



STAT821

Multivariate Analysis

S2 Evening 2014

Statistics

Contents

<u>General Information</u>	2
<u>Learning Outcomes</u>	2
<u>Assessment Tasks</u>	3
<u>Delivery and Resources</u>	6
<u>Unit Schedule</u>	7
<u>Policies and Procedures</u>	8
<u>Graduate Capabilities</u>	9

Disclaimer

Macquarie University has taken all reasonable measures to ensure the information in this publication is accurate and up-to-date. However, the information may change or become out-dated as a result of change in University policies, procedures or rules. The University reserves the right to make changes to any information in this publication without notice. Users of this publication are advised to check the website version of this publication [or the relevant faculty or department] before acting on any information in this publication.

General Information

Unit convenor and teaching staff

Unit Convenor

Georgy Sofronov

georgy.sofronov@mq.edu.au

Contact via email

E4A536

Friday, 11am-1pm

Credit points

4

Prerequisites

Admission to MAppStat or PGDipAppStat or PGCertAppStat

Corequisites

STAT806 or STAT810

Co-badged status

STAT721

Unit description

This 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, and 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) involving 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. Also, students are 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. Use the R package, particularly its functions in multivariate data analysis.

Assessment Tasks

Name	Weighting	Due
<u>Assignment 1</u>	10%	4 September 2014
<u>Assignment 2</u>	15%	9 October 2014
<u>Final Examination</u>	60%	University Examination Period
<u>Assignment 3</u>	15%	30 October 2014

Assignment 1

Due: **4 September 2014**

Weighting: **10%**

For internal students, marked assignments will be available for collection approximately two weeks after the due date. For external students, marked assignments will be sent back by email or post.

Assignments comprise a major part of the learning process. They give each student the opportunity to review his/her understanding of the course material. Assignments are compulsory. Failure to submit any assignment will be taken as an evidence of non-participation in the course and may lead to exclusion from the course. If students need extension on their assignments they must contact the lecturer in writing and provide necessary supporting documents. No extensions will be considered unless satisfactory documentation outlining illness or misadventure is submitted. Assignments must be each student's own work. Discussions are allowed but the final work must be your personal effort.

It is preferable that the assignments are word-processed. If any student wishes to hand in the assignments electronically via email, MS word (or other) documents should be converted to pdf format.

In answering assignment questions, a careful discussion of the problem and its solution is required. Any computer output used should be extracted, clearly headed and placed in the assignment or in an Appendix of the assignment.

The first assignment will be available on the unit webpage in week 3.

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) involving multivariate data.
- Know how to display multivariate data graphically using R. Use the R package, particularly its functions in multivariate data analysis.

Assignment 2

Due: **9 October 2014**

Weighting: **15%**

The second assignment will be available on the unit webpage in week 6.

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) involving 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. Use the R package, particularly its functions in multivariate data analysis.

Final Examination

Due: **University Examination Period**

Weighting: **60%**

The duration of the final examination is three hours plus ten minutes' reading time. An electronic calculator and one A4 sheet of paper (written on one or both sides) may be taken in to the exam room. All material thereon must be in the student's own handwriting and not typed.

Scanned copies are not permitted.

For a passing grade, satisfactory performance is required on both: (i) the average of the assignments; (ii) the final examination. You are expected to present yourself for examination at the time and place designated in the University examination timetable, which will be available at <https://timetables.mq.edu.au>.

Only documented illness or unavoidable disruption may be used as reasons for not sitting an examination at the designated time. In these circumstances you may wish to consider applying for Special Consideration. Information about the special consideration policy and procedure is available at: http://students.mq.edu.au/student_admin/exams/disruption_to_studies/

It is Macquarie University policy not to set early examinations for individuals or groups of students. All students are expected to ensure that they are available until the end of the teaching semester, that is, the final day of the official examination period.

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) involving 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. Also, students are 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. Use the R package, particularly its functions in multivariate data analysis.

Assignment 3

Due: **30 October 2014**

Weighting: **15%**

The third assignment will be available on the unit webpage in week 9.

On successful completion you will be able to:

- Understand the fundamental difference between univariate and multivariate analysis.
- 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. Also, students are 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. Use the R package, particularly its functions in multivariate data analysis.

Delivery and Resources

Teaching staff

Dr Georgy Sofronov

Room: E4A536

Telephone: (02) 9850-8544

Fax: (02) 9850-7669

Email: georgy.sofronov@mq.edu.au

Classes

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

Thursday 6pm – 9pm E4B118

Technologies used and required

We primarily use the software package R in this Unit. R is becoming increasingly important for statisticians. 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 similar to the package S-PLUS and most of its codes 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 Co-op Bookshop has some copies for purchase and the library has some copies for borrowing. The following books may be also useful:

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

What has changed?

The format of the final examination has been changed. Otherwise this offer is similar to the last year's offering.

Unit Schedule

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

Week	Topic
1	<ol style="list-style-type: none">1. Introduction to multivariate analysis2. Overview of matrix algebra
2	<ol style="list-style-type: none">1. Matrix algebra (cont.)2. Basic concepts of multivariate distributions3. Sample statistics
3	<ol style="list-style-type: none">1. Sample statistics (cont.)2. Some useful multivariate distributions
4.	<ol style="list-style-type: none">1. Initial data analysis2. Inferences: Estimation and hypothesis testing
5.	<ol style="list-style-type: none">1. Inferences (cont.)
6.	<ol style="list-style-type: none">1. MANOVA

7.	<ol style="list-style-type: none"> 1. MANOVA (cont.) 2. Multivariate regression
8.	<ol style="list-style-type: none"> 1. Regression (cont.) 2. Principal component analysis (PCA)
9.	<ol style="list-style-type: none"> 1. Factor analysis (FA)
11.	<ol style="list-style-type: none"> 1. Factor analysis (cont.) 2. Discriminant analysis and classification
12.	<ol style="list-style-type: none"> 1. Discriminant analysis (cont.)
13.	<ol style="list-style-type: none"> 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](#). Students should be aware of the following policies in particular with regard to Learning and Teaching:

Academic Honesty Policy http://mq.edu.au/policy/docs/academic_honesty/policy.html

Assessment Policy <http://mq.edu.au/policy/docs/assessment/policy.html>

Grading Policy <http://mq.edu.au/policy/docs/grading/policy.html>

Grade Appeal Policy <http://mq.edu.au/policy/docs/gradeappeal/policy.html>

Grievance Management Policy http://mq.edu.au/policy/docs/grievance_management/policy.html

Disruption to Studies Policy http://www.mq.edu.au/policy/docs/disruption_studies/policy.html *The Disruption to Studies Policy is effective from March 3 2014 and replaces the Special Consideration Policy.*

In addition, a number of other policies can be found in the [Learning and Teaching Category](#) of 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/support/student_conduct/

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://informatics.mq.edu.au/help/>.

When using the University's IT, you must adhere to the [Acceptable Use 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) involving 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. Also, students are 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. Use the R package, particularly its functions in multivariate data analysis.

Assessment tasks

- Assignment 1
- Assignment 2
- Final Examination
- Assignment 3

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) involving 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. Also, students are 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. Use the R package, particularly its functions in multivariate data analysis.

Assessment tasks

- Assignment 1
- Assignment 2
- Final Examination
- Assignment 3

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) involving 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. Also, students are 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. Use the R package, particularly its functions in multivariate data analysis.

Assessment tasks

- Assignment 1
- Assignment 2
- Final Examination
- Assignment 3

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

- Understand the fundamental difference between univariate and multivariate analysis.
- Know how to perform hypothesis testing (mainly the Hotelling T2 test) involving 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. Also, students are 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. Use the R package, particularly its functions in multivariate data analysis.

Assessment tasks

- Assignment 1
- Assignment 2
- Final Examination
- Assignment 3

PG - Engaged and Responsible, Active and Ethical Citizens

Our postgraduates will be ethically aware and capable of confident transformative action in relation to their professional responsibilities and the wider community. They will have a sense of connectedness with others and country and have a sense of mutual obligation. They will be able to appreciate the impact of their professional roles for social justice and inclusion related to national and global issues

This graduate capability is supported by:

Assessment tasks

- Assignment 1
- Assignment 2
- Final Examination
- Assignment 3

PG - Capable of Professional and Personal Judgment and

Initiative

Our postgraduates will demonstrate a high standard of discernment and common sense in their professional and personal judgment. They will have the ability to make informed choices and decisions that reflect both the nature of their professional work and their personal perspectives.

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) involving 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. Also, students are 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. Use the R package, particularly its functions in multivariate data analysis.

Assessment tasks

- Assignment 1
- Assignment 2
- Final Examination
- Assignment 3