Contents

1 Entry Requirements ............................................. 1
2 Course Administration ........................................... 1
3 Structure of Mathematical Statistics .............................. 2
4 Academic Staff in 2008 and their Research Interests .............. 3
5 Available courses .................................................. 4
6 Project ................................................................... 7
   6.1 General information on projects ............................... 7
   6.2 Proposed project topics ........................................ 8
7 Assessment Procedures ............................................... 13
8 Seminars ............................................................... 14
9 Entitlements ........................................................... 14
10 Scholarships, Prizes and Awards .................................... 15
11 Life after Fourth Year .............................................. 16

For more information on honours study, please visit


and

1 Entry Requirements

Entry to Mathematical Statistics 4 is normally granted to anyone satisfying the Faculty regulations (and subject to the approval of the Head of the School), who has completed 24 credit points of Senior Statistics units with a Credit average or better.

The Faculty of Science currently requires each student enrolling as an Honours candidate in Mathematical Statistics 4 to have qualified for the Pass Degree and also to have accumulated a SCIWAM (Science Weighted Average Mark) of 60 or greater, or the equivalent at another institution.

Note: A SCIWAM is a Weighted Average Mark calculated for use by the Faculty of Science, and includes only Intermediate and Senior units of study. Science students can obtain their SCIWAMs by enquiring at the Science Faculty Office after Third Year examination results have been finalised. Students are advised to check the SCIWAM calculation for themselves. For more details on SCIWAM, see


2 Course Administration

Course Coordinators for Mathematical Statistics 4:

• 2007:

  Dr Rafal Kulik,
  Carslaw Building, Room 818, Phone 9036 7878,
  Email: rkuli@maths.usyd.edu.au

• 2008:

  Dr Michael Stewart,
  Carslaw Building, Room 820, Phone 9351 5765,
  Email: M.Stewart@maths.usyd.edu.au

In 2007 and 2008 the director of statistics teaching program will is

Prof. Neville Weber,
Carslaw Building, Room 823, Phone 9351 4249,
Email: N.Weber@maths.usyd.edu.au

The Course Coordinator is the person that students should consult on all matters regarding Mathematical Statistics 4. In particular, students wishing to substitute a course from another Department, School or University must get prior written approval from the Course Coordinator. Matters of ill-health or misadventure should also be referred to the Course Coordinator.
3 Structure of Mathematical Statistics 4

The program consists of seven courses worth a total of 70 and a project which counts as 30. Only 6 courses in Statistics are given from the statistics teaching group in 2006 year and students need to choose another courses. The courses may be chosen from:

- Honours Statistics courses available at the School of Mathematics and Statistics;
- Honours Applied Mathematics and Pure Mathematics courses available at the School of Mathematics and Statistics. Please contact respective coordinators. The courses which may be of your interest include:
  - Advanced Option Pricing (Applied Mathematics, semester 1)
  - Populations and Disease (Applied Mathematics, semester 1)
- third year advanced courses offered at the School of Mathematics and Statistics;
- selected Honours courses offered at the University of New South Wales;
- grid courses;

Details of the available courses in Applied and Pure Mathematics and can be obtained from the Mathematical Statistics 4 Coordinator or directly from the Pure Maths 4 coordinator

Dr Laurentiu Paunescu,
Carslaw Building, Room 816, Phone 9351 2969,
Email: L.Paunescu@maths.usyd.edu.au.

or the Applied Maths 4 coordinator

Dr Martin Weichselberger,
Carslaw Building, Room 628, Phone 9351 3860,
Email: M.Weichselberger@maths.usyd.edu.au.

Students must select their courses after consulting the Honours supervisor and the Honours Coordinator. Each student is also required to write a project on a Mathematical Statistics topic under the supervision of a member of staff from the Statistics Research Group. This project is counted as the equivalent of three options when calculating the overall mark in Mathematical Statistics 4. Due dates for essays and projects are given in Section 6.
## Academic Staff in 2008 and their Research Interests

<table>
<thead>
<tr>
<th>Name</th>
<th>Research Interests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr Jennifer Chan</td>
<td>Generalized linear mixed model, Generalized-t distribution, Geometric process model Likelihood and Bayesian methods, Drop-out &amp; models Scale mixtures distributions, Robustness, Medical and insurance applications.</td>
</tr>
<tr>
<td>Dr Samuel Mueller</td>
<td>Model Selection, Robust Methods, Extreme Value Theory, Applied Statistics.</td>
</tr>
<tr>
<td>Dr Marc Raimondo</td>
<td>Change-point models and wavelet methods. Nonparametric regression and image reconstruction. Inverse problems.</td>
</tr>
<tr>
<td>Dr Shelton Peiris</td>
<td>Time series analysis. Regression models.</td>
</tr>
<tr>
<td>Dr Michael Stewart</td>
<td>Asymptotic methods. Mixture models.</td>
</tr>
<tr>
<td>Dr Jean Yang</td>
<td>Bioinformatics: Data analysis of microarrays, Gene annotation, Sequence information and Clinical data.</td>
</tr>
</tbody>
</table>

Recent publications of these members are available on the Schools website. See the individual staff member for any reprints of their published papers.
5 Available courses

The following courses will be offered in 2008. Note that this list is tentative.

1. **Advanced Time Series Analysis and Forecasting**
   Semester: 2

   *Course description:* This is an advanced course in time series analysis and forecasting.
   *Assumed knowledge:* STAT 3011 Stochastic Processes and Time Series.
   *References:*

2. **Applied Probability and Stochastic Differential Equations**
   Semester: 2

   *Assumed knowledge:* STAT 3911 Stochastic Processes and Time Series (Advanced).
   This course will cover theory of stochastic differential equations. We will discuss properties of Brownian motion, geometric Brownian motion, define stochastic integrals and solve SDEs. Applications to mathematical finance, biology and engineering will be given.

3. **Extreme Value Theory/Asymptotics**
   Semester: 2

   This course has two sections. Section 1 covers applications of extreme value theory (e.g., weather, insurance and finance), extremes of iid sequences (max-stable distributions, the extremal types theorem, domains of attraction), exceedances (Poisson behaviour, generalised Pareto approximations), extremes of dependent Gaussian sequences and extremes in continuous time.

   Section 2 deals with asymptotics: The $\delta$-method, transformations, functional statistics and the influence function, moments, M-estimates, robust statistics.
4. **Generalized Linear Models**  
Semester: 1

This course covers both theory and application of generalized linear models.

*Assumed knowledge:* STAT 3912 Applied Linear Methods (Advanced); STAT 3914 Applied Statistics (Advanced).

*Contents:* Review of standard normal error linear models; model checks; weighted least squares and asymptotic results. General theory: two parameter exponential family; generalized linear model and the algorithm for maximum likelihood estimation; deviance and link functions; model fitting diagnostics. Binary data analysis: application of general theory to logistic and binary response designs; $2 \times 2$ contingency tables; matched case control study models. Survival analysis; Cox regression. Contingency tables: Hierarchical log-linear models; incomplete tables; quasi independence and quasi symmetric models.

*References:*

5. **Probability**  
Semester: 1

This is a rigorous course on probability with a measure theoretic basis.

*Assumed knowledge:* STAT 2911 Probability and Statistical Models (Advanced) + real variable analysis. A knowledge of Measure Theory would be an advantage.

*Contents:* Axiomatic probability: probability space; continuity of probability measures; independence; product spaces; conditional probability and conditional expectations with respect to a given $sigma$-field; inequalities. Modes of convergence: almost sure convergence; convergence in probability; convergence in distribution. Characteristic functions: properties; inversion theorem and continuity. The Helly-Bray lemma; convergence via characteristic functions. Limit theorems: Laws of Large Numbers; Central Limit Theorem (Lindeberg); infinitely divisible distributions.

*References:*
6. **Smoothing and computational statistics**  
Semester: 1

*Outline:* In this course we present an overview of modern statistical methods for curve estimation. This includes the study of kernel smoothing and orthogonal series estimators in nonparametric models (density, regression). We discuss theoretical properties such as consistency, optimality and we illustrate computational properties such as tuning by cross-validation, model selection and thresholding algorithm.

*References:*

6 Project

Each student is expected to have made a choice of a project and supervisor by the beginning of the first semester (or beginning of the second semester for students starting in July), and to commence work on their project immediately. The supervisor must be a member of staff of the Statistics Research Group. Students should consult their appointed supervisor regularly, in both the researching and writing of the work. Three copies of the final project, typed and bound, must be handed to the Course Coordinator by 13 June 2008 (for mid-year completion) or 31 October 2008 (for end-year completion).

6.1 General information on projects

1. Students should choose a topic and a supervisor by January (after consulting Head of Statistics Program / Course Coordinator) and should consult their supervisor weekly during the year.

2. A typed report of about 60 pages counts for 30% of the years assessment.

3. The project will be assessed by three members of staff (including the supervisor). The overall final mark for the project will be a weighted mean of all three grades awarded. A weighting of 50% will be attached to the supervisors original grade, while weight of 25% will be attached to each of the two grades awarded by the other examiners.

4. A list of suggested project topics is available in Section 6.2. However, students are free to choose another topic subject to the availability of a supervisor and the approval of the Course Coordinator.

5. Guidelines to students:
   (a) Have you taken courses closely related to the area of your project?
   (b) Read/study for an adequate coverage of the topic. Consult key sources with the help of your supervisor.
   (c) Integrate the sources in (b) well with your knowledge.
   (d) Original theoretical research is not expected, but is encouraged.

6. Each student will give a 30 minute seminar in October (or May) on their project.

7. Guidelines for the final version: Check the following steps before you submit 3 copies to your supervisor.
   (a) Is there an adequate introduction?
   (b) Have the chapters been linked so that there is overall continuity?
8. Examination of the Project:

I. CRITERIA for each of which marks are to be awarded by each examiner:
   (a) quality of synthesis of material in view of difficulty and scope of topic, and originality, if any.
   (b) evidence of understanding.
   (c) clarity, style and presentation.
   (d) mathematical and/or modelling expertise and/or computing skills.
   (e) presentation of seminar.

II. GUIDELINES to help formulate the assessment: Check whether the criteria above are clearly satisfied. In addition the following points should be addressed by the supervisor:
   (a) Has the student shown initiative hard work which are not superficially evident from the written report?
   (b) Has the student coped well with a topic which is too broad or not clearly defined?

9. Students are advised to read the pamphlet entitled Guide to Essay Writing for Science Students available from the Science Faculty Office.

6.2 Proposed project topics

The following is a list of possible project topics in 2008. Prospective students interested in any of these topics are encouraged to discuss them with the named supervisors as early as possible. However, this list is not exhaustive. Students may wish to suggest their own topics for essays or projects. Before commencing work, however, each student must find a member of staff who will agree to supervise the project.


   Project description: This piece of research relates to the analysis of trend data of counts using a new model called the Geometric Process model. Some time series of count data will be analysed using the models. Different trend patterns will be studied and compared. Students are required to write programs and/or using statistical softwares.
2. **Heavy-tailed distributions for insurance data.** Supervisor: Dr Jennifer Chan.

*Project description:* This piece of research relates to the robust analysis of heavy-tailed distributions, including the Student-t, exponential power and generalized-t distributions, etc. These distributions are applied to the modelling of insurance data that often contain outliers. The robust properties across these distributions will be compared. Students are required to write programs and/or using statistical softwares.

3. **Model Selection in Regression Type Models.** Supervisor: Dr Samuel Mueller.

*Project description:* Model selection is fundamental to the practical application of statistics. The most common models in statistics are regression type models. For such models there are a vast number of different approaches and a substantial literature on the subject. A project in this area could deal with one or more of the recently published model selection criteria. There are always many open questions. E.g. What is the performance compared to other criteria, what are the robustness properties, what is the best choice of the selection criterion parameters?


*Project description:* A project in this area could deal with the analysis and modeling of data from an experiment, a survey, a clinical study, or some other possible source. The data is properly described and insight into the structure of the data is gained. All hypotheses and/or aims are analyzed with at least two appropriate statistical procedures (e.g. parametric and nonparametric, exact and asymptotical, bayesian and frequentist point of view). The results are correctly reported. A discussion is included that presents ideas, problems, suggestions, and other aspects for further studies.

5. **Non-parametric diffusion models with application to financial data.** Supervisor: Dr Marc Raimondo.

*Project description:* We review several non-parametric methods for the estimation of the diffusion coefficient. Both state-smoothing and time-smoothing approaches are studied. The methods are illustrated and compared on a wide range of simulated examples. Some applications to financial data analysis are presented.

6. **Density estimation when the data are contaminated by noise.** Supervisor: Dr Marc Raimondo.
Project description: Wavelet like behaviour of Slepian functions (Prolate Spheroidal Wave functions) allows non-linear multiresolution approximation to a generic density function. The density estimators based on this technique are entire functions of exponential type but nonetheless do have a strong localisation property. This avoids (/reduces) the excessive oscillations (Gibbs phenomenon) commonly observed with wavelet-type density estimation.


Project description: The asymptotic distribution theory for eigenvectors using perturbation methods is given in some work of Tyler. We would examine this technique for principle components. This is a largely theoretical project.


Project description: This project would compare the power of some test statistics used in phylogenetics to test for homogeneity and stationarity of a model. The tests compare some “pivotal” statistics with others which do not have this property. The project is largely in the area of computational and applied statistics.


Project description: Addressing certain open problems concerning the selection of the number of different regimes in a multiple-regime autoregressive model with normal errors where regimes switch independently of the data.

10. Inference about the smoothness of a mixing distribution. Supervisor Dr Michael Stewart.

Project description: An arbitrary mixture of normals with equal variance can be used as a density estimate, however it is not clear how to choose the common variance, which in this setting plays the role of a smoothing parameter. We study this problem by examining various computational and theoretical issues.

11. Limit theorems for self-normalized sums with applications in financial econometrics. Supervisor Dr Qiying Wang

Project description: Sample mean plays a very important role in the development of probability and statistics. The classical limit theorems in connection with sample mean, such as the central limit theorems, Berry-Esseen bound, Edgeworth expansion as well as the large deviations, are
celebrated results in probability theory and have lots of applications in
statistics and other fields. However, many classical results are not so na-
ture from a statistical point of view because the parameters involved in
these classical limit theorems are usually unknown, one has to estimate
them first and then apply the estimators in the classical limit theorems.
Limit theorems for self-normalized sums put a totally new countenance
upon the classical limit theorems. The results in this area not only are
more natural from a statistical point of view because they waive the un-
known parameters, but also they are valid without moment conditions or
under little moment conditions, and they are much neater.

In this project, we will consider the applications of self-normalized limit
theorems in financial econometrics. In view of the fact that, many test
statistics raised in financial econometrics (like unit root tests) may be
written as a function of self-normalized sums, it is expected that self-
normalized limit theorems would be powerful tools to investigate the limit
behaviors of many important statistics in the related areas.

Students are expected to summy the current development in the related
areas, to make some simple extensions and simulations of theoretical re-
sults. The interests in theoretical research is essential for this project.

12. **Nonparametric cointegrating regression: theory and practice**.
Supervisor Dr Qiying Wang

*Project description:* Consider a non-linear cointegrating regression model:
\[ y_t = f(x_t) + u_t, t = 1, ..., n, \]
where \( u_t \) is a stationary error process and \( x_t \) is a regressor. The estimation
of \( f(x) \) is very popular in theoretical research and applications. There are
several books related to the estimation when \( x_t \) is a stationary regressor.
Due to the difficulty in asymptotic results, there are only few papers in the
investigation related to the estimators of \( f(x) \) when \( x_t \) is a nonstationary
process like a random walk. Theoretical research would be very challenge
in this area.

Students are expected to summy the current development in the related
areas, to make some simple extensions and simulations of theoretical re-
sults. The interests in theoretical research is essential for this project.

13. **Residuals for linear models with a general covariance structure**.
Supervisor: Prof. Neville Weber.

*Project description:* Haslett and Haslett (International Statistical Review
(2007), 75, 1-24) gives a survey of 3 types of residuals that can be used
when the error terms in a linear model are not iid. How are these resid-
uals defined and how do they differ? What are the implications of these
different definitions for generalised linear models?

*Project description:* There are various papers using semi-Markov chain models to study different sports (e.g., Meyer, Forbes and Clarke (J. Sports Science and Medicine (2006), 5, 525-532)). How do these models work and can they be applied to other sports? Alternatively Barry and Hartigan (J. Applied Science (1993), 1, 323-336) use a Bayesian approach via Markov Chain Monte Carlo fit a logistic model to predict baseball outcomes. Can this approach be applied to cricket or NRL?

15. **U-statistics based on samples from finite populations.** Supervisor: Prof. Neville Weber.

*Project description:* Review the different techniques used to obtain limit theorems for U-statistics based on samples obtained by sampling without replacement from a finite population. In particular consider the implications of a recent paper on large deviations by Hu, Robinson and Wang (Annals of Statistics (2007), 35, 673-696) for U-statistics.

16. **Meta Analysis for Gene Expression Data.** Supervisor: Dr Jean Yang.

*Project description:* Successful integration of hundreds of experiments obtained by multiple research groups promise to provide new insights into biological systems by exponentially increasing available data. The integration of these data will further enhance our understanding of the biological processes. This research involves evaluating and understanding various integration approaches (meta-analysis) to combine gene expression data from multiple centers.

17. **Gene Set Analysis.** Supervisor: Dr Jean Yang.

*Project description:* One of the major challenges facing biologists is to understand the regulation underlying the biological system of interest. One approach involves jointly analyzing gene expression and miRNA information to find ”gene-module enrichment” among co-regulated expression. This research involves evaluating and understanding various ”gene set analysis” approaches to combine multiple sources of relevant data.

18. **Integrative analysis in bioinformatics.** Supervisor: Dr Jean Yang.

*Project description:* This research relates to Gibbs-sampling, expectation-maximization (EM) techniques for parameter estimation. These are methods used in de-novo motif finding algorithms that model motif sites as missing data in statistical models. Work in this area will lead to development of statistical and computation methods for identifying regulatory element in microarray experiments.
7 Assessment Procedures

Students are required to attend a minimum of 7 courses during the academic year. Only the best 7 results will be included in the overall assessment. These 7 results are weighted equally.

Student performance in each Mathematical Statistics 4 course is assessed by a combination of assignments and examinations. The assignment component is determined by the lecturer of each course and the examination component makes up the balance to 100%. The lecturer converts the resulting raw mark to a mark on the Faculty scale, which indicates the level of Honours merited by performance in that course alone.

<table>
<thead>
<tr>
<th>Grade of Honours</th>
<th>Faculty-Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Class, with Medal</td>
<td>95–100</td>
</tr>
<tr>
<td>First Class (possibly with Medal)</td>
<td>90–94</td>
</tr>
<tr>
<td>First Class</td>
<td>80–89</td>
</tr>
<tr>
<td>Second Class, First Division</td>
<td>75–79</td>
</tr>
<tr>
<td>Second Class, Second Division</td>
<td>70–74</td>
</tr>
<tr>
<td>Third Class</td>
<td>65–69</td>
</tr>
<tr>
<td>Fail</td>
<td>0–64</td>
</tr>
</tbody>
</table>

The examiners recommendation to Faculty of the award of Honours is based on the average mark achieved by each student, over the 7 best courses and the Project. Courses account for 70% of the assessment and the Project for the remaining 30%.

Note: All assessable student work (such as assignments and projects) should be completed and submitted by the advertised date. If this is not possible, approval for an extension should be sought in advance from the lecturer concerned or (in the case of Honours projects) from the Course Coordinator. Unless there are compelling circumstances, and approval for an extension has been obtained in advance, late submissions will attract penalties as determined by the Board of Examiners (taking into account any applications for special consideration).

Appeals against the assessment of any component of the course, or against the class of Honours awarded, should be directed to the Head of School.

Note: Students who have worked on their projects as Vacation Scholars are required to make a declaration to that effect in the Preface of their theses.
8 Seminars

Mathematical Statistics seminars are usually held fortnightly on Friday afternoons. These seminars are an important forum for communicating ideas, developing critical skills and interacting with your peers and senior colleagues. Seminars are usually given by staff members and invited speakers. All full-time research students and all Mathematical Statistics 4 students are expected to attend these seminars.

Each Honours student must give a seminar in their final semester. Mathematical Statistics 4 students deliver a half-hour seminar on their project. Students can obtain advice on seminar presentation from their supervisor. Students are expected to give the title and abstract of their talks to a seminar organizer at least two weeks in advance.

9 Entitlements

Mathematical Statistics 4 students enjoy a number of privileges, which should be regarded as a tradition rather than an absolute right. These include:

- Office space and a desk in the Carslaw building.
- A computer account with access to e-mail and the WorldWideWeb, as well as TEX and laser printing facilities for the preparation of projects.
- A photocopying account paid by the School for assembling project source material.
- After-hours access key to the Carslaw building. (A deposit is payable.)
- A pigeon-hole in room 728 please inspect it regularly as lecturers often use it to hand out relevant material.
- Participation in the Schools social events.
- Class representative at School meetings.
10 Scholarships, Prizes and Awards

The following scholarships are available:

University of Sydney Honours Scholarships

See http://www.usyd.edu.au/scholarships/current/index.shtml#hons

The NSW Institute of Sport Mathematics and Statistics Scholarship

The NSW Institute of Sport Mathematics and Statistics Scholarship was established in 2007 by an offer of funding from the NSW Institute of Sport. The recipient will be undertaking a research topic related to the application of mathematics or statistics to modelling that is relevant to sports. The scholarships shall be awarded on the basis of an application, which shall include a description of the proposed research topic. Selection will be made on the basis of academic merit and relevance of the proposed research topic.

The following prizes may be awarded to Mathematical Statistics 4 students of sufficient merit. Students do not need to apply for these prizes, which are awarded automatically. Have a look at

http://www.maths.usyd.edu.au/u/About/prizes.html

for more details.

The Joyce Prize
Awarded annually to the most outstanding student completing fourth year Honours in Applied Mathematics, Pure Mathematics or Mathematical Statistics (provided the work is of sufficient merit).

George Allen Scholarship
This is awarded to a student proceeding to honours in Mathematical Statistics who has shown proficiency in all Senior units of study in Mathematical Statistics.

University Medal

Awarded to Honours students who perform outstandingly. The award is subject to Faculty rules, which require a mark of at least 90 in Mathematical Statistics 4 and a SCIWAM of 80 or higher. More than one medal may be awarded in any year.

Ashby Prize
Offered annually for the best essay, submitted by a student in the Faculty of Science, that forms part of the requirements of Pure Mathematics 4, Applied Mathematics 4 or Mathematical Statistics 4.

**Barker Prize**

Awarded at the fourth (Honours) year examination for proficiency in Pure Mathematics, Applied Mathematics or Mathematical Statistics.

**Norbert Quirk Prize No IV**

Awarded annually for the best essay on a given mathematical subject by a student enrolled in a fourth year course in mathematics (Pure Mathematics, Applied Mathematics or Mathematical Statistics) provided that the essay is of sufficient merit.

**Australian Federation of University Women (NSW) Prize in Mathematics**

Awarded annually, on the recommendation of the Head of the School of Mathematics and Statistics, to the most distinguished woman candidate for the degree of BA or BSc who graduates with first class Honours in Applied Mathematics, Pure Mathematics or Mathematical Statistics.

## 11 Life after Fourth Year

Students seeking assistance with post-grad opportunities and job applications should feel free to ask lecturers most familiar with their work for advice and written references. The Head of Statistics Programme, the Course Coordinator and the course lecturers may also provide advice and personal references for interested students.

Students thinking of enrolling for a higher degree (MSc or PhD) should see the Director of Postgraduate Studies Dr David Easdown (or Professor N. Weber) for information. Students are also strongly encouraged to discuss potential research topics with individual staff members.

Students who do well in Mathematical Statistics 4 may be eligible for postgraduate scholarships, which provide financial support during subsequent study for higher degrees.

Last but not least, there is a number of jobs for people with good statistical knowledge. Have a look at [http://www.statsci.org/jobs/](http://www.statsci.org/jobs/).