Module Specification

ABOUT THIS DOCUMENT

This module specification applies for the academic year 2019-20
Last revised 10 Sep 2019 by Edmund Njeru Njagi

London School of Hygiene & Tropical Medicine, Keppel St., London WC1E 7HT. www.lshtm.ac.uk

GENERAL INFORMATION

<table>
<thead>
<tr>
<th>Module name</th>
<th>Advanced Statistical Modelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module code</td>
<td>2450</td>
</tr>
<tr>
<td>Module Organiser</td>
<td>Edmund Njeru Njagi and Nick Jewell</td>
</tr>
<tr>
<td>Contact email</td>
<td><a href="mailto:Edmund-Njeru.Njagi@lshtm.ac.uk">Edmund-Njeru.Njagi@lshtm.ac.uk</a> or <a href="mailto:nicholas.jewell@lshtm.ac.uk">nicholas.jewell@lshtm.ac.uk</a></td>
</tr>
<tr>
<td>Home Faculty</td>
<td>Epidemiology &amp; Population Health</td>
</tr>
<tr>
<td>Credit</td>
<td>15 credits</td>
</tr>
<tr>
<td>Accreditation</td>
<td>Not currently accredited by any other body.</td>
</tr>
<tr>
<td>Keywords</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

AIMS, OBJECTIVES AND AUDIENCE

Overall aim
To introduce recent developments in statistical modelling, focussing in the first part of the module on statistical models for drawing causal inferences, and in the second part of the module on models for discrete dependent data.

Intended learning outcomes
By the end of this module, students should:

i. Understand the nature of the scientific enquiry / practical problem that led to the developments and for which they are appropriate.

ii. Understand the concepts on which they are based and their relationship to each other and to the approaches/techniques encountered earlier in the MSc Medical Statistics programme.

iii. Develop the skills to use the methods creatively and independently in practical problems.

iv. Develop the capacity to review critically their own and others’ statistical modelling work in these settings.

Target audience
This module is intended for students with an understanding of linear regression models for hierarchical data and generalized linear models to the
The module is expected to include sessions addressing the following topics:

**Part I:**
- Causal languages in statistics
- Graphical models for statistics and causality
- Causal inference using parametric statistical models: traditional regression
- Causal inference using semiparametric statistical models: propensity scores, inverse probability weighting and doubly robust estimation
- Causal mediation analysis

**Part II:**
- Revision of linear and generalized linear models
- Modelling dependent non-normal data: subject specific and marginal models
- Generalised estimating equations and empirical (sandwich) variance estimates
- Likelihood based analyses for subject-specific models

### Study resources provided or required
Module Information can be found on the Virtual Learning Environment (Moodle) containing information about each session and key references for the module. Copies of all the overheads used, practicals and their solutions; datasets required for the practicals and assessment exercise will be available in Moodle.

### Teaching and learning methods
A combination of lectures and computer-based practicals.

### Assessment details
Students will complete two assignments, one on each part of the module.

The first assignment will involve the application, interpretation and comparison of different statistical approaches to addressing particular causal questions using a dataset from an observational study.

The second assignment will involve the use, interpretation and comparison, of several alternative statistical analyses for a given longitudinal data set from a medical setting.

Resit/deferred/new attempts - The tasks will be similar to the original assessments, but involving a different data set.

### Assessment dates
The hand-in date for the assessment will be the last day of the module.

Resit/deferred/new attempts - the next assessment deadline will be during mid/late September of the current academic year.

### Language of study and assessment
English (please see ‘English language requirements’ below regarding the standard required for entry).
### TIMING AND MODE OF STUDY

<table>
<thead>
<tr>
<th><strong>Duration</strong></th>
<th>5 weeks at 2 days per week</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dates</strong></td>
<td>Thursday morning to Friday afternoon</td>
</tr>
<tr>
<td><strong>Timetable slot</strong></td>
<td>Term 3 – slot E</td>
</tr>
<tr>
<td><strong>Mode of Study</strong></td>
<td>The module is taught face-to-face in London. Both full-time and part-time students follow the same schedule.</td>
</tr>
</tbody>
</table>
| **Learning time** | The notional learning time for the module totals 150 hours, consisting of:  
- Contact time ≈ 42.5 hours  
- Self-directed learning ≈ 57.5 hours  
- Assessment, review and revision ≈ 50 hours |

### APPLICATION, ADMISSION AND FEES

| **Pre-requisites** | This module is intended for students with an understanding of probability (Module 2038), linear regression models, generalized linear models and linear regression models for hierarchical data to the level provided by the modules Generalized Linear Models (2462) and Analysis of Hierarchical & Other Dependent Data (2465). Students require an understanding of matrix algebra and calculus (to first year undergraduate mathematics level). The module Analysis of Hierarchical & Other Dependent Data (2465) is strongly recommended for prior study. |
| **English language requirements** | A strong command of the English language is necessary to benefit from studying the module. Applicants whose first language is not English or whose prior university studies have not been conducted wholly in English must fulfil LSHTM's English language requirements. |
| **Student numbers** | 25 (numbers may be capped due to limitations in facilities or staffing) |
| **Student selection** | Preference will be given to LSHTM MSc students, particularly those registered for specific modules or who have taken specific prior modules, and LSHTM research degree students. Other applicants meeting the entry criteria will usually be offered a place in the order applications are received, until any cap on numbers is reached. Applicants may be placed on a waiting list and given priority the next time the module is run.  
Full Registration (full participation) by LSHTM research degree students is required for this module. |