

Title (Units): SCI7470 BUSINESS STATISTICS AND MODELING (3,3,0)

Syllabus Reviewed by: Michael Ng

Prerequisite/Co-requisite: SCI7430 STATISTICAL SOFTWARE IN BUSINESS AND MANAGEMENT

Objectives: This subject introduces statistical methods for analysing categorical data arisen from qualitative response variables which cannot be handled by methods dealing with quantitative response, such as regression and ANOVA.

Calendar Description: This subject introduces statistical methods for analysing categorical data arisen from qualitative response variables which cannot be handled by methods dealing with quantitative response, such as regression and ANOVA.

References:

- 1) Agresti, A. An Introduction to Categorical Data Analysis, Wiley, 1996.
- 2) Agresti, A. Categorical Data Analysis, 2nd Ed., Wiley, 2000.
- 3) T.J. Santner and D.E. Duffy, The Statistical Analysis of Discrete Data, Springer-Verlag, 1989.
- 4) Christensen, R. Log-Linear Models and Logistic Regression, Springer-Verlag, 1997.
- 5) S.E. Fienberg, The Analysis of Cross-Classified Categorical Data, MIT Press, 2nd Ed., 1980.
- 6) Hosmer, D. W and Lemeshow, S. Applied Logistic Regression, 2nd Ed., Wiley, 2000.

Assessment: Continuous assessment (30%) and Final examination (70%)

Subject Content in Outline:

- I. Two-way Contingency Tables (4 hours of teaching)
 - A. Comparing proportions
 - B. Odds ratio
 - C. Large sample inference
 - D. Small sample exact inference
- II. Three-way Contingency Tables (4 hours of teaching)
 - A. Partial association
 - B. Cochran-Mantel-Haenszel methods
 - C. Exact inference on conditional associations
- III. Generalized Linear Models (8 hours of teaching)
 - A. Binary data
 - B. Poisson counts
 - C. Empty cells and sparseness
- IV. Logistic Linear Models (8 hours of teaching)
 - A. Interpretation and inference
 - B. Qualitative predictors
 - C. Multiple logistic regression
 - D. Other logit functions
- V. Loglinear Models for Contingency Tables (10 hours of teaching)
 - A. Two-way tables

- B. Three-way tables
- C. Tables of higher dimension
- D. Connection between loglinear and logit models

VI. Generalized Estimating Equations (8 hours of teaching)

- A. Repeated Measures/Correlated Data
- B. Longitudinal Data
- C. Alternative for Generalized Linear Model