

STA442 Methods of Applied Statistics

Patrick Brown

Fridays 9-12

- Head TA: Colin Decker colin.decker@utoronto.ca
- TA office hours: TBA
- Course web page: pbrown.ca/teaching/appliedstats
- Instructor's office hour:
 - Tuesdays 8.30-9.30
 - the week an assignment is due thursday 8.30-9.30

Course description

The course will focus the using and interpreting advanced statistical methods with applications in a number of different areas. The course is a mixture of theory and applications, and will include a number of assignments which will involve computing with R.

Statistical methodology to be covered includes:

- Reproducible research with R
- generalised linear models and maximum likelihood estimation
- linear mixed models for longitudinal and hierarchical data
- Bayesian inference using INLA
- generalized linear mixed models
- semi-parametric regression
- survival analysis

Prerequisites

- STA303 is the official prerequisite
- Material from the second year statistical theory courses which are prerequisites to STA302 will be drawn on extensively
- Knowledge of programming with R is essential

Grading

Four homework assignments will each account for 10% of the final grade, with a final exam providing the remaining 60%.

Homework

- 9 Oct: Generalized linear models
- 30 Oct: Mixed models
- 13 Nov: Semi-parametric modelling

- 4 Dec: Survival analysis

Exam

The exam will consist of four questions very similar to the assignments. The exam will be designed to be easy for those who've completed and understood the assignments and extremely difficult for anyone who has not done the assignments themselves.

Practical details

Textbooks

The course will not adhere closely to any single text, though the following will be referred to frequently.

- Maindonald and Braun, *Data Analysis and Graphics Using R: An Example-Based Approach*
- Davison, *Statistical Models*
- Wakefield, *Bayesian and Frequentist Regression Methods*

Computing

Students will need to have access to a computer running a recent version of R, and several additional packages for R will be installed. Those not familiar with R are encouraged to become so within the first few weeks of the course.

Syllabus

- Frequentist inference with Generalized Linear Models
 - Models and inference
 - Applications and interpretation
- Applied statistics in practice

- scientific writing
- reproducible research
- Linear mixed models and longitudinal data
 - Mixed models, maximum likelihood estimation and REML
 - Applications and interpretation
 - Correlation in time
 - Random coefficient models
- Generalized Linear Mixed Models and Bayesian inference
 - Bayesian inference and INLA
 - Random effects models for non-Gaussian data
 - Applications and interpretation
 - Advanced INLA
- Smoothing and semi-parametric models
 - Penalised likelihood and spline smoothing
 - Random walks and Bayesian semi-parametrics
- Survival analysis
 - Parametric event time distributions
 - Censored data
 - Hierarchical models for survival data

References

- Davison, A. *Statistical Models*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2003. URL: <http://books1.scholarsportal.info/viewdoc.html?id=/ebooks/ebooks1/cambridgeonline/2012-11-08/1/9780511815850>.
- Maindonald, J. and W. J. Braun. *Data Analysis and Graphics Using R: An Example-Based Approach*. 3rd ed. Cambridge: Cambridge University Press, May 2010. URL: <https://www.cambridge.org/core/books/data-analysis-and-graphics-using-r/E04AEC5BCEF09D2E51A63EB5A8CB0680>.
- Wakefield, J. *Bayesian and Frequentist Regression Methods*. Springer Series in Statistics. Springer New York, 2013.