MODULE SPECIFICATION

Academic Year (student cohort covered by specification)	2024-25
Module Code	2462
Module Title	Statistical Models for Discrete Outcomes
Module Organiser(s)	Professor Jonathan Bartlett & Professor Ruth Keogh
Faculty	Epidemiology & Population Health
FHEQ Level	Level 7
Credit Value	**CATS:** 15 **ECTS:** 7.5
HECoS Code	101031
Term of Delivery	Term 2
Mode of Delivery	For 2024-25 this module will be delivered by predominantly face-to-face teaching modes. Where specific teaching methods (lectures, seminars, discussion groups) are noted in this module specification these will be delivered by predominantly face-to-face sessions. There will be a combination of live and interactive activities (synchronous learning) as well as self-directed study (asynchronous learning)
Mode of Study	Full-time
Language of Study	English
Pre-Requisites	A good understanding of linear regression (including multiple linear regression models with interaction terms), analysis of variance, likelihood theory and simple methods of analysing quantitative and categorical data is essential.
Accreditation by Professional Statutory and Regulatory Body	None
Module Cap (indicative number of students)	35 (numbers may be capped due to limitations in facilities or staffing)
Target Audience	This module is intended for people with both mathematical (up to first year undergraduate level) and statistical background (undergraduate degree level in joint mathematics/statistics for example) intending to pursue a career in medical statistics.
Module Description	Linear regression models have wide applicability to studies

	where a continuous outcome is related to one or more predictor variables. The family of generalized linear models (GLMs) is a larger class of models that are also applicable to binary and count outcomes. This module gives a methodological introduction to this family of models and others used for the analysis of discrete outcome data. It will equip students with the skills needed to analyse and interpret data from randomised controlled trials and cohort, case-control and cross-sectional studies with binary, count and categorical outcomes. It is assessed through an analysis and reporting exercise.
Duration	5 weeks at 2 days per week
Timetabling slot	Slot C1
Last Revised (e.g. year changes approved)	August 2022

Programme(s) This module is linked to the following programme(s)	**Status**
MSc Medical Statistics	Compulsory

Module Aim and Intended Learning Outcomes

Overall aim of the module
The overall module aim is to: • equip students with the necessary skills to (i) understand the principles underlying statistical models for binary, count and categorical outcomes, and (ii) be able to analyse data using such models.

Module Intended Learning Outcomes
Upon successful completion of the module a student will be able to: 1. Demonstrate an understanding of the theoretical basis of Generalized Linear Models. 2. Use Generalized Linear Models and other models for the analysis of binary, categorical and count data. 3. Present results clearly and accurately in a structured report, such as might form the basis of a report by a statistical consultant. 4. Demonstrate an understanding of the concepts of confounding and associations versus causation, and how to apply these in modelling choices and interpretation. 5. Have an appreciation of different analysis strategies.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Indicative Syllabus

Session Content
The module is expected to be divided into four sections covering the following topics **A. Introductory sessions** 1. Review of linear regression and introduction to logistic regression. 2. Logistic regression and other generalized linear models for Bernoulli outcomes. 3. Types of investigation. **B. Estimating causal effects of exposures and treatments** 4. Estimating treatment effects using observational data. 5. Collapsibility. 6. Analysis strategies for observational studies. 7. Logistic regression in cohort and case-control studies. 8. Matched studies and their analysis. 9. Conditional logistic regression. **C. Prediction modelling** 10. Assessment of prediction models (including discrimination, calibration and validation). 11. Ridge and lasso regression. 12. Non-linear effects using splines. **D. Models for other outcome types** 13. Generalized linear models for grouped binary and count outcomes. 14. Models for rates. 15. Multinomial and ordinal logistic regression.

Teaching and Learning

Notional Learning Hours

Type of Learning Time	Number of Hours	Expressed as Percentage (%)
Contact time	36	24
Directed self-study	74	49
Self-directed learning	10	7
Assessment, review and revision	30	20
Total	**150**	**100**

Student contact time refers to the tutor-mediated time allocated to teaching, provision of guidance and feedback to students. This time includes activities that take place in face-to-face contexts such as lectures, seminars, demonstrations, tutorials, supervised laboratory workshops, practical classes, project supervision as well as where tutors are available for one-to-one discussions and interaction by email.

The division of notional learning hours listed above is indicative and is designed to inform students as to the relative split between interactive and self-directed study.

Teaching and Learning Strategy
The teaching and learning strategy is structured around a combination of live lectures followed by computer or non-computer practical sessions. Practical sessions ensure that students have the opportunity to apply the concepts and methods covered by lecture content. They provide students with “hands on” experience in analysing and interpreting data, using data sets drawn from research work of staff in the faculty. Students are provided with detailed solutions to the tasks set in practical sessions, enabling them to check their understanding of the material. The assessment task, which comes towards the end of the module, is the point at which students demonstrate a consolidation of their learning across the whole module.

Assessment

Assessment Strategy
The assessment for this module has been designed to measure student learning against the module intended learning outcomes (ILOs) as listed above. The grade for summative assessment(s) only will go towards the overall award GPA. The assessment will consist of a data analysis and production of a written report.

Summative Assessment

Assessment Type	Assessment Length (i.e. Word Count, Length of presentation in minutes)	Weighting (%)	Intended Module Learning Outcomes Tested
Coursework	4 pages	100	1 – 5

Resitting assessment
Resits will accord with Chapter 8a of the LSHTM Academic Manual. Resit/deferred/new attempts - The task will be a data analysis report. The next assessment deadline for coursework will be during mid/late September of the current academic year.

Resources

Indicative reading list

Some text books which may be found useful are as follows:

- Dobson, A.J and Barnett, A.G. (2008) An Introduction to Generalized Linear Models, Third Edition. Chapman & Hall.
- Collett, D. (2002) Modelling Binary Data. Chapman & Hall.
- Harrell, F.E. Regression Modeling Strategies (2015). Springer.
- Pearl J, Glymour M, Jewell N.P. (2016) Causal Inference in Statistics: A Primer. Wiley.
- Breslow N.E, Day N.E. (1980) Statistical Methods in Cancer Research Volume I: The Analysis of Case-Control Studies. IARC Scientific Publication No. 32.
- James G, Witten T, Hastie D, Tibshirani R. An Introduction to Statistical Learning. 2nd Edition.m2021. Springer. https://www.statlearning.com/

Other resources

Students will be provided with additional references, and links to resources in the lecture notes.

Teaching for Disabilities and Learning Differences

The module-specific site on Moodle gives students access to lecture notes and copies of the slides used during the lecture. Where appropriate, lectures are recorded and made available on Moodle. All materials posted on Moodle, including computer-based sessions, have been made accessible where possible.

LSHTM Moodle is accessible to the widest possible audience, regardless of specific needs or disabilities. More detail can be found in the Moodle Accessibility Statement which can also be found within the footer of the Moodle pages. All students have access to "SensusAccess" software which allows conversion of files into alternative formats.

Student Support Services can arrange learning or assessment adjustments for students where needed. Details and how to request support can be found on the LSHTM Disability Support pages.