ACE Network Subject Information Guide

STAT4XXX Categorical Data Analysis

Semester 2, 2021

Administration and contact details

Host department	School of Mathematical & Physical Science
Host institution	University of Newcastle
Name of lecturer	Eric Beh
Phone number	02 4921 5113
Email address	eric.beh@newcastle.edu.au
Homepage	https://www.newcastle.edu.au/profile/eric-beh
Name of honours coordinator	Robert King
Phone number	02 4921 5548
Email address	robert.king@newcastle.edu.au
Name of masters coordinator	NA
Phone number	Click here to enter text.
Email address	Click here to enter text.

Subject details

Handbook entry URL	https://www.newcastle.edu.au/degrees/bachelor-of-mathematics-honours
Subject homepage URL	NA
Honours student hand-out URL	Contact the Honours Convenor listed above
Start date:	July 19, 2021
End date:	October 29, 2021
Contact hours per week:	2hrs
Census date:	August 13, 2021
Lecture day(s) and time(s):	TBA
Description of electronic access arrangements for students (for example, WebCT)	Lectures to be held via zoom. Resources distributed via email from eric.beh@newcastle.edu.au

Subject content

1. Subject content description

Categorical data abounds in all disciplines as researchers and analysts search for ways of analysing data collected from surveys or questionnaires. Undergraduate courses only provide a cursory glance at how categorical data can be analysed. In this course we will examine some of the core contributions to categorical data analysis with a focus on measures of association, categorical data visualisation and modelling categorical data.

The course will include the following topics

- Visualisation of categorical data
- History and development of contingency tables
- Pearson's chi-squared statistic and related measures
- Features, and variations of the odds ratio for single and stratified data
- Reciprocal averaging and singular value decomposition
- Correspondence analysis
- Modelling categorical data

2. Week-by-week topic overview

Week 1: An overview of categorical data, its history and standard techniques
Week 2: Analysis of a single categorical variable – goodness-of-fit tests
Week 3: The contingency table, chi-squared statistic & related measures
Week 4: Measures of symmetric association for 2x2 contingency tables
Week 5: Measures of symmetric association for IxJ contingency tables
Week 6: Measures of asymmetric association for IxJ contingency tables
Week 7: Scaling categorical data – reciprocal averaging & canonical correlation
analysis
Week 8: Simple correspondence analysis
Week 9: Non-symmetric correspondence analysis
Week 10: Multiple correspondence analysis
Week 11: Models of correlation and association
Week 12: Log-linear models

3. Assumed prerequisite knowledge and capabilities

Undergraduate major in statistics. Students must have knowledge of common statistica techniques, especially linear statistical models, chi-squared tests, statistical inference and linear algebra (with an emphasis on vectors and matrices). Students must also have an introductory knowledge of R/Rstudio.

4. Learning outcomes and objectives

On successful completion of this course, students will be able to:

1. Gain a deeper understanding of the analysis of categorical data
2. Explore more deeply the issue concerned with Pearson's chi-squared statistic and related measures of association that reflect symmetric and asymmetric association
3. Apply new statistical tools to numerically and visually analyse multiple categorical variables
4. Apply a variety of correspondence analysis techniques
5. Model categorical data using association models and log-linear models
6. Apply their skills to real-life data using R
7. Undertake basic research skills concerned with categorical data analysis

AQF specific Program Learning Outcomes and Learning Outcome Descriptors (if available):

AQF Program Learning Outcomes addressed in this subject	Associated AQF Learning Outcome Descriptors for this subject
Understand some core techniques for analysing categorical data	**K1**
Understand the interconnectedness of techniques used for analysing categorical data	**K1, S1, S2, S3, S5**
Be able to identify the most appropriate statistical technique required to address practical problems involving categorical data	**S1 A2**
Understand the strengths, weaknesses and alternatives of techniques designed for the analysis of categorical data	**A2**
Develop *independent* learning and *problem-solving* skills associated with analysing categorical data	S1, S5
Develop coding skills that produce appropriate (practically relevant) numerical and visual output for the analysis of categorical data	Choose from list below
Insert Program Learning Outcome here	Choose from list below

Learning Outcome Descriptors at AQF Level 8

Knowledge

K1: coherent and advanced knowledge of the underlying principles and concepts in one or more disciplines

K2: knowledge of research principles and methods

Skills

S1: cognitive skills to review, analyse, consolidate and synthesise knowledge to identify and provide solutions to complex problem with intellectual independence

S2: cognitive and technical skills to demonstrate a broad understanding of a body of knowledge and theoretical concepts with advanced understanding in some areas

S3: cognitive skills to exercise critical thinking and judgement in developing new understanding

S4: technical skills to design and use in a research project

S5: communication skills to present clear and coherent exposition of knowledge and ideas to a variety of audiences

Application of Knowledge and Skills

A1: with initiative and judgement in professional practice and/or scholarship

A2: to adapt knowledge and skills in diverse contexts

A3: with responsibility and accountability for own learning and practice and in collaboration with others within broad parameters

A4: to plan and execute project work and/or a piece of research and scholarship with some independence

5. Learning resources

Test & printed notes: Students are not required to purchase any reference book for this course. Instead lecture notes will be provided. Published articles in commonly available, high-profile, journals will also be made available to students for additional insight and context of the weekly topics; they will need to access this material through their own library (full bibliographic information of each article will be provided)

Software: Students will need to have access to R/Rstudio. They can be freely downloaded from the internet.

6. Assessment

Exam/assignment/classwork breakdown					
Exam	55%	**Assignment**	45%	**Class work**	Enter %
Assignment due dates		Assignment 1 Assignment 2 Assignment 3	August 20 (Wk 5) September 17 (Wk 9) October 29 (Wk 13)	Click here to enter a date.	Click here to enter a date.
Approximate exam date			(Week 2 of UoN exam period)	November 8 – 12, 2021	

Assignment 1 and 2 require solving specific questions that are of a conceptual, theoretical, practical and computational nature and are aligned with the weekly lecture material (Weeks 1 – 9 inclusive)

Assignment 3 is a ten-page written assessment that gauges the students broad understanding of the interconnectedness of the weekly topics. This assignment will be based on the full content of the semester (Weeks 1 – 12 inclusive) and will be assessed in terms of ten specific criteria.

Institution honours program details

Weight of subject in total honours assessment at host department	12.5%
Thesis/subject split at host department	Thesis = 37.5%, Course work = 62.5%
Honours grade ranges at host department	
H1	85% or above
H2a	75% to 84%
H2b	65% to 74%
H3	50% to 64%

Institution masters program details

Weight of subject in total masters assessment at host department	Click here to enter text.
Thesis/subject split at host department	Click here to enter text.
Masters grade ranges at host department	
H1	Enter range %
H2a	Enter range %
H2b	Enter range %
H3	Enter range %