

Class 1: An introduction to Missing Data Analysis

Andrew Parnell
andrew.parnell@mu.ie



https://andrewcparnell.github.io/mda_course

PRESS RECORD

Let's get started

- ▶ Introduction from Novartis?
- ▶ Timetable for the week
- ▶ Pre-requisites

How this course works

- ▶ This course lives on GitHub, which means anyone can see the slides, code, etc, and make comments on it
- ▶ The timetable document (`index.html`) provides links to all the pdf slides and practicals
- ▶ The slides and the practicals are all written in `Rmarkdown` format, which means you can load them up in Rstudio and see how everything was created
- ▶ Let me know if you spot mistakes, as these can be easily updated on the GitHub page
- ▶ There is a `mda_course.Rproj` R project file from which you should be able to run all the code

Copyright statement

All the non-GitHub materials provided in the Introduction to Missing Data Analysis course are copyright of Andrew Parnell.

This means:

- ▶ As a user (the student) you have permission (licence) to access the materials to aid and support your individual studies.
- ▶ You are not permitted to copy or distribute any materials without the relevant permission
- ▶ As faculty we may reserve the right to remove a user in the event of any possible infringement

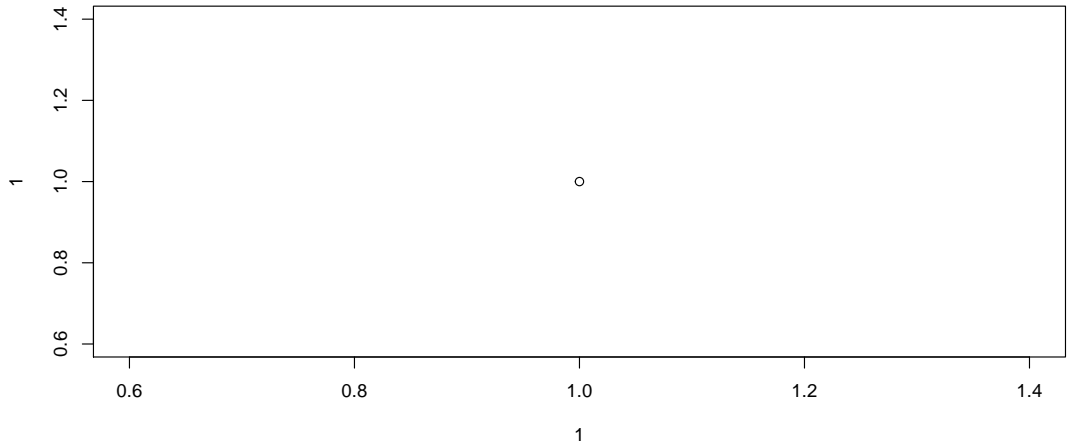
Course format and other details

- ▶ We do two lectures in the morning (separated by a coffee break) then have a practical class where we go through some R code together
- ▶ There is also an (optional) exercise sheet with further questions for those who are particularly keen
- ▶ If you want to send me a private message use the message board and I will try to answer them as we go
- ▶ Please ask lots of questions, but **MUTE YOUR MICROPHONE** when not asking them
- ▶ Some good books:
 - ▶ *Statistical Analysis with Missing Data* by Little and Rubin
 - ▶ *Bayesian Data Analysis* by Gelman et al (chapter on Missing Data Analysis)
 - ▶

Why care about missing data analysis?

Structure of this class

A simple example



Missing data analysis terrible jargon

MCAR MAR NMAR (MNAR)

Other jargon

imputation amputation(!) monotone missing ignorable non-ignorable Complete case (Ch 3) Available case (p61/73)

Missing data analysis common notation

Other examples where missing data occurs

Bivariate MAR

book P22

How people usually analyse missing data

Book P23 (p35)

Available case vs complete case analysis

Complete case (Ch 3) Available case (p61/73)

Simple imputation

P67 Ch 4

Explicit vs implicit imputation

P68/80

A warning about imputation

P67 Ch 4

A simple demonstration of the pitfalls of imputation

P81/69

Regression imputation

P71/83

General guidelines on imputation

P81/93

Estimating uncertainties with imputed data

P90/102

From simple to multiple imputation

P95/107