Faculty of Medicine
School of Medical Sciences

HDAT 9700

Statistical Modelling II

COURSE OUTLINE

Term 2, 2022

CONTENTS PAGE

OBJECTIVES OF THE COURSE **1**

COURSE CO-ORDINATORS and LECTURERS **1**

COURSE STRUCTURE and TEACHING STRATEGIES **2**

COURSE LEARNING OUTCOMES **2**

COURSE EVALUATION AND DEVELOPMENT **3**

ASSESSMENT PROCEDURES **3**

COURSE TIMETABLE **4**

ASSESSMENT TASKS AND TIMETABLE **5**

GENERAL INFORMATION **6**

Special Consideration 6

Student Support Services 6

Academic Integrity and Plagiarism 6

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; methods for drawing causal inferences from observational data; and multiple imputation for missing values.

Content is delivered through a combination of 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.

COURSE CO-ORDINATORS and LECTURERS

Course Coordinators:

Dr Mark Hanly
Centre for Big Data Research in Health, Level 4, Lowy Building (G25), UNSW Sydney.
02 9385 3143
m.hanly@unsw.edu.au

Students wishing to meet with the course coordinator should make an appointment *via* email. Virtual Teams meetings or face-to-face meetings on campus are possible.

Lecturers in this course:

Dr Mark Hanly **m.hanly@unsw.edu.au**
Dr Md Shajedur Rahman Shawon **s.shawon@unsw.edu.au**

COURSE STRUCTURE and TEACHING STRATEGIES

This is a blended learning course comprising of:

- Readings and video recordings
- Weekly online tutorials (Microsoft Teams Meetings)
- Interactive R learnR tutorials

The weekly online tutorial sessions will be recorded and made available for those who cannot make the allotted time. If you do not consent to being recorded, please contact the course convenors ahead of the session.

Students are reminded that UNSW recommends that a 6 units-of-credit course should involve about 150 hours of study and learning activities, **equating to 15 hours per week**. 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. Read/watch the recommended readings/videos
2. Attend weekly tutorial sessions
3. Complete the interactive learnR tutorial
4. Ask questions and join discussions in Teams
5. Work on assessments

The course is accessed via the course website https://cbdrh-hdat9700.github.io. Core material will be delivered as R learnr documents with explanatory text, embedded videos and interactive coding activities using Rstudio. Assessments will be distributed, completed and submitted using Git and GitHub.

Assumed knowledge

This course assumes proficiency in fitting and interpreting Generalised Linear Models using R or a similar statistical analysis package (e.g. Stata/SAS/Python), at a level covered in HDAT9600 Statistical Modelling I. The course is delivered using R so some experience with R programming is expected. Assessments are submitted using Git so familiarity with the Git and GitHub workflow is useful, although not a formal prerequisite.

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. Student feedback is taken seriously, and continual improvements will be made to the course based, in part, on such feedback.

ASSESSMENT PROCEDURES

- Assessment 1: Short exercises **Weight 50%**

Throughout the course, students will complete summative assessments. In these tasks, students will answer questions, complete coding exercises and interpret results to reinforce the learning covered in the course. Assessments will be distributed, and solutions submitted, using GitHub. Links to the assessment will be posted on the course website. Students are expected to use the standard git workflow to commit and push their completed assessment. Feedback on code may be provide directly through GitHub.

- Assessment 2: Individual final project **Weight 50%**

Students are provided with a health research question and a relevant dataset. The assignment requires the production of a 1500-word report, implementing statistical modelling techniques covered in the course to address the research question. This requires developing, coding and interpreting appropriate analysis for the context. Emphasis is given to both the implementation and interpretation of the findings. The assignment is assessed via a rubric.

COURSE TIMETABLE

UNSW Term week	Week staring	Chapter topic	Lecturer	Online tutorial session
1	Mon 30 May	Directed Acyclic Graphs (DAGs)	Dr Mark Hanly	5-7pm, Mon 30 May
2	Mon 6 Jun	Matching for causal inference from observational data	Dr Mark Hanly	5-7pm, Mon 6 Jun
3	Mon 13 Jun	Multilevel modelling I (Introduction)	Dr Mark Hanly	5-7pm, Tue 14 Jun*
4	Mon 20 Jun	Multilevel modelling II (Beyond the basics)	Dr Mark Hanly	5-7pm, Mon 20 Jun
5	Mon 27 Jun	Multilevel modelling III (Repeated measures analysis)	Dr Mark Hanly	5-7pm, Mon 27 Jun
6	Mon 4 Jul	Reading week	NA	NA
7	Mon 11 Jul	Time series analysis	Dr Mark Hanly	11-1pm, Mon 11 Jul**
8	Mon 18 Jul	Interrupted time series analysis	Dr Mark Hanly	5-7pm, Mon 18 Jul
9	Mon 25 Jul	Missing data and multiple imputation	Dr Mark Hanly	11-1pm, Mon 25 Jul**
10	Mon 1 Aug	Presenting and summarising model results	Dr Md Shajedur Rahman Shawon	5-7pm, Mon 1 Aug

* Tutorial moved to Tuesday in Week 3 because Monday 13 June is a public holiday in NSW.

** Note mid-morning tutorial in Week 7 and 9.

ASSESSMENT TASKS AND TIMETABLE

Assessment	Weight	Release date	Due Date	Feedback date
1A: Weeks 1-2	50%	Mon 6 Jun	9am, Mon 27 Jun	Week starting Mon 11 Jul
1B: Weeks 3-5		Mon 20 Jun	9am, Mon 11 Jul	Week starting Mon 25 Jul
1C: Weeks 7-8		Mon 18 Jul	9am, Mon 1 Aug	Week starting Mon 15 Aug
2. Short report	50%	Week starting Mon 18 Jul	9am, Mon 15 Aug	Week starting Mon 29 Aug

If you submit assessments late without special consideration, a 5% penalty deduction for every day late will be applied. For example, if you submit an assessment 3 days late, then 15% (5% x 3 days) will be deducted from the assessment mark. Thus, if your assessment was marked as 80% but was submitted 3 days late, then your final mark will be 65% only. Assessments will not be marked if submitted 5 or more days after the assessment due date and will receive a score of 0.

GENERAL INFORMATION

Special Consideration

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

If you find you are going to miss an assessment due to illness, misadventure or circumstances beyond your control, you need to apply for special consideration. To do this:

1. Tell the course convenors as soon as possible via email. Applications for special consideration will not normally be received more than 3 days after the assessment due date.
2. Submit supporting evidence. This may include a medical certificate or a supporting email from your supervisor for extenuating work commitments.

If your application for special consideration is approved, your course convenors will discuss with you how you can complete your assessment. Note that all special considerations are handled centrally through the processes described in the links above.

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.