Statistical Analysis for Engineers

Module Information

"Statistical testing is the process of making decisions based on numbers calculated from available data."

- Module title: Statistical Analysis for Engineers
- Module Code: MIOT H6014
- Lecturer: Damian Cox, damian.cox@itb.ie
- Moodle page: http://moodle.itb.ie/course/view.php?id=1566
- The module will be delivered over 8 weeks, June and July.

Revision/Prerequisites

It would be very useful and helpful if you have done some kind of mathematics recently, especially linear algebra. If you haven't try:

Linear algebra and Calculus (Keyword LAC2018)

http://moodle.itb.ie/course/view.php?id=1007

It would be even more helpful if you have done some Statistics and Probability recently. If not, try

Statistics and Probability: (SP2018)

http://moodle.itb.ie/course/view.php?id=485

Module Content

Descriptive Statistics

We will review the basics of descriptive statistics and introduce linear regression (using a matrix formulation) and correlation.

Probability

The fundamental ideas of probability. Bayes Theorem.

Random Variables

Random Variables and Distributions, expected values, the definition of population parameters as sample values. The important discrete and continuous probability distributions.

Statistical Testing

The concept of a Statistical Test. The Central Limit Theorem. Standard tests and their origins.

The limitations of Statistical testing; type I and II errors.

The Design of Experiments

Comparison of group means by analysis of variance (ANOVA).

Relating an experiment to the form of the data collected.

Multiple regression and the General Linear Model.

Analysis of Covariance, including Multivariate models

(MANOVA, ANCOVA, MANCOVA).

Other Topics

Parameter Estimation

Reporting on the results of Experiments and Studies and interpreting research work.

Statistical Packages

Assessment

The module will be 60 % on the exam, 40 % on Assignments.

All assignments will on Moodle. They will probably be on:

- Examples of the basic statistical tests (10%).
- Bootstrapping in correlation (10%).
- Confidence Intervals and their interpretation (10%).
- An Analysis on a large data set (5%), where possible, based on something you are doing in a different module or project
- A case study on interpreting a reported statistical result (5%)

All deadlines will be generous.

Module Delivery

Here is the way I intend to deliver the content of the module.

Lectures

The lectures will consist of roughly 2.5 – 3 hours of recorded lectures each week.

It will be available by 6pm every Tuesday on the Moodle page.

It should take you 4 hours to make your way through each one.

If you have a fundamental question on the content of a lecture, drop it to me in an email and I will reply or use it in a workshop.

Frames and Notes

The lecture frames will be available on the Moodle page in advance of the lectures. These will eventually be replaced by a set of notes expanding on the same topic, with the same section structure.

Worksheets

Every so often I will put up a worksheet with a few questions on the topics we are working on. Solutions will be provided in due course. They will resemble exam questions.

On-Site Workshops

We will have an occasional workshop to review the material we have covered and some of the more difficult worksheet questions.

- I will be hoping you will be trading your observations on the lectures I have delivered; questions you had and answered or deeper questions we can try and answer together.
- They will be two more on campus in June, on 12th and 17th.
- They will resume on 17th, 24th and 31st July.
- They will all be on Tuesday evenings, typically 4pm.

Some Reading

My notes will be fairly comprehensive, but here is some interesting
background reading and textbooks.

Recommended Reading

Chris Chatfield 1983, Statistics for Technology.

Henk Tijms, Understanding Probability.

David A. Freedman, Statistical Models, Theory and Practice.

Leonard Mlodinow, The Drunkards Walk

Stephen M. Stigler, The Sevens pillars of statistical wisdom.

Textbooks

GM Clarke & D Cooke 2004, A Basic Course in Statistics,

Rouncefield & Holmes 1993, Practical Statistics,

W Mendenhall, D Wackerly & R Scheaffer 1990, Mathematical Statistics with Applications,

D Montgomery 1991, Design and Analysis of Experiments,

Module Aims

"To often, statistics can be like magic; you think you have a handle on a dataset, what the control variables are and the confounding

variables. But … your analysis becomes more and more complex, you trip up over assumptions, then there's the alternative analysis which gives you a different answer and Lords paradox."

"Too many people are developing huge data sets, carrying out statistical analyses with a package or trying to reproduce them with mathematical models without knowing what the fuck they are doing." Robert May, chief science advisor to HMG, 1990s.

"You'll never be lonely as long as you know Statistics."