

ST346-15 Generalised Linear Models for Regression and Classification

21/22

Department

Statistics

Level

Undergraduate Level 3

Module leader

Martyn Plummer

Credit value

15

Module duration

10 weeks

Assessment

Multiple

Study location

University of Warwick main campus, Coventry

Description

Introductory description

This module runs in Term 1 and is optional for students with their home department in Statistics . It may be possible for students from other courses to take this module as an Unusual Option, subject to permission from the module leader and the home department.

This module has ST221 as a pre-requisite.

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

Module aims

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.

The linear modelling course, ST221, gave an introduction to this problem. This course extends this 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, random effect models, 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 from ST221 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. random effects and 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

TBC

Transferable skills

TBC

Study

Study time

Type	Required	Optional
Lectures	30 sessions of 1 hour (20%)	2 sessions of 1 hour
Private study	90 hours (60%)	
Total	150 hours	

Type	Required	Optional
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.

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 D1

	Weighting	Study time	Eligible for self-certification
Assignment 2	10%	15 hours	Yes (extension)
Due in week 9 of term 1.			
The assignment will contain a number of questions for which solutions and / or written responses will be required.			
The number of words noted 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. 500 words is equivalent to one page of text, diagrams, formula or equations; your ST346 Assignment 2 should not exceed 15 pages in length.			
Assignment 1	10%	15 hours	Yes (extension)
Due in week 4 of term 1.			
The assignment will contain a number of questions for which solutions and / or written responses will be required.			
The number of words noted 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. 500 words is equivalent to one page of text, diagrams, formula or equations; your ST346 Assignment 1 should not exceed 15 pages in length.			
On-campus Examination	80%		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
~Platforms - Moodle			
	<ul style="list-style-type: none"> • Answerbook Pink (12 page) • Students may use a calculator 		

Assessment group R1

	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.			
~Platforms - Moodle			

Feedback on assessment

Marked assignments will be available for viewing at the support office within 20 working days of the submission deadline. Cohort level feedback and solutions will be provided, and students will be given the opportunity to receive feedback via face-to-face meetings.

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

[Past exam papers for ST346](#)

Availability

Courses

This module is Optional for:

- 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-G304 Undergraduate Data Science (MSci)
- Year 4 of USTA-G303 Undergraduate Data Science (with Intercalated Year)
- Year 4 of USTA-G300 Undergraduate Master of Mathematics, Operational Research, Statistics and Economics

- USTA-G1G3 Undergraduate Mathematics and Statistics (BSc MMathStat)
 - Year 3 of G1G3 Mathematics and Statistics (BSc MMathStat)
 - Year 4 of G1G3 Mathematics and Statistics (BSc MMathStat)
- USTA-G1G4 Undergraduate Mathematics and Statistics (BSc MMathStat) (with Intercalated Year)
 - Year 4 of G1G4 Mathematics and Statistics (BSc MMathStat) (with Intercalated Year)
 - Year 5 of G1G4 Mathematics and Statistics (BSc MMathStat) (with Intercalated Year)
- Year 3 of USTA-GG14 Undergraduate Mathematics and Statistics (BSc)
- Year 4 of USTA-GG17 Undergraduate Mathematics and Statistics (with Intercalated Year)
- Year 3 of USTA-Y602 Undergraduate Mathematics, Operational Research, Statistics and Economics

This module is Option list B for:

- Year 3 of USTA-G302 Undergraduate Data Science
- Year 3 of USTA-G300 Undergraduate Master of Mathematics, Operational Research, Statistics and Economics

This module is Option list D for:

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

This module is Option list E for:

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

This module is Option list F for:

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