Subject Information Guide

Advanced Data Analysis

Semester 1, 2016

Administration and contact details

Host Department	**School of Mathematics and Applied Statistics**
Host Institution	**University of Wollongong**
Name of lecturer	**Pavel N. Krivitsky**
Phone number	**(2) 4221 3713**
Email Address	**pavel@uow.edu.au**
Homepage	
Name of Honours coordinator	Mark Nelson
Phone number	02-42214400
Email Address	Click here to enter text.

Subject details

Handbook entry URL	Click here to enter text.
Subject homepage URL	**https://moodle.uowplatform.edu.au/**
Honours student hand-out URL	Click here to enter text.
Start date:	**1 March 2015 (first lecture)**
End date:	**31 May 2015 (last lecture)**
Contact hours per week:	**2**
Lecture day and time:	**Tuesdays, 10:30–12:30**
Description of electronic access arrangements for students (for example, WebCT)	Click here to enter text.

Subject content

1. Subject content description

STAT902 introduces a variety of techniques for advanced data analysis, particularly regression, for handling categorical data, dependent data, nonlinear data, and situations where parts of the model are unknown or misspecified. Generalised linear models are considered in detail, as well as other ways of modelling nonlinearity, such as nonlinear models and nonparametric analysis. A variety of ways to model dependent (e.g., repeated

measures) data is considered, particularly, linear mixed models, generalised linear mixed models, generalised estimating equations, and latent variable models. Frequentist and Bayesian approaches to inference are considered, including likelihood, quasilikelihood, bootstrap, and sandwich estimation, conjugate priors, Monte-Carlo methods and Markov chain Monte-Carlo, as well as some graphical models, numeric integration, and prior elicitation.

2. Week-by-week topic overview

Note that the following schedule is tentative:
Weeks 1–2: Subject overview; revision of matrix algebra and vector calculus, distributions, and maximum likelihood.
Weeks 3–4: Estimating functions, quasi-likelihood, sandwich estimation, mean-variance misspecification, and bootstrap techniques.
Weeks 4–5: Bayesian inference and computation, prior elicitation, and model selection.
Weeks 6–7: Generalised linear models and nonlinear least-squares.
Weeks 8–9: Linear models for dependent data: linear mixed models and generalised estimating equations.
Weeks 10–11: Nonlinear models for dependent data: generalised linear mixed models and nonlinear mixed models.
Weeks 12–13: Nonparametric methods.

3. Assumed prerequisite knowledge and capabilities

- Equivalent of University of Wollongong's STAT332/STAT921 and STAT333/STAT922: statistical distribution theory, maximum likelihood estimation, fundamentals of statistical inference, multiple regression (linear), logistic and/or Poisson regressions, basic matrix algebra. These topics will be reviewed early in the subject, but not introduced.
- Familiarity with *R* will be very helpful.

4. Learning outcomes and objectives

1) Analyse complex data, particularly data with dependence, repeated measures, categorical response variables, overdispersion, and nonlinearity; and interpret the analyses.
2) Select and apply sophisticated statistical techniques (mostly regression-related) to answer substantive questions. Techniques include generalised linear models, Bayesian inference, mixed models, quasilikelihood, sandwich estimation, bootstrap, generalised estimating equations, model selection, and nonlinear and nonparametric models.
3) Recognise assumptions and limitations of statistical techniques considered, and diagnose their suitability for the data and the research question.
4) Derive expressions for point estimates, variances, and other quantities of interest for the techniques considered.
5) Implement and perform data analyses and diagnostics using R and BUGS/JAGS, including writing independent code in R language.

5. Learning resources

Text: Jon Wakefield. *Bayesian and Frequentist Regression Methods*. New York : Springer, **2013**. ISBN: 978-1-4419-0924-4 (Print) 978-1-4419-0925-1 (Online). doi:10.1007/978-1-4419-0925-1

Miscellaneous papers: Linked from the Moodle site.

Lecture notes: Posted on the Moodle site.

R examples: Posted on the Moodle site.

R: Freely available at http://www.r-project.org/; graphical frontends are available as well.

6. Assessment

Exam/assignment/classwork breakdown					
Exam	**70%**	**Assignment**	**30%**	**Class work**	**0%**
Assignment due dates		**Set Tuesday each week**	**Due Sunday of subsequent week**	Click here to enter a date.	Click here to enter a date.
Approximate exam date				**Mid June 2016**	

Institution Honours program details

Weight of subject in total honours assessment at host department	1/8
Thesis/subject split at host department	BMath (Hons): Thesis worth 25% BMathAdv (Hons): Thesis worth 37.5%
Honours grade ranges at host department:	
H1	85-100 %
H2a	75-84 %
H2b	65-74 %
H3	50-64 %