ST346-15 Generalised Linear Models for Regression and Classification

26/27

Department
Statistics
Level
Undergraduate Level 3
Module leader
Mark Steel
Credit value
15
Module duration
10 weeks
Assessment
Multiple
Study location
University of Warwick main campus, Coventry

Description

Introductory description

Regression is a typical problem in statistical learning and it involves explaining the behaviour of some response variable in terms of its relationship to a collection of possible explanatory variables. The ultimate aim is to do this via a modelling framework that provides both interpretability as well as uncertainty quantification.

This module is optional for students with their home department in Statistics. It may be possible for students from other courses to take this module as a listed module or an Unusual Option, subject to permission from the module leader and the home department.

Pre-requisites:

- Statistics students: ST231 Linear Statistical Modelling with R.
- Non-Statistics students: ST240 Linear Statistical Modelling or ST351 Linear Statistical Modelling (For Finalists).

Anti-requisites: Please check for anti-requisites under the availability tab and the course handbook.

Leads to: further modules that require statistical modelling, especially ST332 Medical Statistics.

Module web page

Module aims

This course extends the linear modelling framework to a broader range of problems and data types, e.g. binary and count data. The course will begin by introducing the Generalised Linear Model (GLM) and describe the model fitting process, variable selection, model validation, interpretation and uncertainty quantification.

This will be followed by topics that further extend the modelling possibilities to cases where model assumptions are not met e.g. overdispersion, quasi-likelihood methods.

Outline syllabus

This is an indicative module outline only to give an indication of the sort of topics that may be covered. Actual sessions held may differ.

1. Introduce Generalised Linear Models as an extension of the standard linear model under the umbrella of maximum likelihood estimation.
2. Covering model selection, parameter uncertainty and model validation for GLMs.
3. Extend this framework further to models that can handle overdispersion and be more robust to model selection e.g. quasi-likelihood models.

Learning outcomes

By the end of the module, students should be able to:

- identify an appropriate modelling approach for different types of response variable.
- select a model and analyse it in a principled way and verify the model assumptions.
- give clear uncertainty quantification for both parameter estimates as well as for model predictions.
- understand the key modelling assumptions and have a good understanding of how to make basic extensions to the standard GLM when issues like overdispersion are present.
- both apply and understand the theory and computation of GLMs

Subject specific skills

- Demonstrate facility with rigorousstatistical modelling methods.
- Evaluate, select and apply appropriate statistical modelling techniques.
- Demonstrate knowledge of and facility with formal statistical modelling concepts, both explicitly and by applying them to the solution of problems.
- Create structured and coherent arguments communicating them in written form.

- Construct logical arguments with clear identification of assumptions and conclusions.
- Reason critically, carefully, and logically.

Transferable skills

- Problem solving: Use rational and logical reasoning to deduce appropriate and well-reasoned conclusions. Retain an open mind, optimistic of finding solutions, thinking laterally and creatively to look beyond the obvious. Know how to learn from failure.
- Self awareness: Reflect on learning, seeking feedback on and evaluating personal practices, strengths and opportunities for personal growth.
- Communication: Present arguments, knowledge and ideas, in a range of formats.
- Professionalism: Prepared to operate autonomously. Aware of how to be efficient and resilient. Manage priorities and time. Self-motivated, setting and achieving goals, prioritising tasks.

Study

Study time

Type	Required
Lectures	30 sessions of 1 hour (20%)
Seminars	5 sessions of 1 hour (3%)
Private study	85 hours (57%)
Assessment	30 hours (20%)
Total	150 hours

Private study description

Weekly revision of lecture notes and materials, wider reading, practice exercises and preparing for examination.

Other activity description

Revision support.

Costs

No further costs have been identified for this module.

Assessment

You do not need to pass all assessment components to pass the module.

Students can register for this module without taking any assessment.

Assessment group D6

	Weighting	Study time	Eligible for self-certification
Assignment 1	10%	15 hours	No

The assignment will contain a number of questions for which solutions and / or written responses will be required.
The study time refers to the amount of time in hours that a well-prepared student who has attended lectures and carried out an appropriate amount of independent study on the material could expect to spend on this assignment. Your ST346 Assignment should not exceed 15 pages in length.

Assignment 2	10%	15 hours	No

The assignment will contain a number of questions for which solutions and / or written responses will be required.
The study time refers to the amount of time in hours that a well-prepared student who has attended lectures and carried out an appropriate amount of independent study on the material could expect to spend on this assignment. Your ST346 Assignment should not exceed 15 pages in length.

Centrally-timetabled examination (On-campus)	80%		No

The examination paper will contain four questions, of which the best marks of THREE questions will be used to calculate your grade.

- Students may use a calculator
- Answerbook Pink (12 page)

Assessment group R6

	Weighting	Study time	Eligible for self-certification
In-person Examination - Resit	100%		No

The examination paper will contain four questions, of which the best marks of THREE questions will be used to calculate your grade.

	Weighting	Study time	Eligible for self-certification

- Students may use a calculator
- Answerbook Pink (12 page)

Feedback on assessment

Assignments are marked and given feedback online within 20 working days of the submission deadline. Where appropriate, model solutions will be provided.

Solutions and cohort level feedback will be provided for the examination.

Past exam papers for ST346

Availability

Anti-requisite modules

If you take this module, you cannot also take:

- ST426-15 Generalised Linear Models with Advanced Topics

Courses

This module is Core for:

- Year 3 of USTA-G300 Undergraduate Master of Mathematics,Operational Research,Statistics and Economics

This module is Optional for:

- Year 3 of UCSA-G4G1 Undergraduate Discrete Mathematics
- UCSA-G4G3 Undergraduate Discrete Mathematics
  - Year 3 of G4G1 Discrete Mathematics
  - Year 3 of G4G3 Discrete Mathematics
- Year 4 of UCSA-G4G4 Undergraduate Discrete Mathematics (with Intercalated Year)
- Year 4 of UCSA-G4G2 Undergraduate Discrete Mathematics with Intercalated Year
- USTA-G300 Undergraduate Master of Mathematics,Operational Research,Statistics and Economics
  - Year 3 of G300 Mathematics, Operational Research, Statistics and Economics
  - Year 4 of G300 Mathematics, Operational Research, Statistics and Economics

This module is Option list A for:

- Year 3 of USTA-G300 Undergraduate Master of Mathematics,Operational Research,Statistics and Economics

- Year 4 of USTA-G1G3 Undergraduate Mathematics and Statistics (BSc MMathStat)
- Year 3 of USTA-GG14 Undergraduate Mathematics and Statistics (BSc)
- Year 4 of USTA-G140 Undergraduate Mathematics and Statistics (MMathStat)
- Year 4 of USTA-GG17 Undergraduate Mathematics and Statistics (with Intercalated Year)
- Year 3 of USTA-Y602 Undergraduate Mathematics,Operational Research,Statistics and Economics
- Year 4 of USTA-Y603 Undergraduate Mathematics,Operational Research,Statistics,Economics (with Intercalated Year)

This module is Option list B for:

- Year 3 of USTA-G302 Undergraduate Data Science
- Year 3 of USTA-G304 Undergraduate Data Science (MSci)
- UMAA-G105 Undergraduate Master of Mathematics (with Intercalated Year)
    - Year 4 of G105 Mathematics (MMath) with Intercalated Year
    - Year 5 of G105 Mathematics (MMath) with Intercalated Year
- Year 3 of UMAA-G100 Undergraduate Mathematics (BSc)
- UMAA-G103 Undergraduate Mathematics (MMath)
    - Year 3 of G100 Mathematics
    - Year 3 of G103 Mathematics (MMath)
    - Year 4 of G103 Mathematics (MMath)
- Year 4 of UMAA-G107 Undergraduate Mathematics (MMath) with Study Abroad
- Year 3 of USTA-G1G3 Undergraduate Mathematics and Statistics (BSc MMathStat)
- Year 3 of USTA-G140 Undergraduate Mathematics and Statistics (MMathStat)
- Year 4 of UMAA-G101 Undergraduate Mathematics with Intercalated Year

This module is Option list E for:

- Year 3 of USTA-G300 Undergraduate Master of Mathematics,Operational Research,Statistics and Economics

This module is Option list F for:

- Year 4 of USTA-G300 Undergraduate Master of Mathematics,Operational Research,Statistics and Economics