Faculty of Medicine
School of Medical Sciences

HDAT 9700

Statistical Modelling II

COURSE OUTLINE

SEMESTER 2, 2018
HDAT9700 Course Information

OBJECTIVES OF THE COURSE
Sophisticated modelling techniques are essential for the analysis of real-world health data. Building on Health Data Analytics: Statistical Modelling I (HDAT9600), this course expands the statistical toolkit and broadens students’ understanding of relevant statistical approaches for the analysis of realistically complex data structures and research questions. The course is aimed at those currently working or planning on working in health or a health-related field, and who are interested in applying advanced statistical methods to analyse complex data.

Topics covered in this course include multilevel models for hierarchical data; analysis of time series and longitudinal data; quasi-experimental approaches for drawing causal inferences from observational data; multiple imputation for missing values; and simulation approaches for study planning and model evaluation.

Content is delivered through a combination of online readings, expert guest lectures and practical hands-on tutorials. Statistical concepts are illustrated with a variety of health examples, and students will learn how to implement methods using leading statistical software. Lectures are followed by weekly exercises, which reinforce the learning and programming skills covered in the face-to-face tutorials.

COURSE CO-ORDINATORS and LECTURERS
Course Coordinators:
Dr Maarit Laaksonen
Centre for Big Data Research in Health, Level 1, AGSM Building (G27), UNSW Sydney:
02 9385 1402       m.laaksonen@unsw.edu.au

Dr Mark Hanly
Centre for Big Data Research in Health, Level 1, AGSM Building (G27), UNSW Sydney:
02 9385 3143       m.hanly@unsw.edu.au

Students wishing to see the course coordinators should make an appointment via email as our offices are not readily accessible. We will organize to meet you in a convenient location elsewhere in the building.

Lecturers in this course:
Dr Andrew Blance            a.blance@unsw.edu.au
Dr Andrea Schaffer         andrea.schaffer@unsw.edu.au
Dr Kylie-Ann Mallitt       k.mallitt@unsw.edu.au
Dr Mark Hanly              m.hanly@unsw.edu.au
Ms Sanja Lujić             s.lujic@unsw.edu.au
COURSE STRUCTURE and TEACHING STRATEGIES

This is a blended learning course comprising of:

- 10 face-to-face sessions: Tuesday 1pm-4pm Mat 228
  Friday 2pm-5pm Goldstein G05

- Interactive learnr tutorials

Students are expected to attend all face-to-face sessions for their full duration (3 hours per chapter). Students are reminded that UNSW recommends that a 6 units-of-credit course should involve about 150 hours of study and learning activities. The formal learning activities are approximately 100 hours throughout the semester and students are expected (and strongly recommended) to do additional hours of self-study.

The workflow of a typical week includes the following activities:

1. Independent work prior to attending face-to-face activities (e.g. online/textbook reading)
2. Face-to-face activities: expert guest lectures and practical hands-on tutorials
3. Weekly exercises (Assessment 1)
4. Work towards the final project (Assessment 2)

The course is accessed via UNSW Moodle and www.openlearning.com. Core material will be delivered as learnr documents with explanatory text, embedded videos and interactive coding activities using Rstudio. Assessments will be distributed, completed and submitted using Git and GitHub Classroom.

Assumed knowledge: This course is taught using basic R programming. For students not familiar with R, a beginner’s introduction to R can be completed before the start of the course. Contact MScHDS@unsw.edu.au about how to gain basic R proficiency.

COURSE LEARNING OUTCOMES

The course aims to provide students with both conceptual understanding and hands-on experience with a range of statistical modelling techniques relevant to the analysis of health data. On completion of this course students should be able to:

1. Critique the relative merits of a range of statistical models for analysing health research data
2. Construct statistical models with appropriate data structures to address distinct health research questions
3. Compose code using appropriate statistical software to run a range of sophisticated analytic techniques
4. Appraise model fit using a variety of model diagnostics
5. Interpret the fitted model parameters from a range of statistical models.
COURSE EVALUATION AND DEVELOPMENT

For course evaluation, feedback will be gathered at the completion of the course using the myExperience online student survey in Moodle. Student feedback is taken seriously, and continual improvements will be made to the course based, in part, on such feedback.

ASSESSMENT PROCEDURES

• Assessment 1: Individual exercises 40%

After each chapter, the student will complete a summative assessment task reinforcing the learning from that chapter. Assessments will be distributed, and solutions submitted, using GitHub Classroom. Links to the assessment will be posted on Open Learning. Students are expected commit and push their completed assessment using Git.

• Assessment 2: Individual final project 60%

A contextual health scenario will be posed, and a relevant dataset provided. Students are required to analyse the data using the appropriate statistical method(s) covered in the course. The assignment requires the production of a 1500-word report, outlining the (health) issue, describing the analyses undertaken, presenting the results and interpreting the findings. The assessment will be marked using a rubric.
**COURSE TIMETABLE**

<table>
<thead>
<tr>
<th>UNSW term week</th>
<th>Chapter</th>
<th>Date</th>
<th>Chapter topic</th>
<th>Face-to-face</th>
<th>Assessment due date</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
<td>Tue 11 Sep</td>
<td>Review of research designs and statistical methods</td>
<td>1-4pm, Matthews 228</td>
<td>11:59pm Mon 24 Sept</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>Tue 18 Sep</td>
<td>Statistical issues in the analysis of baseline</td>
<td>1-4pm, Matthews 228</td>
<td>11:59pm Mon 1 Oct</td>
</tr>
<tr>
<td>Break</td>
<td></td>
<td>Tue 25 Sep</td>
<td>Mid-semester break</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>Tue 2 Oct</td>
<td>Multilevel modelling (Part A)*</td>
<td>1-4pm, Matthews 228</td>
<td>*</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>Fri 5 Oct</td>
<td>Multilevel modelling (Part B)*</td>
<td>2-5pm, Goldstein G05</td>
<td>11:59pm Thu 18 Oct</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>Tue 9 Oct</td>
<td>Repeated measures analysis</td>
<td>1-4pm, Matthews 228</td>
<td>11:59pm Mon 22 Oct</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>Fri 12 Oct</td>
<td>Time series analysis</td>
<td>2-5pm, Goldstein G05</td>
<td>11:59pm Thu 25 Oct</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
<td>Tue 16 Oct</td>
<td>Interrupted time series analysis</td>
<td>1-4pm, Matthews 228</td>
<td>11:59pm Mon 29 Oct</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>Fri 19 Oct</td>
<td>Causal inferences from observational data</td>
<td>2-5pm, Goldstein G05</td>
<td>11:59pm Thu 1 Nov</td>
</tr>
<tr>
<td>13</td>
<td>9</td>
<td>Tue 23 Oct</td>
<td>Missing data and multiple imputation</td>
<td>1-4pm, Matthews 228</td>
<td>11:59pm Mon 5 Nov</td>
</tr>
<tr>
<td>13</td>
<td>10</td>
<td>Fri 26 Oct</td>
<td>Estimating sample size for intractable scenarios via simulation</td>
<td>2-5pm, Goldstein G05</td>
<td>11:59pm Thu 8 Nov</td>
</tr>
</tbody>
</table>

*Note there will be a joint assignment for chapters 3 and 4
A penalty of 5% per day will apply to late submissions of assessment tasks. For example, if your assessment was marked as 60%, but was handed in two days late, you would instead receive \(60\times(1-0.05\times2) = 54\%\).

<table>
<thead>
<tr>
<th>Task</th>
<th>Release date</th>
<th>Due Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment 1: Chapter 1</td>
<td>Tue 11/09/2018</td>
<td>Mon 24/09/2018</td>
</tr>
<tr>
<td>Assessment 1: Chapter 2</td>
<td>Tue 18/09/2018</td>
<td>Mon 07/10/2018</td>
</tr>
<tr>
<td>Assessment 1: Chapter 3 &amp; 4</td>
<td>Fri 5/10/2018</td>
<td>Thu 18/10/2018</td>
</tr>
<tr>
<td>Assessment 1: Chapter 5</td>
<td>Tue 9/10/2018</td>
<td>Mon 22/10/2018</td>
</tr>
<tr>
<td>Assessment 1: Chapter 6</td>
<td>Fri 12/10/2018</td>
<td>Thu 25/10/2018</td>
</tr>
<tr>
<td>Assessment 1: Chapter 7</td>
<td>Tue 16/10/2018</td>
<td>Mon 29/10/2018</td>
</tr>
<tr>
<td>Assessment 1: Chapter 8</td>
<td>Fri 19/10/2018</td>
<td>Thu 1/11/2018</td>
</tr>
<tr>
<td>Assessment 1: Chapter 9</td>
<td>Tue 23/10/2018</td>
<td>Mon 5/11/2018</td>
</tr>
<tr>
<td>Assessment 1: Chapter 10</td>
<td>Fri 26/10/2018</td>
<td>Thu 8/11/2018</td>
</tr>
<tr>
<td>Assessment 2</td>
<td>Fri 26/10/2018</td>
<td>Fri 16/11/2018</td>
</tr>
</tbody>
</table>
GENERAL INFORMATION

Special Consideration
Please see UNSW-Special Consideration and Student Advice-Special Consideration

If you unavoidably miss assessment deadlines, you must lodge an application with UNSW Student Central for special consideration within 3 days after the assessment due date. You should also inform the course conveners by email as soon as possible.

If your request for consideration is granted an alternative assessment deadline will be organised with the course convener.
See: Student-Advice-Reviews and Appeals

Student Support Services
See: Student Advice-Student support services.

Academic Integrity and Plagiarism
The UNSW Student Code outlines the standard of conduct expected of students with respect to their academic integrity and plagiarism.

More details of what constitutes plagiarism can be found here