Faculty of Science

School of Mathematics and Statistics

ZZSC5855

Multivariate Analysis for Data Scientists

Hexamester 6 2022

Course Outline

Description

This course will give you a solid methodological background in multivariate analysis as a backbone of applied statistics. You will learn the theoretical foundations of the most commonly applied multivariate techniques such as mean vector and covariance matrix estimation and testing, estimation and testing of correlations, multivariate linear models, discriminant analysis, classification and support vector machines, principal components, canonical correlations analysis, cluster analysis, factor analysis and structural equations. You will study the properties and the importance of the multivariate normality assumption in thecontext of each of these methods. R-based computing will feature prominently in the course. At the end ofthe course, you should be able to use all the above techniques in your work as an applied statistician, for practical analysis of real datasets.

Pre-requisites and Assumed Knowledge

It is assumed that you would have successfully completed 'ZZSC5905 Statistical Inference' and 'ZZSC9001 Foundations of Data Science' or equivalent before enrolling in this course. Additionally, it is strongly recommended you have successfully completed 'ZZSC5806 Regression Analysis for Data Scientists'.

Units of Credit

The course is worth 6 units of credit.

Learning Outcomes

On successful completion of the course, you should be able to:

1. Use the general terminology, notation and concepts in the theory, methods and applications of multivariate analysis.
2. Use the properties of the multivariate normal distribution to justify statistical inference proceduresbased on multivariate normality.
3. Formulate and solve inference problems that use one or a combination of multivariate statisticalprocedures.
4. Perform multivariate inference procedures by using the R suite.
5. Write simple R programs for data input and output and implement simple multivariate analyses by using R.
6. Apply the multivariate techniques and models to the analysis of datasets, interpret the results and draw conclusions.

Schedule

Date	Week	Topic	Assessment Due
17 Oct	0	Orientation Week	
24 Oct	1	Foundations of Multivariate Statistics	
31 Oct	2	Exploring Correlations	Assessment 1: Quiz
7 Nov	3	Testing Multivariate Relationships	Assessment 2: Quiz

ZZSC5855 Multivariate Analysis for Data Scientists, Hexamester 6 2022

UNSW
SYDNEY

14 Nov	4	Analysing the Covariance Matrix	Assessment 3: Quiz
21 Nov	5	Multivariate analysis for categorical outcomes	Assessment 4: Quiz
28 Nov	6	Clustering and Copula methods	Assessment 5: Quiz
5 Dec	7	Examination Week	Assessment 6: Quiz Assessment 7: Project

Assessment

In order to pass the course, you must:

- achieve a total mark of at least 50;
- meet any additional requirements of the assessment

tasks. The assessment tasks are:

No.	Type	Weight	Due*	Learning Outcomes
1	Quiz	10%	Week 2, Monday, 11:00 pm (31 October)	1,2,3,4,5,6
2	Quiz	10%	Week 3, Monday, 11:00 pm (7 November)	1,2,3,4,5,6
3	Quiz	10%	Week 4, Monday, 11:00 pm (14 November)	1,2,3,4,5,6
4	Quiz	10%	Week 5, Monday, 11:00 pm (21 November)	1,2,3,4,5,6
5	Quiz	10%	Week 6, Monday, 11:00 pm (28 November)	1,2,3,4,5,6
6	Quiz	10%	Week 7, Monday, 11:00 pm (5 December)	1,2,3,4,5,6
7	Project	40%	Week 7, Monday, 11:00 pm (5 December)	1,2,3,4,5,6

*All dates and times are Sydney (Australia) dates and times

If you submit after the due date, you will be penalised 5% per day capped at five days (120 hours) from the assessment deadline, after which you cannot submit an assessment unless you have an approved extension through Special Consideration or Equitable Learning Services.

Teaching Staff

Your teaching staff are:

- Pavel Krivitsky, p.krivitsky@unsw.edu.au (Online Lecturer and Course Convenor)

Teachers are your main point of contact. Their consultation times will be advised on the course website.

Resources
The website for the course is on Moodle, at: http://moodle.telt.unsw.edu.au.

Technical Requirements
The course is fully online. You will need:

3

UNSW SYDNEY

- A fast and reliable computer (or equivalent device), with an up-to-date operating system
- A fast and reliable internet connection
- The latest version of a modern browser (e.g., Edge, Chrome, Firefox or Safari)
- A reliable way to store your files - either on your computer with a backup routine, or in the cloud (e.g., using Dropbox)
- There is no specific software requirement as all activities can be done in Ed, though you might prefer to work locally, in which case you would need to install the latest version of R and RStudio.

Learning and Teaching Activities

The course contains a variety of resources and activities that are carefully designed to enhance your learning.

Some activities require you to work and think alone, by reading some text, listening to a recording or watching a video. You might be asked to engage with the material and explore interactive elements by clicking to reveal content, to help you better absorb and process the concepts. Some activities require you to produce work of your own. You might be answering a question, writing code to solve a problem, or posting to a forum, for example. Some activities are assessment tasks, which have been carefully designed to measure how well you have achieved the learning outcomes of the course. Typically, you will get feedback on your work, either from yourself (by checking your work with models that are provided), or from an automatic marking process, or from your peers, or from your teacher.

You also have access to a variety of ways to communicate with your peers and with the teaching staff. The general discussion forums are a place for you to ask and answer questions, to interact with your peers, and to be challenged by your teachers. Getting involved in these forums will enhance your learning experience and make it more enjoyable. Your course may include Webinars, which provide an opportunity to hear directly from your Online Lecturers and ask questions in real time. All webinars are recorded so you can access them at any time. Online Lecturers are available for consultations and will post information about how to access consultations on the course website. You can also contact your Online Lecturer by email using the email address in the teaching staff section of this outline.

It is up to you how much work you do. The more time and effort that you can dedicate to the course, thebetter will be your learning and your results.

Special Consideration

If illness or other circumstances beyond your control interfere with your assessment performance then you can apply for special consideration, to get an extra opportunity to demonstrate your level of performance.

You must make your application online, through the [Special Consideration portal on myUNSW](#). Do not apply to your course teaching staff — they will be notified automatically.

You must apply before the assessment task is due or the exam is held - if you submit the assessment or sit the exam then you are declaring yourself well enough to do so and are unable to subsequently apply for special consideration. If illness or misadventure prevent you from applying in advance, then you must apply as soon as possible, and provide evidence that you could not apply sooner. If you become unwell on the day of the exam, you must provide evidence dated within 24 hours of the exam, with your application.

Your application will be considered centrally, by a case review team, and they will notify you of the outcome.

If your application is successful, then an alternative means of fulfilling the assessment requirements of the course will be provided.

You can read more about special consideration at: https://student.unsw.edu.au/special-consideration.

Academic Integrity

UNSW values academic integrity and has strict rules against cheating. In particular, it has strict rules against trying to get credit that you don't deserve. Thus, you should not plagiarise — i.e., present someone else's work as if it's your own. This could be the work of an academic, or a peer, or a contract writer, and it includes work of all kinds - exact words, general ideas, designs, drawings, software, and so on. Nor should you recycle your own work — i.e., submit it for credit in multiple courses. UNSW also has strict rules against helping others to cheat — e.g., by giving someone your work to copy, or doing someone's work for them, and so on.

For further information about academic integrity and plagiarism at UNSW go to:https://student.unsw.edu.au/plagiarism

For information about acknowledging your sources and referencing go to: https://student.unsw.edu.au/referencing. If you are not sure what referencing style to use in this course, youshould ask your Online Lecturer.

Evaluation and Development

Toward the end of the hexamester you will be asked to give feedback about the course, via UNSW's MyExperience survey. Your feedback will be used, along with feedback from other stakeholders, to help improve the course. You can also contact your Course Convenor any time you have suggestions or other feedback.

Quality Assurance

UNSW actively monitors student learning and quality of the student experience in its programs. A random selection of completed assessment tasks may be used for quality assurance, such as determining the extent to which program learning goals are being achieved. The information is required for accreditation purposes, and aggregated findings will be used to inform changes aimed at improving the quality of programs. All material used for such processes will be treated as confidential.