

AMS 573 (3 credits, Spring 2022) Categorical Data Analysis

Instructor	Changsoon(C.S.) Park Research Professor, Department of Applied Mathematics & Statistics, SUNY Korea
Class	Mon, Wed 10:30PM~11:50PM
Office	A611
Office Hour	Mon, Wed 09:00-10:00AM
Phone	1916
E-Mail	Changsoon.park@stonybrook.edu
Text	An Introduction to Categorical Data Analysis, 3rd Edition, Wiley Alan Agresti ISBN 9781119405269
Grading	ABCDF grading

Course Description:

Measuring the strength of association between pairs of categorical variables. Methods for evaluating classification procedures and inter-rater agreement. Analysis of the associations among three or more categorical variables using log linear models. Logistic regression.

Learning Outcomes:

- 1) Demonstrate skills of working with various categorical data, including binary, nominal, ordinal and count data:
 - * Expectation, variance, covariance and probability density function;
 - * Point estimation with maximum likelihood method;
 - * Hypothesis testing with Wald, score and likelihood ratio tests;
 - * Constructing confidence intervals based on Wald, score and likelihood ratio test statistics.
- 2) Demonstrate skills with statistical inference for contingency tables (joint distribution of categorical variables):
 - * Difference of proportions, relative risk and odds ratio;
 - * Chi-squared tests;
 - * Fisher's exact test;
 - * McNemar test for matched pairs.
- 3) Demonstrate skills with statistical modeling for binary/nominal/ordinal response:
 - * Build and apply logistic regression, baseline category and cumulative logit models;
 - * Maximum likelihood fitting and goodness of fit tests;
 - * Model diagnostic and model selection;
- 4) Demonstrate skills with statistical modeling for count data:
 - * Build and apply log-linear models;
 - * Connection between log-linear and logit models;

- * Model fitting and goodness of fit tests;
- 5) Demonstrate skills with proficient usage of standard statistical software tools for categorical data analysis:
 - * Understanding of the assumptions, derivation and interpretation of results from statistical analysis;
 - * Proficient in R functions using libraries for categorical data analysis. Optional use of SAS procedures: FREQ, GENMOD, GLM and LOGISTIC.

Grades:

Class attendance – 10%
Homework – 10%
Midterm 1,2 – 20% each (Apr 5, May 8)
Final – 40% (TBA)

Homework will be assigned at the end of each chapter, and is due one week from the day it is assigned on. Each student must turn in the homework at the beginning of the lecture on the due date. Late homework will not be accepted. Solutions will be posted on the webpage after the due date. The first page of the homework should contain the following information:

AMS 573
Homework #
(Your name)

All exams are closed-book tests. You may bring a calculator and an equation paper (one page of the letter size).

Final grade will be given according to the distribution of sum of five items.

Lectures:

Lecture notes will be available to be downloaded from the course webpage. It would be a good idea to print the notes and bring them to the class.

No reference is needed, and examples and problems only in the textbook will be explained.

Bring the textbook to the class.

Academic Integrity:

Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty are required to report any suspected instances of academic dishonesty to the Academic Judiciary. For more comprehensive information on academic integrity, including categories of academic dishonesty, please refer to the academic judiciary website at <http://www.stonybrook.edu/uaa/academicjudiciary/>

School Policy on Attendance:

1. If a student has over 20% unexcused absences, the student's final course grade will be an F.
2. Students should report the reason of absence to the professor in advance, or immediately after the absence.
3. When a student excuses his/her absence, the student must provide documentation of the reason for the absence to the professor.
4. The professor of the course reserves the right to excuse absences.
5. The professor may excuse the absence if the submitted documentation fulfills the following conditions: extreme emergencies, severe medical reasons with doctor's note, very important events.

Tentative Course Schedule

Week	Dates	Chapter	Topic
1	2/21 2/23	Chapter 1	Introduction Review of Regression Theory
2	2/28, 3/2	Chapter 2	Analyzing Contingency Tables
3	3/7, 3/9	Chapter 2 No Class	
4	3/14, 3/16	Chapter 3	Generalized Linear Models
5	3/21, 3/23	Chapter 3	
6	3/28 3/30	Review Midterm I	
7	4/4, 4/6	Chapter 4	Logistic Regression
8	4/11, 4/13	Chapter 4	
9	4/18, 4/20	Chapter 5	Binary Regression Models
10	4/25 4/27	Chapter 5 Review	
11	5/2, 5/4	No class	Adjustment Day
12	5/9 5/11	Midterm II Chapter 6	Multicategory Logit Models
13	5/16, 5/18	Chapter 6	
14	5/23 5/25	Chapter 7	Contingency and Counts
15	5/30, 6/1	Chapter 8 No Class	Matched Pairs
16	6/7	Review	