ST505 and ST697R: Applied Regression Analysis
Fall 2012

Meets: MWF 12:20 - 1:10.
Instructor: John Buonaccorsi **Office:** LGRT 1435K
Phone: 545-2809, **email**:johnpb@math.umass.edu

TEXT (Required): *Applied Linear Regression Models* by Kutner, Nachsteim and Neter (4th edition) or
Applied Linear Statistical Models by Kutner, Nachtsteim, Neter and Li (5th edition)

NOTE The first 14 chapters of Applied Linear Statistical Models (ALSM) is what makes up Applied Linear Regression Models, 4th ed. The second half of ALSM covers experimental design and the analysis of variance and is used in our ST506. If going to take ST506 or you want to have the added design material for a modest additional course you should buy the Applied Linear Statistical Models.

PREREQUISITES: Previous coursework in Probability and Statistics, including knowledge of estimation, confidence intervals, and hypothesis testing and its use in at least one and two sample problems. ST516, or equivalent, is optimal; ST501 is sufficient. **ST515 by itself is NOT a sufficient background for this course.** Some familiarity with basic matrix notation and operations is helpful.

Description. This is an primarily an applied statistics course. While models and methods are written out carefully, the focus of the course is on the understanding and presentation of regression models and associated methods, data analysis, interpretation of results, statistical computation and model building. Topics covered include simple and multiple linear regression; correlation; the use of dummy variables; residuals and diagnostics; model building/variable selection, regression models and methods in matrix form; an introduction to weighted least squares, regression with correlated errors and nonlinear regression. There will be abundant data analysis with examples from many different disciplines, using SAS or R (prior experience not assumed). A matrix formulation of the linear regression model is given partway through the course. This is for ease in presenting models and results and understanding some of the computational documentation, not for proving regression results using matrix theory; something which is done in ST705.

Note: The course is co-listed. Graduate students need to register for ST697R, undergraduates for ST505. This will accommodate differences in the two groups and the fact that ST505 serves as an IE (Integrative Experience) course for the undergraduate math majors. The two classes will have common lectures, but there will be some differences in outside work.

Tentative Syllabus

1. An introduction to regression. Motivating examples, an overview of the objectives of regression analysis.

2. Simple Linear Regression. (Most of Chapters 1 - 4: Explicit readings assigned as we move through the material)
   - The regression model (1.1-1.5)
   - Estimation of the regression coefficients and error variance (1.6-1.8).
   - Inferences for the regression coefficients. (2.1-2.3, 4.1)
   - Estimating the expected response at a particular x; one-at-a-time and simultaneous confidence intervals. (2.4, 2.6, 4.2)
   - Predicting future observations. (2.5,4.3)
   - Inverse prediction and regulation (regulation in notes only)(4.6)
   - The Analysis of Variance approach to regression and general linear tests (2.7, 2.8))
   - Measures of association, correlation and random predictors. (2.9 - 2.11)
   - Assessing model assumptions and a first look at remedial measures. (Parts of Chapter 3)

3. An introduction to multiple linear regression models (section 6.1, description of models in 8.1 and 8.2)

4. Regression models in matrix form. Section 5.1-5.4, the definition of an inverse in 5.6, 5.8-5.9, 6.2)

5. Multiple Linear Regression (5.10-5.13, much of Chapters 6, 7 and 8)
   - Estimation of the regression coefficients and error variance.
   - Inferences for the regression coefficients.
   - Estimating the expected response at a particular x; one-at-a-time and simultaneous confidence intervals.
   - Predicting future observations.
   - The Analysis of Variance approach to regression and general linear F-tests

6. More on diagnostics and tests for assessing model assumptions with some on measures to accommodate violations of usual assumptions. (Parts of Chapters 10 and 11)

7. Model building/variable selection (Ch. 8)

8. Autocorrelation in time series data. (Ch. 12)

9. An introduction to nonlinear regression models. (Some of 13 and 14)