# UNSW SCIENCE

# SCHOOL OF MATHS AND STATISTICS

# MATH2801
# Theory of Statistics

## Term 2, 2019

Cricos Provider Code: 00098G

# MATH2801 – Course Outline

**Course authority/lecturers:**

Dr Jakub Stoklosa (RC-3071)
j.stoklosa@unsw.edu.au

Regular **consultation times** with lecturers will be announced on Moodle and in lectures. Other times may be arranged by appointment. Please use email to arrange an appointment and make sure you use your UNSW email account.

* MATH2901 (Higher Theory of Statistics) is a more theoretical/advanced version of MATH2801 which spends more time on proof and theoretical considerations. Please contact student services if you would like to switch to MATH2901

**Credit/prerequisites/exclusions:** This course counts for 6 Units of Credit (6UOC).

First year mathematics is assumed knowledge for this course:
MATH1231 or MATH1241 or MATH1251 (or, in program 3653, MATH1131 or MATH1141) or MATH1031(CR) (for MATH2801 only).

Excluded: introductory stats courses with theoretical focus: ECON2215, MATH2089, MATH2099, MATH2829, MATH2839, MATH2841, MATH2859, MATH2899.

**Assumed knowledge:** First year probability theory and integration. Probability theory revision notes and exercises (available from the web-page) will be revised briefly early in the course.

**UNSW Moodle:** Course notes, tutorial material, announcements, additional resources and Internet links copies of other essential information will be provided on the course web pages via UNSW Moodle.

**Students are recommended to download the lecture notes and bring them to lectures.**

**Lectures/tutorials:** Lectures and tutorials start in Week 1 and continue until the end of Week 10.

For times and locations of lectures and tutorials please see your myUNSW timetable.

There will be five hours of lectures per week. Each student will have one tutorial a week. The Week 2 tutorial will require the use of your laptop as we will be learning how to use a well-known statistical software called `R/RStudio`.

Tutorial questions are available from the course web page and students are strongly recommended to attempt them before class.

**Computing:** We will use the (free) statistics package `RStudio`.

This is available in the computer labs within the School of Mathematics and Statistics and is also downloadable from `https://www.rstudio.com`, or you can just Google "`RStudio`" into your search engine and open the first link. From there just follow the usual download and setup.

## Course aims

This course is an introduction to the theoretical underpinnings of statistics, essential knowledge for anyone considering a career in quantitative modeling or data analysis. You will learn probability and distribution theory on which modern statistical practice is founded, and how to apply this theory to answer important practical questions raised in medical research, ecology, the media and more.

**Relation to other mathematics courses:**

MATH2801 is the entry-point for a statistics major and a prerequisite for most higher level statistics courses. It is also compulsory for mathematics majors to ensure an introduction to statistics as a discipline for studying stochastic (random) systems, as opposed to the deterministic. Statistics has important connections with many branches of mathematics and offers an interesting career path for the mathematically minded.

## 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.

We believe that effective learning is best supported by a climate of inquiry, in which students are actively engaged in the learning process. This course has a strong emphasis on problem-solving tasks in tutorials and in assessments. Students are expected to devote the majority of class and study time to the solving of such tasks.

## Student Learning Outcomes

By the end of this course you should be able to:

- Apply probability rules in a given setting to calculate key quantities
- Use key theoretical tools to explore the properties of random variables
- Apply key methods of statistical inference in applied settings
- Use `RStudio` to summarise data using descriptive statistics
- Use `RStudio` to perform arithmetical and statistical computations
- Use `RStudio` to perform simulations of random variables to explore and demonstrate fundamental concepts and theorems
- *Derive fundamental results in the theory of probability and random variables

* Higher-order skills only expected of Distinction/High Distinction students.

**Relation to graduate attributes** – The learning outcomes for MATH2801 impact on the development of most Science Faculty Graduate Attributes.

**Attribute 1**: Coursework will develop your analytical skills, hence there is a major focus on the attribute *Research, inquiry and analytical thinking abilities.*

**Attribute 2**: Foundation skills in theoretical statistics are essential for higher-level learning in statistics and many aspects of Science, so you will improve the attribute *Capability and motivation for intellectual development.*

**Attribute 4** : Discussions in class and written submissions for assessment will develop your skills at the attribute of *Communication.* The communication of statistical ideas, of the results of statistical inference and the implications of these results are crucial in statistics and throughout science.

**Attribute 6**: Computing skills developed in this course will explicitly develop the attribute of *Information Literacy.*

# Assessment

UNSW assesses students under a standards based assessment policy. For how this is applied in the School of Mathematics and Statistics, see `http://www.maths.unsw.edu.au/currentstudents/assessment-policies`

Assessment in MATH2801 consists of:
an online introductory statistics quiz (10%), an assignment (10%), a mid-session test (20%), and a final examination (60%).

| Assessment: | Due date: | Weight |
|---|---|---|
| Online stats quiz | Week 4 details: TBA | 10% |
| Mid-session (class) test | Week 7 details: TBA | 20% |
| Assignment (take home) | Week 9 details: TBA | 10% |
| Final exam | During UNSW exam period | 60% |

*Every class is different and some variation from the above assessment schedule may be prudent. In the case of assessment dates, no changes will be made without consultation with the class as well as confirmation being announced on Moodle.*

**Online statistics quiz:** The rationale for the quiz is to give students some revision on introductory statistical concepts covered in Weeks 1 to 3. The online quiz will be done on Maple TA (a platform which students can access through Moodle) and will be open for a few days. **No extensions or supplementary quizzes will be permitted**.

**Mid-session test**: The mid-session test will be held in Week 7. More information about **when and where** the class test will be held and **what will be assessed** will be made available closer to the time in lectures and on Moodle.

Students may provide their own UNSW-approved calculator for the mid-session test (calculators will not be provided).

The mid-session test is held under exam conditions. It is designed to give students feedback on progress and mastery of the first parts of the course, **under exam conditions** and to evaluate progress towards the stated learning outcomes.

*Illness and misadventure and the mid-session:* If you miss the mid-session test due to illness or misadventure, then ensure you read the UNSW procedure for applying for Special Consideration. Please see the "Applications for special consideration" section below.

**Assignment:** The assignment will be due in Week 9. The rationale for the assignment is to give students feedback on their progress and mastery of the material, and to obtain measures of student progress towards the stated learning outcomes.

Assessing using take-home assignments rather than under exam conditions offers the opportunity to assess more challenging questions, and gives you the opportunity to think more deeply about your responses.

It also enables the assessment of computer-aided data analysis and problem solving.

Some questions may involve a computing component, for which you are required to use the (free) statistics package `RStudio`, downloadable from `https://www.rstudio.com`

The assignment will be available on Moodle. Precise details of when and how the assignment be presented/submitted will be given on Moodle.

Students are strongly encouraged and expected to attempt the assignment as they are important part of the course at the time when they are due.

*Late assignments*: Late assignments may be accepted up to a few days after the assignment is due, however any late assignments will occur a late penalty in the marks awarded. Please ensure you read the UNSW procedure for applying for Special Consideration, and see the "Applications for special consideration" section below.

## Final examination

A 2 hour examination held during the examination period. The final exam is designed to assess student progress and mastery of the entire course. Further details about the final exam will be available closer to the time in lectures and on Moodle. Please see the "Applications for special consideration" section below if applying for special consideration for the final exam.

### Assessment criteria

The main criteria for marking all assessment tasks will be clear and logical presentation of correct solutions. You will be assessed on the process by which you arrive at solutions as well as the solution itself, so it is important to include your working, and to set it out in a logical fashion. Some of the assessment in MATH2801 and MATH2901 will involve common tasks.

## Additional resources and support

**Textbooks:** The content of the course will be defined by the lectures. The following are recommended additional references.

• Robert V. Hogg, Joseph W. McKean and Allen T. Craig (2005). "Introduction to mathematical statistic", sixth edition. Pearson Education, Upper Saddle River NJ. (In the library, Call number: `S 519.5/98 D` (High Use Collection)).

Hogg *et al.* not only covers MATH2801/2901 content, but also MATH3811, so it would be a useful resource for you throughout a statistics major.

• John A. Rice (2007). "Mathematical Statistics and Data Analysis", third edition. Duxbury, Belmont CA. (Call number: 519.9/569 L (High Use Collection)).

The book by Rice does not follow the content of MATH2801/2901 as closely as Hogg *et al.* but would be a useful source for alternative explanations of key concepts and practice exercises.

**Discussion forum:** Please use the MATH2801 discussion forum on the Moodle course page. You can post your MATH2801-related queries here in this discussion board. Feel free to answer each others' questions.

**School's Student Services:** For any administrative problems please contact either:

Markie Lugton (RC-3072)

Julie Hebblewhite (RC-3088)

Email: ug.mathsstats@unsw.edu.au

# Course schedule

Note that not all topics are exactly a week long but they will correspond approximately to the weeks as indicated.

- Week 1: Introduction to statistics. Includes: course outline; key themes; descriptive statistics.
  The first tutorial will revise some probability concepts.

- Week 2: Random variables. Includes: expectation; moments; standard deviation and variance; Chebychev's inequality and applications; transformations of a single variable.
  **Please bring your laptop to the Week 2 tutorial as we will be learning about `RStudio`.**

- Week 3: Common distributions. Includes: discrete and continuous distributions; quantiles; quantile plots.

- Week 4: Bivariate distributions. Includes: Survey designs and experiments (MATH2801 only); bivariate transformations (MATH2901 only).

**Online Stats Quiz test this week.**

- Week 5: Distribution of sums and averages. Includes: convergence in probability, convergence in distribution, Slutsky's theorem, distributions of sums and averages; convolution formulae; moment generating functions; Central Limit Theorem (CLT); Delta Method.

- Week 6: Estimators and their properties. Includes: parametric models and parameter estimators. Bias, consistency, mean square error, standard error of an estimator, confidence intervals via asymptotic normality, sample size determination.

- Week 7: Parameter estimation and inference. Includes: method of moments and maximum likelihood estimation (mle) and their properties; asymptotic normality of mle; likelihood based confidence intervals.

**Midsession test this week.**

- Week 8: Hypothesis testing. Includes: Wald tests and CLT based tests, Power.

- Week 9: Small sample inference from normal samples. Includes: examples of normal and non-normal sample inference.

**Assignment due this week.**

- Week 10: Inference for categorical data. Includes: chi-squared test an paired samples for Bernoulli variables.

# Applications for special consideration

Please adhere to the University special consideration policy and procedures provided on the web page below.Special Consideration web site: `https://student.unsw.edu.au/special-consideration`

For final exams with special consideration granted, the Exams Unit will email the rescheduled supplementary exam date, time and location to your student zID email account directly.

For successful applications for special consideration only. Please ensure you regularly check your student email account (zID account) for the information on new dates to attend an assessment, or dates for any supplementary exam both in Term and Final.

The supplementary exam period/dates for the final exam can be found at this web site: `https://student.unsw.edu.au/exam-dates`

Please ensure you are aware of these dates and that you are available during this time.

# Administrative matters

**School Rules and Regulations**: See the School of Mathematics and Statistics web page for general policy on additional assessment, and for fuller details of the general rules regarding attendance, release of marks, special consideration etc.

`http://www.maths.unsw.edu.au/currentstudents/assessment-policies`

# Academic integrity, referencing and plagiarism

Academic integrity is fundamental to success at university. Academic integrity can be defined as a commitment to six fundamental values in academic pursuits: honesty, trust, fairness, respect, responsibility and courage.

At UNSW, this means that your work must be your own, and others ideas should be appropriately acknowledged. If you don't follow these rules, plagiarism may be detected in your work.

Issues you must be aware of regarding the university's policies on academic honesty and plagiarism can be found at `https://student.unsw.edu.au/plagiarism` and at the *ELISE* site `http://subjectguides.library.unsw.edu.au/elise`.

Additional resources concerning conduct obligations of students can be found on the *Conduct and Integrity Unit* site `https://student.unsw.edu.au/conduct`