A Handbook of Statistical Analyses using R (second edition), Brian S. Everit and Torsten Hothorn, CRC Press, Taylor and Francis Group, Chapman & Hall, Boca Raton, ISBN 978-1420079333 (hardcover, $57.95), xx+355 pages, by Christian P. Robert, Université Paris-Dauphine, Institut Universitaire de France, and CREST, Paris.

Table of contents

Readership: Statistics students with some background in statistics and researchers and practitioners looking for an introduction to statistical modelling via R.

This book is the second edition of a successful handbook that can benefit a wide audience interested in using R for its data analysis. It covers most of non-Bayesian statistical methods, with forays into exploratory data analysis with tools like principal components, clustering and bagging/boosting. As reflected in the list of chapters given above, the coverage is quite extensive and only missing more specialised statistical domains like time-series (apart from longitudinal data), econometrics (except for generalised linear models), and signal processing. Beside the absence of a Bayesian perspective (only mentioned in connection with BIC and the mclust package, while it would be a natural tool for analysing mixed models), I miss some material on simulation, the only entry found in the book being bootstrap (pages 153-154).

Given its title and emphasis on analyses, the book is logically associated with an R package HSAUR2 and works according to a fixed pattern: each chapter starts with a description of a few datasets, summarises the statistical main issues in one or two pages, and then engages into an R analysis. As the complexity increases along chapters, the authors are relying more and more on specialised packages that need to be downloaded by the reader. I have no objection with this pedagogical choice, especially when considering that

the packages are mostly recent. I would however have like a bit more details about the packages or at least about their main function, as the reader is left to experiment from the line of code provided in the handbook. (In contrast, a few passages are a bit "geeky" and require a deeper understanding of R objects than casual readers master.) My only criticism of the book at this level is the puzzling insistance on including all the datasets used therein in the form of tables. I frankly fail to see the point in spending so many pages on those tables given that they all are available from the HSAUR2 package. A page of further explanation, background or statistical theory would have been much more beneficial to any reader, in my opinion! The same criticism applies to the few exercises found at the end of each chapter.

In conclusion, I find the book by Everitt and Hothorn quite pleasant and bound to fit its purpose. The layout and presentation is nice (with a single noticeable mishap on page 332 caused by Darwin's tree of life.) It should appeal to all readers as it countains a wealth of information about the use of R for statistical analysis. Included seasoned R users: When reading the first chapters, I found myself scribbling small light-bulbs in the margin to point out features of R I was not aware of. (In particular, the authors mentioned the option type="n" for plot that *R-bloggers* signaled as the most useful option for plotting.) In addition, the book is quite handy for a crash introduction to statistics for (well-enough motivated) non-statisticians.