UNSW SYDNEY

Faculty of Science

School of Mathematics and Statistics

MATH2831/2931

Linear Models/Higher Linear Models

Term 3, 2019

MATH2831/2931 – Course Outline

Information about the course

Course Authority/Lecturer: Dr. Eka Shinjikashvili, RC-2054, email: `eka@unsw.edu.au`

Consultation: Tuesday 12noon (other times may be arranged by appointment).

Credit, Prerequisites, Exclusions: This course counts for 6 Units of Credit (6UOC).

Prerequisites: MATH2801 (for MATH2831) or MATH2901 (for MATH2931)

Exclusions: BEES2041, BIOS2041, MATH2831 (for MATH2931), MATH2931 (for MATH2831).

Lectures: There will be four hours of lectures per week running from week 1 to 5 and then from week 7 to 10 (no lectures in week 6):

Tuesday 9am to 11am	**Rex Vowels Theatre**
Friday 11am to 1pm	**Ainsworth G03**

Tutorials: The tutorials will be held on the odd weeks (w1,3,5,7,9). The initial tutorial classes are scheduled at the Red Centre tutrial rooms:

MATH2931	Tuesday 4pm	RC-1043
MATH2931	Wednesday 3pm	RC-2060
MATH2931	Thursday 11am	RC-1043
MATH2931	Thursday 12pm	RC-1043
MATH2931	Friday 2pm	RC-1043
MATH2831	Tuesday 4pm	RC-2063
MATH2831	Thursday 11am	RC-2063
MATH2831	Thursday 12pm	RC-2063
MATH2831	Friday 2pm	RC-2063

Lab Classes: The lab classes will be held on the even weeks (w2,4,6,8,10). The initial lab classes are scheduled at the Red Centre computer labs:

MATH2931	Tuesday 4pm	RC-G012A
MATH2931	Wednesday 3pm	RC-G012A
MATH2931	Thursday 11am	RC-G012A
MATH2931	Thursday 12pm	RC-G012A
MATH2931	Friday 2pm	RC-G012C
MATH2831	Tuesday 4pm	RC-G012C
MATH2831	Thursday 11am	RC-G012C
MATH2831	Thursday 12pm	RC-G012C
MATH2831	Friday 2pm	RC-G012A

Moodle: Further information, announcements, lecture notes, tutorial/lab problems and other material will be provided via *UNSW Moodle* (`https://moodle.telt.unsw.edu.au/`).

Course aims

Statistics is about using probability models to make decisions from data in the face of uncertainty. This course gives an introduction to the process of building statistical models using an important class of models (linear models). In a linear model we try to predict or explain variation in a response variable in terms of related quantities (predictors). The relationship between the expected response and predictors is linear in unknown model parameters. Topics covered in the course include how to estimate parameters in linear models, how to compare models using hypothesis testing, how to select a good model or models when prediction of the response is the goal, and how to detect violations of model assumptions and observations which have undue influence on decisions of interest. Concepts are illustrated with applications from finance, economics, medicine, environmental science and engineering.

Relation to other statistics courses

Building on the foundation statistical inference course MATH2801/2901, linear models are considered a fundamental component of statistical practice. Satisfactorily completing this course will provide a solid background for more advanced statistical courses.

Student learning outcomes

This course is expected to give students an understanding of the fundamentals of regression modelling, which is essential for anyone contemplating a career as a professional statistician or higher study in statistics for students majoring in mathematics and statistics.

Statistical software R/RStudio will be used to demonstrate how to produce and interpret the computer output when applying the concepts of linear modelling.

Relation to graduate attributes

- The problem-solving activities in tutorials, labs and assignments will improve your research, enquiry and analytical thinking abilities (Science Graduate Attribute 1) and your capacity and motivation for intellectual development (Science Graduate Attribute 2);
- Regular coursework assignments will provide you with timely feedback on your progress and improve your Communication skills (Science Graduate Attribute 4);
- Computing skills developed in this course will improve your Information Literacy (Science Graduate Attribute 6)

Teaching strategies underpinning the course

New ideas and skills are introduced and demonstrated in lectures, then students develop these skills by applying them to specific tasks in tutorials and assessments.

Rationale for learning and teaching strategies

We believe that effective learning is best supported by a climate of inquiry, in which students are actively engaged in the learning process. Hence this course is structured with a strong emphasis on problem-solving tasks. Students are expected to devote the majority of their class and study time to the solving of such tasks. Computing skills are developed and practiced in regular computer laboratory sessions. This course has a major focus on research, inquiry and analytical thinking as well as information literacy. We will also explore capacity and motivation for intellectual development through the solution of both simple and complex mathematical models of problems arising in finance, economics, medicine, environmental science and engineering, and the interpretation and communication of the results.

Assessment

Assessment in this course will consist of two group assignments (each worth 20%) and a final examination (60%).

Knowledge and abilities assessed: All assessment tasks will assess the learning outcomes outlined above, specifically, the ability to derive logical and coherent proofs of relevant results, and the ability to solve a variety of regression problems, both theoretical and in practice.

Assessment criteria: The main criteria for marking all assessment tasks will be clear and logical presentation of correct solutions.

MATH2931 assessment: Students in the higher level course (MATH2931) will receive additional, and more challenging assignment and examination problems.

Assessment policies

- UNSW assesses students under a standards-based assessment policy. For how this policy is applied within the School of Mathematics and Statistics, please visit the web site:
  `https://www.maths.unsw.edu.au/currentstudents/assessment-policies`
- For information on how the School implements special consideration policies for assessments during the term and the final examination, refer to the School's website: `https://www.maths.unsw.edu.au/currentstudents/special-consideration-illness-misadventure`

Assignments

Rationale: Assignments will give an opportunity for students to try their hand at more difficult problems requiring more than one line of argument and also introduce them to aspects of the subject which are not explicitly covered in lectures. Assignments are also intended to give regular feedback on a student's progress and mastery of the material, to identify as soon as possible any problems that students may have. Assessment in this course will use problem-solving tasks of a similar form to those practiced in tutorials and labs, to encourage the development of the core skills underpinning this course and the development of analytical thinking.

Task	Available	When Due	Form of Submission	Weighting
Assignment 1	Week 4	Week 6	Written	20%
Assignment 2	Week 7	Week 9	Written	20%

The content of the assignments will be available on *Moodle* at least two weeks before the assignment is due.

Academic integrity, plagiarism and referencing

- Plagiarism is the presentation of the thoughts or work of another person as one's own. It is a serious breach of ethics at UNSW and is not taken lightly. You should consult the University web page on academic integrity and plagiarism at:
  `https://student.unsw.edu.au/plagiarism`
- The Learning Centre website offers you the resources and support to help you develop and refine your academic skills. It can be located at:
  `https://student.unsw.edu.au/skills`

Final examination

Duration: Two hours.

Rationale: The final examination will assess student mastery of the material covered in the lectures, assignments and tutorials.

Weighting: 60% of your final mark.

Further details about the final examination will be available in lectures and on *Moodle* closer to the time.

Resources for students

Lecture notes

A set of weekly lecture notes will be provided on *Moodle*. It is recommended to download these notes and bring them to lectures.

Course pack

The lecture notes are not sufficient to understand the course material. There is a coursepack "*MATH2831/2931* L*inear Models*" containing course notes covering all the topics. It will be available on *Moodle*. **Students are strongly advised to read the course pack.**

Tutorial and lab exercises

All tutorial and lab classes will be held in the Red Center. Weekly exercises and datasets will be provided on *Moodle*. These problems are for YOU to do to enhance mastery of the course, even if you require longer than the allocated tutorial to complete them.

Computing and R/RStudio statistical software

We will be using RStudio, which is a user friendly environment to the R programming language for statistical computing and graphics. Both R and RStudio can be freely downloaded and installed from:

`http://www.r-project.org`

and

`http://www.rstudio.com/products/rstudio/download/`

RStudio is available in the computer labs within the School of Mathematics and Statistics.

The Red Centre computer laboratories (RC-M020 and RC-G012) are open 8:00am – 9:00pm Monday – Friday on teaching days. You are welcome to use the labs when they are not used for classes or tests. Information about the labs can be found at:

`http://www.maths.unsw.edu.au/currentstudents/computing-information`

Moodle

Many additional course materials will be available on *Moodle*. You should check regularly for new materials.

Some useful textbooks are listed below for your reference.

Textbooks

- Draper, Norman R., and Harry Smith (2014). Applied Regression Analysis, John Wiley and Sons
- Raymond H. Myers (1990). Classical and Modern Regression of Applications. PWS-KENT.
- R. Dennis Cook and Sanford Weisberg (1999). Applied Regression including Computing and Graphics. Wiley.
- Raymond H. Myers and Janet S. Molton (1991). A First Course in the Theory of Linear Statistical Models. PWS-KENT.
- Alan Stuart and J. Keith Ord (1987). Kendall's Advanced Theory of Statistics Vol. 2 (5th Edition). Edward Arnold.

Detailed course schedule

Weeks 1 and 2 – Simple linear regression

Formulation of the simple linear regression model; least squares estimation; estimation of error variance; maximum likelihood; confidence intervals; hypothesis testing; prediction; model criticism.

Weeks 3 and 4 – The general linear model

Formulation of the general linear model; least squares; maximum likelihood; estimation of error variance; interval estimation; hypothesis testing; multicollinearity.

Week 5 – Model selection

The model selection problem; PRESS; cross validation, C_p, sequential procedures; limitations of automated model selection procedures.

Week 6 – No lectures

Weeks 7 and 8 – Residuals and diagnostics

Residual plots; outlier detection; partial regression and residual plots; testing for normality; influence measures; transformations.

Week 9 – Categorical predictors

Categorical predictor variables; dummy variables; hypothesis testing; interactions.

Week 10 – Less than full rank models and other regression Models

One-way classification model with fixed effects; logistic regression;

Course evaluation and development

The School of Mathematics and Statistics evaluates each course each time it is run. We carefully consider the student responses and their implications for course development. It is common practice to discuss informally with students how the course and their mastery of it are progressing.

School rules and regulations

It is the student's responsibility to be familiar with UNSW and School of Mathematics and Statistics policies.

Fuller details of the general rules regarding attendence, release of marks, special consideration etc are available via the School of Mathematics and Statistics web page at:
`http://www.maths.unsw.edu.au/currentstudents/student-services`