This is archived information. Please visit http://maths.manchester.ac.uk for current course unit information.

MATH38052 - 2010/2011

General Information

- Title: Generalized Linear Models
- Unit code: MATH38052
- Credits: 10
- **This course unit cannot be taken as well as MATH48052 which is a level 4 version of the same course unit.**
- Prerequisites: MATH20701, MATH20802. MATH20812 and MATH38011 are helpful but are not strictly required.
- Co-requisite units: None
- School responsible: Mathematics
- Members of staff responsible: Dr. J. Yuan

Specification

Aims

To study an important aspect of modern statistical modelling in an integrated way, and to develop the properties and uses of GLM, focusing on those situations in which the response variable is discrete. To explore some of the wide range of real-life situations occurring in the fields of agriculture, biology, engineering, industrial experimentation, medicine and social science that can be investigated using GLM.

Brief Description of the unit

As an important modelling strategy Linear Models is concerned with investigating *whether*, and *how*, one or more so-called explanatory variables, such as age, sex, blood pressure, etc., influence a response variable, such as a patient's diagnosis, by taking random variations of data into account. In Linear Models, linear regression technique and Normal distribution are used to explore the possible linear relation between a continuous response and one or more explanatory variables. In this course unit we depart from linearity and normality, the very strict limitation in Linear Models. We study the extension of linearity to non-linearity and normality to a commonly encountered distribution family, called the *exponential family of distributions*. This extension forms Generalized Linear Models (GLM). The GLM, on the one hand, unifies linear and non-linear models in terms of statistical modelling. On the other hand, it can be used to analyze discrete data, including binary, binomial, counted and categorical data that arise very often in biomedical and industrial applications.

Learning Outcomes

On successful completion of this course unit students will have a good understanding of

- the principles and methods of statistical modelling for GLM: response and explanatory variables, maximum likelihood estimation, confidence interval and hypothesis testing, goodness of fit, etc.;
- the use of the computer statistical software R or S-Plus, which is available on the Mathematics PC Cluster and does not require any previous programming experience;
- the statistical analysis of both continuous and discrete data arising in practice through using the statistical software R or S-Plus.

Future topics requiring this course unit

This course unit is part of the 4th year/MSc Generalised Linear Models and Survival Analysis. It is naturally related to another 4th year unit, Longitudinal Data Analysis.

Syllabus

1. **Introduction:** background, review of linear models in matrix notation, model assessment, some pre-required knowledge. [2]
2. **The exponential family of distributions:** Definition and examples. Mean and variance, variance function and scale parameter. [2]

3. **Generalized linear models (GLM):** linear predictor, link function, canonical link, maximum likelihood estimation, iterative reweighted least squares and Fisher scoring algorithms, significance of parameter estimates, deviance, Pearson and deviance residuals, Pearson's chi-square test and the likelihood ratio test, model fitting using R or S-Plus. [7]
4. **Normal linear regression models:** least squares, analysis of variance, orthogonality of parameters, factors, interactions between factors. [2]
5. **Binary and Binomial data analysis:** distribution and models, logistic regression models, odds ratio, one- and two-way logistic regression analysis. [5]
6. **Poisson count data analysis:** Poisson regression models with offset, two-dimensional contingency tables, log-linear models. [4]

Textbooks

- Dobson, A. J., *An Introduction to Generalized Linear Models*, Chapman & Hall 2002.
- Krzanowski, W., *An Introduction to Statistical Modelling*, Edward Arnold 1998.
- McCullagh, P. and Nelder, J. A., *Generalized Linear Models*, Chapman & Hall 1990.

Teaching and learning methods

Two lectures and one examples class each week. In addition students should expect to spend at least four hours each week on private study for this course unit.

Assessment

Coursework: 20%
End of semester examination: two hours weighting 80%

Arrangements