



MACQUARIE UNIVERSITY

FACULTY OF SCIENCE

**MACQUARIE UNIVERSITY
FACULTY OF SCIENCE
DEPARTMENT OF STATISTICS
STAT821: MULTIVARIATE ANALYSIS
UNIT OUTLINE**

Year and Semester: Second semester, 2012
Convenor/Lecturer-in-charge: Associate Professor Jun Ma
Corequisites: STAT810

Students in this unit should read this unit outline carefully at the start of the semester. It contains important information about the unit. If anything in it is unclear, please consult the teaching staff in the unit.

ABOUT THIS UNIT

- **Unit value:** Four (4) credit points.
- **Brief introduction:** 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. Multivariate statistical techniques constitute an important part of modern statistical analysis.
- **Unit rationale:** This unit offers student both practical skills and theoretical understanding on multivariate data analysis. Students will find that the skills they learned in this unit useful for their future career development.

TEACHING STAFF

Associate Professor Jun Ma
Room: E4A511
Telephone: (02) 9850-8548
Fax: (02) 9850-7669
Email: jun.ma@mq.edu.au

CLASSES

Lectures

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

Monday 6.00pm – 9.00pm E4B 102

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	Multivariate statistical modelling based on generalized linear models (QA 278 .F34)
FLURY, B	A first course in multivariate statistics
FLURY, B	Multivariate statistics: A practical approach
MORRISON, D	Multivariate statistical methods

UNIT WEB PAGE

The unit webpage is located on iLearn at <http://ilearn.mq.edu.au>. You can only access the material on iLearn if you are enrolled in this unit. All lecturing materials are available at this webpage.

GRADUATE CAPABILITIES

In addition to discipline specific learning outcomes, all academic programs at Macquarie assist students to develop or enhance capabilities in a range of areas. The Macquarie Graduate Capabilities are listed below:

1. Discipline specific knowledge and skills
2. Critical, analytical and integrative thinking
3. Problem solving and research capability
4. Effective communication
5. Socially and environmentally active and responsible
6. Creative and innovative

7. Commitment to continuous learning
8. Capable of professional and personal judgement and initiative
9. Engaged and ethical local and global citizens

The relationship between Graduate Capabilities and student learning outcomes are shown under Learning Outcomes as a reference to Graduate Capabilities above.

LEARNING OUTCOMES

The aim of this unit is to gain familiarity with modern concepts and procedures of multivariate statistical analysis. The unit will begin with reviewing basic matrix algebra and multivariate distributions as they are required throughout this course. Afterwards, key methods in multivariate analysis will be introduced and studied, including multivariate linear models, multivariate analysis of variance (MANOVA), discriminant analysis, principal components analysis and factor analysis.

By the end of this unit students should be able to:

- Understand the fundamental difference between univariate and multivariate analysis.
- Know how to display multivariate data graphically using R.
- Know how to perform hypothesis testing (mainly the Hotelling T^2 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.
- Use the R package, particularly its functions in multivariate data analysis.

In addition to the discipline-based learning objectives, all academic programs at Macquarie seek to develop students' generic skills in a range of areas. One of the aims of this unit is that students develop their skills in the following:

- *Foundation skills of literacy, numeracy and information technology;*
- *Communication skills;*
- *Critical analysis skills;*
- *Problem-solving skills;*
- *Creative thinking skills.*

SOFTWARE USED IN TEACHING

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.

RELATIONSHIP BETWEEN ASSESSMENT AND LEARNING OUTCOMES

While attendance at classes is important, for internally enrolled students, it is only a small proportion of the total workload for the unit: reading, completing assignments and other assessments, using the computer and private study are all part of the work involved. At Macquarie it is expected that on average students should spend four hours per week per credit point including attending lectures and tutorials.

The assessment is based on the performance in the assignments and exam:

Assignment 1	10%
Assignment 2	15%
Assignment 3	15%
Take home exam	60%

Satisfactory performance (i.e. at least 50%) is required in each of these assessment aspects.

Assignments

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 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. Late submission without extension approval will be penalized at the rate of 20% deduction per day. Assignments must be each student's own work. Discussions are allowed but the final work must be your personal effort.

We prefer that 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.

There will be **three** assignments for this unit. They will be available on the unit webpage in, approximately, weeks 3, 6 and 10. Students are expected to hand in assignments on the due dates.

Early Assessment

We will use assignment 1 as an early assessment.

Final Examination

The examination will be designed to test your knowledge and understanding of the materials discussed in this unit. The final exam will be a **take-home** exam and students will have **THREE** days to complete their exam paper. The designated exam date for this unit is from 10.00 am on Friday Nov 9 2012 to 10.00 am on Monday Nov 12 2012.

There is a University policy regarding requests for Special Consideration for examinations and the granting of supplementary examinations, which can be found at:
http://www.mq.edu.au/policy/docs/special_consideration/policy.html.

The University Examination period in Second Half Year 2011 is from 14th Nov to 2nd Dec.

The only exception to not sitting an examination at the designated time is because of documented illness or unavoidable disruption. Information about unavoidable disruption and the special consideration process is available at <http://www.reg.mq.edu.au/Forms/APSCon.pdf>. For special consideration, the Faculty of Science web page for Special Consideration Applications is http://web.science.mq.edu.au/undergraduate_programs/current/admin_central/?, and the special consideration policy can be found at http://www.mq.edu.au/policy/docs/special_consideration/policy.html.

If a Supplementary Examination is granted as a result of the Special Consideration process the examination will be scheduled after the conclusion of the official examination period.

You are advised that 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.

ACADEMIC HONESTY

You should familiarise yourself with University policy on **academic honesty**:
http://www.mq.edu.au/policy/docs/academic_honesty/policy.html .

Academic honesty is an integral part of the core values and principles contained in the Macquarie University Ethics Statement: <http://www.mq.edu.au/ethics/ethic-statement-final.html> . Its fundamental principle is that all staff and students act with integrity in the creation, development, application and use of ideas and information.

ASSESSMENT POLICY

Your final grade in STAT821 will be based on your work during semester and in the final examination as specified above. The grades allocated are as set out in the Grading Policy (<http://www.mq.edu.au/policy/docs/grading/policy.html>) as follows:

High Distinction (HD): Provides consistent evidence of deep and critical understanding in relation to the learning outcomes. There is substantial originality and insight in identifying, generating and communicating competing arguments, perspectives or problem solving approaches; critical evaluation of problems, their solutions and their implications; creativity in application as appropriate to the discipline.

Distinction (D): Provides evidence of integration and evaluation of critical ideas, principles and theories, distinctive insight and ability in applying relevant skills and concepts in relation to learning outcomes. There is demonstration of frequent originality in defining and analysing issues or problems and providing solutions; and the use of means of communication appropriate to the discipline and the audience.

Credit (Cr): Provides evidence of learning that goes beyond replication of content knowledge or skills relevant to the learning outcomes. There is demonstration of substantial understanding of fundamental concepts in the field of study and the ability to apply these concepts in a variety of contexts; convincing argumentation with appropriate coherent justification; communication of ideas fluently and clearly in terms of the conventions of the discipline.

Pass (P): Provides sufficient evidence of the achievement of learning outcomes. There is demonstration of understanding and application of fundamental concepts of the field of study; routine argumentation with acceptable justification; communication of information and ideas adequately in terms of the conventions of the discipline. The learning attainment is considered satisfactory or adequate or competent or capable in relation to the specified outcomes.

Fail (F): Does not provide evidence of attainment of learning outcomes. There is missing or partial or superficial or faulty understanding and application of the fundamental concepts in the field of study; missing, undeveloped, inappropriate or confusing argumentation; incomplete, confusing or lacking communication of ideas in ways that give little attention to the conventions of the discipline.

Please note that a student must meet the performance standard outlined above in both the coursework and the examination sections of this unit in order to be awarded a particular grade. Your final result will include one of these grades plus a standardised numerical grade (SNG).

STUDENT SUPPORT SERVICES

Macquarie University provides a range of Academic Student Support Services. Details of these services can be accessed at <http://www.student.mq.edu.au>.

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 Blackboard 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. Sample statistics (cont.) 2. Some useful multivariate distributions
4.	1. Initial data analysis 2. Inferences: Estimation and hypothesis testing
5.	1. Inferences (cont.)
6.	1. 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