

Syllabus

Course and Contact Information:

Course: Statistics, Categorical Data Analysis, STAT 6231, 10

Semester: Spring, 2022

Meeting time: Wed, from 06:10PM to 8:40PM,

Location: Rome 351

Instructor:

Name: Hua Liang

Campus Address: Rome Hall, Room 754

Phone: 202-994-7844

E-mail: hliang@gwu.edu

Office hours: Wed 5:30pm-6:00pm or by appointment

TA: Yufei Gao (guoyufei@gwu.edu)

Course Description

The purpose of this course is to provide a basic broad overview of the statistical procedures for analyzing categorical data. We will begin with analysis of contingency tables with emphasis on testing independence. We will learn traditional methods for two-dimensional tables and then generalize to multidimensional tables. Theoretical bases underlying the analysis of categorical data will be covered. We will move to advanced topics such as measures and tests of association, partial table, conditional independence; Cochran-Mantel-Haenszel procedure; maximum likelihood estimators in generalized linear models; estimating equations; logistic regression; loglinear models, variable selection and model diagnostics. Various examples with detailed analyses in R will be provided.

Prerequisite

Basic knowledge for linear models, estimation, hypothesis testing, and central limit theory. Stat 6201, 6202.

Goals of Course

- Remember some basic formulae like odds ratio, chi-squared test, likelihood ratio test
- Know what assumptions are required
- Know how to use the methods for data analyses, theoretical basis of these methods and any potential limitations
- Develop strong data analysis skills
- Use R for implementation of various methods

Course Structure

The class consists of one 2.5 hours lecture per week. The grade is based on homework, one midterm and the final exam/final project/presentation.

Textbook:

1. Categorical Data Analysis by Alan Agresti (required)
2. Analysis of Categorical Data with R by Christopher R. Bilder and Thomas M. Loughin (2014, recommended)
3. An Introduction to Categorical Data Analysis by Alan Agresti (recommended)

Exams

There will be one midterm examination and one final examination. These will contribute the majority of your final grade. The midterm exam will cover contingency tables. The final examination will be held during the final exam week, and covers the entire course with an emphasis on the latter part of the course. Each examination will have both mathematical and conceptual (written) components.

Home Work

Homework will be given on a biweekly basis. Regularly homework is due every other Wed, or otherwise the date announced. Some will be of a theoretical basis, and some will be real data analysis. Data analysis should be used R, or Splus, or Matlab, or SAS. If you miss the class, you can ask/copy the assignment.

Final Grade

Your final grade will depend on the following components with these proportions: homework (30%), midterm (20%), final exam/project/presentation (50%). Last, poor attendance of the class will be taken into account for you final grade. The percentage grades needed to achieve an A, B, C, or D will follow approximately the following scale: $90 - 100 = A$, $80 - 89 = B$, $70 - 79 = C$, $55 - 69 = D$, $0 - 54 = F$.

University policies:

University policy on observance of religious holidays

In accordance with University policy, students should notify faculty during the first week of the semester of their intention to be absent from class on their day(s) of religious observance. For details and policy, see: students.gwu.edu/accommodations-religious-holidays.

Academic Integrity

Academic dishonesty is defined as cheating of any kind, including misrepresenting one's own work, taking credit for the work of others without crediting them and without appropriate authorization, and the fabrication of information. For details and complete code, see: studentconduct.gwu.edu/code-academic-integrity

Safety and security

In the case of an emergency, if at all possible, the class should shelter in place. If the building that the class is in is affected, follow the evacuation procedures for the building. After evacuation, seek shelter at a predetermined rendezvous location.