FIT5197
Modelling for data analysis

Unit Guide

Monash Online Teaching Period 6, 2015

Copyright © Monash University 2014. All rights reserved. Except as provided in the Copyright Act 1968, this work may not be reproduced in any form without the written permission of the host Faculty and School/Department.

The information contained in this unit guide is correct at time of publication. The University has the right to change any of the elements contained in this document at any time.

Last updated: 13 Oct 2015
# Table of Contents

**FIT5197 Modelling for data analysis - Monash Online Teaching Period 6, 2015**

- Mode of Delivery ................................................................. 1
- Workload Requirements ............................................................ 1
- Unit Relationships .................................................................. 1
  - Prerequisites ...................................................................... 1
- Chief Examiner ...................................................................... 2
- Campus Lecturer .................................................................... 2
  - Clayton ............................................................................. 2
- Your feedback to Us .............................................................. 2

## Academic Overview ................................................................. 3
- Learning Outcomes ............................................................... 3

## Unit Schedule ............................................................................ 4
- Teaching Approach ............................................................... 4
- Assessment Summary ............................................................ 4

## Assessment Requirements ....................................................... 6
- Assessment Policy ............................................................... 6
- Assessment Tasks .................................................................. 6
  - Participation ...................................................................... 6
- Learning resources ............................................................... 9
- Reading list .......................................................................... 9
- Feedback to you .................................................................. 9
- Extensions and penalties ..................................................... 9
- Returning assignments ......................................................... 9
- Referencing requirements .................................................... 9
- Assignment submission ....................................................... 9
- Online submission .............................................................. 10
- Required Resources ............................................................ 10
  - Prescribed text(s) ............................................................... 10
- Technological Requirements ................................................ 10
- Recommended Resources .................................................... 10

## Other Information ..................................................................... 11
- Policies .............................................................................. 11
  - Faculty resources and policies ........................................... 11
    - Graduate Attributes Policy ............................................. 11
- Student Charter ..................................................................... 11
- Student services .................................................................... 11
- Monash University Library .................................................... 11
- Disability Liaison Unit .......................................................... 11
FIT5197 Modelling for data analysis - Monash Online Teaching
Period 6, 2015

This unit explores the statistical modelling foundations that underlie the analytic aspects of Data Science. Motivated by case studies and working through real examples, this unit covers the mathematical and statistical basis with an emphasis on using the techniques in practice. It introduces data collection, sampling and quality. It considers analytic tasks such as statistical hypothesis testing and exploratory versus confirmatory analysis. It presents basic probability distributions, random number generation and simulation as well as estimation methods and effects such as maximum likelihood estimators, Monte Carlo estimators, Bayes theorem, bias versus variance and cross validation. Basic information theory and dependence models such as Bayesian networks and log-linear models are also presented, as well as the role of general modelling such as inference and decision making, predictive models, experts and assessing probabilities.

Mode of Delivery

Monash Online (Online)

Workload Requirements

Minimum total expected workload equals 144 hours per semester comprising:

1. Contact hours for on-campus students:
   - Two hours/week lectures
   - Two hours/week laboratories

Contact hours for Monash Online students:

- Two hours/week online group sessions.
- Online students generally do not attend lecture, tutorial and laboratory sessions, however should plan to spend equivalent time working through resources and participating in discussions.

Additional requirements (all students):

- A minimum of 8 hours per week of personal study (22 hours per week for Monash online students) for completing lab/tutorial activities, assignments, private study and revision, and for online students, participating in discussions.

See also Unit timetable information

Unit Relationships

Prerequisites

Students need to have the equivalent of first year undergraduate university mathematics as taught in an analytics degree such as Engineering, Finance, Physics and some Computer Science degrees.
Your feedback to Us

Monash is committed to excellence in education and regularly seeks feedback from students, employers and staff. One of the key formal ways students have to provide feedback is through the Student Evaluation of Teaching and Units (SETU) survey. The University’s student evaluation policy requires that every unit is evaluated each year. Students are strongly encouraged to complete the surveys. The feedback is anonymous and provides the Faculty with evidence of aspects that students are satisfied and areas for improvement.

For more information on Monash's educational strategy, see:

www.monash.edu.au/about/monash-directions/ and on student evaluations, see:
www.policy.monash.edu/policy-bank/academic/education/quality/student-evaluation-policy.html
Academic Overview

Learning Outcomes

Upon successful completion of this unit, it is expected that students will be able to:

1. perform the general roles of exploratory, confirmatory and decision analysis applied to data;
2. explain how the source of data affects analysis;
3. summarise the role of domain experts in supporting analysis and the difficulties they may have;
4. implement a computational model for statistical analysis of simple problems and evaluate the results;
5. conduct statistical analysis using the concepts of entropy, likelihood, correlation, and independence;
6. interpret the challenges involved in estimation from data;
7. describe basic methods of random sampling, simulation, and hypothesis testing;
8. write basic programs for analysing data.
## Unit Schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Activities</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No formal assessment or activities are undertaken in week 0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Module one - Introduction to Modelling for Data Science</td>
<td>Programming Assignment 1</td>
</tr>
<tr>
<td>2</td>
<td>Module two - Data analysis</td>
<td>Programming Assignment 2</td>
</tr>
<tr>
<td>3</td>
<td>Module three - Dependence, regression and clustering</td>
<td>Programming Assignment 3; Quiz</td>
</tr>
<tr>
<td>4</td>
<td>Module four - Statistical inference and evaluation</td>
<td>Programming Assignment 4</td>
</tr>
<tr>
<td>5</td>
<td>Module five - Simulation</td>
<td>Programming Assignment 5</td>
</tr>
<tr>
<td>6</td>
<td>Module six - Modelling and validation</td>
<td>Programming Assignment 6; Quiz; Report</td>
</tr>
<tr>
<td>7</td>
<td>No formal assessment is undertaken during SWOT VAC</td>
<td></td>
</tr>
</tbody>
</table>

*Unit Schedule details will be maintained and communicated to you via your learning system.*

## Teaching Approach

### Other

This subject is learned online, using Moodle and Alexandria as the primary methods of delivering information and guidance, drawing upon text readings, videos, exercises and activities, as well as online contact with instructors.

## Assessment Summary

In-semester assessment: 100%

<table>
<thead>
<tr>
<th>Assessment Task</th>
<th>Value</th>
<th>Due Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming Assignment 1</td>
<td>5%</td>
<td>1/11/15</td>
</tr>
<tr>
<td>Programming Assignment 2</td>
<td>5%</td>
<td>8/11/15</td>
</tr>
<tr>
<td>Programming Assignment 3</td>
<td>5%</td>
<td>15/11/15</td>
</tr>
<tr>
<td>Quiz I</td>
<td>25%</td>
<td>15/11/15</td>
</tr>
</tbody>
</table>
### Unit Schedule

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Percentage</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming Assignment 4</td>
<td>5%</td>
<td>22/11/15</td>
</tr>
<tr>
<td>Programming Assignment 5</td>
<td>5%</td>
<td>29/11/15</td>
</tr>
<tr>
<td>Programming Assignment 6</td>
<td>5%</td>
<td>6/12/15</td>
</tr>
<tr>
<td>Quiz II</td>
<td>25%</td>
<td>6/12/15</td>
</tr>
<tr>
<td>Report</td>
<td>20%</td>
<td>6/12/15</td>
</tr>
</tbody>
</table>
Assessment Requirements

Assessment Policy

Faculty Policy - Unit Assessment Hurdles

Academic Integrity - Please see resources and tutorials at
http://www.monash.edu/library/skills/resources/tutorials/academic-integrity/

Assessment Tasks

Participation

- Assessment task 1

  Title: Programming Assignment 1

  Description: Students are required to write and submit a program to sample from two binomial distributions, with and without a dependency between them.

  Learning outcomes 1, 2, 5, 8

  Weighting: 5%

  Criteria for assessment: Correctness of result; program commentary/explanation.

  Due date: 1/11/15

- Assessment task 2

  Title: Programming Assignment 2

  Description: Students are required to write and submit a program that visualizes sample data and their relation to probability models.

  Learning outcomes 1, 2, 8

  Weighting: 5%

  Criteria for assessment: Correctness of result; program commentary/explanation.

  Due date: 8/11/15

- Assessment task 3

  Title: Programming Assignment 3

  Description:
Assessment Requirements

Students are required to write and submit a programming exercise that produces a multiple regression model for a multivariate data set and computes the coefficient of determination.

Learning outcomes 1, 4, 5, 8

Weighting:
5%

Criteria for assessment:
Correctness of result; program commentary/explanation.

Due date:
15/11/15

• Assessment task 4

Title:
Quiz I

Description:
Students are required to demonstrate their understanding of theoretical, practical and algorithmic methods of data analysis by completing a quiz made up multiple choice and short-answer questions.

Learning outcomes 1, 4, 5

Weighting:
25%

Criteria for assessment:
Correctness

Due date:
15/11/15

• Assessment task 5

Title:
Programming Assignment 4

Description:
Students are required to write and submit a programming exercise that tests a null hypothesis and an alternative on a data set and performs a confidence interval estimation of effect size.

Learning outcomes 1, 4, 6, 7, 8

Weighting:
5%

Criteria for assessment:
Correctness of result; program commentary/explanation.

Due date:
22/11/15

• Assessment task 6

Title:
Programming Assignment 5

Description:
Students are required to write and submit a programming exercise that tests the claim that a particular statistical test has the power needed to decide that an alternative model is true and the null model false.
Learning outcomes 1, 2, 4, 7, 8

Weighting:
5%

Criteria for assessment:
Correctness of result; program commentary/explanation.

Due date:
29/11/15

• Assessment task 7

Title:
Programming Assignment 6

Description:
Students are required to write and submit a programming exercise that selects variables relevant to predicting a target variable in a data set.

Learning outcomes 1, 3, 4, 5, 7, 8

Weighting:
5%

Criteria for assessment:
Correctness of result; program commentary/explanation.

Due date:
6/12/15

• Assessment task 8

Title:
Quiz II

Description:
Students are required to demonstrate their understanding of theoretical, practical and algorithmic methods of data analysis by completing a quiz made up multiple choice and short-answer questions.

Learning outcomes 1, 2, 4, 6, 7

Weighting:
25%

Criteria for assessment:
Correctness

Due date:
6/12/15

• Assessment task 9

Title:
Report

Description:
Students are required to produce a report analysing a given data set and solving a scientific problem about it.

Learning outcomes 1, 2, 3, 4, 5, 6, 7

Weighting:
20%

Criteria for assessment:
Assessment Requirements

Students will be assessed according to: the clarity & organization of the report; the appropriateness of tools used; the explanation of analytic choices made; the quality of the interpretation of results; the correctness of conclusions drawn and support provided for those conclusions; the individual contribution made to the overall report for group work.

Due date:
6/12/15

Learning resources

Reading list


Monash Library Unit Reading List (if applicable to the unit)
http://readinglists.lib.monash.edu/index.html

Feedback to you

Types of feedback you can expect to receive in this unit are:

• Informal feedback on progress in labs/tutes
• Graded assignments with comments
• Graded assignments without comments
• Quiz results
• Solutions to tutes, labs and assignments

Extensions and penalties

Submission must be made by the due date otherwise penalties will be enforced.

You must negotiate any extensions formally with your campus unit leader via the in-semester special consideration process: http://www.monash.edu.au/exams/special-consideration.html

Returning assignments

Students can expect assignments to be returned within two weeks of the submission date or after receipt, whichever is later.

Referencing requirements

Students should use APA reference style, as explained at
http://intranet.monash.edu.au/infotech/resources/students/style-guide/referencing.html

Assignment submission

It is a University requirement (http://www.policy.monash.edu/policy-bank/academic/education/conduct/student-academic-integrity-managing-plagiarism-collusion-procedures.html) for students to submit an assignment coversheet for each assessment item. Faculty Assignment
Assessment Requirements

coversheets can be found at http://www.infotech.monash.edu.au/resources/student/forms/. Please check with your Lecturer on the submission method for your assignment coversheet (e.g. attach a file to the online assignment submission, hand-in a hard copy, or use an electronic submission). Please note that it is your responsibility to retain copies of your assessments.

Online submission

Assessment work should be submitted via Moodle.

Required Resources

Please check with your lecturer before purchasing any Required Resources. Limited copies of prescribed texts are available for you to borrow in the library, and prescribed software is available in student labs.

R and RStudio

Prescribed text(s)

Limited copies of prescribed texts are available for you to borrow in the library.


Technological Requirements

Students may use Windows, Linux or Mac environments for this subject. R and/or RStudio must be used for programming assignments. Other tools will be useful or necessary, but the choice of a particular tool (e.g., Weka, MATLAB, or a spreadsheet) is not mandated.

Recommended Resources

Weka. See http://www.cs.waikato.ac.nz/ml/weka/
Other Information

Policies

Monash has educational policies, procedures and guidelines, which are designed to ensure that staff and students are aware of the University’s academic standards, and to provide advice on how they might uphold them. You can find Monash’s Education Policies at: www.policy.monash.edu.au/policy-bank/academic/education/index.html

Faculty resources and policies

Important student resources including Faculty policies are located at http://intranet.monash.edu.au/infotech/resources/students/

Graduate Attributes Policy

http://www.policy.monash.edu/policy-bank/academic/education/management/monash-graduate-attributes-policy.html

Student Charter


Student services

The University provides many different kinds of support services for you. Contact your tutor if you need advice and see the range of services available at http://www.monash.edu.au/students. For Malaysia see http://www.monash.edu.my/Student-services, and for South Africa see http://www.monash.ac.za/current/

Monash University Library

The Monash University Library provides a range of services, resources and programs that enable you to save time and be more effective in your learning and research. Go to www.lib.monash.edu.au or the library tab in my.monash portal for more information. At Malaysia, visit the Library and Learning Commons at http://www.lib.monash.edu.my/. At South Africa visit http://www.lib.monash.ac.za/

Disability Liaison Unit

Students who have a disability or medical condition are welcome to contact the Disability Liaison Unit to discuss academic support services. Disability Liaison Officers (DLOs) visit all Victorian campuses on a regular basis.

- Website: http://www.monash.edu/equity-diversity/disability/index.html
- Telephone: 03 9905 5704 to book an appointment with a DLO; or contact the Student Advisor, Student Commuity Services at 03 55146018 at Malaysia
- Email: dlu@monash.edu
- Drop In: Equity and Diversity Centre, Level 1, Building 55, Clayton Campus, or Student Community Services Department, Level 2, Building 2, Monash University, Malaysia Campus