Module Code | CS7DS3
---|---
Module Name | Applied Statistical Modelling
ECTS Weighting | 5 ECTS
Semester taught | Semester 2
Module Coordinator/s | Dr. Arthur White

Module Learning Outcomes

On successful completion of this module, students will be able to:

- **LO1.** Define a Markov chain and describe its theory;
- **LO2.** Identify an appropriate Monte Carlo simulation method for a given probability distribution and implement it;
- **LO3.** Describe and interpret complex statistical models in terms of a graphical model framework;
- **LO4.** Describe and implement the state of the art methodology in several topical applications in data science;
- **LO5.** Describe and discuss different approaches for model selection;
- **LO6.** Complete a data science project that applies the methods of this and other modules to a real data set.

Module Content

This module continues on from CS7CS4 (Machine Learning) with a focus on sampling methods and topical applications. It also gives an opportunity for students to apply, through a large project, the methods that they have explored in CS7DS1 (Data Mining & Analytics) and that they are currently exploring in CS7DS2 (Optimisation Algorithms for Data Analysis).

Teaching and Learning Methods

Mainly lectures, and some case studies and other supplemental reading will also be provided.

Assessment Details

<table>
<thead>
<tr>
<th>Assessment Component</th>
<th>Brief Description</th>
<th>Learning Outcomes Addressed</th>
<th>% of total</th>
<th>Week set</th>
<th>Week due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short assignment 1</td>
<td>Problem sets</td>
<td>LO1, LO2</td>
<td>15%</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Main assignment</td>
<td>Detailed report of a statistical analysis of a real data set.</td>
<td>LO2, LO3, LO4, LO5, LO6</td>
<td>70%</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Short assignment 2</td>
<td>Problem sets</td>
<td>LO3, LO5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reassessment Details

Analysis of main assignment repeated and re-written. Worth 100% in reassessment.

---

1. TEP Glossary
2. TEP Guidelines on Workload and Assessment
# Contact Hours and Indicative Student Workload

<table>
<thead>
<tr>
<th>Contact Hours (scheduled hours per student over full module), broken down by:</th>
<th>22 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>lecture</td>
<td>22 hours</td>
</tr>
<tr>
<td>laboratory</td>
<td>0</td>
</tr>
<tr>
<td>tutorial or seminar</td>
<td>0</td>
</tr>
<tr>
<td>other</td>
<td>0 hours</td>
</tr>
</tbody>
</table>

**Independent study (outside scheduled contact hours), broken down by:**

<table>
<thead>
<tr>
<th>Activity</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>preparation for classes and review of material (including preparation for examination, if applicable)</td>
<td>42 hours</td>
</tr>
<tr>
<td>completion of assessments (including examination, if applicable)</td>
<td>52 hours</td>
</tr>
</tbody>
</table>

**Total Hours**

116 hours

# Recommended Reading List


# Module Pre-requisites

**Prerequisite modules:** CS7DS1 (Data Mining & Analytics), CS7CS4 (Machine Learning)

**Other/alternative non-module prerequisites:** Knowledge of basic probability and statistical inference, and a familiarity with R will be beneficial. Some background knowledge of supervised and unsupervised machine learning methods are also assumed.

# Module Co-requisites

# Module Website

[scss.tcd.ie/~arwhite/Teaching/CS7DS3.html](scss.tcd.ie/~arwhite/Teaching/CS7DS3.html)

# Last Update

18/07/2019 by Your Name