) Achen, Christopher H. (1982) *Interpreting and Using Regression* Sage Series no. 29. This monograph considers "how and under what circumstances regression analysis is actually put to good use in the social sciences." It is a very good question to ask, and Achen provides nice answers while considering important issues of inference and interpretation. This is sort of a "thinking human's" introduction to regression.


3) King, Gary (1989) *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. This serves well as an introduction to more advanced estimation problems where linear techniques perform poorly, if at all (event count models, discrete, censored, or truncated dependent variables, etc.)
SYLLABUS

I. Review of Bivariate Regression and Correlation

Required Reading:

Gujarati, Chapters 1-6.


Suggested Reading:


Applications:


II. Scalar Introduction to Multiple Regression and Analysis of Variance

Required Reading:

Gujarati, chapters 7 and 8.

Applications:


### III. A Matrix Introduction to Multiple Regression Analysis

**Required** Reading:

Gujarati, Appendices B and C.

### IV. Regression Diagnostics: Finding and Treating Estimation Problems in Multiple Regression Analysis

**A. Heteroscedasticity, Autocorrelation**

**Required** Reading:

Gujarati, Chapters 11 and 12.


**Suggested Reading:**


Applications:


**B. Collinearity, Specification Error, Measurement Error**

**Required** Reading:

Gujarati, Chapters 10 and 13.

Fox, John (1992) *Regression Diagnostics* Section 3, Appx. 3.1, 3.2.


**Required** Reading:

Fox, John (1992) *Regression Diagnostics* Sections 4, 5 and 7 and Appx.4.1-4.4, 6.1.

**Suggested Reading:**


V. Topics in Regression: Dichotomous Variables, Pooled Time-Series, Introduction to Maximum Likelihood

A. **Dichotomous Independent Variables** (a.k.a. "Dummy" Variables).

Gujarati, Chapter 9.

B. **Pooled Cross-Section Time-Series Analysis: Confronting a Variety of Estimation Problems All at Once**

**Theory:**


**Applications:**


Weede, Erich (1986) "Sectoral Reallocation, Distributional Coalitions, and the Welfare


C. A Basic Intro to Maximum Likelihood Estimation and Some of its Uses


Greene, William H. Econometric Analysis (2nd ed.) Chapter 13, sections 1-4.

i) Discrete Dependent Variables: Logit and Probit Techniques

Gujarati, Chapter 16.


Greene, William H. Econometric Analysis (2nd ed.) Chapter 21, sections 1-4.

Applications:


**ii) Event Count Models and Poisson Regression**


Application:


**Important Considerations (applies in particular to students taking the course for credit):**

*Academic Honesty.* All of the work you do in this course is expected to be your own. Absolutely no cheating or plagiarism (using someone else's words or ideas without proper citation) will be tolerated. Any cases of cheating or plagiarism will be reported to the university committee on academic misconduct and handled according to university policy.

*And a Special Note Regarding Assignments.* The above wording is a statement of official University policy. You should be aware that I do allow students to work together while they progress through each assignments. But each student must work alone as they write up their own final copy for submission. This means that all work you submit must be yours alone. Each student should show all work progressing toward the solution, and I ask you to highlight your final answer to each problem. In other words, it is fine with me if you are stuck on a problem, and ask a fellow student to explain a point that will help you get over that obstacle. However, I do not allow one student to take another by the hand through many steps of a problem, effectively solving the problem for them. It is fine with me if two students compare answers to a problem in order to check their work. When their answers do not match, it is fine with me if they discuss the process by which they arrived at their answer, in order to discover who is wrong. I do not allow one student to copy the solution to a problem from another student, whatever the reason. Nor do I allow two or more students to divide up the work on any problem or problem set and share their results with each
other. Any and all work you submit with your name on it must be entirely your own work. If any part of this policy is not clear to you, consult with me or the teaching assistants.