



University of Maryland School of Public Health

EPIB 652 – Categorical Data Analysis

Semester: Fall 2020
Section: 0101
Classroom and Time: Wednesday 4:00-6:45pm , Online (Zoom)
Course webpage: <https://umd.instructure.com/courses/1287830>
Instructor: Charles Ma, Ph.D. **Office Hours:** Wednesday 2:00-3:00pm or by appointment
Office: SPH 2234M
Phone: 301-405-6421
Email: tma0929@umd.edu

Course Description:

This course provides an introduction to methods for analyzing categorical data, with focus on public health research. It emphasizes the ideas behind the methods and their applications. The course covers descriptive and inferential methods for contingency tables and generalized linear models for binary, nominal, ordinal, and count data. It also covers methods for matched-pairs data. Computations are illustrated using SAS.

Course Pre- and Co-requisites:

Required: SPHL 602 and EPIB 651 or equivalent, or permission of instructor.

Recommended: Previous experience with at least one statistical software (e.g. SAS*, R, STATA)

*: SAS is the main software used for in-class demonstration

Course Learning Objectives:

Upon completing this course, the student will be able to:

1. Understand and describe the major types of categorical data and discrete probability distributions.
2. Identify, describe, and apply appropriate descriptive and inferential methods for contingency tables (e.g., odds ratio, independence test).
3. Identify, describe, and build appropriate models for different types of categorical data (e.g. logistic regression model, loglinear model).
4. Interpret results of categorical data analyses in oral and written forms.
5. Analyze categorical data using statistical software.

Program Competencies Addressed in this Course:

The following competencies for the *Master of Public Health with concentration in Biostatistics* are addressed in this course:

1. Distinguish among the different measurement scales or types of variables and select appropriate descriptive statistical methods for summarizing public health data.
2. Select appropriate inferential statistical methods to answer research questions relevant to public health research.

3. Conduct descriptive and inferential statistical analyses that are appropriate to different basic study designs used in public health research.
4. Critically evaluate statistical analyses presented in public health literature.
5. Use statistical analytical software packages (e.g. SAS, R, STATA) to describe, explore, and summarize data as well as perform statistical procedures.
6. Communicate results of statistical analyses to lay and professional audiences.
7. Analyze quantitative and qualitative data using biostatistics, informatics, computer-based programming and software, as appropriate.
8. Interpret results of data analysis for public health research, policy or practice.

Required Texts and Other Readings:

Required:

Agresti, Alan. (2019). *An Introduction to Categorical Data Analysis*, 3rd Edition. John Wiley & Sons. ISBN: 978-1119405269.

- Companion website for the textbook: <http://bcs.wiley.com/he-bcs/Books?action=index&bcsId=11293&itemId=1119405262>
- Data files used in the textbook: <http://users.stat.ufl.edu/~aa/intro-cda/data/>
- SAS files containing datasets in the Appendix of the textbook: <http://users.stat.ufl.edu/~aa/intro-cda/appendix.html>

Recommended:

Agresti, Alan. (2013). *Categorical Data Analysis*, 3rd Edition. John Wiley & Sons. ISBN: 978-0470463635.

- Website for Categorical Data Analysis by Alan Agresti (Wiley, 2013): <http://users.stat.ufl.edu/~aa/cda/cda.html>
- Github site for Categorical Data Analysis by Alan Agresti (Wiley, 2013): <https://github.com/cjgeyer/CatDataAnalysis>

Penn State Online STAT 504: Analysis of Discrete Data:
<https://online.stat.psu.edu/stat504/node/49/>

Required Technology and Other Materials: Zoom Application

Course Communication: *See announcements at the course website.*

Course Requirements:

Homework:

There will be four homework assignments for this course, and each will be due at the beginning of the due date class. You may discuss with your classmates or work in teams for the homework,

but each student has to submit their own homework. **Late homework will NOT be accepted without a reasonable and advance notice.**

Exam:

There will one midterm exam, **open book and open note**. The contents of the exam will be cumulative. The midterm will be made available on the course website on the test date and you can access and download the exam any time on that day. Once accessed, you will have **three hours to complete the exam and submit back to the course website**. No makeup exams are allowed in general. **You CANNOT discuss the exam with your classmates**. Exceptions to this rule are evaluated on a case-by-case basis. Students must submit a request before the exam with valid supporting documents. No post-exam request will be considered except the student is hospitalized during the exam period.

Project:

By the end of the semester (12/18/2020), you will be asked to turn in a final project report. The purpose of this project is to gain experience in applying some of the statistical methods that you have learned in class, and in communicating the results of these analyses. For details, please refer to the [Final Project guideline](#).

Covid-19 related policies:

For unanticipated absences during the pandemic, **self-certified notes** will serve as documentation for COVID-19-related absences or missed course expectations. To provide the academic accommodations and help students who face challenging issues related to the pandemic, we may consider late assignment and make-up work, detailed exemption will be evaluated on a case-by-case basis.

Academic Integrity: The University of Maryland [defines academic dishonesty](#) as committing or facilitating cheating, fabrication, plagiarism, or self-plagiarism. Academic integrity is expected and students who do not uphold the UMD Code of Academic Integrity will be referred to the Office of Student Conduct.

Use of Course Assistance Websites and Online Group Forums: Course assistance websites, such as CourseHero and others, are not permitted resources for SPH courses, unless the professor explicitly gives permission for you to use one of these sites. Material pulled from these sites can be deemed unauthorized material and a violation of academic integrity. These sites offer information that might not be accurate and more generally stifle the learning process. In addition, it is understandable that students may use one of a variety of online or virtual forums for course-wide discussion (e.g., GroupME, WeChat, etc.). Collaboration in this way regarding concepts discussed in this course is permissible. However, collaboration on graded assignments is strictly prohibited. Examples include: asking classmates for answers on quizzes or exams, asking for access codes to clicker polls, etc. Additional information on academic integrity is found in **University Course Related Policies**, below.

University Course Related Policies:

All University of Maryland-approved graduate course policies are provided here:
<https://gradschool.umd.edu/course-related-policies>

Policy descriptions, resources, and links to official policy documents are provided for:

Academic Integrity: What is cheating? What is plagiarism? What is the Honor Pledge?

Code of Student Conduct: What behavior is prohibited?

Sexual Misconduct: What to do in case of sexual harassment or sexual assault.

Non-Discrimination: Procedures to prohibit discrimination, complaints about discrimination, harassment, and retaliation.

Accessibility: Information about disability support services (DSS) and accommodations.

Attendance, Absences, or Missed Assignments: The student must notify the instructor in a timely manner (typically first week of class). Read this prior to Schedule Adjustment date.

Course and Credit Changes and Withdrawals

Reasonable Expectations of Faculty in Conducting Academic Courses

Official UMD Communication: Use of email, communication with faculty, communication about cancelled class meetings, and weather-related or other urgent notifications.

Arbitrary and Capricious Grading Policy

Incomplete Grades

Good Standing and Academic Probation and Dismissal

Leave of Absence Policies

Graduate Student Rights and Responsibilities

Grievance Procedure

Other Resources: Ombuds Office, counseling, learning workshops, tutoring, writing help, questions about graduation, etc.

Course Procedures and Policies:

Inclement Weather / University Closings / Emergency Procedures:

In the event that the University has a delayed opening or is closed for an emergency or extended period of time, the instructor will communicate to students regarding schedule adjustments, including rescheduling of examinations and assignments due to inclement weather and campus emergencies.

Grading Procedures:

Grade of this course will be based on determined as follows:

- Homework 40%
- Midterm Exam 30%
- Project 30%

Course Outline / Course Calendar:

Course Schedule Summary

Session	Date	Topic	Reading/ Assignments
# 1	09/02/2020	Introduction and review: - Categorical response data - Distributions of categorical data (Bernoulli, Binomial/Multinomial, Poisson/Negative Binomial)	Chapters 1.1-1.2
# 2	09/09/2020	Introduction and review: - Likelihood function and maximum likelihood estimation - Inference and hypothesis test (Wald, Likelihood-ratio and Score tests)	Chapters 1.3-1.4
# 3	09/16/2020	Two-way contingency tables I: - Table structure, Joint/Marginal/Conditional distribution - Measures of association (difference in proportion, relative risk, odds ratio)	Chapters 2.1-2.3
# 4	09/23/2020	Two-way contingency tables II: - From 2x2 to IxJ table - Test of independence (X^2 and G2), residuals - Testing independence for ordinal data	Chapters 2.1-2.5
# 5	09/30/2020	Three-way contingency tables I: - Fisher's exact test (two-way table cont'd) - Partial and conditional tables, Marginal and conditional odds ratio - Simpson's paradox	Chapters 2.6-2.7 HW1 due
# 6	10/07/2020	Three-way contingency tables II: - Different types of independence (Marginal, Joint and Conditional independence) - Cochran-Mantel-Haenszel (CMH) test - Homogeneous association and Breslow-Day statistics	Chapters 2.7, PSU STAT 504 Lesson 5
# 7	10/14/2020	Generalized linear models (GLM): - Components of GLM - GLMs for binary data - GLMs for count data - Statistical inference and model checking	Chapters 3.1-3.5 HW2 due
# 8	10/21/2020	Logistic regression I: - Interpreting logistic regression - Inference for logistic regression - Categorical predictors - Multiple logistic regression	Chapters 4.1-4.6
# 9	10/28/2020	Logistic regression II: - Strategies in model selection - Model check and diagnostics - Conditional logistic regression	Chapters 5.1-5.3
# 10	11/04/2020	Multicategory logistic regression models:	Chapters 6.1-

		- Baseline-category logit models for nominal responses - Cumulative logit model for ordinal responses	6.2 HW3 due
# 11	11/11/2020	Midterm Exam	
# 12	11/18/2020	Loglinear model I: - Loglinear models for two-way table - Loglinear models for three-way table - Connection to logistic model - Independence graph	Chapters 7.1-7.4
# 13	11/25/2020	Thanksgiving	
# 14	12/02/2020	Loglinear model II: - Modeling ordinal associations and Linear-by-linear association model - Poisson regression - Negative Binomial regression	Chapters 7.5-7.6
# 15	12/09/2020	Models for matched pairs: - Comparing dependent proportions (McNemar's test) - Marginal vs. Subject-specific model for matched pair - Kappa measure of agreement - Bradley-Terry model for paired preference - Generalized Estimating Equations (GEE)	Chapters 8.1-8.2, 8.5.5, 8.6 HW4 due

* Please note that this is a tentative schedule, and the actual materials covered in each lecture might not be exactly the same

Basic Needs Security: If you have difficulty affording groceries or accessing sufficient food to eat every day, or lack a safe and stable place to live and believe this may affect your performance in this course, please visit <http://go.umd.edu/basic-needs> for information about resources the campus offers you.

Student Name Change in ELMS-

Canvas: https://go.umd.edu/change_name_in_ELMS_Canvas

Campus Building Amenities: <https://maps.umd.edu/map/> (follow the prompts below and amenities will populate on the interactive campus map)

All Gender Restrooms: Click: Layers, Building Amenities; Select All Gender or Gender Inclusive Restrooms

Family Restrooms: Click: Layers, Building Amenities; Select All Gender or Gender Inclusive Restrooms

Lactation/Feeding Room Locations: Click: Layers, Building Amenities; Select Lactation/Feeding Rooms