“An engineer who does not know experimental design is not an engineer.”
(Remark made by an executive of the Toyota Motor Company to one of the authors of Statistics for Experimenters.)
STAFF CONTACT DETAILS

Course Convenor  Ross Odell
and lecturer  Room 507, Samuels Building
Phone:  9385-3920
Email:  r.odell@unsw.edu.au
Consultation times:  whenever my office door is open. Or by phone.

COURSE INFORMATION

Lecture  Biomedical Theatre E
Time  Thursday 6-7:30 pm
Laboratory  GS8mE Green Room (Samuels Room 518/519)
Time  Thursday 7:30-9 pm or
       Friday 10 – 11 am

Units of credit:  6

This course is offered in the Master of Engineering Science and related programs. It is part of the Engineering and Technical Management specialisation. Its content is not specific to any engineering field and it is potentially of interest to students in all specialisations. It is also open to students in the Master of Biomedical Engineering program and to students in the fourth or fifth years of the various concurrent BE/MBiomedE programs.

The course is also available to research students, Masters and PhD, who will find the coverage of design of experiments useful.

This course covers selected topics in statistics that are of relevance for engineers in general. Topics include: a brief review of basic statistical concepts and methods (random variables, statistical models and hypothesis tests); linear models (analysis of variance and regression); the design of efficient experiments, including factorial experiments; models for discrete data; and the effective display of data using tables and graphs.

This course will cover

- The principles of efficient experiment design, including multifactorial experiments and response surface methods
- The common statistical models that are used in the analysis of data from designed experiments in engineering. These include linear statistical models such as simple and multiple regression and analysis of variance.
- An introduction (only) to statistical models that are used for special types of data. An example would be logistic regression for binary data.
- Effective communication of experimental results using graphs and tables
- Other topics that might be of interest to student in the class. If there is a topic that you think we should cover, please make a suggestion in the first weeks of the course. It may be possible to cover it, given advanced notice.
This course may be of interest to

- Postgraduate research students who need to know how to design effective experiments and how to analyse the results of those experiments. Where do you begin? What will you measure? What factors might affect the outcome? Which factors will you control? Which will you vary?
- Postgraduate coursework engineering students who want to learn the statistical methods commonly used in their field and how to design experiments involving many factors, perhaps in the context of process optimisation. If you have 6 factors that you can adjust, how do you approach this problem efficiently?
- Other engineering students who are interested in statistics and experiment design.

Presumed knowledge and skills.

Mathematics:
It is expected that you have had some university level mathematics and an introductory statistics course, although the latter is not an absolute necessity. If you have not taken a statistics course previously, you must be prepared to put in extra effort for the first few weeks to make sure that you understand the basic concepts.

Aims:
On completion of the course you should be familiar with the language of statistics and have a solid understanding of the fundamental concepts. You should be able to design an efficient experiment; to analyse correctly the results of that experiment; to choose the appropriate model and method for the analysis of engineering data in general; and to understand the statistics encountered in the engineering literature.

Learning outcomes – knowledge

(Following each outcome is a list of the assessment tasks most related.)

On completion of this course, you should:

- Understand concepts such as random variable, probability, distribution, independence. (A1, Exam)
- Understand the basic linear statistical models such as those used in regression and analysis of variance. (A2 - A5, Exam)
- Understand the logic of hypothesis tests and the meaning of the terms "significance level" and "power". (A2 - A5, Exam)
- Have an appreciation of the wide variety of statistical models that are used in engineering and an understanding of some of them. (A2 - A5, Exam)
- Understand the principles of good experiment design (A2,A3,A5,Exam)

Learning outcomes - skills

On completion of this course, you should be able to:

- Design an efficient experiment involving several factors. (A3,A5,Exam)
- Estimate the power of the experiment and how many subjects would be needed to achieve a specified power. (A2,Exam)
- Analyse correctly the results of the experiment. (A2-A5,Exam)
- Choose the appropriate model and method for the analysis of engineering data in general. (A2 - A5,Exam)
- Understand the statistics encountered in the engineering literature. (A2 - A5,Exam)
- Be able to perform the requisite analyses using a statistics program. (A2 - A5, Exam)

These learning outcomes relate most strongly to the following UNSW graduate capabilities:

**Scholars who are:**

- understanding of their discipline in its interdisciplinary context
- rigorous in their analysis, critique, and reflection
- able to apply their knowledge and skills to solving problems
- digitally literate

**Teaching strategies**

Each week there will be a lecture that covers the relevant statistical model or models that underlie the methods used to analyse a particular type of experimental data. This course is not intended to be strongly theoretical but you must understand the basics of the models and their assumptions.

Each week several problems will be assigned, generally from the textbook. Answers/solutions will be supplied. At the end of lecture, there will be time to ask questions about the questions from the previous week. You are expected to have attempted the problems and the selection of problems to review in class will be student driven - you tell me which problems you want me to do. Assignments will also be discussed during the tutorial following the submission date.

The final component will be a 1-1.5 hour computer laboratory session using the statistics program Minitab®. Detailed step-by-step lessons will be supplied and tutorial assistance will be available. The computer laboratory is also available outside of class time and it is expected that you will continue the tutorial exercises and use Minitab for problems from the textbook on your own time. Educational prices for Minitab are quite low and if you feel that you do not need tutorial assistance, you can do the exercises at home.

The approach taken in this course is that statistics is a tool to be used by the engineer/researcher. It will be used correctly if the methods are understood and if they are implemented correctly. This requires some theoretical understanding and extensive practice in analysing data produced in engineering experiments.

**Suggested approach to learning**

Statistics should not be a difficult subject but there is no denying that it often is difficult. It demands a mode of thinking that is different from that used in most other subjects. It extremely important to grasp the fundamental concepts as soon as possible and to ask for help if you do not understand.

Attend the lectures. If something is unclear, ask questions. It is difficult, because you think that you are the only one who doesn't understand, but you are probably not alone. And lecturers appreciate questions, believe it or not.

Review the lecture notes and read any other material that is distributed. There will be points in the lecture that were not clear. Read the assigned section of the textbook to clarify. Attempt the assigned textbook problems. If you get the same answer as in the book, good. If not, raise questions in class.
The tutorials using Minitab in the computer laboratory are extremely important. This program will probably be new to you. If you have trouble, it is important that you ask for help as soon as possible. A few extra tutoring sessions will probably get you going, but you must ask for help.

The University expects that you will devote 10-11 hours per week to this course. In addition to the 3 hours in class, you should spend 7-8 hours reading and working on problems and assignments.

Course improvement
This course is periodically evaluated by student surveys (CATEI) and other means and changes are made in response to the feedback we get. These will be discussed in the first lecture. A continuing issue is that not enough time is spent on example problems. It has proved difficult to problems in the lecture setting. In session 2 we will try other strategies to solve this problem.

Topics

Introduction (1 week)
  - Probability
  - Random variables and their distributions
  - Summarising data
  - Probability plots

Estimation (1 week)
  - Point estimation
  - Interval estimation

Hypothesis testing, significance, power (1 week)
  - Comparing means - the t-test
  - Comparing proportions (z-test, chi-squared test)
  - Goodness of fit tests

Linear models (4 weeks)
  - Simple linear regression
  - Analysis of variance, one-way, two-way
  - Multiple linear regression

Design of experiments (3 weeks)
  - What is good design? Basic principles
  - The randomised block design, (Latin square, split plot)
  - Factorial designs
  - Fractional factorial designs
  - Response surface methods, optimisation

Other
  - Non-parametric methods
  - Graphing
  - Modelling dichotomous outcomes (logistic regression)
Students are encouraged to contribute their own problems from research, work, or from reading the literature in their field. There is scope in the course plan for covering special topics, if desired.

**Schedule**

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Lecture topics</th>
<th>Assignment due</th>
<th>Quiz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31-Jul</td>
<td>Probability, random variables, distributions, summarising data, probability plots</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7-Aug</td>
<td>point estimation, interval estimation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>14-Aug</td>
<td>t-test, test of proportions, Goodness of fit tests</td>
<td>Assignment 1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>21-Aug</td>
<td>Bivariate distributions, Linear regression, Chi-squared test</td>
<td>Quiz1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>28-Aug</td>
<td>1-way anova</td>
<td>Assignment 2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4-Sep</td>
<td>2-way anova, higher order models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>11-Sep</td>
<td>Multiple regression model building</td>
<td>Assignment 3</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>18-Sep</td>
<td>Design of experiments 1</td>
<td>Quiz2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>25-Sep</td>
<td>Design of experiments 2</td>
<td>Assignment 4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-Oct</td>
<td>(Break)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>9-Oct</td>
<td>Design of experiments 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>16-Oct</td>
<td>Nonparametric methods, graphing</td>
<td>Assignment 5</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>23-Oct</td>
<td>Other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Assessment**

**Assignments (40%)**

There will be 5 assignments, not of equal weight, during the term. The weights will be (in order) 5, 6, 9, 5 and 15% of the final mark. Each will require the analysis of a set of data. In some assignments, you will acquire the data using a simulator. The last assignment will require the design of a multifactor experiment and the generation of the data to be analysed using a simulator.

The assignments are a direct test of the degree to which the learning outcomes listed above have been achieved, as indicated by the list of assignments following each outcome.

**Expectations regarding the assignments**

Full marks will be given if the analysis is appropriate, if the interpretation of the data is correct, if any supplementary tasks are completed correctly, and if the submitted document is clear and contains a limited (but sufficient) amount of computer output and graphs. Inappropriate analysis will be penalised heavily; incorrect interpretation will be penalised less. Minor errors will be penalised lightly or not at all.
Minitab output should be copy/pasted from Minitab into a word processing program. It should be formatted as Courier New font. Any Minitab output must be accompanied by an explanation and highlighting indicating the relevant parts. Excessive computer output without explanation will be penalised.

All submissions must be accompanied by a Graduate School of Biomedical Engineering (GSBmE) non-plagiarism declaration, which will include a statement about its originality.

Late submission will draw a penalty of 10% of the assigned weight per day unless prior approval of late submission has been granted.

**Quizzes (10%)**
These will test understanding of a selection of the concepts that have been covered. Their purpose is to give me a measure of how well you understand the material and to give the you a measure of your understanding relative to my expectation.

**Final examination (50%)**
The exam will comprise around 10 questions. Most will relate to the analysis of a set of data. A correct answer will include the appropriate model/statistical method along with assumptions and either correct calculations or correct interpretation of computer output.

The exam will be closed book. A page of formulas and all required statistical tables will be supplied. A mark of 40% or greater must be achieved in the final exam in order to pass the course.

**RESOURCES**

**Statistical Computing**
The statistics package Minitab® will be used for this course. You will have access to the GSBmE computing laboratory in the Samuels Building room 518/519(fifth floor). You might want to buy or rent a copy for use on your own computer. Prices are reasonable: $30 (USD) for 6-month rental, $50 (USD) for one year.

**Textbook (required)**
Montgomery DC, Runger GC, Hubele NF: *Engineering Statistics* 5ED
ISBN 0470631473.

**Notes**
Extensive written statistics notes will be supplied along with the lecture notes.

**Learning management system**
We will use Moodle to supplement the course. It can be accessed from [http://moodle.telt.unsw.edu.au/](http://moodle.telt.unsw.edu.au/)
Notes, tutorials and other materials will be posted on Moodle.

**School policies on academic matters and plagiarism**
Please consult the policies of the University on [academic misconduct](http://). You should pay particular attention to the University's policy on plagiarism. If you are at all uncertain about what constitutes plagiarism and how to avoid it, The Learning Centre has excellent advice on [recognising and avoiding plagiarism](http://).
All assignments submitted in this course must be accompanied by a signed GSBmE cover sheet from the School’s web page.