Math 362: Statistical Methods II **(4-0-4)**
January 28, 2014

Catalog Description: Review of inferential statistics, analysis of variance one factor and two factor, simple and multiple regression, analysis of categorical data using tests and confidence intervals for proportion and chi-square tests, correlation, goodness of fit, non-parametric tests. Data sets used will come from various fields including: business, psychology, biology, environmental science, engineering, manufacturing, and communication.
Prerequisite: Math 361 or instructor consent.

Course Objectives: After completing this course, students will be able to:

1. Calculate and interpret parameter estimates and their errors.
2. Describe data using graphical methods.
3. Use parametric and non-parametric hypothesis tests to address statistical questions and form statistically sound conclusions.
4. Perform simple hypothesis testing on more than two groups using one or two way ANOVA.
5. Perform simple and multiple regressions.

Learning Outcomes and Performance Criteria

1. Calculate and interpret parameter estimates and their errors.

   Core Criteria:

   (a) Generate summary statistics for standard measures of center and spread with confidence intervals.
   (b) Calculate estimates of parameters for simple and multiple linear regression.
   (c) Calculate linear correlation coefficients.
   (d) Interpret the coefficient of determination as the percentage of variation in the response variable explained by a general linear model.

   Additional Criteria:

   (a) Generate confidence intervals using bootstrapping techniques for small samples
   (b) Calculate estimates of parameters for logistic regression.
   (c) Calculate transformations of data and apply regression methods to appropriately transformed data.
   (d) Explain the difference between the results of models built non-transformed and transformed data when interpreting the results of a model (including back transformation from log transformed data).

2. Describe data using graphical methods.

   Core Criteria:

   (a) Create and interpret box plots with multiple factors
   (b) Create and interpret scatter-plots with estimated simple linear regression lines.

(c) Create and evaluate histograms and boxplots for general identification of outliers, skewness, spread, and center of data.

(d) Create and analyze residual plots of a statistical model.

Additional Criteria:

(a) Explore graphical techniques such as three dimensional scatterplots, heat maps, and other graphical techniques.

(b) Create and interpret bar graphs with multiple variables

3. Use parametric and non-parametric hypothesis tests to address statistical questions and form statistically sound conclusions.

   Core Criteria:

   (a) List and verify the assumptions behind the following tests: one and two sample t-test(dependent and independent), binomial, chi-squared test, the sign test, One Sample Wilcoxon Test, Kolmogorov-Smirnov Test, and the Mann-Whitney-U Test and Fisher's exact test.

   (b) Apply the tests listed above.

   (c) Apply graphical or summary techniques to identify what violations of assumptions may be in the data.

   (d) State the null and alternative hypothesis to a given parametric or non-parametric test.

   (e) Utilize software to estimate sample size or power for a planned experiment under a simple parametric test.

   (f) Communicate results in the context of the hypothesis test for a given problem.

   (g) Create efficient multiway tables in order to summarize categorical data.

4. Perform simple hypothesis testing on more than two groups using one or two way ANOVA.

   Core Criteria:

   (a) Apply Exploratory Data Analysis to graphically evaluate the differences in center.

   (b) List and verify the assumptions behind standard ANOVA.

   (c) Create and interpret the components of an ANOVA table.

   (d) With the aid of software, apply parametric ANOVA methods to data and interpret the results.

   (e) Apply and analyze post-hoc tests.

   (f) Generate and analyze the results of residual plots to assess goodness of fit.

   (g) Apply two way ANOVA to a dataset and interpret confounding, main and interaction effects.

   (h) Block design.

   Additional Criteria:

   (a) Identify when there are problems with running ANOVA and what the potential fixes are (transformation of data, removal of outliers, Welsh ANOVA for unequal variances, etc).

   (b) With the aid of software, apply non-parametric ANOVA methods to data and interpret the results.

(c) Interpret factor choice criterion using the likelihood ratio test for nested models.

5. Perform simple and multiple regressions.

   Core Criteria:

   (a) Apply least squares methods to estimate the "line of best fit" from a simple or multiple linear regression.
   (b) Interpret the results of a simple linear regression in terms of the F test for determining if any of the covariates are significant.
   (c) Interpret the R^2 values for a simple or multiple linear regression within the context of the problem.
   (d) Identify appropriate variables to determine significance in a given model.
   (e) Identify the effects that correlation of covariates creates in multiple linear regression.
   (f) Apply criterion to find a parsimonious model.
   (g) Utilize residual plots to assess the appropriateness of a model.
   (h) Estimate confidence bounds on a simple linear regression.

   Additional Criteria:

   (a) Center data and fit a polynomial regression (third order or below).
   (b) Apply multiple categorical factors within a linear regression.
   (c) Apply prediction intervals (with the aid of software) when estimating a new value.
   (d) Generate confidence and/or prediction intervals for multiple linear regression.
   (e) Code factors to be included in a linear model.
   (f) Generate and interpret the ANOVA table associated with a multiple linear regression with categorical predictors.

6. (Optional) Perform logistic regression.

   Core Criteria: none

   Additional Criteria:

   (a) Identify confounding and effect modification between potential predictors for a logistic regression.
   (b) Run multiple logistic regression in order to predict the odds/probability of an event.
   (c) Graph logistic regression results.
   (d) Understand the concept of Odds and its role in logistic regression.
   (e) Demonstrate why linear regression is inappropriate when the response variable is binary.
   (f) Fit a logistic regression and interpret the results.
   (g) Identify and apply parameters of a logistic regression to predict the odds and probability of an event.
   (h) Assess goodness of fit for a logistic regression through the use of Concordance and/or AUC/ROC analysis.