

Liverpool John Moores University

Title: Statistical Methods in R
Status: Definitive
Code: **7011DATSCI** (125176)
Version Start Date: 01-08-2021

Owning School/Faculty: Astrophysics Research Institute
Teaching School/Faculty: Astrophysics Research Institute

Team	Leader
Daniel Perley	Y

Academic Level: FHEQ7 **Credit Value:** 20 **Total Delivered Hours:** 60
Total Learning Hours: 200 **Private Study:** 140

Delivery Options

Course typically offered: Semester 1

Component	Contact Hours
Lecture	20
Practical	40

Grading Basis: 50 %

Assessment Details

Category	Short Description	Description	Weighting (%)	Exam Duration
Portfolio	Assignment	Computer Assignments	70	
Test	Test	Timed In-Class Test	30	

Aims

The module aims to provide an introduction to students from differing academic disciplines to key concepts in statistics and statistical computing using the R programming language, with an emphasis on the informed interpretation of the results of statistical testing.

Learning Outcomes

After completing the module the student should be able to:

- 1 Critically apply a variety of statistical techniques to problems in data science, using the R computer language.
- 2 Critically evaluate the appropriateness of various techniques for a particular problem.
- 3 Critically analyse the outcomes of hypothesis testing in an informed fashion
- 4 Synthesize a combination of statistical methods to test a complex problem.

Learning Outcomes of Assessments

The assessment item list is assessed via the learning outcomes listed:

Computer Assignments	1	2	3	4
In Class Test	1	2	3	4

Outline Syllabus

1. *Statistical fundamentals (significance, inference and hypotheses, replication, models, confidence intervals)*
2. *Introduction to the R language (dataframes, functions, loops, I/O, plotting, packages)*
3. *Descriptive statistics (measures of central tendency and variability)*
4. *Probability density, central limit theorem, normal distribution*
5. *Single sample tests (Z-test, t-test, testing for normality, nonparametric tests, jackknife/bootstrap)*
6. *Two sample testing (F-test, ECDF-based tests, Fisher's exact test, Chi-squared test, correlation and covariance)*
7. *Regression (maximum likelihood, least-squares, linear and nonlinear regression, nonparametric curve-fitting)*
8. *Analysis of variance (ANOVA inference, post-hoc tests, multifactor analysis, structured error, goodness of fit / model comparison)*
9. *Multiple regression (ANCOVA, stepwise regression, cross-validation)*
10. *Count data (Binomial/Poisson errors, logistic regression, generalized linear models)*
11. *Survival analysis (Hazard, censorship, distribution fitting)*
12. *Bayesian inference (Likelihood, priors and posteriors, credibility intervals)*

Learning Activities

Lectures
Directed Reading
Computer Exercises

Notes

