STAT4040
Statistical Learning

Course Description

This course provides an introduction to statistical learning and aims to develop skills in modern statistical data analysis. There has been a prevalence of "big data" in many different scientific fields. In order to tackle the analysis of data of such size and complexity, traditional statistical methods have been reconsidered and new methods have been developed for extracting information, or "learning", from such data. This course will cover a range of topics in statistical learning including linear regression, classification techniques, resampling methods (e.g., the bootstrap), regularisation methods, tree based methods, and unsupervised learning techniques (e.g., clustering). As the extensive use of statistical software is integral to modern data analysis, there will be a strong computing component in this course.

Mode of Delivery

On campus by way of lectures and tutorials

Prerequisites

Contact RSFAS to request a permission code to enrol in this course.

Incompatible Courses

STAT7040

Co-taught Courses

STAT3040 and STAT7040. Graduate students attend joint classes with undergraduates but are assessed separately.

Course Convener:

Dr Yanrong Yang
Phone: +61 2 612 58975
Email: yanrong.yang@anu.edu.au
Office hours for student consultation: To be advised on Wattle in due course

Research Interests

High dimensional statistics; large dimensional random matrix theory; large panel data analysis.

Relevant administrator if any (optional)

Maria Lander
Phone:
Email: maria.lander@anu.edu.au
Lecturer(s)

Dr Yanrong Yang
Phone(s): +61 2 612 58975
Email(s): yanrong.yang@anu.edu.au
Office hours for student consultation: To be advised on Wattle in due course

Tutor(s)

To be advised on Wattle in due course
COURSE OVERVIEW

Course Description
This course provides an introduction to statistical learning and aims to develop skills in modern statistical data analysis. There has been a prevalence of "big data" in many different scientific fields. In order to tackle the analysis of data of such size and complexity, traditional statistical methods have been reconsidered and new methods have been developed for extracting information, or "learning", from such data. This course will cover a range of topics in statistical learning including linear regression, classification techniques, resampling methods (e.g., the bootstrap), regularisation methods, tree based methods, and unsupervised learning techniques (e.g., clustering). As the extensive use of statistical software is integral to modern data analysis, there will be a strong computing component in this course.

Learning Outcomes

Upon successful completion of the requirements for this course, students should have the knowledge and skills to:

- **LO1**: Discuss in detail the rationale behind the formulation and components of a statistical model.
- **LO2**: Analytically describe and implement approaches to compare and contrast statistical models in the context of a particular scientific question.
- **LO3**: Communicate complex statistical ideas and heuristics to a diverse audience.
- **LO4**: Develop, analytically describe, and implement a statistical solution to real-data research problems.
- **LO5**: Demonstrate an in-depth level interpretation of modelling results.
- **LO6**: Discuss in detail the theoretical and computational underpinnings of various statistical procedures, including common classes of statistical models.
- **LO7**: Demonstrate computational skills to implement various statistical procedures.

Assessment Summary

<table>
<thead>
<tr>
<th>Assessment Task</th>
<th>Value</th>
<th>Due Date</th>
<th>Date for Return of Assessment</th>
<th>Linked Learning Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Assignment 1</td>
<td>10%</td>
<td>12 noon Wed 21 March</td>
<td>by Wed 11 April</td>
<td>LO1; LO2</td>
</tr>
<tr>
<td>2. Assignment 2</td>
<td>10%</td>
<td>12 noon 16 May</td>
<td>by Wed 30 May</td>
<td>LO3; LO6</td>
</tr>
<tr>
<td>3. Mid-semester exam</td>
<td>20%</td>
<td>Probably in Week 7</td>
<td>by Week 9</td>
<td>LO4; LO5; LO7</td>
</tr>
<tr>
<td>4. Final exam</td>
<td>60%</td>
<td>Final exam period</td>
<td></td>
<td>LO1 to LO7</td>
</tr>
</tbody>
</table>

Research-Led Teaching

If time permits, the lecturer may illustrate selected topics by discussing relevant examples from papers she has published. New material in these examples will not be assessable.

Feedback

Staff Feedback
Students will be given feedback in the following forms in this course:
• Verbal communication from lecturers and tutors, individually upon request.
• Marks and summaries for the assignments and mid-semester exam.

Student Feedback
ANU is committed to the demonstration of educational excellence and regularly seeks feedback from students. One of the key formal ways students have to provide feedback is through Student Experience of Learning Support (SELS) surveys. The feedback given in these surveys is anonymous and provides the Colleges, University Education Committee and Academic Board with opportunities to recognise excellent teaching, and opportunities for improvement.

For more information on student surveys at ANU and reports on the feedback provided on ANU courses, go to http://unistats.anu.edu.au/surveys/selt/students/ and http://unistats.anu.edu.au/surveys/selt/results/learning/

Policies
ANU has educational policies, procedures and guidelines, which are designed to ensure that staff and students are aware of the University’s academic standards, and implement them. You can find the University’s education policies and an explanatory glossary at: http://policies.anu.edu.au/

Students are expected to have read the Academic Misconduct Rule before the commencement of their course.

Other key policies include:
• Student Assessment (Coursework)
• Student Surveys and Evaluations

Recommended Resources
Prescribed Text (not compulsory but highly recommended)
• Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. (2013). An Introduction to Statistical Learning (with applications in R). Springer Texts in Statistics.

Other Recommended Text

COURSE SCHEDULE

<table>
<thead>
<tr>
<th>Week</th>
<th>Summary of Activities</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Orientation Week</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Chapter 1: Introduction to statistical learning</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Chapter 2: Linear Regression</td>
<td></td>
</tr>
</tbody>
</table>
ASSESSMENT REQUIREMENTS

The ANU is using Turnitin to enhance student citation and referencing techniques, and to assess assignment submissions as a component of the University's approach to managing Academic Integrity. For additional information regarding Turnitin please visit the ANU Online website.

Students may choose not to submit assessment items through Turnitin. In this instance you will be required to submit, alongside the assessment item itself, copies of all references included in the assessment item.

Assessment for this course consists of two assignments and two exams as detailed above and below. Attendance and participation in lectures and tutorials is not assessable.

STAT4040 is co-taught with STAT3040 and STAT7040. STAT3040 students will have slightly different learning outcomes and assessment. They will also be allocated to separate tutorial groups. Tutorial questions will be the same for both cohorts.

Assessment Tasks

Participation

Attendance and participation in lectures and tutorials is not assessable.

Assessment Task 1: Assignment 1

Details of task: This assignment covers Chapters 1–4 and is worth 10% of the total assessment. The assignment will be due in Week 6 and returned to students by Week 7.

Assessment Task 2: Assignment 2
**Details of task:** This assignment covers Chapters 5—6 and is worth 10% of the total assessment. The assignment will be due in Week 11 and returned within two weeks.

**Assessment Task 3: Mid-semester exam**

**Details of task:** This exam covers Chapters 1–4 and is worth 20% of the total assessment. The exam will probably be in Week 7 (with the exact date and venue to be announced and placed on Wattle in due course).

**Assessment Task 4: Final exam**

**Details of task:** This exam covers Chapters 1 to 8 and is worth 60% of the total assessment. The exam will be in the final examination period (with the exact date and venue to be announced and placed on Wattle in due course).

**Examination(s)**

Both the mid-semester and final exams will be open book, with permitted materials being a "Non-programmable calculator" and otherwise "No restrictions", apart from items excluded by general ANU examinations policy (such as mobile phones). The mid-semester exam will cover Chapters 1 to 4 only. The final exam will cover Chapters 1 to 8. These details may change. Students should check them prior to each exam, at [http://timetable.anu.edu.au/](http://timetable.anu.edu.au/)

**Assignment submission**

**Online Submission:** Assignments are submitted using Turnitin in the course Wattle site. You will be required to electronically sign a declaration as part of the submission of your assignment. Please keep a copy of the assignment for your records.

**Extensions and penalties**

Extensions and late submission of assessment pieces are covered by the Student Assessment (Coursework) Policy and Procedure.

The Course Convener may grant extensions for assessment pieces that are not examinations or take-home examinations. If you need an extension, you must request it in writing on or before the due date. If you have documented and appropriate medical evidence that demonstrates you were not able to request an extension on or before the due date, you may be able to request it after the due date.

No submission of assessment tasks without an extension after the due date will be permitted. If an assessment task is not submitted by the due date, a mark of 0 will be awarded.

**Returning assignments**

Through Turnitin.

**Resubmission of assignments**

It will not be possible for assignments to be resubmitted.

**Referencing requirements**
In assignments and exams, students must appropriately reference any results, words or ideas that they take from another source which is not their own. A guide can be found at: https://academicskills.anu.edu.au/resources/handouts/referencing-basics

Scaling

Your final mark for the course will be based on the raw marks allocated for each of your assessment items. However, your final mark may not be the same number as produced by that formula, as marks may be scaled. Any scaling applied will preserve the rank order of raw marks (i.e. if your raw mark exceeds that of another student, then your scaled mark will exceed the scaled mark of that student), and may be either up or down.

Privacy Notice

The ANU has made a number of third party, online, databases available for students to use. Use of each online database is conditional on student end users first agreeing to the database licensor’s terms of service and/or privacy policy. Students should read these carefully.

In some cases student end users will be required to register an account with the database licensor and submit personal information, including their: first name; last name; ANU email address; and other information.

In cases where student end users are asked to submit ‘content’ to a database, such as an assignment or short answers, the database licensor may only use the student’s ‘content’ in accordance with the terms of service – including any (copyright) licence the student grants to the database licensor.

Any personal information or content a student submits may be stored by the licensor, potentially offshore, and will be used to process the database service in accordance with the licensors terms of service and/or privacy policy.

If any student chooses not to agree to the database licensor’s terms of service or privacy policy, the student will not be able to access and use the database. In these circumstances students should contact their lecturer to enquire about alternative arrangements that are available.

Tutorial Seminar Registration

Tutorial signup for this course will be done via the Wattle website. Detailed information about signup times will be provided on Wattle or during your first lecture. When tutorials are available for enrolment, follow these steps:

1. Log on to Wattle, and go to the course site
2. Click on the link “Tutorial enrolment”
3. On the right of the screen, click on the tab “Become Member of…..” for the tutorial class you wish to enter
4. Confirm your choice

If you need to change your enrolment, you will be able to do so by clicking on the tab “Leave group…..” and then re-enrol in another group. You will not be able to enrol in groups that have reached their maximum number. Please note that enrolment in ISIS must be finalised for you to have access to Wattle.

SUPPORT FOR STUDENTS

The University offers a number of support services for students. Information on these is available online from http://students.anu.edu.au/studentlife/