Syllabus for BST 623 **General Linear Models**
Fall 2004 **T. Mark Beasley, Ph.D.**

Purpose: BST 623 is an intermediate-level course in regression analysis methods, intended to familiarize students with simple and multiple linear regression estimated by least squares. These introductory topics are covered in more detail in advanced courses. Students are taught to conduct the relevant analyses using current computer software such as the Statistical Analysis System SAS).

Prerequisites: BST 621, BST 622, BST 626 or equivalent. Other courses recommended, but not required, include the following: BST 631 and 632 (Introductory Probability and Inference). It may be taken by other graduate students with appropriate backgrounds in calculus and matrix algebra.

Objectives: This course is designed to provide students with a fundamental understanding of the theory and application of the general linear model (GLM). At the end of the course, students are expected to be able to derive basic regression equations, understand and verify the assumptions of all models, perform appropriate transformations and diagnostics, and interpret and explain the application of the GLM using real data.

Required Text Book: Muller, K.E. and Fetterman, B.A. (2002). Regression and ANOVA: An Integrated Approach Using SAS Software. SAS Publishing.
Supplement: Draper, N. R., and Smith, H. (1998). Applied Regression Analysis. 3rd Edition. Wiley

Grading Policy:
During the course there will be homework problems assigned to be handed in for review. Most problems can be done by computer; however, some may require hand calculations for help in identification and understanding of computer results. Some problems which can be done by computer may require coding in the SAS matrix manipulation procedure IML, again so that the output of standard SAS procedures can be more readily understood. This course will have an in-class midterm with a take home component as well as an in-class final exam with a take home component. Homework will account for 20% of the final grade, midterm will account for 40% of the final grade, and the final will account for approximately the last 40% of the grade.

Final Grade Assignments:	A	90% - 100%
	B	80% - 89%
	C	70% - 79%

Note: All assignments must be individually performed by the students submitting them. Collaborative work on graded assignments is not allowed unless specifically stated by the instructor for a particular assignment. Work must be turned in on time unless special permission is granted by the instructor due to truly extenuating circumstances. Requests for time extensions will be considered on a case-by-case basis.

Disability Student Services

Any student with a disability that may need accommodations in order to successfully complete all requirements for this course should visit the Office of Disability Support Services, located in Room 516 of the Hill University Center (205-934-4205). This office is responsible for registering students and ensuring the University's compliance with Section 504 of the Rehabilitation Act. Once registered, this office will then inform faculty members of all courses in which the student is enrolled, of the student's status, and the specific nature of any accommodations required. Any student requiring such accommodation should discuss this with the course master and assure that the appropriate correspondence is sent from the Office of Disability Support Services.

Lecture 1 - 2: Chapter 2: Statement of the Model, Estimation, and Testing

- Motivation
- Statement of the Model
- Least Squares Assumptions
- Discussion of Homogeneity
- Gaussian Errors Assumption
- Estimation of the GLM
- Hypothesis Testing for the GLM

Lecture 3-4: Chapter 3: Some Distributions for the GLM

- Motivation
- A Full-Rank Basis for Less-than-Full-Rank Models
- Definitions and Theorems
- GLM Distributions
- Definitions and Properties of Residuals

Lectures 5-7: Chapter 4: Multiple Regression: General Considerations>

- Motivation
- Definitions of Basic Sums of Squares
- The Nature of the Intercept
- Models that Span buy May Not Include an Intercept
- Corrected SS
- Intercept-Only Model
- Null Model
- Overall ANOVA Table for Multiple Regression
- Usual ("Corrected") Overall Test for Regression
- "Uncorrected" Overall Test for Regression
- Strength of Association

Lectures 8-10: Chapter 5: Testing Hypotheses in Multiple Regression

- Motivation
- Choosing an Error Term
- Review of GLH Concepts
- Model Pools
- Test Class 1: Overall
- Test Class 2: Addition of One Variable
- Test Class 3: Tests of the Intercept
- Test Class 4: Addition of a Group of Variables
- Test Class 5: GLH Tests
- The Multiple Testing Issue
- Interaction

Lecture 11-12: Chapter 6: Correlations

- Motivation
- Interpreting p2
- Correlation Formulas
- Partial Correlation
- Semipartial Correlation
- Relating Semipartial Correlations to Standardized Regression Coefficients
- Multiple Partial Correlation
- Multiple Semipartial Correlation
- Hypothesis Tests for Correlations
- Relating Multiple Partials and Semipartials to Regression Coefficient Tests
- Using Correlations to Interpret Added-in-Order Tests
- Computing Partial Correlations
- Some Useful Properties of Correlations
- The Importance and Utility of Correlations
- Homework Exercises

Lecture 12-13: Chapter 7: GLM Assumption Diagnostics

- Motivation
- The First Step: Get to Know Your Data
- Residual Analysis
- Outliers

Lecture 14-15: Chapter 8: GLM Computation Diagnostics

- Motivation
- Single Variable Problems and Solutions
- Collinearity Definitions and Concepts
- Matrices of Interest
- Models Corresonding to Cross-Products Matrices
- Eigenanalysis
- R^2_j, Tolerance and VIF
- Detecting Numerical Inacuracy
- Treating Regression Problems

Lectures 16: Midterm

Lecture 17: Chapter 10: Transformations

- Motivation
- General Principles
- Power Transformation of the Response
- Other Comments
- Transformations of Predictors
- Pitfalls
- Weighted Least Squares

Lecture 18: Chapter 12: Coding Schemes for Regression

- Motivation
- Reference Cell Coding
- Cell Mean Coding
- Classical ANOVA Coding
- Effect Coding
- Polynomial Coding
- The Essence Matrix
- Comments on Coding Schemes
- Relationships among Coding Schemes

Lecture 19-20: Chapter 13: One-Way ANOVA

- Motivation
- Specification of the Model
- (Usual) Overall Test
- Defining and Estimating Cell Means
- Which Means Are Different
- Contrasts
- Conducting Multiple Comparisons
- Welch ANOVA

Lectures 22-24: Chapter 14: Complete, Two-Way Factorial ANOVA

- Motivation
- Model Concepts
- Coding Schemes
- Geneating Cell Means
- Computing Estimates and Tests
- Contrast Matrices for Marginal Means
- Choosing and Interpreting Tests
- Step-Down Tests
- Missing Data

Lectures 25-26: Chapter 15: Special Cases of Two-Way ANOVA and Random Effects Basics

- Motivation
- Blocking Variables and Block Designs
- Fixed Block Design
- Introduction to Random Effects Models
- The Classical Approach to a Random Block Design
- Role of Nonindependence of Observations
- Computations for Mixed Models
- Generalized Least Squares
- Review Comments

Lectures 27-28: Chapter 16: The Full Model in Every Cell (ANCOVA as a Special Case)

- Motivation
- Cell Mean Style Coding of Full Model
- Properties of the Model
- Testing Strategies
- Implementing Strategy 1, Adjusted ANOVA
- Implementing Strategy 2, GLM Testing
- Models of Interest
- Implementing Strategy 3: Backwards Groupwise
- Difference Scores: A Special Case of ANCOVA
- Other Contrasts of Interest in ANCOVA
- Other Tests of Interest in the Full Model
- Regression (Reference Cell) Style Coding
- Effect Style Coding
- Modeling a Baseline Covariate
- Comparing Coding Schemes
- Generalizations

Lectures 29-30: Other Topics

- Polynomial Regression
- Continuous Interactions
- Response Surface Analysis and Polynomial Interactions
- Model Selection (Chapter 11)
- Generalized Least Squares
- Errors-in-Variables Regression
- Seemingly Unrelated Regression
- Path Analysis
- Repeated Measures
- Longitudinal Analysis
- Missing Data
- Mixed Models – Random Coefficient Model
- Generalized Linear Models
- Power (Chapter 17)