The Chinese University of Hong Kong (Shenzhen)
Fall Term, 2020/2021

1. Course Identity

Course code	STA4030
Course title (English)	Categorical Data Analysis
Course title (Chinese)	范畴性数据分析
Units	3
Description (English)	This course deals with major statistical techniques in analyzing categorical data. Topics include measures of association, inference for two-way contingency tables, log-linear models, logit models and models for ordinal variables. The use of related statistical packages will be demonstrated as well.

2. Instructor

Name: Dr. Bojun Lu (吕伯君)
Email: bojunlu@cuhk.edu.cn
Office: Room 413, Chengdao Building
Office hours: 13:00-- 14:00PM, every Thursday or by appointments

3. Teaching Assistant

Name: Mr. Ceyao Zhang
Email: 218019058@link.cuhk.edu.cn
Office: Room 410, TD
Office hours: 15:00--17:00PM, every Tuesday or by appointments

Name: Mr. Chi Li
Email: 220019044@link.cuhk.edu.cn
Office: Room 320A, Chengdao Building
Office hours: 15:00—17:00PM, every Monday or by appointments

4. Lecture and Tutorial Information

	Time Slot	***Venue***
Lectures	Tuesday 1:30PM—3:00PM Friday 10:30AM—12:00PM	Room 110 Zhixin Building

Tutorial	Thursday 7:00—7:50PM Thursday 8:00—8:50PM Friday 8:00—8:50PM	Room 307 Teaching A Building

5. Prerequisites / Co-requisites

STA2001 Probability and Statistics I or ECO2121 Basic Statistics, STA2002 Probability and Statistics II

6. Learning Outcomes

After completing the course, students should be able to,

1. Analyze and interpret categorical data accurately;
2. Appreciate and understand the applicability of different techniques for analyzing categorical data;
3. Understand the relationships between different methods of analyzing categorical data;
4. Model categorical data.

6. Assessment Scheme

Component/ method	**% weight**
Problem Sheets	20
Mid-term exam	30
Final exam	50

✓ Problem Sheets: Four problem sheets will be set, with the deadlines for each made clear in advance. No late submissions will be accepted; late submissions will receive a zero mark. Students may discuss set problems with others, but their final submissions must be their own work.

7. Feedback for evaluation

- Formal course and teaching evaluation
- Feedback from office hour discussions
- Feedback after class
- Feedback from tutorial sessions

8. Reading

A. Required

a) A. Agresti, An Introduction to Categorical Data Analysis, 2nd Edition
b) A. Agresti, Categorical Data Analysis, 3rd Edition.
c) Instructor's notes (Available at Blackboard System online).

B. Recommended

a) R. B. Christopher, T. M. Loughin, Analysis of Categorical Data with R, 1st Edition.

9. Course components

Activity	**Hours/week**
Lectures	3 hours
Tutorial	1 hour
Voluntary exercises (and their solutions) for practice	2 hours

10. Indicative teaching plan

Week	**Content/ topic/ activity**	**Comments**
1	Preliminaries (Chapter 1)	
2	One-way Tables (Chapter 1)	
3	2-way Contingency Tables (Chapter 2)	Problem Sheet 1 handed out
4	2-way Contingency Tables (Chapter 2)	
5	3-way Contingency Tables (Chapter 3)	Solutions to Problem Sheet 1 collected, Problem Sheet 2 handed out
6	3-way Contingency Tables (Chapter 3)	
7	Generalized Linear Models (Chapter 4)	Solutions to Problem Sheet 2 collected,
8	Midterm	

9	Generalized Linear Models (Chapter 4)	Problem Sheet 3 handed out
10	Logistic Regression (Chapter 5)	
11	Multicategory Logit Models (Chapter 6)	Solutions to Problem Sheet 3 collected, Problem Sheet 4 handed out
12	Multicategory Logit Models (Chapter 6)	
13	Loglinear Models for Contingency Tables (Chapter 7)	Solutions to Problem Sheet 4 collected
14	Final	