Introduction to Statistics Using R

Paul Gribble

Winter, 2019

Administrivia

- Tuesdays lecture 2:00–3:30pm in WIRB 1130
- Thursdays lab 1:30pm–3:00pm in WIRB 1130
- No formal office hours, contact me with Qs or to setup a meeting
  - paul [at] gribblelab [dot] org
  - Office: WIRB 4122
  - our TA is Kathleen Lyons: (klyons8 [at] uwo [dot] ca)
- All materials will be posted on the course website
  - `https://www.gribblelab.org/stats2019`
- we will use OWL for submitting assignments and managing grades
- Check your email regularly in case of updates/corrections/class cancellations

Course Goals

- understand the logic of statistics
- understand the rationale behind statistical tests
- learn about common statistical tests & procedures
- given a dataset and a question, how to proceed?
- limitations of statistical tests
- learn to use R for data analysis & statistical tests
- develop some degree of statistical / computational independence

Topics

- logic & rationale of different approaches to statistical analysis of data
- null hypothesis significance testing (NHST)
- we'll choose from a list of commonly taught techniques
- numerical / computational approaches (e.g. bootstrapping & resampling)
- maximum likelihood estimation (MLE)

Evaluation

- 70% assignments (7 marks x 10 assignments)
- 15% take-home midterm exam (15 pts)
- 15% take-home final exam (15 pts)

Resources

- Designing Experiments and Analysing Data: A Model Comparison Perspective (3rd Edition) by Scott E. Maxwell, Harold D. Delaney and Ken Kelley. Routledge (2017). ISBN: 978-1138892286
- There is a website accompanying the textbook with datasets from the book, and sample R code for many of the chapters: https://designingexperiments.com
- There is a package for R on the CRAN site that contains datasets from the Maxwell & Delaney text: `https://cran.r-project.org/web/packages/AMCP/`

Software

- The R Project for Statistical Computing: `http://www.r-project.org`
- RStudio `https://www.rstudio.com` (an R IDE)

Lectures

- in Lectures I'll highlight the main concepts covered in the readings
- my role is to introduce you to the basics of each topic and get you started with examples
- it's your responsibility to dig deeper when you need and/or want to

Sadistics

- statistics has a bad rap
- courses can be boring stressful, confusing
- often focused on rote memorization of recipies for statistical tests
- one can end up with little understanding of how and why (even after getting high grades in the course)

Sadistics

- limited statistical repertoire
- inability and/or little desire to confront new statistical challenges encountered in your own research
- "I was never taught to do X, so I can't do X"
- science is about exploration, you should be able to (and you should want to) adapt what you know and learn new things

Our Goals

- understand the rationale behind various statistical approaches
- enable you to reason your way out of any statistical jam you find yourself in
- it's the difference between only knowing how to heat up a frozen dinner versus figuring out for yourself how to cook new food from scratch based on a description of the meal

Our Approach

- understand how and why things are happening, rather than on low-level arithmetic
- I don't care if you memorize equations; personally I can never remember them
- what I do remember is the underlying logic and rationale
- this is far more important

Your Goals

- learning about ways to explore, visualize, and model your data should not be a chore—this is the fun part of science!
- You should be excited by:
  - your data
  - the idea of obtaining more data
  - looking at data in many different ways
  - thinking about a model for your data
  - thinking about what your data means
  - exploring and analysing your data in new ways
  - thinking about how to change your experiments to ask new questions

To Do Next

1. Install R & RStudio
2. Read Chapters 1 and 2 of Maxwell & Delaney