STATISTICS 544
Categorical Data Analysis
Fall 2017
Course Syllabus

PREREQUISITES: STAT 512, 514

INSTRUCTOR: Prof. Le Bao (lebao@psu.edu)

CLASSROOM: Willard Bldg 365, M W F, 11:15AM - 12:05PM

OFFICE HOURS: Wartick Lab 514C, Tuesday 2:00 PM - 3:30 PM

CLASS WEBPAGE: available through CANVAS
check regularly for lecture notes and updated announcements

HOMEWORK: Given out biweekly. Late homework assignments will not
be accepted, unless the student obtains the permission in advance.
You need work on the problem sets independently.
Plagiarism on any assignments will be dealt with as per
the policy of the Eberly College of Science on Academic Integrity
(modification of Senate Policy 49-20), available at
http://www.science.psu.edu/academic/Integrity/Policy.html

Wiley, New York.

EVALUATION: Homework assignments: 30%
Midterm exam (in class): 25%
Project and presentation: 35%
Class participation: 10%

COMPUTER: Most of the examples we talk about in class have been
implemented in R (http://www.r-project.org/).
If you have other software packages available from your work
or research that can implement the proposed methods, then that
will be fine as well.
Learning Objectives

Upon entering this course, you are expected to have a basic knowledge of applied statistical methods. You should know how to fit basic descriptive statistics and linear regression models to datasets and to have an understanding of hypothesis testing, confidence intervals either through coursework or practical experience.

After completing the course, the student can ordinarily expect to be able to:

1. Identify and understand the structure of categorical data and be able to phrase the appropriate scientific questions in terms of parameters of interest.
2. Understand the various assumptions needed for the various methodologies covered in the class as well as their implementation.
3. Read the scientific literature and comprehend the use (and misuse) of categorical data analysis methodologies reported by study authors.

Course Topics

1. Review of basic probability distributions
2. Contingency table analysis, inference, and model evaluation
3. Loss function for categorical data and decision theory
4. Generalized linear model (Logistic regression and Poisson regression)
5. Discriminant analysis
6. Generalized linear mixed model and Bayesian hierarchical model
7. Tree structured classification models
8. Special topics e.g. data compression, loglinear models for contingency table, analysis of rank data, causal inference, Bayesian network, latent variable approach for ordinal data, GEE, etc, based on majority votes.
ACADEMIC INTEGRITY:
The conduct of a student registered or taking this course should be consistent with that of
a professional person. Courtesy, honesty and respect should be shown by students toward
faculty members, guest lecturers, administrative support staff and fellow students. Similarly,
students should expect faculty to treat them fairly, showing respect for their ideas and
opinions and striving to help them achieve maximum benefits from their experience in the
course.
Student academic misconduct refers to behavior that may include plagiarism, cheating, fab-
rication, falsification of records or official documents, intentional misuse of equipment or
materials (including library materials), and aiding and abetting the perpetration of such
acts. The preparation of reports, papers, and examinations, assigned on an individual ba-
sis, must represent each student’s own effort. Reference sources should be indicated clearly.
The use of assistance from other students or aids of any kind during a written examination,
except when the use of aids such as electronic devices, books or notes has been approved by
an instructor, is a violation of the standard of academic conduct. All violations of academic
integrity will be assessed and dealt with based on the policy of the Eberly College of Science
on Academic Integrity, available for reading at

http://www.science.psu.edu/academic/Integrity/Purpose.html

Physically disabled and learning disabled students
It is Penn State’s policy to not discriminate against qualified students with documented dis-
abilities in its educational programs. If you have a disability related need for modifications
in this course, contact your instructor and the Office for Disability Services (located in 116
Boucke Building) or the Disability Contact Liaison at your Penn State location. Instructors
should be notified as early in the semester as possible. You may refer to the Nondiscrimina-