City University of Hong Kong
Course Syllabus

offered by School of Data Science
with effect from Semester A 2024/25

Part I   Course Overview

Course Title:	Statistical Methods for Categorical Data Analysis
Course Code:	SDSC8013
Course Duration:	One semester
Credit Units:	3
Level:	R8
Medium of Instruction:	English
Medium of Assessment:	English
Prerequisites: *(Course Code and Title)*	Nil
Precursors: *(Course Code and Title)*	Nil
Equivalent Courses: *(Course Code and Title)*	Nil
Exclusive Courses: *(Course Code and Title)*	Nil

Part II   Course Details

1.   Abstract

In this course, students will learn how to use descriptive and test statistics, statistical models and statistical inferences to analyse categorical data. Topics include statistical inference using odds ratios and relative risks, multi-way contingency tables, tests for stratified analysis, odds ratio and relative risk, generalized linear models for discrete data, multi-category logit model for nominal and ordinal responses, inference for matched-pairs, and loglinear models. Students will learn where these methods may be applied, how to apply them, and how to properly interpret the results. The course is appealing to those interested in categorical data analysis, and examples are based on case studies. Students will learn statistical way of thinking in analyzing categorical data, and the methodology can be applied to different fields to solve similar problems.

2.   Course Intended Learning Outcomes (CILOs)

No.	CILOs	Weighting (if applicable)	Discovery-enriched curriculum related learning outcomes		
			A1	*A2*	*A3*
1.	Describe different study designs and data types in categorical data analysis	15%	✔		
2.	Explain different descriptive and test statistics and generalized linear models for different types of outcomes	20%	✔	✔	
3.	Utilize software packages and programs to conduct statistical tests and fit statistical models	30%	✔	✔	
4.	Interpret model results properly and draw conclusions in case studies	20%		✔	✔
5.	Implement taught statistical methods to analyze real-world datasets	15%		✔	✔
		100%			

A1:   *Attitude*
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.
A2:   *Ability*
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to self-life problems.
A3:   *Accomplishments*
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

3. Learning and Teaching Activities (LTAs)

LTA	Brief Description	CILO No.					Hours/week (if applicable)
		1	2	3	4	5	
Lectures	Students will gain knowledge on different categorical data types and statistical tests and models through lectures with examples of case studies.	✔	✔	✔	✔		26 hours/sem
Tutorial Sessions	Students will engage in tutorial activities to utilize computer software tools for implementing statistical tests and models in real-world data analysis.	✔	✔	✔	✔	✔	13 hours/sem

4. Assessment Tasks/Activities (ATs)

Assessment Tasks/Activities	CILO No.					Weighting	Remarks
	1	2	3	4	5		
Continuous Assessment: __100__ %							
Assignments Questions are designed to evaluate how well students are able to apply statistical methods and computational tools learnt in class to analyze datasets and solve problems.	✔	✔	✔	✔	✔	30%	
Midterm Exam Examination will assess whether students get a deep understanding of statistical methods and tools and their applications in solving different problems.	✔	✔	✔	✔	✔	30%	
Project A final project using real data will assess students' abilities to choose appropriate statistical tests and models to solve real-world problems, interpret results properly, draw conclusions, and summarize findings in the write-up.	✔	✔	✔	✔	✔	40%	
						100%	

5. **Assessment Rubrics**

Applicable to students admitted from Semester A 2022/23 to Summer Term 2024

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B)	Marginal (B-, C+, C)	Failure (F)
1. Assignments	Ability to choose appropriate models and tests for given tasks and apply statistical tools to solve problems and draw conclusions based on data.	High	Significant	Basic	Not even reaching marginal levels
2. Midterm Exam	Students' level of achievements of the intended learning outcomes, with emphases on understanding, correct application of statistical methods, correct interpretation, and conclusions.	High	Significant	Basic	Not even reaching marginal levels
3. Project	Ability to demonstrate a deep understanding of the concepts, choose appropriate tests and models for solving problems, apply methods correctly, interpret results and draw conclusions properly, and summarize key findings.	High	Significant	Basic	Not even reaching marginal levels

Applicable to students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Marginal (D)	Failure (F)
1. Assignments	Ability to choose appropriate models and tests for given tasks and apply statistical tools to solve problems and draw conclusions based on data.	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. Midterm Exam	Students' level of achievements of the intended learning outcomes, with emphases on understanding, correct application of statistical methods, correct interpretation, and conclusions.	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Project	Ability to demonstrate a deep understanding of the concepts, choose appropriate tests and models for solving problems, apply methods correctly, interpret results and draw conclusions properly, and summarize key findings.	High	Significant	Moderate	Basic	Not even reaching marginal levels

Part III   Other Information (more details can be provided separately in the teaching plan)

1. **Keyword Syllabus**

   - Odds ratios and relative risks
   - Wald tests and likelihood ratio tests
   - Stratified analysis, multi-way contingency tables
   - Generalized linear models for discrete data
   - Multi-category logit model for nominal and ordinal responses
   - Matched-pairs
   - Loglinear models

2. **Reading List**
2.1 **Compulsory Readings**

1.	Agresti, A. (2007). An Introduction to Categorical Data Analysis, 2nd Edition. New York: Wiley.

2.2 **Additional Readings**

1.	Agresti, A. (2013). Categorical Data Analysis, 3rd Edition. New York: Wiley.