Advanced Data Analysis: 4H
General Information 2012

The purpose of this course is to prepare you for a possible future role as a practising statistician and to develop advanced expertise in formulating and implementing statistical approaches to practical problems in a wide variety of subject areas.

Aims

- To develop advanced expertise in formulating and implementing statistical approaches to practical problems in a wide variety of subject areas.
- To integrate material covered in various lecture courses with skills developed through practical work in order to solve real-world problems.
- To train students in more advanced aspects of statistical computing through the statistical packages R and SAS.
- To further develop written skills of presentation and communication.
- To provide the opportunity for critical thinking and independent learning.

Intended Learning Outcomes

By the end of this course students will be able to:

- formulate questions of interest and identify relevant informal and formal statistical methodology for a wide variety of practical contexts
- implement the various stages of advanced statistical analysis appropriately in R and/or SAS
- interpret the output of R and SAS procedures
- critically collate results and conclusions
- present the main results and conclusions in the form of concise summaries
- present results of analyses in the form of written reports
- critically assess published applications of statistical analysis
- work independently on practical data analysis problems

Format of course:

- Course instructor: Dr. Mayetri Gupta, Reader of Statistics (mayetri.gupta@glasgow.ac.uk)
- Semester 1: September - December
- 10 laboratory sessions of 2 hours each
- All assessments based on course work: NO FINAL EXAMINATION

Assessment:

4H Single and Combined Honours students

3 in-class assessments of 10 points each
1 poster: 10 points
Final project: 10 points

Assessment Schedule

Date/Due Date	Assessment type	Percent of total marks
Oct 9	in-class 1	20%
Oct 30	in-class 2	20%
Nov 13	in-class 3	20%
Nov 20	poster due & viva scheduled	20%
Nov 27	final project report due	20%

MSc students

For MSc and MSci students, the above assessments and schedule will hold, but will count for 85% of their total marks. For the remaining 15% of their marks, they will be required to do an individual project that involves critically assessing a statistical analysis carried out in a journal article, provided by the instructor. Although this is not due until the end of the semester, please arrange to meet with the instructor as soon as possible to discuss this, especially if there is a potential future chance of scheduling conflicts with other courses.

Format of assessment

The assessment sessions will focus on material that has been covered in the practical classes associated with this course. The assessments will examine your ability to analyse data and interpret results in `R` (mainly) and any of the methods that you meet in the course may be covered in the in-class assessments. Practice test questions (with solutions) will be handed out occasionally (or put on moodle), throughout the course. You should also aim to develop an understanding of using the help files in `R` for commands that are less familiar to you.

Instructions for final project

For details, please see the separate document "Project Report Guidelines".

Course topics (lab sessions)

1. (Sep 25) Revisiting R: data exploration, logistic regression models and beyond
2. (Oct 2) Analysis of repeated measures data
3. (Oct 9) Multi-level and hierarchical modelling
4. (Oct 16) SAS Part 1: Data importing, manipulation, and basic statistical analyses
5. (Oct 23) SAS Part 2: Advanced statistical analyses, including programming and simulation
6. (Oct 30) Statistical methods for clustering and classification
7. (Nov 6) Computation-based statistical modelling and analysis
8. (Nov 13) Variable selection in high-dimensional regression models
9. (Nov 20) Nonlinear regression
10. (Nov 27) Smoothing and additive models