



---

# *Journal of Statistical Software*

February 2013, Volume 52, Book Review 4.

<http://www.jstatsoft.org/>

---

Reviewer: Norman Matloff  
University of California, Davis

---

## **The R Student Companion**

Brian Dennis

Chapman and Hall/CRC, Boca Raton, FL, 2012.

ISBN 978-1439875407. 360 pp. USD 39.95 (P).

<http://www.crcpress.com/product/isbn/9781439875407>

---

An R book for high schoolers! This is an excellent idea, and the quality of the product is equally excellent. It may be suitable for non-calculus-based introductory courses at the college level as well.

I believe there is a fairly broad consensus that the teaching of introductory statistics often shortchanges the students. Course use unmotivating examples are unmotivating, principles are reduced to mere formulas, and even the word *statistics* evokes an image of tedious computation. Though it evoked controversy, I side with Xiao-Li Meng's comments (Meng 2009) that the teaching of AP Statistics generally suffers from such problems.

Dennis' point is that one partial remedy is to modernize the manner in which computation is done. Use of pocket calculators for teaching high school statistics is anachronistic, given the ubiquity of computers and the availability of open-source software such as R. Not only is R free, opposed to expensive calculators, but R has an overwhelming advantage over calculators in aspects such as graphics and ability to handle real data sets. The AP Statistics Development Committee, as well as the ASA Statistics Education Section, should find this notion of interest.

This is a book about statistics and data, not programming. Loops are used only occasionally, and vectorization is barely mentioned. Functions are mainly called, rarely written. This de-emphasis of programming is the proper approach in my view, but a chapter previewing advanced topics in both statistics and programming might have been useful.

There is more about mathematics than one would expect. Each of Chapters 8–12 has a mathematical title, such as Chapter 9's "Trigonometric Functions". This might be viewed as a negative by students and even some teachers. The author explains that a good math grounding is necessary for learning R. While this is true at some level, there is actually more math content than is typical even in an introductory college course.

For example, I applaud Dennis for including material on multiple regression. As a topic that it is arguably the gateway to "real" statistics, this was an excellent choice for inclusion in the book. But I would question whether students at this level need to see the matrix formulation of linear models.

On the other hand, there is very little on statistical inference. For most students, introductory statistics will be the only exposure to the field they ever get. This is a shame, given how much statistics impacts their daily lives. For example, they will hear of margins of error in election surveys, and of one drug for hypertension being “significantly” better than another. They should be introduced to the meaning of such terms (and of the dangers in the latter one). Though there are some excellent discussions on the modeling process, I could find nothing on how models can mislead.

Dennis does a good job dispelling the “steep learning curve” myth concerning R, starting in the preface and culminating in a final chapter titled “It Doesn’t Take a Rocket Scientist” (a clever play on the astronomical content of that chapter).

The writing style is clear and lively, and the examples should appeal to high school students. It is high time that introductory statistics be taught in an engaging manner that reflects our own enthusiasm for the subject, with meaningful data sets, attractive graphics and so on. Dennis’ book is a fine contribution toward that goal.

## References

Meng XL (2009). “Statistics: Your Chance for Happiness (or Misery).” *Amstat News*.

## Reviewer:

Norman Matloff  
Department of Computer Science  
University of California, Davis  
Davis, CA, 95616, United States of America  
E-mail: [matloff@cs.ucdavis.edu](mailto:matloff@cs.ucdavis.edu)  
URL: <http://heather.cs.ucdavis.edu/matloff.html>

Computing  
Reviews


**Association for  
Computing Machinery**  
  
**ThinkLoud**

TODAY'S ISSUE
HOT TOPICS
SEARCH
BROWSE
RECOMMENDED
MY ACCOUNT
LOGOUT

REVIEW
?

## Review

Search  

### The R student companion

Dennis B., Chapman & Hall/CRC, Boca Raton, FL, 2013. 360 pp. Type: Book (978-1-439875-40-7)

Date Reviewed: Apr 3 2013

[Full Text](#)

Over the last few years, the programming language R has become something like the darling of open-source data analysis. The R website itself (<http://www.r-project.org/>) lists 121 books at the time of writing this review (February 2013), and more are sure to follow. What does this particular book have to offer? It is unique in that it does not present R for statistics. Rather, it presents R as a generic programming tool that will enable high school and college students to use computers to solve quantitative problems. It assumes no computer programming background, and it attempts to show that R can be used both to introduce students to programming and to enable them to tackle interesting real-life problems.

R has much going for it for such a role: it is highly interactive and interpreted; it can produce fascinating graphics; and it allows for the exploration of data with short sequences of commands. The fact, however, that R is not widely acknowledged as a means to introduce computer programming points to its potential shortcomings: R is markedly different from other, mainstream programming languages; the basic R building blocks, vectors, are intuitive to use for handling data, but not for general-purpose computation; and most material on R is aimed at statisticians and programmers dealing with advanced data analysis, not computing novices.

To address these issues, the author takes a carefully circumscribed approach. The book covers material for doing computations and writing scripts in R, but barely touches on doing statistics with it. This might make it seem that the author is twisting R, taking a tool for one thing and turning it into something else. That is true to some extent. However, it is equally true that to use R well, one must know not only its statistical libraries, but its core concepts as well. The book does a good job of presenting these core concepts in a crisp, educational way, while leaving no doubt that R remains a serious tool for solving real, not toy, problems.

The book kicks off with the very basics: how to start R and how to interact with it. The author then describes the fundamentals of R: scripts and functions. Readers are treated to a brief introduction to graphics with R, which will surely whet the appetite of young students. More R elements follow, such as input and output, loops, logic, and control statements. In tune with its target audience, the book goes on to cover quadratic functions, trigonometry, exponential and logarithmic functions, matrix arithmetic, and systems of linear equations. Advanced graphs follow, and only near the end of the book do we come across probability, statistics, and fitting models to data. The book then backs away from statistics, and concludes with an extended example of a script that calculates the trajectory of the earth.

Each chapter includes real-world examples, summaries with the key points learned, exercises, some references, and occasionally very interesting notes that provide some scientific background, historical material, and hints for further study.

Does this book succeed in what it set out to do? Motivated students will appreciate the power of R and learn how to use it to explore problems in other courses. They will even glimpse what's behind the furor over big data, data analytics, and all that. Instructors will find this an enjoyable book, and they may discover interesting ideas to pursue with their students. The book will fare well if it finds its way into the hands of willing young people and eager teachers.

Reviewer: [Panagiotis Louridas](#)

Review #: CR141101 (1307-0581)

**Recommendations**

 **Featured Reviewer**

---

**Related Topics**

| Browse                                       | Alerts              |
|--|---------------------|
| <a href="#">Mathematical Software (G.4 )</a> | <a href="#">Add</a> |
| <a href="#">General (D.3.0 )</a>             | <a href="#">Add</a> |
| <a href="#">Programming Languages (D.3 )</a> | <a href="#">Add</a> |
| <a href="#">Reference (A.2 )</a>             | <a href="#">Add</a> |

[Manage Alerts](#)
[More Alerts](#)



Would you recommend this review?  yes  no

Other reviews under "**Mathematical Software**":

|   | <b>Date</b> |
|---|-------------|
| <a href="#">R graphics cookbook</a><br>Chang W., O'Reilly Media, Inc., Sebastopol, CA, 2013. 416 pp. Type: Book (978-1-449316-95-2)   | Feb 27 2014 |
| <a href="#">Pattern recognition of big nutritional data in RCT</a><br>Wang J., Fang H., Wang H., Olendzki G., Wang C., Ma Y. BodyNets 2013 (Proceedings of the 8th International Conference on Body Area Networks, Boston, MA, Sep 30-Oct 2, 2013) 394-400, 2013. Type: Proceedings | Feb 18 2014 |
| <a href="#">Pro data visualization using R and JavaScript</a><br>Barker T., Apress, Berkeley, CA, 2013. 216 pp. Type: Book (978-1-430258-06-3)  | Dec 17 2013 |
| <a href="#">more...</a>   |             |

[E-Mail This](#)

[Printer-Friendly](#)

**REVIEWER'S AREA**

**MASTHEAD**

**SUBSCRIBE**

**PRESS**

**TIPS**

**HELP**

**CONTACT US**

Select Language

Powered by Google Translate

Reproduction in whole or in part without permission is prohibited. Copyright © 2000-2014 ThinkLoud, Inc.

[Terms of Use](#) | [Privacy Policy](#)

**The R Student Companion**

Brian Dennis

Chapman and Hall/CRC, 2013, xvii + 339 pages, £25.99/\$39.95, softcover

ISBN: 978-1-4398-7540-7

*Table of Contents*

- |   |   |
|---|---|
| 1. Introduction: getting started with R   | 11. Matrix arithmetic                               |
| 2. R scripts                              | 12. Systems of linear equations                     |
| 3. Functions                              | 13. Advanced graphs                                 |
| 4. Basic graphs                           | 14. Probability and simulation                      |
| 5. Data input and output                  | 15. Fitting models to data                          |
| 6. Loops                                  | 16. Conclusion – it doesn't take a rocket scientist |
| 7. Logic and control                      | Appendix A: Installing R                            |
| 8. Quadratic functions                    | Appendix B: Getting help                            |
| 9. Trigonometric functions                | Appendix C: Common R expressions                    |
| 10. Exponential and logarithmic functions |   |

*Readership:* Beginners in scientific computation, such as high school and undergraduate students, learning to use R.

This book requires no prior knowledge of calculus, programming, or statistics. The author seems to have an aim at providing an accompanying book for high school and college science courses that use (or should use) R as a substitute for graphing calculators. A typical chapter in this book presents a review of the underlying mathematical concepts that are to be implemented in R, complete with a 'what we learned' section and some quite compelling computational challenges.

In the preface, the author seems confident that his rather lengthy introductory R book is of considerable help to the students. This is true at least in the sense that basically all the commands and real-world examples are explained very thoroughly. This should make the book suitable for self-study and hold interest for the target group (high school and college level students), but it also makes this book less attractive to more advanced students, for whom a less detailed presentation should be sufficient. Especially for statistics students, there are more appropriate alternatives, such as Dalgaard's 'Introductory Statistics with R' (2nd ed., Springer, 2008) and Albert & Rizzo's 'R by Example' (Springer, 2012) to name but a few.

Joonas Kauppinen: [joonas.kauppinen@uta.fi](mailto:joonas.kauppinen@uta.fi)  
 School of Information Sciences  
 FI-33014 University of Tampere, Finland

**The Black Swan: The Impact of the Highly Improbable, Second Edition**

Nassim Nicholas Taleb

Random House, 2010, xxxii + 444 pages, \$17.00, softcover

ISBN: 978-0-8129-7381-5



**The R Student Companion.** By B. Dennis. Boca Raton, FL: Chapman Hall/CRC. 2013. 339 pages. UK£26.99 (paperback). ISBN 978-1439875407.

My approach to this book is to treat it as an introductory text for learning the R language. Consequently I am more concerned with how it presents the R language than with its Mathematical or Statistical content. Norman Matloff's (2013) review provides useful comments on the latter.

One of the positive features of this book is its fearlessness. The message is that you are never too young to learn R (the book is aimed at high school and college students) and that the best way to learn is to jump straight into writing R expressions. The author's enthusiasm for tackling challenges head on also shines off the page. There is a concern that sometimes the book may get a little too fearless; for example, some of the exercises are quite ambitious. The monthly mortgage payments calculations starting on page 25 is quite a lot to take in for someone just getting started.

I am also in favour of the hands-on approach of the book; it is clearly written to be read while sitting at a keyboard. However, the approach does sometimes stray into rough-and-ready coding. For instance, starting on page 55, the author presents an example from economics and political science wherein the reader is encouraged to type a substantial number of data values manually within R expressions in a text file. The reader is then encouraged to save a copy of this text file under a new name before adding further R code to work with the data. It would be better practice to read data into R from something pre-prepared, like a CSV file, and it would be better practice to `source()` common R code rather than creating multiple copies of the R code. It must be acknowledged that the author faces non-trivial problems in respect of how to order topics, and these problems include the risk of frustrating the reader by placing file input/output before more interesting material. Nonetheless there is a danger in the current approach of encouraging poor code management habits.

The problem-based style of the book is another plus, where learning is motivated by first setting up an interesting question to answer. For example, I enjoyed the 'molar mass' example starting on page 46. The book also elevates R above just its use in Statistics and emphasises the fact that it is an enormously powerful tool for Scientific and Mathematical exploration in general. The examples also come from a very wide range of areas of application. One downside is that this approach can sometimes lead to a somewhat chaotic ordering of topics as the reader encounters R concepts in the order that the problems present them. An example is the sudden introduction of `NaN` values early on in the discussion of functions in Chapter 3.

A major disappointment with the text is the fact that the R code samples do not provide a good role model for code layout: they lack indenting and do not (consistently) use white space around operators or between function arguments.

Overall, this book provides a lively and interesting introduction to R. It furnishes a lot of useful material for introducing a very powerful tool to enquiring young minds. However, the successful use of this text within the classroom may require the firm hand of an instructor with a solid understanding of both R and code management practices.

PAUL MURRELL

The University of Auckland  
e-mail: paul@stat.auckland.ac.nz

#### Reference

MATLOFF, N. (2013). The R Student Companion: Book Review 4. *J. Statist. Soft.* **52**, 1–2.

Dennis, Brian

\* 1260.00016

**The R student companion.**

**Boca Raton, FL: CRC Press (ISBN 978-1-4398-7540-7/pbk). xvii, 339 p. £ 25.99 (2013).**

R is a free open access software package for scientific computations. It can be used for simple calculations as well as for advanced programming applied in scientific simulations. This book provides the easy introduction on how to use R in typical applications taken from different sciences.

The book consists of sixteen chapters and three appendices. Throughout the chapters the following topics are explained: writing R scripts, writing functions, drawing graphs, data input and output, quadratic, trigonometric, exponential and logarithmic functions, matrix arithmetic, linear equations systems, probability and simulation, fitting models to data. Many mathematical problems are explained as the practical applications of the presented theory as well. In appendices the advices on how to install R and getting help are gathered.

The book is not a typical R manual, it describes how to solve common mathematical problems drawn from all sciences, discussed in the book, using R in a simple way. Each chapter consists of a short description of an R aspect, typical mathematical problems to be solved and R solutions of these problems. Each chapter ends with a summary and exercises. The book is full of practical examples. It is suitable for students and all who are familiar with some high school algebra and has no computer programming background.

*Agnieszka Lisowska (Sosnowiec)*

## The R Student Companion



**Brian Dennis**

|                          |                    |
|--------------------------|--------------------|
| <b>Publisher:</b>        | Chapman & Hall/CRC |
| <b>Publication Date:</b> | 2016               |
| <b>Number of Pages:</b>  | 339                |
| <b>Format:</b>           | Paperback          |
| <b>Price:</b>            | \$44.95            |
| <b>ISBN:</b>             | 9781439875407      |
| <b>Category:</b>         | Manual             |

MAA REVIEW

TABLE OF CONTENTS

[Reviewed by Jason M. Graham, on 08/10/2016]

Beginning in the 1980s, commercially developed calculators with graphing and programmable capabilities have been available to students and others for aid in solving various mathematical problems. In parallel, many mathematics software packages were developed for use by both students and professionals in mathematics, science and related fields. Examples of such software are the popular [MATLAB](#) for numerical computing with some symbolic capability and [Mathematica](#) for symbolic mathematical manipulation with some numerical capability. Both MATLAB and Mathematica as well as a variety of other mathematical software packages have in addition the ability to produce high quality graphics for plotting functions, etc. One can also write and execute programs using for-loops and other standard computer science techniques in MATLAB and Mathematica. Other popular tools of this kind are [Maple](#), [Mathcad](#), and statistical software such as [SPSS](#) and [SAS](#). The accessibility of sophisticated calculators and mathematical software has revolutionized both teaching and problem solving in mathematics, science, and engineering. Of course dramatic advances in both hardware and software continue to occur that makes mathematical software ever better.

But there is another revolution in mathematics related computing that is well in motion. This is the development, improvement, and increasing ubiquity of freely available and powerful open access mathematical computing tools. Some important examples of this are [Octave](#), [Sage](#), [Julia](#), [Python](#) with its packages such as NumPy, SciPy and SymPy; and [R](#). The fact that all of these are free is impressive enough. What is more, they all have a widespread, diverse, and most of all, very dedicated community supporting them in order to insure a high quality product available to anyone anywhere at no financial cost. This is certainly the case with R.

Originally developed for statistical analyses and graphics, R is among the most popular and most powerful of the open source tools for mathematical computing. Together with an amazing number of

user developed add-on packages and a variety of very well developed integrated development environments, R is as powerful and useful as any commercial mathematics or statistics software package. However, a newcomer to R, and in particular a student, may find it a bit intimidating to use. Brian Dennis's *The R Student Companion* is meant to address this issue. And it does this successfully, especially for students with little background or experience in advanced mathematics or computer programming. A senior in high school or a first-year college student could easily pick up this book and immediately start doing interesting things with R.

What do you get out of Dennis's book? You learn how to install R and add-on packages. You learn the basic syntax and discover how to make plots and do basic computations using R. A reader of this book will learn how to do some essential programming in R, and also see it applied to solve interesting example of real scientific problems. Indeed, *The R Student Companion* makes picking up R quick and easy.

Some of the nicest features of *The R Student Companion* are the end of chapter summaries and exercises, the available R code, and of course the already mentioned real world applications. I would strongly recommend for any student entering college intending to study a science-oriented field to learn R. If you want a resource to help, then *The R Student Companion* is certainly recommended.

[Jason M. Graham](#) is an assistant professor in the department of mathematics at the University of Scranton, Scranton, Pennsylvania. His current professional interests are in teaching applied mathematics and mathematical biology, and collaborating with biologists specializing in the collective behavior of groups of organisms.

**Tags:** [Mathematical Software](#)  
[Statistical Software](#)  
[Log in](#) to post comments

---

**Mathematical Association of America**  
**P: (800) 331-1622**  
**F: (240) 396-5647**  
**E: [maaservice@maa.org](mailto:maaservice@maa.org)**

Copyright © 2017

## Book review

Statistical Methods in Medical Research  
2016, Vol. 25(5) 2394–2395  
© The Author(s) 2013  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/0962280213487424  
smm.sagepub.com



Brian Dennis, *The R student companion*. CRC Press: Boca Raton, 2013; 339 pp.  
ISBN 978-1-4398-7540-7, \$39.95 (pbk).

**Reviewed by:** Virgilio Gómez-Rubio, Department of Mathematics, University of Castilla-La, Mancha

The R programming language is widely used for teaching and research in statistics and other related fields nowadays. This book is an accessible manual for high school and college students to learn the R language. Throughout the book, the author has assorted a wealth of examples on general science and mathematics that are developed using R.

The first five chapters of the book explain the basics of R such as main data types, how to edit and run scripts, basic control structures, and simple plots. In Chapter 6, loops are explained, whilst Chapter 7 covers logic structures and controls. The contents are clearly presented and they could also be used for teaching a generic programming course with R.

Next, the book devotes different chapters to important mathematical functions, such as quadratic functions (Chapter 8), trigonometric functions (Chapter 9), and exponential and logarithmic functions (Chapter 10). This is seldom discussed in other R manuals, but these topics are of interest if the focus is on applied mathematics rather than statistics.

The book also provides a good introduction to matrices (Chapter 11) and solving linear equations (Chapter 12). Fitting linear models is explained using simple matrix algebra. R graphics are revisited in Chapter 13. Here, previously used graphical functions are described in more detail and some new other are introduced.

The author provides an introduction to probability in Chapter 14, where he works with the main probability distributions with R. He also shows how to perform simple simulations from different distributions using R. How to fit general linear models and non-linear models is explained in Chapter 15. General linear models are fitted using simple linear algebra, whilst for non-linear models the `nls` R function is employed.

Each chapter is full of real life examples and includes some final remarks with a summary of the main points in the chapter. There are also a good number of extra exercises (to highlight particular computational and mathematical challenges). This is interesting for students, who can test their skills, and lecturers, who have plenty of activities for the classroom. A reference list is also provided with relevant papers and books.

Finally, the author develops a full example in the last chapter. This example is based on computing Earth's orbit around the Sun. This problem is solved using basic mathematics and physics, and it is suitable for a high school student. Computations are dealt with R, and a script with the solution is provided.

The book includes appendices with information on how to install R, getting help and a summary of R expressions, which can be very helpful when getting started with R.

---

In a nutshell, the *The R Student Companion* provides a nice and accessible guide to the R language. The book does not focus on statistics (apart from some basic probability) but there are plenty of examples from applied mathematics and the level is suitable for high school and college students. In addition, the contents can be used to teach a course on general programming for science and engineering students.

statistics and specific plots of data are discussed as the initial stage of the larger data analysis process (including inference) for that type of data.

Any qualms that I have with the text are generally minor. The authors have a tendency to mention topics before they have been introduced; for example, on pp. 22–23 when describing JMP's emphasis on building context for data analysis, they mention histograms, *t*-tests, scatterplots, and nonparametric methods before any statistical topics have been introduced. Additionally, I do not care for the fact that the book (and JMP) uses the term "histogram" as the name of the displays for both a single numerical and a single categorical variable (e.g., p. 293 and 303).

My biggest qualm (and the basis for my reservation to suggest that this be used as a standalone teaching text) is that the authors tend to skip on details in some places. (I realize that, in some instances, this is by necessity, otherwise the text would be gigantic.) The most troubling offense is the discussion of the plots generated as part of a logistic regression analysis (p. 318). These plots consisted of the fitted logit curve as well as the "data points." The values plotted on the *x*-axis made sense for the example—they were in the dataset. I could not, however, make heads or tails of what was being plotted on the *y*-axis—they were values between 0 and 1, but were nowhere to be found in the data table nor could they be constructed from information in the data table or output. What was more troubling was that when I repeated the analysis for the example dataset a second time, different "data points" were plotted. To be fair, neither the text nor the JMP Help menu gives a clear explanation of what is being plotted. (It appears that the *y*-coordinates are randomly generated.)

On the whole, the organization, layout, and clear step-by-step instructions make the fifth edition of *JMP Start Statistics* a great reference text. I would recommend it as a supplement to a more standard statistical text (possibly in a graduate course that uses JMP, since more material is likely to be covered) or for self-learners interested in using JMP (while it would not be necessary to have a background in statistics, some prior exposure would be helpful). I would also say it is a must-have for instructors who want to start using JMP in their courses; it has some great examples and scripts to use for teaching! An e-book version (ISBN 978-1-61290-307-1) is also available from the publisher.

Jessica L. CHAPMAN  
St. Lawrence University

#### R Statistical Application Development by Example: Beginner's Guide.

Prabhanjan Narayanachar TATTAR. Birmingham, UK: Packt Publishing Ltd., 2013, vi + 324 pp., \$44.99(P), ISBN: 978-1-84951-944-1.

*R Statistical Application Development by Example* authored by Prabhanjan Narayanachar Tattar is a nice compact book that integrates R programming with popular statistical methods. For each topic introduced, an example is provided as well as code implementing the example with R, which is followed by discussion. The cover of the book states: "Learn by doing: less theory, more results," and this book accomplishes exactly that.

The book serves well for both undergraduate and graduate students who are inexperienced with R. It does not devote much space developing the mathematical background for each topic but instead briefly introduces what the reader needs to know, keeping notation to a level that is easily understandable. Each chapter begins with a description of the overall objective, the topics that will be covered, and what the reader can expect to learn. After a topic has been introduced, there is a section called: "Time for action." Here, the authors carefully walk the reader through an example and provide detailed R code. Snapshots of output and graphs are provided to further enhance the readers' comprehension. Following the example, there is a "What just happened" section, where the authors summarize what the reader learned and accomplished. The book also contains "Have a go hero" sections that challenge the reader to carry out exercises on their own.

Chapter 1 begins with an introduction to different types of variables (categorical and continuous), R installation, and its various packages. It then proceeds to give a crash course on some of the more popular discrete and continuous distributions, where relevant R code is given to familiarize the reader with these distributions in an R environment. Chapter 2 goes over the fundamentals of objects, vectors, matrices, and lists. Next, the chapter details how to store variables in data frames followed by a section on how to import data from various types

of external files and the various types of R functions available to accomplish this task. There is then a short discussion on the programming language SQL and the chapter concludes with a section on exporting data and graphs and how to manage an R session. Chapter 3 discusses graphical techniques for categorical variables and continuous variables. Chapter 4 focuses on exploratory analysis, with a special attention to methods that are robust to outliers, such as quantiles, the median, and interquartile range. It proceeds with stem-and-leaf plots, letter values, and bagplots (bivariate boxplots) and then it discusses resistant lines and smoothing techniques. The focus of chapter 5 is statistical inference. The chapter opens with the definition of the likelihood function, presents them for some of the well-known distributions, and provides code for how to plot them. It then discusses how to find maximum likelihood estimators by hand and using R. The chapter wraps up with interval estimation and hypothesis testing. Chapter 6 covers linear regression analyses and provides an example complete with code, output, and interpretation. The chapter also discusses ANOVA within the regression context, multiple regression, model building, confidence intervals, diagnostics, and model selection. Chapter 7 is concerned with methods when the outcome variable is binary, including probit and logistic regression, model validation, diagnostics, and ROC curves. Chapter 8 discusses alternative regression methods, such as polynomial regression, splines, piecewise linear regression, and ridge regression as well as associated issues such as over fitting and model assessment. Chapters 9 and 10 introduce the reader to classification and regression trees (CART), and related topics such as pruning, bagging, bootstrapping, and random forests.

Overall the book is a great resource. It clearly spans a wide variety of topics and they are all supplemented with detailed examples. It does have a few problems, however. There are a number of typographical errors, including notational errors, which caused me some confusion. (Presumably, these will be corrected in later printings or editions.) Some of the snapshots of R output and graphs were blurred and/or used print that was too small making them difficult to read. The "Pop quiz" sections described in the book's preface as being intended to test the readers understanding, located in chapters 8 and 9, were not multiple choice as the preface stated they would be.

In summary, this book should nicely fit the needs of the beginning R programmer who wants to develop a working knowledge of the software. Although other quality options such as the cookbooks of Crawley (2013) and Teetor (2011) cover a greater range of statistical topics, Tattar's book has a friendly format, numerous examples, detailed code, and commentary that make it a useful addition and a valuable learning resource.

Roberto C. CRACKEL  
University of California, Riverside

#### REFERENCES

Crawley, M. J. (2013), *The R Book* (2nd ed.), West Sussex, UK: Wiley.

Teetor, P. (2011), *R Cookbook*, Sebastopol, CA: O'Reilly.

#### The R Student Companion.

Brian DENNIS. Boca Raton, FL: CRC Press, 2013, xvii + 339 pp., \$41.95(P), ISBN: 978-1-439-87540-7.

R is a widely used, "free software environment for statistical computing and graphics" (R Core Team 2013). However, as the author of this book points out, "...R goes way beyond statistics. It is a comprehensive software package for scientific computations of all sorts, with many high-level mathematical, graphical, and simulation tools built in" (p. xiv). As this statement may suggest, this book does not focus on the statistical capabilities of R. Rather, it provides an introduction to basic R skills for novice programmers with the idea that by mastering these skills, students will be well prepared to use R in upper-level high school and college mathematics, statistics, and other science courses.

This book is written for readers with only a moderate knowledge of high school algebra, and focuses on using R for mathematical and scientific calculations that precede calculus and statistics courses. The author states, "Anything in science, mathematics, and other quantitative courses for which a calculator is used is better performed in R" (p. xiv). The instruction in most chapters is two-fold: a review of a mathematical, scientific, or statistical concept paired

with a hands-on demonstration of R using data that have been published in a scientific journal article. The reader is encouraged to practice coding in R using the examples that are given in the text and using the Computational Challenges that are listed at the end of each chapter. Both the mathematical and coding instruction is written in a conversational tone, which is well suited for the intended high school and undergraduate audience.

Chapters 1–5 introduce fundamental concepts of using R. Chapter 1 shows how basic mathematical operations are performed in the console window and demonstrates how vectors are used in R. During this introduction, the reader is reminded of the algebraic order of operations. This chapter concludes with simple graphing commands and a real-life ecology example. Chapter 2 demonstrates some “best-practice” coding techniques while introducing the use of R scripts. The example for this chapter is a compound interest equation, which is derived and then coded in an R script. Built-in and user-defined functions are introduced in Chapter 3, and an example is given of both types of functions to calculate the molar mass of a compound. The graphical tools described in Chapter 1 are enhanced in Chapter 4, where various one- and two-variable graphical techniques are demonstrated using an economic/political dataset. Data input and output is the topic of Chapter 5, where a “large” dataset ( $n = 90$ ) relating university grade point averages to other variables is employed to demonstrate the use and manipulation of data frames in R.

Chapters 6 and 7 focus on programming tools such as loops, logical and Boolean operators, and conditional statements. Loops are demonstrated using a population growth example, and the other tools are used to simulate baseball at-bats, and to construct a profile plot for a blood pressure dataset. Chapters 8–10 mathematically describe quadratic, trigonometric, and exponential functions, and then R code is used in conjunction with real-life data to explore these functions further. Data relating fishing effort to fishing yield are modeled in Chapter 8, distances to nearby stars are calculated in Chapter 9, and examples such as radioactive decay and population growth are examined in Chapter 10. Chapters 11 and 12 cover simple matrix algebra concepts, with Chapter 11 focusing on vector and matrix operations and Chapter 12 focusing on solving systems of linear equations using matrices. Population growth is again studied, this time using matrices, and the time to the next Old Faithful eruption is solved using a system of linear equations. Advanced tools used to customize graphs are described in Chapter 13, as are three-dimensional plots. The latter are used to show sums of squares for the Old Faithful eruption data using a range of possible slope and intercept values. To prepare the reader for simulation capabilities of R, probability is discussed in Chapter 14. This discussion is followed by a simulation of randomly fluctuating stock prices. Linear and nonlinear statistical models are discussed in Chapter 15, and the methods are demonstrated using the fishing effort to fishing yield data and population growth data of earlier chapters. In Chapter 16, all of the R and mathematical tools are combined in a large example that models the Earth’s trajectory around the Sun.

The book concludes with three appendices. Appendix A informs the reader how to install R. Appendix B details the various methods for obtaining help in R. Appendix C is a short (14 page) mini-manual detailing R commands. Many of these commands appear in the main text of the book, while other commands allow the user to begin to explore additional R capabilities.

With the Computational Challenge problems given at the end of each chapter, the format of this book is like a textbook. In fact, the book itself mentions presenting what is learned in front of the class (see, e.g., p. 1), and in the Preface the author gives ideas of how this book could be used in a classroom setting. However, due to the assortment of mathematical/scientific topics, this book may not fit well into a typical high school or college curriculum. While the wide variety of mathematical and scientific topics make the book interesting, they also make the book difficult to integrate into a standard course. With that being said, this is a good book for high school or college students wanting to learn R on their own. Complete mathematical explanations paired with computational examples in R provide an excellent tool for these students to obtain a solid foundation in R.

There are many other books and online resources that provide an introduction to R. However, unlike many of the other books, this book does not target an upper-level statistical audience. Instead, it attempts to engage the mathematically less-mature reader at a level that is appropriate for him/her to form a solid understanding of R. The author acknowledges that, “... R is huge, and this book is not meant as a comprehensive guide to R” (p. xiii). The premise is, that by allowing students to learn how to handle quantitative problems that arise in the high school and college curriculum, they will be well prepared to handle more

difficult material in the future. By having a solid understanding of R tools for the more simple material, they will be able to extend their R knowledge to handle more complicated scientific computations with minimal difficulty.

Erin R. LEATHERMAN  
West Virginia University

## REFERENCE

R Core Team, (2013), *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing. Available at <http://www.R-project.org>.

Reproducible Research With R and RStudio.

Christopher GANDRUD. Boca Raton, FL: CRC Press, 2014, xxv + 288 pp., \$69.95(P), ISBN: 978-1-466-57284-3.

This book fills a niche in the sense that it provides some broader context to the overall data analysis process than Yihui Xie’s *Dynamic Documents with R and Knitr*. Xie’s (2013) book is technical, and gives many specific examples and explanations, with particular focus on the R package, knitr. Gandrud’s book, on the other hand, is more general and gives a basic overview of the entire research process. Gandrud has written a great outline of how a fully reproducible research project should look from start to finish, with brief explanations of each tool that he uses along the way. Something that Gandrud does very well is to provide sources that expand upon his skeletal introductions to topics. It is a good place to start when you need to know where to go next, kind of like how people tend to use Wikipedia.

The book has four parts. The first tells you how to get set up to begin conducting reproducible research, including an explanation of the topic, starting to use R, RStudio and knitr, and file management. The second part focuses on data, collecting, storing, sharing, and cleaning data. This section even introduces version control systems using github (<https://github.com/>) for collaborative projects. The third part focuses on analysis and results. The fourth part wraps up with producing pdf and web reports, and presentations.

This book is very readable. Readers with little-to-no experience with R and RStudio may find it difficult at times to get through the code, but usually there is an explanation of the purpose or function after each code chunk.

There are several typographical errors (e.g., “e” or “c” is omitted from “cache” in the arguments to a chunk of knitr code), which would cause problems if the reader were to blindly type in the code and expect it to run. Wait! Is not the purpose of reproducible research that code chunks embedded in the document will actually run? Here is where there is a small difficulty in using knitr versus explaining knitr. To explain knitr one needs to write the knitr commands “<< ... >>= ...@” into the document:

```
<<example, cache=TRUE>>=
# Create data
sample <- rnorm(n=1000, mean=5, sd=2)

# Save sample
save(sample, file='sample.Rdata')
@
```

but when the document is processed these lines are part of the control commands, and only the lines corresponding to R code will actually be printed in the resulting document. So to explain how to use knitr requires writing lines that look like code but are not executed, or to use some very clever tricks. Gandrud has clearly used the former approach, which has led to the typos. However, the author provides links to the github pages where errors/typos are addressed. If the reader finds a mistake, she can easily check this site to see if it has been reported, and if not, she can submit the error report. As for quality of the book, it covers a lot of good material, but could be improved and made more relevant with additional examples and figures.

If this book goes into a second edition, a recommendation is to include more figures! Good graphics are an important part of data analysis and particularly plots made with the ggplot2 package (Wickham 2009) assist with reproducibility. Graphics produced with this package have an advantage of being described



Brian Dennis <dr.of.chaos@gmail.com>

---

**i think there's must be something wrong in chinese edition (R student companion)**

1 message

---

**王博** <imcarrot@yeah.net>  
To: "dr.of.chaos" <dr.of.chaos@gmail.com>

Sat, Sep 24, 2016 at 7:17 PM

Dear Doctor Brian Dennis,

I am a Chinese reader of The R Student Companion, I want to say I really love this book. Unfortunatly, I've found something wrong in chinese edition (R student companion) at the page of P77, the part of 6.2, the 3<sup>rd</sup> part,,the vector's name should be num.fibs, but it's num.bifs.

I want to say thanks to you , because I've learned a lot from your book.,not only R knowledge ,but also my math skill. It helps me a lot ,and as a chinese I feel the AMERICAN HUMMOR.

Wish you happy.

Wang Bo from china.

发送自 Windows 10 版邮件应用