Journal of Statistical Software

August 2015, Volume 66, Book Review 5. http://www.jstatsoft.org/

Reviewer: Robin K. S. Hankin
Auckland University of Technology

Circular Statistics in R

Arthur Pewsey, Markus Neuhäuser, Graeme D. Ruxton
Oxford University Press, Oxford, 2013.
ISBN 978-0-19-967113-7. 183 pp. GBP 24.99 (P).
http://ukcatalogue.oup.com/product/9780199671137.do

Circular Statistics in R by Pewsey, Neuhäuser and Ruxton provides an accessible and R-centric introduction to circular statistics, together with an online resource furnishing the R code used in the book. Circular statistics is a topic whose global usefulness is matched by its poor coverage in statistical texts: just now I scanned my bookshelf of undergraduate-level statistics textbooks but not one even mentioned circular data and indeed the authors claim that only six books have been published that treat circular data in depth. Perhaps only compositional data (Aitchison 1986) shares this peculiar status of being ubiquitous yet so frequently overlooked.

In today's world of instant online access and Wikipedia pages on everything, one has to wonder what the role of a textbook is; after all, one can search for, and instantly find, in-depth articles on virtually any topic imaginable. When pressed to list advantages of books over internet sources, typical responses include the following: A coherent provision of material at a uniform level; some expectation of global – or at least broad – coverage of a topic; uniform notation and style; structured development of ideas, following a logical order; and just perhaps one might wish for a sprinkling of authors' idiosyncracies, the better to learn from.

I am happy to report that the book satisfies all these needs and more, and, with a few easily accommodated caveats, would be happy to recommend it.

The book has two somewhat overlapping audiences: Firstly, those new to R and who wish to apply it in an area familiar to them (in this case, circular statistics); and secondly, those familiar with R and who wish to learn something about circular statistics. It seems to me that the first group is well catered-for: Assumptions about the reader's mathematical background are reasonable and well executed. The second group, however, is not so well-served.

The authors explain that they wish to "promote a modern, computer-based approach to the analysis of circular data". The book's approach is that of a new breed of statisticans (of which I am one) for whom the measure of understanding of a statistical concept is the ability to manipulate that concept in R.

The book relies heavily on the **circular** package which, though generally well-written, does not (as of July 2015) pass the CRAN quality control (QC) procedures, and was last modified in 2011. Packages failing QC are sometimes orphaned and if this were to happen the book's long-term usefulness would be severely diminished.

The book has some interesting coverage imbalances: The von Mises distribution is treated only very briefly, justified by the authors on the reasonable grounds that they wished to focus on more computationally accessible concepts such as asymmery and peakedness, of interest to modern practitioners. Directional data (points confined to the unit hypersphere) is omitted entirely on the grounds that computational support for such data is not well-supported in R. The authors give only a very brief discussion of axial data.

The book was let down by poor readability and quality of the code fragments; also, the use of a non-fixed-width font for the R idiom made the structure of the coding difficult to see. Indenting was not used as a visual aid, and this made it hard for me to discern the overall structure. There seemed to be a large number of clumsy "for" loops lurking about, and annoying repeated use of a "stopgo" construction, where I would have expected clean vectorized code or at least standard do-while or equivalent idiom. The authors could have defined and used helper functions to perform acceptance-rejection.

The code on pages 63–73 was particularly bad: It used unintelligible variable and function names, and together with the inappropriate font, lack of indentation, and unvectorized code, was effectively incomprehensible. It left me wondering what the value of the printed code was if deciphering it was such a difficulty.

Also, it would have been nice to include the R output of the code in verbatim format (as in an `Sweave` package vignette, for example). This, together with instructive use-cases would have allowed me to understand, and perhaps modify, the idiom. The authors could have provided Rd files documenting their functions and showing exemplary and informative use-cases; this would have greatly increased the utility of the work.

But all in all, this was a great little book. I found a great deal to praise and, apart from the R code appearing in the text, not very much to criticize.

References

Aitchison J (1986). *The Statistical Analysis of Compositional Data*. The Blackburn Press.

Reviewer:

Robin K. S. Hankin
Auckland University of Technology
Private Bag 92006

Auckland 1142, New Zealand
E-mail: hankin.robin@gmail.com
URL: http://www.aut.ac.nz/

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume 66, Book Review 5 *Published:* 2015-08-06
August 2015