Journal of Statistical Software

December 2016, Volume 75, Book Review 3. doi: 10.18637/jss.v075.b03

Reviewer: Tim Downie
Beuth University of Applied Sciences

Using the R Commander: A Point-and-Click Interface for R

John Fox
Chapman & Hall/CRC, Boca Raton, 2017.
ISBN 978-1-4987-4190-3. 213 pp. USD 47.96 (P).
http://socserv.mcmaster.ca/jfox/Books/RCommander/

The R Commander (**Rcmdr**, Fox 2005) is a graphical user interface (GUI) front end for R (R Core Team 2016). The aims of the R Commander are to provide a cross-platform GUI front end to R that is easy to use. The range of commands available are appropriate for a basic statistics course and some intermediate subjects, e.g., generalized linear models, are covered. Currently most of the resources for the R Commander are online tutorials (Karp 2010) and short online articles (Fox and Bouchet-Valat 2015). There are some books, that dedicate a chapter to this topic (Larson-Hall 2015), but *Using the R Commander: A Point-and-Click Interface for R* is the first book in English,[1] that I am aware of, which is dedicated specifically to the R Commander.

This book, part of *The R Series* by Chapman & Hall/CRC Press, is written by John Fox, the R Commander software developer. Fox explains his motivation in the preface, *"I felt that R's command line was an obstacle. I expected that eventually someone would introduce a graphical interface to R, but none materialized."* So he started on the project himself.

Chapter 1: The book starts with a quick explanation of R and the R Commander, including a historical context and an explanation of the book's overall structure.

The next chapter gives detailed instructions to downloading and installing the R Commander on Windows, Mac and Linux systems. This includes troubleshooting advice and running tips for each system. These tips are useful, however this is the kind of information that will soon be out of date.

Chapter 3 gives *A Quick Tour of the R Commander*. This chapter introduces the GUI menu system to new users. The book kicks off in earnest with the importing of data in delimited text format. There follows a worked example including data manipulation, summary statistics and graphics. A key feature of the R Commander is that all of the point-and-click commands generate R code. The resulting R commands and output can be saved and edited and re-run later and can be used as a stepping stone to learning R commands. At the end of the chapter

[1] Two books, which use the R Commander to demonstrate the statistical methods, are available in German: Wagner and Reisinger (2015) and Kronthaler (2015).

is a section on the RZarkdown interface, which generates a report of the data analysis in PDF or HTML format. The lack of presentation quality results had long been a weak point of S-PLUS and R. Now report generation is available, it is good that this subject gets a substantial airing early on in the book.

Chapter 4 covers data input and data management. The point-and-click approach to data handling is intuitive and much easier for the beginner than using R code. The subject is approached with practical data analysis in mind; for example, the default ordering of a variable's factor levels is often inconvenient. Fox comments on this and shows how to specify a user-defined ordering; many other statistical texts overlook such practical inconveniences. From this chapter on all examples are driven by real datasets. When the original data have to be checked, corrected or transformed then this is included in the example, rather than using datasets already in ideal form. The datasets are interesting albeit somewhat dated for current undergraduates. An unfortunate problem is that two different datasets on the theme "occupational prestige" are used and are confusingly similar. In particular one dataset records the variable `income` on a dollar scale, whereas the other has `income` as a percentage.

Chapter 5 introduces summarizing and graphing data with the R Commander. This covers standard methods such as contingency tables and bar charts. Throughout the book, all figures are printed in black and white, except for a small color section on high-quality paper at the end of this chapter. The color figures have been chosen carefully, using diagrams, for which the use of color is important to the book reader, rather than to the user.

Chapter 6 introduces simple statistical tests, including t tests, ANOVA (analysis of variance) tests, tests of a proportion, tests on variances and nonparametric tests. The author presupposes a good knowledge of the principles of hypothesis testing and confidence intervals. p values are reported without explanation of their meaning and the first mentioned confidence interval is a one-sided interval; a difficult concept for statistical beginners to understand. Due primarily to the structure of the the R Commander menus, the section on ANOVA models is here rather than in the fitting linear models chapter. The dataset for the one-way ANOVA example is unnecessarily complicated. The dependent variable is a mark out of 40, which is transformed into a proportion and then the logit function is applied. It is a nice example but inappropriate for a first ANOVA model.

Chapter 7 is called *Fitting Linear and Generalized Linear Models.* The notation used is not for beginners. The chapter starts with the algebraic definition of a linear model with multiple dependent variables. This will be familiar to statistics graduates but no doubt confusing to someone with a weak statistics background wanting to carry out a simple linear regression. A nice feature of the R Commander is that the output of a fitted model is saved as an R object and is available in the form of a "model". The model information e.g., residuals, diagnostic plots etc. can then be accessed and the book documents this feature well. The statistical modeling in this chapter is more thorough than in many books of this type. Someone with experience of another statistical program could learn the topics in the previous chapters on their own using the point-and-click approach directly. In this chapter there are explanations of R model formulae, contrast factors and R effect plots, aspects of the software, that are more efficiently learnt in a course or using a text book such as this.

Chapter 8 introduces *Probability Distributions and Simulation.* All of the commonly used R functions for probability distributions are available in the R Commander and are easy to use. Plotting the density, probability and distribution functions is clearly explained and an

introduction to random number seeds is given.

Chapter 9 covers plug-in packages for the R Commander. Like R, the R Commander allows programmers to contribute packages for other users. The section concentrates on just two packages **TeachingDemos** (Snow 2016; Fox 2007) and **survival** (Therneau and Grambsch 2000; Fox and Carvalho 2012). This chapter primarily describes the functionality of these two packages, but also underlines that it is worth investigating the further capabilities of the R Commander packages.

The appendix consists of screen shots of the R Commander menus and submenus. It is not clear to me to whom this section would be useful. The screen shots are version-dependent so will be out of date before too long.

On reading, I kept returning to the question; *who would benefit from this book?* Fox makes his intended target audience for the R Commander clear – mainly for use in basic statistics classes, whilst giving students of such courses an insight into more advanced statistical themes. Many users of the R Commander will be students taking a first statistics course, either as part of a mathematics/statistics degree or in a service course. I would not recommend this book as the primary course book for students to work through, even though the author directs many comments to this group. *"... if you're unfamiliar with least squares don't worry: You'll almost surely study the topic in your basic statistics course"* states a footnote in the introduction. The book is not too mathematical, but does often presuppose statistical knowledge, e.g., the normal distribution, p values and confidence intervals. The thread of the book is heavily driven by the design of the software and, as a result, it reads more like a user guide rather than a course text.

There will be three broad groups for whom the book *will* be useful. Firstly, those teaching statistics using the R Commander in a lecture course or short course. The breadth of the software capabilities is well-documented and there are many nice examples, succinctly described, which could be incorporated into computer workshops. Another target group would be applied statisticians, who are already familiar with one or more statistical programs and wish to learn R. This type of user already has a good working knowledge of the statistical procedures and can use the R Commander as a helpful stepping-stone to learning R. The third group will be those who have already learnt the R Commander as part of a statistics course and who move on to analyzing their own data. This book is well suited as a reference and a "refresher course".

The limitations of the R Commander are barely mentioned. There are some unnecessary limitations and restrictions in the R Commander, which are annoying to an experienced R user and perplexing to a new user. I feel that a comprehensive book on this subject should address the software limitations, if only to teach what the practical workarounds are.

In summary, this is a comprehensive book written in an approachable style, that serves as a good user guide to the R Commander but is less appropriate as a primary course companion. It gives comprehensive demonstration of the capabilities of the R Commander, which will be of benefit to those who are teaching the software, moving over from another statistical program or seeking a detailed revision.

References

Fox J (2005). "The R Commander: A Basic Statistics Graphical User Interface to R." *Journal*

of Statistical Software, **14**(9), 1–42. doi:10.18637/jss.v014.i09.

Fox J (2007). "Extending the R Commander by "Plug-In" Packages." *R News*, **7**(3), 46–52. URL https://www.R-project.org/doc/Rnews/.

Fox J, Bouchet-Valat M (2015). *Getting Started with the R Commander*. Version 2.2-1, URL http://socserv.socsci.mcmaster.ca/jfox/Misc/Rcmdr/Getting-Started-with-the-Rcmdr.pdf.

Fox J, Carvalho MS (2012). "The **RcmdrPlugin.survival** Package: Extending the R Commander Interface to Survival Analysis." *Journal of Statistical Software*, **49**(7), 1–32. doi:10.18637/jss.v049.i07.

Karp NA (2010). "R Commander an Introduction." URL https://CRAN.R-project.org/doc/contrib/Karp-Rcommander-intro.pdf.

Kronthaler F (2015). *Statistik Angewandt: Datenanalyse ist (k)eine Kunst mit dem R Commander*. Springer-Verlag, Berlin. doi:10.1007/978-3-662-47118-0.

Larson-Hall J (2015). *A Guide to Doing Statistics in Second Language Research Using SPSS and R*. Second Language Acquisition Research. Routledge.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Snow G (2016). *TeachingDemos: Demonstrations for Teaching and Learning*. R package version 2.10, URL https://CRAN.R-project.org/package=TeachingDemos.

Therneau TM, Grambsch PM (2000). *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, New York.

Wagner G, Reisinger CM (2015). *AlleR Anfang ist Leicht: Datenanalyse mit dem R Commander*. Facultas, Wien.

Reviewer:

Tim Downie
Beuth University of Applied Sciences
Department II
D-13353, Berlin, Germany
Email: tim.downie@beuth-hochschule.de

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
December 2016, Volume 75, Book Review 3 *Published:* 2016-12-06
doi:10.18637/jss.v075.b03