



# COURSE OUTLINE

Never Stand Still

Science

School of Mathematics and Statistics

# MATH5916

# SURVIVAL ANALYSIS

## Semester 1, 2016

# MATH5916 – Course Outline

## Information about the course

**Course Authority:** Jake Olivier

**Lecturer:** Jake Olivier  
Room RC-2051  
Phone 9385 6656  
email [j.olivier@unsw.edu.au](mailto:j.olivier@unsw.edu.au).

**Consultation:** Please use e-mail to arrange an appointment.

### **Credit, Prerequisites, Exclusions:**

This course counts for 6 Units of Credit (6UOC).

There are no prerequisites for this course.

Exclusion: MATH5915.

**Lectures:** There will be one three-hour lecture per week: Tuesday 5.00-8.00pm. Most classes will be in RC-2060, but some (to be advised) will be held in the computer labs.

**Tutorials:** There are no separate tutorials.

**UNSW Moodle:** Further information and course material will be provided via UNSW Moodle.

## Course aims

The aim of this course is to introduce the main statistical concepts, methods and models used in the analysis of survival data.

### **Relation to other mathematics and statistics courses**

Survival Analysis is an elective course for the Master of Statistics program. Whilst there are no formal pre-requisites for the course, it does make use of techniques based on statistical theory such as maximum likelihood estimation and hypothesis testing, and some familiarity with these ideas is assumed. In this sense, the course is related to MATH5905: Statistical Inference which covers the theory behind some of the techniques we will apply. The course is also related to MATH5806: Applied Regression Analysis in dealing with modelling of relationships between variables. The focus here is specifically on models for which the response variable is survival time, however. The course also has links with MATH5906: Design and Analysis of Clinical Trials and MATH5826: Statistical Methods in Epidemiology since survival

time is often a response variable in clinical trials and epidemiological studies. Consideration of multivariate failure time data has some links with longitudinal data analysis in that the survival times may be correlated. Methods for longitudinal studies are covered more generally in MATH5885: Longitudinal Data Analysis. Some of the techniques described in MATH5945: Categorical Data Analysis may also be useful for this course.

## Student Learning Outcomes

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

- Explain the key features of survival data and different types of censoring and truncation
- Define hazard and other basic concepts of survival analysis
- Fit parametric models such as exponential, Weibull, and the general accelerated failure time model to survival data and interpret the results
- Use non-parametric methods such as the Kaplan-Meier estimator and the log-rank test to analyze survival data
- Use the Cox proportional hazards model to examine the effect of covariates on survival
- Extend the Cox model to allow for stratification and time-dependent covariates
- Implement hypothesis tests for survival data
- Calculate residuals and influence for survival models, and assess whether the proportional hazards assumption is justified
- Explain the counting process formulation for survival data
- Explain the notion of frailty in the context of survival analysis
- Solve theoretical problems related to survival analysis
- Use the computer language R to analyze survival data, and interpret R output.

### Relation to graduate attributes

Computing skills developed in this course will improve *information literacy* (Science Graduate Attribute 6).

Assignments, problems, and lab exercises will develop *research, inquiry and analytical thinking abilities* (Science Graduate Attribute 1).

# Teaching strategies underpinning the course

To support the learning outcomes, the course will use the following teaching strategies:

- Lectures, explaining the necessary statistical concepts and theory underlying survival analysis
- Computer labs, providing essential practice in applying the techniques explained in lectures
- Independent study of the course notes and readings, to reflect more deeply on ideas introduced in lectures
- Problems and assignments, giving you an opportunity to independently solve theoretical problems, analyse datasets using statistical software, reflect on aspects of the course, and evaluate your understanding
- Assessment in this course will use problem-solving tasks of a similar form to those in practice problems and computer labs, to encourage the development of the core analytical and computing skills underpinning this course.

## 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 in computer labs and in assessment tasks, and students are expected to devote the majority of their class and study time to the solving of such tasks.

## Assessment

Assessment in this course will consist of two assignments (10% each), a mid-session test (20%), and a final examination (60%).

**Knowledge and abilities assessed:** All assessment tasks will assess the learning outcomes outlined above, specifically, the ability to explain the theory underlying survival analysis techniques, to apply the techniques in analysing real datasets and critically interpret the results of analyses, and to solve theoretical problems related to survival analysis.

**Assessment criteria:** The main criteria for marking assessment tasks involving the explanation of theory and solution of theoretical problems will be clear and logical presentation of correct solutions. In the case of assessment tasks involving

the application of techniques to the analysis of real datasets, the main criteria will be selection and justification of appropriate analysis methods; clear, logical, and well-documented computer code; well-organised output giving evidence of successful implementation; correct interpretation of results; and clear, complete, and fully justified conclusions.

## Assignments

**Rationale:** Assignments will give students an opportunity to practice solving theoretical problems related to survival analysis, and to apply survival analysis techniques in analysing real datasets and interpreting the results of those analyses.

Assignments must be YOUR OWN WORK, or severe penalties will be incurred.

You should consult the University web page on plagiarism  
[student.unsw.edu.au/plagiarism](http://student.unsw.edu.au/plagiarism)

### Schedule and weighting:

Task	Date Avail.	Date Due	Form of Submission	Weighting
Asst 1	Tue 15 March Week 3	5pm Tue 5 April Week 5	Written	10%
Asst 2	Tue 3 May Week 9	5pm Tue 17 May Week 11	Written	10%

Assignments must be submitted by 5.00pm (at the start of class). In general, late assignments will NOT be accepted without good documented reasons (such as illness or misadventure).

## Mid-session test

**Duration and schedule:** 1 hour 15 mins, during the lecture in week 8 (27 April).

**Rationale:** The mid-session test will assess student mastery of the material covered up to that point in the lectures.

**Weighting:** 20% of your final mark.

## Examination

**Duration:** Two hours.

**Rationale:** The final examination will assess student mastery of the material covered in the lectures.

**Weighting:** 60% of your final mark.

Further details about the final examination will be available in class closer to the time.

## **Additional resources and support**

### **Lecture notes**

Lecture notes (including computer lab exercises and practice problems) will be available on UNSW Moodle.

### **Textbooks**

There are no set textbooks.

### **References**

The following references are available in the library:

1. Hosmer, D.W., Lemeshow, S. and May S. (2008). *Applied Survival Analysis*. Wiley.
2. Collett, D. (1994 or 2003). *Modelling survival data in medical research*. Chapman & Hall/CRC.
3. Klein, JP and Moeschberger, ML (1997 or 2003). *Survival analysis*. Springer.
4. Cox, DR and Oakes, D. (1984). *Analysis of survival data*. Chapman & Hall.

Other references will be suggested in class.

### **UNSW Moodle**

All course materials (excluding photocopied handouts) will be available on UNSW Moodle. You should check regularly for new materials.

### **Computer laboratories**

Computer laboratories (RC-M020 and RC-G012) are open 9-5 Monday-Friday on teaching days. RC-M020 has extended teaching hours (usually 8:30-9pm Monday-Friday, and 9-5 Monday-Friday on non-teaching weeks).

## **Library information**

The library website is [library.unsw.edu.au](http://library.unsw.edu.au).

## **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.

## **Administrative matters**

### **Important Information for Postgraduate Course Work Students**

See [www.maths.unsw.edu.au/sites/default/files/pginfo\\_s1\\_2014.pdf](http://www.maths.unsw.edu.au/sites/default/files/pginfo_s1_2014.pdf)

### **School Rules and Regulations**

Fuller details of the general rules regarding attendance, release of marks, special consideration etc are available via the School of Mathematics and Statistics Web page at

[www.maths.unsw.edu.au/currentstudents/assessment-policies](http://www.maths.unsw.edu.au/currentstudents/assessment-policies).

### **UNSW Occupational Health and Safety policies and expectations:**

<https://safety.unsw.edu.au/staff-student-resources/students>

**Student equity and diversity issues:** should be directed to the Student Equity Officers in the Student Equity and Disabilities Unit: phone 9385-4734, website:

<http://www.studentequity.unsw.edu.au/>

**Students with Disabilities:** Further information for students with disabilities is available at

<http://www.studentequity.unsw.edu.au/disability-services>

### **Plagiarism and academic honesty**

Plagiarism is the presentation of the thoughts or work of another as one's own. Issues you must be aware of regarding plagiarism and the university's policies on academic honesty and plagiarism can be found at [student.unsw.edu.au/plagiarism](http://student.unsw.edu.au/plagiarism) and <https://student.unsw.edu.au/conduct>.

## Course schedule

An approximate schedule for the course is given below.

Date of class	Topic	Computer lab	Assessment
Week 1 1 March	Definition, features, and examples of survival data, censoring and truncation, hazard and survival functions. Parametric models: exponential.		
Week 2 8 March	Parametric models: exponential (continued)	Computer lab 1	
Week 3 15 March	Weibull and other parametric models.		Assignment 1 distributed
Week 4 22 March	Non-parametric methods: Kaplan-Meier estimator.	Computer lab 2	
Mid-session break			
Week 5 5 April	Non-parametric methods: log-rank and other tests. Cox PH model: partial likelihood.		Assignment 1 due 5.00pm*
Week 6 12 April	Cox proportional hazards model: applications, model selection.		
Week 7 19 April	Cox model: time-dependent covariates.	Computer lab 3	
Week 8 26 April	Mid-session test		Mid-session test 5.00pm
Week 9 3 May	Introduction to counting processes. Residuals and influence for Cox model.		Assignment 2 distributed
Week 10 10 May	Cox model: residuals and influence (continued)	Computer lab 4	
Week 11 17 May	Non-proportional hazards. Sample size determination.		Assignment 2 due 5.00pm*
Week 12 24 May	Introduction to multivariate survival models.		

\* At the **START** of the lecture