Advanced Methods for Reproducible Science

Funded by the BBSRC and the European College of Neuropsychopharmacology

Programme 2019

Cumberland Lodge, Windsor Great Park
http://www.cumberlandlodge.ac.uk/
Sunday 6th to Friday 11th January

Learning Goals and Objectives

On completion of this course, student will:

1. Understand the basics of good experimental design, including features designed to avoid bias (e.g., random allocation, blinding, and controls).
2. Be able to apply these principles to their own research projects: students will have help in drafting a protocol for a Registered Report for a study of their own.
3. Be more confident in critical analysis of existing literature, and be aware of methods for detecting publication bias.
4. Have the opportunity to join a collaborative project by coding papers during the course, with the goal of producing a published piece of work.
5. Know how to use version control and scripting to ensure research pipelines / workflows are reproducible.
6. Know how to simulate data in order to try out analyses before collecting data, and to assess statistical power.
7. Understand the importance of collaborative working and open science practices, and how these support reproducibility.
8. Be familiar with the advantages and disadvantages of conventional and Bayesian approaches to statistical analysis, and be more confident in using these.
9. Join a peer-led network of other early-career researchers to take forward reproducibility initiatives into the future.

Overview

Students will be sent detailed joining instructions prior to the course, which will include instructions of what software and packages to upload.

They will also be asked to work through this R introduction by Software Carpentry before they come to the course if they are unfamiliar with R:

http://swcarpentry.github.io/r-novice-gapminder/

They will also be asked to come with a draft outline of a study that they would like to pre-register.

There will also be an opportunity to join an exercise that involves collaborating on a meta-analysis, with the goal of achieving a publishable paper.

Programme

Sunday 6th January

INTRODUCTION

14.00-14.30 Welcome, introductions, assignment of tutors **Chris Chambers**

14.30-16.00 Using Github for reproducible research **Laura Fortunato**

This session will give students the chance to do hands-on exercises to build skills needed for the rest of the course, with supervision by course tutors.

16:00-16.30 Tea

16.30-18.00 Getting to grips with R Markdown **Mike Smith**

Students will learn to create a R Markdown file. This will be build up during the week as we create a joint project. Those who wish to do so can also create a Markdown file for their own project.

PART A: WHAT IS THE PROBLEM?

18.30-19:15 Reproducibility – what is the scale of the problem? **Marcus Munafò**

19.15-20.15 Dinner

20.15-20.45 Welcome talk. **Cumberland Lodge**

Monday 7th January

08.00-08.45 Breakfast

09.00-10.00 Pre-registration and Registered Reports. **Chris Chambers**

10.00-10.30 Is pre-registration feasible for your study? **Discussion w. tutors**

10.30-11.00 Coffee

PART B: EXPERIMENTAL DESIGN FOR REPRODUCIBILITY

11.00-12.30 How to design a good experiment. (seminar with breakout discussion sessions) **Natasha Karp**

13.00-14.00 Lunch

14.00-16.00 'Surgery' sessions – individual students can book a slot to discuss specific issues with their assigned tutor. Tutors may run additional sessions on specific topics as requested

16.00-16:30 Tea

16.30-17.15 The Open Science Framework. **Courtney Soderberg**

17.15-18.30 Making your research open. **Exercise**

Participants work on their own protocol for pre-registration and learn how to prepare data files and meta-data files and meta-data for open data release using OSF

19.15-20.15 Dinner

PART C: BARRIERS TO REPLICABILITY

Tuesday 8th January

08.00-08.45 Breakfast

09.00-09.45 Diagnosing publication bias and other anomalies. **Daniel Lakens**

09.45-10.30 Methods for exploring published datasets **Exercise**

10.30-11.00 Coffee

11.00-12.00 Using R to simulate data sets. **Dorothy Bishop**

This session has two goals: (i) to show students how to simulate data so they can anticipate how to analyse data before conducting a study; (ii) to demonstrate how easy it is to obtain 'significant' results if hypotheses are not pre-specified – i.e., p-hacking.

12.00-12.30 Simulating data for your project. **Exercise**

Students work in small groups: course tutors available to help.

13.00-14.00 Lunch

14.00-15.00 Simulation / other methods for power calculations. **Natasha Karp**

15.00-15.30 Tea

15.30-16.30 Assessing and reducing risk of bias in the literature **Emily Sena & Kate Button**

16.30-17.30 Exercise in critical appraisal: Assessing methodological quality of experimental studies

Participants will work in pairs to rate the methodological quality of a range experimental studies spanning biomedicine using a new outcome evaluation platform currently under development for use in an RCT. The goals are (i) for students to gain an understanding of the core principles of good experimental design that span disciplines (ii) to participate in the development of a robust generic measure for assessing methods rigour.

17:30 – 19.00 Small group session with tutors: 2 options

a) Collaborative project with Kate Button and Emily Sena: Assessing methodological quality of experimental studies

b) 'Surgery' sessions with tutors.

19.15-20.15 Dinner

20.15-21.00 Discussion session: Q&A. **t.b.c.**

PART D: COLLABORATIVE WORKING

Wednesday 9th January

08.00-08.45 Breakfast

9.00-10.00 Transforming research with collaborative working. **Kirstie Whitaker**

10.00-11.00. R, Github and R markdown refresher.

Opportunity to consolidate skills using either own pre-registered study, or the collaborative project.

11.00-11.30 Coffee

11.30-12.15 Collaborative student projects. **Kate Button**

12.15-12.30 Discussion

13.00-14.00 Lunch

Afternoon free for networking, independent study and reflection

19.15-20.15 Dinner
20.15-21.00 Reflections from a previous student. **Amy Orben**

What I learned on the course last year – and how it influences how I work now.

PART E: DATA PROCESSING AND ANALYSIS

Thursday 10th January *(precise details still to be confirmed for this day)*

08.00-08.45 Breakfast

09.00-09.45 Seminar: Using statistics to answer questions and claim discoveries.
David Spiegelhalter

09.45-10.30 Using R for data cleaning and data wrangling **Paul Thompson**

10.30-11.00 Coffee

11.00-11.45 Why use Bayesian analysis? **Alex Etz**

11.45–13.00 Small group sessions.

a) Using JASP.
b) Statistics consulting.
c) Problems encountered in coding studies/writing a pre-registration document.

13.00-14.00 Lunch

14.00 – 14.45 The experimental design assistant (EDA). **Nathalie Percie du Sert**

14.45 – 16.00 Design your experiment, using the EDA Tutorial. **Exercise (optional)**

16.00-16.30 Tea

19.15-18.15 Dinner

PART E: MANAGING A CAREER IN REPRODUCIBLE SCIENCE

Friday 11th January

08.00-08.45 Breakfast

09.00-09.45 Selfish reasons to do reproducible science**.** **Florian Markowetz**

09.45-10.30 Initiatives for reproducible science. **Marcus Munafo**

10.30-11.00 Student-led discussion.

11.00-11.30 Coffee

11.30-13.0 Student-led round-table session.

a) Barriers to doing reproducible science.
b) What further mentoring and training do researchers need?
c) What they did/didn't like about the course.
d) Suggestions for next course in 2020.

13.00-14.00 Lunch and course ends