MATHEMATICS 373
LINEAR STATISTICAL MODELING

I. Introduction

A. Catalog Description

Using multiple regression as a unifying theme, the student will learn the theoretical foundations of regression, many real-world applications of regression, the underlying algorithms and their limitations. The student will learn when regression is and is not appropriate, and what alternatives are available in the latter case. Prerequisites: MATH 271 and MATH 232 or permission of the instructor. Offered every three years; offered Spring 2004.

B. Objectives

The main objective of this course is to provide mathematics majors (and other qualified students) with a course in applied statistics. In creating the course the guidelines given by the American Statistical Association in distinguishing between mathematics and statistics as disciplines (or sub-disciplines within mathematical sciences) played a major role. (See appendix).

This main objective suggests four specific requirements.

- (Mathematical content) The student will be introduced to a brief history of the subject and the major theoretical results.

- (Computational content) The student will use either a programming language with statistical capabilities (for example S-Plus) or the more advanced programming features of a statistical software (for example Minitab).

The student will learn the underlying algorithms and potential numerical problems associated with statistical procedures.

- (Communication skills) The student will work with real-world data, and will learn to analyze data properly, and accurately communicate the findings in non-technical language both in writing and orally. The data should be drawn from a wide variety of disciplines, emphasizing the interdisciplinary nature of applied statistics.

- (Statistical content) The student will see the standard topics considered under the heading “regression and time series” and learn all the issues required to do a complete analysis of a data set.

II. Topics

A. Simple linear regression

1. Gauss, Galton and the least squares idea
2. Fitting the model
3. Estimation of parameters
4. Inference
5. Prediction
6. Correlation

* * 7. Randomization and non-parametric tests

B. History and theoretical foundations

1. The central limit theorem for regression parameters
2. The Gauss – Markov theorem
3. Likelihood functions and maximum likelihood
4. Least squares and maximum likelihood

* * 5. Nonlinear least squares

C. Basic multiple regression

1. Fitting the model
2. Interaction, indicators, curvature
3. Statistical inference
4. Model building
5. Model validation

D. Advanced multiple regression

1. Residuals Analysis
2. Diagnostics and model checking
3. Common pitfalls

* * 4. Weighted least squares

* * 5. Ridge regression

* * 6. Logistic regression

II. Topics (continued)

D. Time series and forecasting

1. What is a time series?
* * 2. Simple smoothing
3. ARIMA models
4. Diagnostics
5. Model selection
6. Trend and seasonality

* * Optional topics

III. Bibliography

Box, G. & G. Jenkins. *The Statistical Analysis of Time Series* (Wiley)
Daniel, *Applied Non-Parametric Statistics* (PWS-KENT)
Dielman, *Applied Regression Analysis for Business and Economics* (PWS-KENT)
Mendenhall, Wm. & T. Sincich *A Second Course in Statistics: Regression Analysis* (Prentice-Hall)
Neter, Wasserman, & Kutner, *Applied Linear Regression Models* (Irwin)
Weisberg, , S. *Applied Linear Regression* (Wiley)

IV. Assessment:

In addition to in-class examinations and homework exercises, students will perform data analysis through case studies and report their findings both orally and in writing. The homework assignments may consist of computational problems, programming assignments, conceptual problems, mathematical derivations, and (occasionally) mathematical proofs.

Appendix:

The following information was taken from the address below:

http://www.amstat.org/education/Curriculum_Guidelines.html

The American Statistical Association identifies these as the distinctive qualities of an (applied) statistician. Although the reference is to statistics programs, it applies equally well to an applied statistics course within a mathematics department.

Skills Needed

Effective statisticians at any level display a combination of skills that are not exclusively mathematical. Programs should provide some background in these areas:

Statistical - Graduates should have training and experience in statistical reasoning, in designing studies (including practical aspects), in exploratory analysis of data by graphical and other means, and in a variety of formal inference procedures.

Mathematical - Undergraduate major programs should include study of probability and statistical theory along with the prerequisite mathematics, especially calculus and linear algebra. Programs for non-majors may require less study of mathematics. Programs preparing for graduate work may require additional mathematics.

Computational - Working with data requires more than basic computing skills. Programs should require familiarity with a standard statistical software package and should encourage study of data management and algorithmic problem-solving.

Non-mathematical - Graduates should be expected to write clearly, to speak fluently, and to have developed skills in collaboration and teamwork and in organizing and managing projects. Academic programs often fail to offer adequate preparation in these areas.

Substantive area - Because statistics is a methodological discipline, statistics programs should include some depth in an area of application.