Unit Lecturer: John Lau (john@maths.uwa.edu.au or Room 1.07).

Contact Hours: You are welcome to pass at any time at my office to discuss unit related problems that you may have. Monday and Thursday will usually be the best days to catch me in my office. If you wish you can book an appointment by emailing me.

Lectures: Scheduled for 9.00am-9.45am on Mondays and Tuesdays, 11.00am-11.45am on Thursdays. All lectures are in Maths Lecture Room 1.

Practicals: 10.00-10.45am Tuesdays in the Mathematics Computing Laboratory (MCL), starting in week 2. These sessions will give you experience using a modern statistics package (R) and will also help re-inforce concepts learnt in lectures. All of the assessment in this unit will require you to use R, so the practical sessions are vital.

Prerequisites: Either,

- MATH1020 Calculus, Statistics and Probability (formerly 530.101 Mathematics 101)
- STAT1160 Statistics A (formerly 530.160 Statistics 160) and, at least concurrently, MATH 2030 Calculus and Matrix Methods or MATH2210 Calculus and Algebra (formerly 530.210 Mathematics 210)
- STAT1106 Economic and Business Statistics (formerly 535.106 EBS 106) and, at least concurrently, MATH2030 Calculus and Matrix Methods or MATH2210 Calculus and Algebra (formerly 530.210 Mathematics 210)

WWW: The unit web page is at


On this web site I will place all lecture notes, assignments etc for this course. For the first lecture I will bring copies of this course outline and the 1st week’s lecture notes. After that it will be your responsibility to print these out. I will attempt to post the following week’s lecture notes by Friday at the latest.

Unit Aims and Objectives: The overall aim of this unit is to develop skills in the statistical analysis of data from designed experiments and observational studies. Specific learning objectives are

- Understanding the fundamentals of good design in experiments and studies;
- Understanding the elements of a linear model;
- Ability to develop and apply linear models for data from real-world experiments and studies;
- Proficiency with a statistical computer package for linear modelling
Suggested Reading: There are no required texts for this unit. However, the books and WWW resources listed below may prove helpful.

- Everitt, B.S., Northorn, T. (2006). A handbook of Statistical Analyses Using R. Chapman and Hall. (while this book is not in the library, it is a good book and would serve you well for most further courses in statistics. It can be ordered through the bookshop.)
- Faraway, J.J. (2002). Practical Regression and ANOVA using R.
  http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf
- R Development Core Team: An Introduction to R.
  http://cran.r-project.org/doc/manuals/R-intro.pdf
  http://www.math.montana.edu/Rweb/Rnotes/R.html

In addition, The Comprehensive R Archive Network site at http://cran.r-project.org/ contains lots of useful material. In particular, you can download your own free copy of the R package, and browse through the FAQs (Frequently Asked Questions).

Computing: The computer package R will be used during this unit. It can be freely downloaded from The Comprehensive R Archive Network site given above.

Assessment: The following assessments take place:

<table>
<thead>
<tr>
<th>method of assessment</th>
<th>tentative timing</th>
<th>weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>take home assignment 1</td>
<td>about the mid semester break</td>
<td>10%</td>
</tr>
<tr>
<td>take home assignment 2</td>
<td>about week 11</td>
<td>10%</td>
</tr>
<tr>
<td>quiz 1</td>
<td>about week 5</td>
<td>10%</td>
</tr>
<tr>
<td>quiz 2</td>
<td>about week 11</td>
<td>10%</td>
</tr>
<tr>
<td>3h examination in MCL</td>
<td>June examination period</td>
<td>60%</td>
</tr>
</tbody>
</table>
Details on the Faculty’s policies on assessment can be viewed at

- http://www.ecm.uwa.edu.au/studentnet/exams/assessment, which contains information on scaling of marks, amongst other things;
- http://www.ecm.uwa.edu.au/studentnet/exams/dishonesty, which contains information on academic dishonesty;
- http://www.ecm.uwa.edu.au/studentnet/exams, which contains information on examinations and appeals;

Unit Synopsis: The unit will cover:

- **Introduction**: Unit information; motivational examples; basic concepts in design of experiments and studies; fundamentals of linear models; the perils of bad design and statistical analysis.

- **Introduction to the Software Package R**

- **Simple Linear Regression**: Revision of the basics; simple linear regression model; least squares parameter estimation; sampling distributions of parameter estimates; estimation of error variance; assessing the fit of a regression line; decomposition of sums of squares and ANOVA tables for simple linear regression; the F distribution and its use in testing; connection between F and t tests; model diagnostics (residual plots, normal probability plots, outlier detection); leverage; transforming the data; prediction (prediction intervals versus confidence intervals); examples and implementation using R.

- **Multiple Linear Regression**: Introduction to the multiple linear regression model; parameter estimators and their sampling distributions; F-testing overall model fit; what makes a ‘good model’ and the principle of parsimony; nested models and their comparison; variable selection; ANOVA tables for multiple linear regression models; the stepwise algorithm for variable selection; polynomial regression; prediction and prediction intervals; model diagnostics; multicollinearity and use of orthogonal polynomials; models that include categorical variables; dummy variables; matrix formulation of the general linear model; examples and implementation in R.

- **Analysis of Experimental Data**: Revision of the basics (factors, levels and experimental units; randomization); blocking; model parameterization and interpretation, viewing as multiple regression model; contrasts; two-way ANOVA models; crossed factors and interactions; interaction plots; testing for interactions and main effects; testing for linear, quadratic and higher order polynomial effects of a quantitative factor; many-way balanced ANOVA; introduction to complete and incomplete designs; analysis of covariance; brief overview of unbalanced designs; nesting of factors; model diagnostics, testing for homogeneity of variance; examples and implementation in R.

- **Outlook on Logistic Regression and Generalized Linear Models**