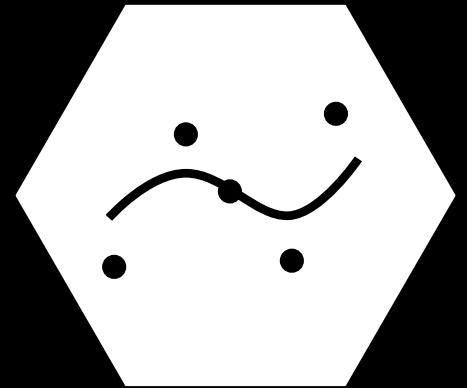# University of Glasgow

Advanced Predictive Models
Course information sheet 2022-23

Full course, 10 weeks

This course is concerned with models which can account for a non-normal distribution of the response and/or the fact that data is not independent, but correlated.

Prerequisite Knowledge

Learners should have prior experience of linear modelling and basic experience with the R programming language (e.g. data management and plotting).

Intended Learning Outcomes
By the end of this course learners will be able to:

- explain and derive key aspects of the theory of exponential families and generalised linear models;
- make correct use of models with various link functions and link distributions such as models for discrete data;
- determine whether a time series exhibits any evidence of a trend, seasonality or short-term correlation;
- define the class of ARIMA probability models;
- determine an appropriate model for a data set from the class of ARIMA models;
- predict future values for a given time series;
- make correct use of regression models assuming correlated residuals as well as models based on generalised estimation equations;
- explain the notion of a random effect, why and when it is useful and, in particular, how it differs from a fixed effect;
- make correct use of hierarchical models with random effects.

Syllabus

Week 1
- Introduction to advanced predictive models

Week 2 (sample material)
- Introduction to generalised linear models (GLMs)
- Exponential family of distributions
- Inference for GLMs

Week 3
- Models for binary/binomial response
- Fitting models for binary/binomial data in R
- Interpreting logistic regression coefficients
- Issues with models for binary/binomial data

Week 4
- Models for categorical responses
- Identifing categorical responses as nominal or ordinal
- Fitting models for ordinal/nominal data in R
- Interpreting model coefficients in terms of odds ratios

Week 5
- Models for count responses
- Fitting models for count data in R
- Interpreting model coefficients in terms of rate ratios

Mid-term week break

Week 6
- Recognising time series data
- Describing the main features of a time series
- Removing the trend seasonality or both from time series data

Week 7
- Autoregressive processes
- Moving average processes
- Fitting autoregressive and moving average processes in R

Week 8
- ARIMA processes
- Fitting ARIMA processes in R
- Forecasting future values of a time series

Week 9
- Linear mixed models
- Fixed and random effects
- Fitting linear mixed models in R

Week 10
- Generalised linear mixed models (GLMM) and fitting them in R
- Generalised estimating equations (GEE) and fitting them in R
- Appreciating the difference between the GLMM and GEE approaches

"A good follow-on from predictive modelling which almost feels like going back to the start, but covering things we missed. Interesting perspectives were given on time series data analysis. Very glad fixed, mixed and random effects models were included. Some of the logistic models towards the beginning were interesting. I have no idea ordinal regression modelling was a thing."

Online Learning
- Weekly live sessions with tutor(s)
- Weekly learning material (reading material, videos, exercises with model answers)
- Bookable one-to-one sessions with tutor(s)

Textbooks
Faraway, J (2006) Extending the linear model with R
Hyndman, R and Athanasopoulos G (2013) Forecasting: Principles and Practice.

Assessment
(for credit only)

This will typically be made up of online quizzes and individual assignments.

DATA ANALYTICS
GLASGOW

School of Mathematics and Statistics
University of Glasgow
http://gla.ac.uk/mdatagov
http://gla.ai
Email:
maths-stats-analyticscpd@glasgow.ac.uk

Software
To take our courses please use an up-to-date version of a standard browser (such as Google Chrome, Firefox, Safari, Internet Explorer or Microsoft Edge) and a PDF reader (such as Acrobat Reader). Learning material will be distributed through Moodle. We encourage all learners to install R and RStudio and we provide detailed installation instructions, but learners can also use free cloud-based services (RStudio Cloud). Learners need to install Zoom for participating in video conferencing sessions. We recommend the use of a head set for video conferencing sessions.