Reviewer: Jan de Leeuw
University at California at Los Angeles

Robust Statistical Methods with R

There is a rapidly increasing number of books with titles "Something with R", where "Something" is some area of statistics. Clearly this is a good development from the point of view of JSS: statistical software gets more attention than it did in the "Without R" era. I think it is also good from a somewhat broader perspective: paying more attention to software blends applied and theoretical aspects of statistics, and illustrates the fact that statistics is properly defined as the development and study of techniques for data analysis. For those of us who are so inclined source code for a working algorithm is a precise and reproducible way to explain what a technique actually does. And finally it is good that the books use R, and not something else, because R is the lingua franca of computational statistics.

Of course, the title "Something with R" does not guarantee that the "Something" and the "R" are integrated in some interesting and appropriate way. The book by Jurečková and Picek under review is a case in point. It is, to be sure, an excellent introduction to robust statistical methods, eminently suited for an upper division undergraduate course. It discusses the basic ideas of Huber, Hampel, Bickel and Hajek in an accessible and rigorous form. For a beautiful introduction to the theory of M, L, and R estimation, there is no need to look any further.

From the point of view of computational statistics, and of statistical software development, the book is disappointing. It has a seven-page appendix with an introduction to R. My feeling is that seven pages is not enough, even in a book of this form. Publishers such as Springer or Chapmann & Hall/CRC should prepare a standard 20 page introduction and distribute that, in some form, with all of their "Something with R" books. Or, even better, a CD.

Most chapters in the Jurečková and Picek book have a "Software Notes" section of about 5 pages which has some data analysis in R, some R code, a small number of R functions, and the smallest possible number of R packages. The main part of the book, over 90% of the pages, does not refer to computation or software at all. Thus it seems safe to conclude that there is very little integration between the theoretical and computational parts, and that the amount of R used and presented is minimal. There is no attempt to organize the software into a package, and there is no reference to the substantial number of existing R packages

for robust data analysis (except **MASS**, which is part of the standard R distribution). This is something of a missed opportunity, and one hopes that it will not be a model for a whole sequence of "Something with a Minimal Dash of R" books. Adding R to the title of a book is easy, but integrating the theoretical and computational aspects of statistical techniques with software is much harder.

Reviewer:

Jan de Leeuw
University of California at Los Angeles
Department of Statistics
Los Angeles, CA 90095-1554
E-mail: deleeuw@stat.ucla.edu
URL: http://gifi.stat.ucla.edu/