Stat 140A: Linear Models

California State University, Sacramento · Department of Mathematics & Statistics

CATALOG DESCRIPTION

Practical statistical modeling of data using professional-grade statistical software. Simple and multiple linear regression; analysis of variance; logistic regression; variable transformation; variable selection; model selection; residual analysis. **3.0 Units; Letter Graded**

PREREQUISITES

Stat 115A and Stat 128 and (Math 35 or Math 100)

LEARNING OUTCOMES

Students will be able to:

1. Apply regression techniques by fitting linear models, conducting hypothesis tests, constructing confidence intervals, and making predictions.
2. Evaluate the appropriateness of models using diagnostic plots, statistics, and residual analysis.
3. Interpret model parameters by communicating results for a general audience.
4. Select appropriate models by considering the bias variance tradeoff.
5. Use statistical software by fitting models to real world data.

SAMPLE TEXT AND MATERIALS

Linear Models with R by Julian Faraway
Extending the Linear Model with R by Julian Faraway for content on logistic regression

Technology

- Students will need a laptop or desktop computer to access the course materials online.
- The course will make extensive use of statistical software such as R.

METHODS OF EVALUATION (GRADING SYSTEM, EXAMINATIONS, ETC)

The course grade will be determined by assignments, projects, midterms, and a comprehensive final examination.

Specific grading breakdown is instructor dependent. A suggested breakdown is as follows.

- Homework 15%
- Project(s) 40%
- Midterm Exams 25%
- Final Exam 20%

TOPICS LIST

1. Simple linear regression (4 weeks)
    a. Introduction to software
    b. Least squares regression line
    c. Properties of least square estimators
    d. Confidence and prediction intervals
2. Model diagnostics (2 weeks)
    a. Hypothesis tests
    b. Correlation coefficients
    c. Coefficient of determination
    d. Residual analysis
3. Analysis of variance (2 weeks)
    a. Experimental design
    b. One-way and two-way ANOVA
    c. Additive, nonadditive models
4. Multiple linear regression (3 weeks)
    a. Variable transformation
    b. Variable selection
    c. Interaction effects
    d. Model selection (AIC, BIC, etc.)
5. Logistic regression (4 weeks)
    a. Link functions
    b. Prediction and interpretation
    c. ROC curves
    d. Model performance metrics