STAT721

Multivariate Analysis

S2 Evening 2018

Dept of Statistics

Contents

General Information

Unit convenor and teaching staff
Unit Convenor
Jun Ma
jun.ma@mq.edu.au
Contact via email
12 Wally's Walk (E7A), room 526
TBA

Credit points
4

Prerequisites
Admission to MRes

Corequisites
STAT710

Co-badged status
STAT721

Unit description
This unit studies basic methods of multivariate statistical analysis. Multivariate data arise when each unit of observation in the sample has more than one variable measured. Multivariate statistical analysis provides ways to analyse dependence structures within multivariate data, as well as to meaningfully simplify, classify and group such data. The unit introduces methodologies and techniques for the exploration and analysis of multivariate data. Topics include graphical displays, discriminant analysis, principal components analysis, multivariate normal distribution, multivariate linear models, cluster analysis.

Important Academic Dates

Information about important academic dates including deadlines for withdrawing from units are available at https://www.mq.edu.au/study/calendar-of-dates

Learning Outcomes

On successful completion of this unit, you will be able to:

Understand the fundamental difference between univariate and multivariate analysis.

Know how to perform hypothesis testing (mainly the Hotelling T2 test) using multivariate data.

Understand and be able to apply MANOVA.

Understand multivariate regression.

Know the theories of PCA and factor analysis, and be able to apply these methods to real data.

Understand likelihood based, as well as minimum expected cost based, discriminant analysis. Be able to apply these discriminant analysis methods to real data.

Understand the principles of classification tree and canonical analysis.

Know how to display multivariate data graphically using R.

Be able to use the R package for multivariate data analysis.

General Assessment Information

Assignment late submission

- Late submissions without approval will be penalized at a rate of 5% of earned mark per day, up to maximum of 50%. This implies that submissions will not be accepted more than 10 days late.
- Late submissions due to unavoidable disruptions must be approved and students are required to submit evidence for the disruptions.

Assessment Tasks

Name	Weighting	Hurdle	Due
Assignment 1	15%	No	6pm on Sep 5, 2018
Assignment 2	15%	No	6pm on Oct 31, 2018
Takehome examination	45%	No	10am Nov 12, 2018
Written Examination	25%	No	University Exam period

Assignment 1

Due: **6pm on Sep 5, 2018**
Weighting: **15%**

Assignment 1 will be available on the unit webpage in week 3 and due in the week 6 lecture. Assignments may be handwritten or word-processed. Students can submit their assignment in person to the lecturer, or electronically via email to "jun.ma@mq.edu.au".

On successful completion you will be able to:

- Understand the fundamental difference between univariate and multivariate analysis.

- Know how to perform hypothesis testing (mainly the Hotelling T2 test) using multivariate data.
- Understand and be able to apply MANOVA.
- Know how to display multivariate data graphically using R.
- Be able to use the R package for multivariate data analysis.

Assignment 2

Due: **6pm on Oct 31, 2018**
Weighting: **15%**

Assignment 2 will be available in week 9 and due in the week 12 lecture. Assignments may be handwritten or word-processed, and can be submitted in person, or electronically via email to "jun.ma@mq.edu.au".

On successful completion you will be able to:

- Understand multivariate regression.
- Know the theories of PCA and factor analysis, and be able to apply these methods to real data.
- Understand likelihood based, as well as minimum expected cost based, discriminant analysis. Be able to apply these discriminant analysis methods to real data.
- Understand the principles of classification tree and canonical analysis.
- Know how to display multivariate data graphically using R.
- Be able to use the R package for multivariate data analysis.

Takehome examination

Due: **10am Nov 12, 2018**
Weighting: **45%**

For the **take-home** exam students will have THREE days to complete their exam papers. The exam paper will be available on unit web page at 10am on Friday 9th November, 2018 and the students must submit their answers before 10am, Monday November 12, 2018. This examination involves mainly analysis of real data sets and some simple theoretical questions. Computer coding in R is required.

On successful completion you will be able to:

- Understand the fundamental difference between univariate and multivariate analysis.
- Know how to perform hypothesis testing (mainly the Hotelling T2 test) using multivariate data.
- Understand and be able to apply MANOVA.
- Understand multivariate regression.

- Know the theories of PCA and factor analysis, and be able to apply these methods to real data.
- Understand likelihood based, as well as minimum expected cost based, discriminant analysis. Be able to apply these discriminant analysis methods to real data.
- Understand the principles of classification tree and canonical analysis.
- Know how to display multivariate data graphically using R.
- Be able to use the R package for multivariate data analysis.

Written Examination

Due: **University Exam period**
Weighting: **25%**

This is a written exam and it is to be scheduled in the university exam period. This examination mainly involves conceptual questions or simple calculation questions. For example, it may ask students to identify an appropriate hypothesis testing method for a particular data set. For this exam, students are allowed to bring into the exam room TWO A4 paper written/typed on both sides; photocopies **are not** allowed. Only non-programmable calculators that do not have text retrieval capacity are allowed. Students who apply Supplementary Exams must make themselves available during supplementary exam period.

If you receive special consideration for the final exam, a supplementary exam will be scheduled in the interval between the regular exam period and the start of the next session. By making a special consideration application for the final exam you are declaring yourself available for a resit during the supplementary examination period and will not be eligible for a second special consideration approval based on pre-existing commitments. Please ensure you are familiar with the policy prior to submitting an application. You can check this supplementary exam information page for dates, and approved applicants will receive an individual notification one week prior to the exam with the exact date and time of their supplementary examination.

If you are given a second opportunity to sit the final examination as a result of failing to meet the minimum mark required in a hurdle assessment, you will be offered that chance during the same supplementary examination period and will be notified of the exact day and time after the publication of final results for the unit.

Supplementary exams for Session 2, 2018 will be held in the week of December 17-21, 2018.

On successful completion you will be able to:

- Understand the fundamental difference between univariate and multivariate analysis.
- Know how to perform hypothesis testing (mainly the Hotelling T2 test) using multivariate data.
- Understand and be able to apply MANOVA.
- Understand multivariate regression.

- Know the theories of PCA and factor analysis, and be able to apply these methods to real data.
- Understand likelihood based, as well as minimum expected cost based, discriminant analysis. Be able to apply these discriminant analysis methods to real data.
- Be able to use the R package for multivariate data analysis.

Delivery and Resources

Classes

You are required to attend a 3-hour lecture each week; the time and room are:

Wednesday	6pm – 9pm	6 Eastern Rd (E4B), 208 Faculty PC Lab

Technologies used and required

We primarily use the software package R in this Unit. R is becoming increasingly important for statisticians and other scientists. More information about R can be found at the web site http://www.r-project.org/ and the package can be downloaded **free of charge** from there. R is very similar to the package S-PLUS and most of its codes will also work in S-Plus. From week 2, students will be given exercises each week covering materials from the lectures, and most exercises require using R.

Recommended texts

Prescribed textbook: "**Applied Multivariate Statistical Analysis**" by Richard A. Johnson, Dean W. Wichern (6th edition)

Students are expected to possess a copy of this textbook and are required to read certain book chapters each week. The following books may be used as other references for this unit:

DILLON & GOLDSTEIN	Multivariate Analysis – Methods and applications(QA 278 .d55)
FAHRMEIR & TUTZ (QA 278 .F34)	Multivariate statistical modelling based on generalized linear models
FLURY, B	A first course in multivariate statistics
FLURY, B	Multivariate statistics: A practical approach
MORRISON, D	Multivariate statistical methods

Unit Schedule

The following is a detailed list of the topics covered in this Unit, together with the planned delivery time. All lecture notes will be available on iLearn prior to the lecture.

Week	Topic
1	1. Introduction to multivariate analysis 2. Overview of matrix algebra
2	1. Matrix algebra (cont.) 2. Basic concepts of multivariate distributions 3. Sample statistics
3	1. Multivariate sample statistics (cont.) 2. Some useful multivariate distributions
4.	1. Initial data analysis 2. Inferences: Estimation and hypothesis testing
5.	1. Inferences (cont.)
6.	2. MANOVA
7.	1. MANOVA (cont.) 2. Multivariate regression
8.	1. Regression (cont.) 2. Principal component analysis (PCA)
9.	1. Factor analysis (FA)
11.	1. Factor analysis (cont.) 2. Discriminant analysis and classification
12.	1. Discriminant analysis (cont.)
13.	1. Brief introduction to canonical correlation analysis 2. Brief introduction to cluster analysis

Policies and Procedures

Macquarie University policies and procedures are accessible from Policy Central (https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/policy-central). Students should be aware of the following policies in particular with regard to Learning and Teaching:

- Academic Appeals Policy
- Academic Integrity Policy
- Academic Progression Policy
- Assessment Policy
- Fitness to Practice Procedure
- Grade Appeal Policy
- Complaint Management Procedure for Students and Members of the Public
- Special Consideration Policy *(**Note:** The Special Consideration Policy is effective from 4 December 2017 and replaces the Disruption to Studies Policy.)*

Undergraduate students seeking more policy resources can visit the Student Policy Gateway (https://students.mq.edu.au/support/study/student-policy-gateway). It is your one-stop-shop for the key policies you need to know about throughout your undergraduate student journey.

If you would like to see all the policies relevant to Learning and Teaching visit Policy Central (https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/policy-central).

Student Code of Conduct

Macquarie University students have a responsibility to be familiar with the Student Code of Conduct: https://students.mq.edu.au/study/getting-started/student-conduct

Results

Results shown in *iLearn*, or released directly by your Unit Convenor, are not confirmed as they are subject to final approval by the University. Once approved, final results will be sent to your student email address and will be made available in eStudent. For more information visit ask.mq.edu.au.

Student Support

Macquarie University provides a range of support services for students. For details, visit http://students.mq.edu.au/support/

Learning Skills

Learning Skills (mq.edu.au/learningskills) provides academic writing resources and study strategies to improve your marks and take control of your study.

- Workshops

- StudyWise
- Academic Integrity Module for Students
- Ask a Learning Adviser

Student Services and Support

Students with a disability are encouraged to contact the Disability Service who can provide appropriate help with any issues that arise during their studies.

Student Enquiries

For all student enquiries, visit Student Connect at ask.mq.edu.au

IT Help

For help with University computer systems and technology, visit http://www.mq.edu.au/about_us/offices_and_units/information_technology/help/.

When using the University's IT, you must adhere to the Acceptable Use of IT Resources Policy. The policy applies to all who connect to the MQ network including students.

Graduate Capabilities

PG - Discipline Knowledge and Skills

Our postgraduates will be able to demonstrate a significantly enhanced depth and breadth of knowledge, scholarly understanding, and specific subject content knowledge in their chosen fields.

This graduate capability is supported by:

Learning outcomes

- Understand the fundamental difference between univariate and multivariate analysis.
- Know how to perform hypothesis testing (mainly the Hotelling T2 test) using multivariate data.
- Understand and be able to apply MANOVA.
- Understand multivariate regression.
- Know the theories of PCA and factor analysis, and be able to apply these methods to real data.
- Understand likelihood based, as well as minimum expected cost based, discriminant analysis. Be able to apply these discriminant analysis methods to real data.
- Understand the principles of classification tree and canonical analysis.
- Know how to display multivariate data graphically using R.
- Be able to use the R package for multivariate data analysis.

Assessment tasks

- Assignment 1
- Assignment 2
- Takehome examination
- Written Examination

PG - Critical, Analytical and Integrative Thinking

Our postgraduates will be capable of utilising and reflecting on prior knowledge and experience, of applying higher level critical thinking skills, and of integrating and synthesising learning and knowledge from a range of sources and environments. A characteristic of this form of thinking is the generation of new, professionally oriented knowledge through personal or group-based critique of practice and theory.

This graduate capability is supported by:

Learning outcomes

- Understand the fundamental difference between univariate and multivariate analysis.
- Know how to perform hypothesis testing (mainly the Hotelling T2 test) using multivariate data.
- Understand and be able to apply MANOVA.
- Understand multivariate regression.
- Know the theories of PCA and factor analysis, and be able to apply these methods to real data.
- Understand likelihood based, as well as minimum expected cost based, discriminant analysis. Be able to apply these discriminant analysis methods to real data.
- Understand the principles of classification tree and canonical analysis.
- Know how to display multivariate data graphically using R.

Assessment tasks

- Assignment 1
- Assignment 2
- Takehome examination
- Written Examination

PG - Research and Problem Solving Capability

Our postgraduates will be capable of systematic enquiry; able to use research skills to create new knowledge that can be applied to real world issues, or contribute to a field of study or practice to enhance society. They will be capable of creative questioning, problem finding and problem solving.

This graduate capability is supported by:

Learning outcomes

- Understand the fundamental difference between univariate and multivariate analysis.
- Know how to perform hypothesis testing (mainly the Hotelling T2 test) using multivariate data.
- Understand and be able to apply MANOVA.
- Understand multivariate regression.
- Know the theories of PCA and factor analysis, and be able to apply these methods to real data.
- Understand likelihood based, as well as minimum expected cost based, discriminant analysis. Be able to apply these discriminant analysis methods to real data.
- Understand the principles of classification tree and canonical analysis.
- Know how to display multivariate data graphically using R.
- Be able to use the R package for multivariate data analysis.

Assessment tasks

- Assignment 1
- Assignment 2
- Takehome examination

PG - Effective Communication

Our postgraduates will be able to communicate effectively and convey their views to different social, cultural, and professional audiences. They will be able to use a variety of technologically supported media to communicate with empathy using a range of written, spoken or visual formats.

This graduate capability is supported by:

Learning outcomes

- Know how to perform hypothesis testing (mainly the Hotelling T2 test) using multivariate data.
- Understand and be able to apply MANOVA.
- Understand multivariate regression.
- Know the theories of PCA and factor analysis, and be able to apply these methods to real data.
- Know how to display multivariate data graphically using R.
- Be able to use the R package for multivariate data analysis.

Assessment tasks

- Assignment 1
- Assignment 2
- Takehome examination
- Written Examination