STAT2008/6038
Regression Modelling
COURSE OUTLINE for Semester 1, 2014

STAT2008/STAT6038 is a course in applied statistics that studies the use of linear regression techniques for examining relationships between variables. The course emphasizes the principles of statistical modelling through the iterative process of fitting a model, examining the fit to assess imperfections in the model and suggest alternative models, and continuing until a satisfactory model is reached. Both steps in this process require the use of a computer: model fitting uses various numerical algorithms, and model assessment involves extensive use of graphical displays. The R statistical computing package is used as an integral part of the course.

<table>
<thead>
<tr>
<th>Mode of Delivery</th>
<th>On campus, see <a href="http://timetable.anu.edu.au/">http://timetable.anu.edu.au/</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Prerequisites</td>
<td>To enrol in STAT2008 as an undergraduate student, you MUST have completed STAT1003 or STAT1008. STAT6038 is a transitional Graduate Studies course, offered as part of programs such as the Graduate Certificate in Applied Statistics, Master of Applied Statistics or Master of Actuarial Studies. To enrol in STAT6038 you must have satisfied the admission requirements to the relevant program, which will typically involve completion of the equivalent of a bachelor's degree including an introductory statistics course equivalent to STAT1003/STAT1008.</td>
</tr>
</tbody>
</table>
Co-teaching

STAT2008 and STAT6038 share the same lecture and tutorial content and will have the same examination. However, the two cohorts will be treated separately in grading and any scaling that is applied.

Course Convenor and Lecturer:

Ian McDermid  
Room 3.08, CBE Building (26c)

Phone: 02 612 51084  
Email: ian.mcdermid@anu.edu.au

Office consultation: Mondays to Thursdays, preferably by prior appointment (made and confirmed via e-mail).

Research Interests
Population health and mortality  
Sample survey analysis and design

Student administrator
Patty Penm, RSFAS Office (details above)

Phone: 02 612 51526  
Email: patricia.penm@anu.edu.au

Tutors
A full list of contact details for course tutors and their consultation arrangements will be posted and regularly updated on the Wattle site for this course:

Course URL (Wattle)

All class materials will be made available via the Wattle site and any course announcements will be made in lectures (generally in the Tuesday lecture at the start of each week) and also posted on Wattle. You should check the Wattle site regularly.

If I, or anyone in the School, College or University administration, need to contact you, we will do so via your official ANU student e-mail address, which you also need to check regularly.

COURSE OVERVIEW

Learning Outcomes

Upon successful completion of this course, students will be able to:

1. Demonstrate a working knowledge of the R statistical computing language, particularly the graphical capabilities.

2. Fit Simple Linear regression models and interpret model parameters.

3. Summarise and analyse relationships between a response variable and a covariate.

4. Summarise and analyse relationships between a response variable and several covariates.

5. Assess and refine simple and multiple linear regression models based on diagnostic measures. Identify outlying and influential data points.

Summary of Proposed Assessment: Assessment for this course will be confirmed after consultation with students at the first lecture of the semester. If there are any changes to the assessment, those changes will be publicised on Wattle.

<table>
<thead>
<tr>
<th>Assessment Task</th>
<th>Value</th>
<th>Due Date</th>
<th>Learning Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Assignment 1</td>
<td>20%</td>
<td>5pm 3 April 2014</td>
<td>1, 2, 3 &amp; 5</td>
</tr>
<tr>
<td>2. Assignment 2</td>
<td>20%</td>
<td>5pm 22 May 2014</td>
<td>All</td>
</tr>
<tr>
<td>3. Final Examination</td>
<td>60%</td>
<td>Exam period</td>
<td>All</td>
</tr>
</tbody>
</table>

Research-Led Teaching
My teaching in this introductory course in statistical modelling will draw on examples from my extensive experience in applied statistical research and consulting.

Feedback
Students will be given feedback by their tutor who will mark their assignments. Students are also welcome to ask questions of myself or their tutor during class or at a consultation. I am also happy to answer SHORT questions on the course material sent via e-mail or posted on Wattle. If you send me a question via e-mail, I will, unless you specifically ask me not to, post your question (anonymously) and my answer on Wattle.

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 Code of Practice for Student Academic Integrity before the commencement of their course.

Key policies include:
- Student Assessment (Coursework)
- Student Surveys and Evaluations
- Assessment of Student Learning
## COURSE SCHEDULE

<table>
<thead>
<tr>
<th>Week</th>
<th>Lecture Topics</th>
<th>Tutorials / Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (18 February)</td>
<td>Introduction. Getting started with R. Simple Linear Regression (revision).</td>
<td>No tutorials in week 1</td>
</tr>
<tr>
<td>3 (4 March)</td>
<td>Properties of least squares estimators.</td>
<td>R Worksheets 3 &amp; 4</td>
</tr>
<tr>
<td>4 (11 March)</td>
<td>ANOVA. Hypothesis testing and interval estimation in a SLR context.</td>
<td>Tutorial 1</td>
</tr>
<tr>
<td>5 (18 March)</td>
<td>Prediction intervals. Regression diagnostics (residual plots).</td>
<td>Tutorial 2</td>
</tr>
<tr>
<td>6 (25 March)</td>
<td>Outliers and influential observations. Scale transformations.</td>
<td>Tutorial 2 continued</td>
</tr>
<tr>
<td>7 (1 April)</td>
<td>Introduction to Multiple Regression. Model interpretation and estimation.</td>
<td>Assignment 1 due (discussed in tutorials)</td>
</tr>
<tr>
<td>8 (22 April)</td>
<td>ANOVA. Sequential sum of squares.</td>
<td>Tutorial 3</td>
</tr>
<tr>
<td>9 (29 April)</td>
<td>Hypothesis testing, confidence intervals and prediction for multiple regression.</td>
<td>Tutorial 3 continued</td>
</tr>
<tr>
<td>10 (6 May)</td>
<td>Model diagnostics. Outlier detection. Types of residuals.</td>
<td>Tutorial 4</td>
</tr>
<tr>
<td>11 (13 May)</td>
<td>Influence diagnostics. Multicollinearity.</td>
<td>Tutorial 4 continued</td>
</tr>
<tr>
<td>12 (20 May)</td>
<td>Model selection and criteria for comparing models. Stepwise procedures.</td>
<td>Assignment 2 due (discussion in tutorials)</td>
</tr>
<tr>
<td>13 (27 May)</td>
<td>Revision for Final Examination.</td>
<td>Tutorial 5</td>
</tr>
<tr>
<td>(5 to 21 June)</td>
<td>Examination period</td>
<td>Final Examination</td>
</tr>
</tbody>
</table>

## ASSESSMENT REQUIREMENTS

**Assignments 1 (Simple Linear Regression) and 2 (Multiple Regression)**

Detailed assignment specifications will be handed out at least three weeks prior to the due dates. Assignments will involve using R to analyse data from a case study, then organising and editing the R output and preparing a written report on your analyses. These reports should be submitted to the appropriate box at the RSFAS School Office by the due date (see page 3). As a general rule, late assignments will NOT be accepted.

**Final Examination**

Permitted materials and other conditions for the Final Examination will be discussed with students and the outcome advised on Wattle. The Final Examination will be centrally timetabled and the details released via [http://timetable.anu.edu.au/](http://timetable.anu.edu.au/)

**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.
**Workloads**
Students taking this course are expected to commit at least 10 hours a week to completing the work. This will include:

- attendance at 3 lectures per week and 1 tutorial (sign-up details below)
- 6 or more hours of private study, including time spent reviewing course materials and completing computing, tutorial work and assignments.

**Tutorial signup**
Enrolment in tutorials will be completed online using the CBE Electronic Teaching Assistant (ETA). To enrol, follow these instructions:

2. You will see the Student Login page. To log into the system, enter your University ID (your student number) and password (your ISIS password) in the appropriate fields and hit the Login button.
3. Read any news items or announcements.
4. Select "Sign Up!" from the left-hand navigation bar.
5. Select your courses from the list. To select multiple courses, hold down the control key. On PCs, this is the Ctrl key; on Macs, it is the ⌘ key. Hold this key down while selecting courses with the mouse. Once courses are selected, hit the SUBMIT button.
6. A confirmation of class enrolments will be displayed. In addition, an email confirmation of class enrolments will be sent to your student account.
7. For security purposes, please ensure that you click the LOGOUT link on the confirmation page, or close the browser window when you have finished your selections.
8. If you experience any difficulties, please contact the School Office (see page 1 for contact details).
9. Students will have until the end of week 2 to finalise their enrolment in tutorials. After this time, students will be unable to change their tutorial enrolment.

**Reference Materials**
Class materials, including detailed lecture notes, Power-point slides, lecture demonstrations, tutorials, assignments and other relevant materials, will be made available on the class web page on Wattle at [https://wattle.anu.edu.au/](https://wattle.anu.edu.au/)

To log on to Wattle, you need to have your ANU ID (student number) and password. In order to access the class web page within Wattle, you will need to be formally enrolled in the course. It is essential that you visit the class web page regularly. STAT6038 students should use the STAT2008 web page.
Prescribed Text
As we have a lot of detailed course material already available, there is no prescribed text for this course. However, I recommend the following text for students who would like an even more detailed discussion of the course contents:


There is a one copy of this text on 2 hour reserve in the ANU Hancock library (Call # QA279.F37) and the ANU Co-op Bookshop has copies available for purchase ($113.00 or $105.09 with the member's discount).

I will recommend other references during the course and possibly also make additional material available in the library e-Reserve. Use the following links to find course material in the ANU library:


Technology and Software
The application of modern statistical techniques requires familiarity with a statistical computing package and the assignments for this course will require some data analysis on a computer.

This course makes extensive use of the R computing package, which is freely available to download at [http://www.r-project.org](http://www.r-project.org). Further instructions on R, including a series of "getting started" workshops will be made available on the Wattle site for this course. R is also available on all InfoCommons computers on the ANU campus.

You will need also need access to a scientific calculator for the Final Examination.

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/](http://students.anu.edu.au/studentlife/)