



MATH40752 - 2006/2007

General Information

- Title: Longitudinal Data Analysis
- Unit code: MATH40752
- Credits: 10
- Prerequisites: Probability and Statistics I (Year 1 Semester 1) Probability and Statistics II (Year 2 Semester 1), Linear Models (Year 3 Semester 1)
- Co-requisite units:
- School responsible: Mathematics
- Member of staff responsible: Dr Jianxin Pan

Specification

Aims

- To study advanced statistical techniques, and to develop statistical skill of analyzing longitudinal data and clustered data.
- To explore a wide range of real-life examples arising many areas such as biological, medical and social sciences.

Brief Description of the unit

In longitudinal studies, repeated measurements are made on subjects over time and responses within a subject are likely to be correlated, although responses between subjects may be independent. Data such as these are very common in practice, for example, in quality control in industry, panel data analysis in economics, growth curve analysis in biology and agriculture, randomized controlled trials in medical sciences, etc. When modelling such data without accounting for the correlation, statistical inferences can be very biased. The primary objective of longitudinal data analysis is to study how a response variable is related to explanatory variables of interest, by taking into account the within-subject correlation. The second objective is to quantify random variations in different sources and to characterize the within-subject correlation structures, which are very useful in many fields such as finance and demography.

Learning Outcomes

- the principles and methods of longitudinal data analysis
- the use of the statistical software R or S-PLUS to analyze longitudinal data arising in different areas

Future topics requiring this course unit

This course is naturally related to some other 4th year courses, including Linear and Generalized Linear Models, Survival Analysis, etc.

Syllabus

1. Introduction: real data examples, longitudinal data exploration [2]
2. Ordinary linear regression model for longitudinal data: a brief review of statistical inferences, effects of ignoring the within-subject correlation [4]
3. General linear models for longitudinal data: typical covariance structures - compound symmetry and AR(1) structures, maximum likelihood estimation, restricted maximum likelihood estimation [8]
4. Linear mixed models: fixed effects, random effects, random variation in different sources, variance components, maximum likelihood estimation, EM-algorithm, restricted maximum likelihood estimation, prediction of random effects, goodness of fit [14]

5. Generalized linear mixed models: extension of linear mixed models to generalized linear mixed models, exponential family of distributions, discrete longitudinal data, penalized quasi-likelihood estimates, variance components estimates, goodness of fit [8]

Textbooks

Davis, C. S. (2002). Statistical methods for the analysis of repeated measurements. Springer, New York
Diggle, P. J., Liang, K Y. and Zeger, S. L. (1994). Analysis of longitudinal data. Oxford University Press
Pan, J. and Fang, K. T. (2002). Growth curve models and statistical diagnostics. Springer, New York

Teaching and learning methods

Three lectures per week plus one weekly examples class.

Assessment

Coursework 20%. End of semester examination (2 hours and 30 mins) 80%.

Arrangements