Using the R Commander in Basic Statistics Courses

John Fox

McMaster University
Hamilton, Ontario, Canada

American Statistical Association
Teaching of Statistics in the Health Sciences Section
March 2018

Introduction

- In statistics courses for undergraduates (and most graduate students) in disciplines other than statistics, teaching statistical computing isn't an end in itself.
- Statistical software should support the fundamental goals of the course.
- Because (what I take to be) the fundamental course goals differ by the level of the course, use of statistical software should differ as well.
- This webinar focuses on a typical introductory statistics course and explains why the R Commander graphical user interface (GUI) for R is a reasonable software choice.
- The R Commander is also a reasonable interface in other contexts, such as for occasional users of R.
- After a general introduction, the majority of the webinar will be devoted to a demonstration of the R Commander for a typical data-analysis exercise.

An Introductory Statistics Course

Course Goals

- The emphasis (in my opinion) should be on statistical reasoning and basic statistical concepts.
- Teaching data analysis is a secondary goal.
- Students completing the course should be able to read and critically assess research reports using basic statistical methods; use basic statistical methods in their own work; and pursue further course-work in applied statistics.

An Introductory Social Statistics Course

Basic Content (based, e.g., on Moore et al., 2015, *The Basic Practice of Statistics*)

- Examining Data.
  - Categorical vs. numerical data.
  - Distribution and variation.
  - Association (marginal vs. partial).
  - Visualizing data with statistical graphs.
- Producing Data.
  - Observational vs. experimental data (random selection vs. random assignment; causal inference).

An Introductory Social Statistics Course
Basic Content

- Basics of Statistical Inference.
  - Sampling variation and sampling distribution.
  - Confidence intervals.
  - Hypothesis tests.
  - Possibly likelihood, Bayesian inference.

An Introductory Statistics Course
Role/Properties of Appropriate Statistical Software

- Simplicity and transparency of use.
- Inexpensive availability on all commonly used platforms (Windows, macOS, Linux/Unix; in future maybe Chrome OS, iOS, web browser).
- "Low threshold/high ceiling" (borrowed from Seymour Papert and **Logo**).
- Consistency with course goals, integration into course content.
- Ability to use on students' own computers and after course ends.
- Encourage good habits (e.g., reproducible research).

R Commander Design Goals

Many present at the origin of the project, circa 2003 — see Fox (2005) and Fox (2017)

- Familiar menu/dialog-box interface.
- Simple work-flow, based on an "active," modifiable data set.
- Where appropriate, an active statistical model on which the user can compute; models are associated with and synchronized with data sets.
- Keep menus and dialog boxes simple, with uniform (tabbed) structure.
- Make it difficult to do unreasonable things (e.g., take the mean of a categorical variable).
- Simple to install on all platforms — implemented as a standard R package, the **Rcmdr** package (but needs XQuartz on macOS).
- Generate visible, editable, reusable scripts of R commands.
- Generate dynamic, editable, reusable documents with executable R code (in RMarkdown or LaTeX, compiled to HTML, PDF, or Word) to encourage reproducible research.

R Commander Design Goals

- Cover the content of a basic-statistics course and beyond, including statistical graphics.
- Extensibility (via plug-in packages).

Beyond the Basics
Goals for a Second Statistics Course (Briefly)

- The emphasis (again in my opinion) should be on practical data analysis.
- Teach statistical modeling as part of the work-flow of data analysis.
- Encourage sound data-analysis practices (e.g., examination of data, model criticism, reproducible research).
- Provide basis for further study, including self-study.
- A command-line interface with a programming editor and literate-programming tools, such as RStudio, is more appropriate than a GUI to these goals.
- The R Commander, however, is capable of supporting a second statistics course focusing on applied regression analysis, generalized linear models, and (via **Rcmdr** plug-in packages) beyond.

References

Fox, J. (2005). The R Commander: A basic statistics graphical user interface to R. *Journal of Statistical Software*, 14(9):1–42.

Fox, J. (2017). *Using the R Commander: A Point-and-Click Interface for R*. Chapman & Hall/CRC Press, Boca Raton.

Moore, D. S., Notz, W. I., and Fligner, M. A. (2015). *The Basic Practice of Statistics*. Macmillan, New York, seventh edition.

Other Resources

Website for R Commander book: `tinyurl.com/RcmdrBook`
Installation instructions: `tinyurl.com/Rcmdr-install`