Journal of the
Royal Statistical Society

SERIES A
Statistics
in Society

Book Reviews   697

work for likelihood inference and in particular for construction of profile likelihood confidence intervals for general smooth parameter functions in a very flexible class of regression-type models, including models where observations do not all follow the same distribution type.

The first two chapters give the required theoretical background. In Chapter 1, a compact and very general notation for inference is introduced. The meaning of common terms such as 'statistical evidence' and 'evidential meaning' is discussed. Likelihood concepts are introduced in a nice fashion, including an explicit reference to the accuracy of measurement for quantitative observations. Confidence intervals based on the profile likelihood function are compared with the traditional Wald regions. However, the illustrations showing the iteration process that is involved in the construction of these profile likelihood confidence intervals are close to incomprehensible and seem to have been written for teaching purposes only, to be accompanied by oral explanations which have unfortunately been left out here. Chapter 2 introduces the flexible regression model framework with a very formal definition of a 'generalized regression model' which can accommodate 'almost any' model for independent observations.

In Chapter 3, traditional theory on linear (normal) models is presented in brief. There is no mention of the geometry that is involved in such models; nor are there any comments on the relationship between F-regions and individual t-regions. Chapter 4 deals with non-linear models and looks like a rewriting of Chapter 3, only with the word 'approximative' added in the relevant places. Geometrical arguments are now used, although without graphical illustrations. No effort has been made to judge the adequacy of the approximations; neither in the theoretical results nor in the practical examples, where the results are barely commented on.

Chapter 5 gives traditional mathematical background for generalized linear models and, in Chapters 6–9, these are exemplified by almost identical chapters on the binomial, Poisson, multinomial, negative binomial and gamma distributions. Long explicit formulae for the moment- and cumulant-generating functions are given, apparently with the sole aim of establishing an approximation to the normal distribution. Numerous sloppy details (undefined symbols, uninformative matrices, undiscussed pictures etc.) indicate that the writing process has been speeded up.

Chapter 10 (and Appendix B) further broadens the class of models that are considered to include for example modelling of the variance function and lifetime regression models.

All the models that are described in this book

are implemented by the author in the statistical inference package collection of Mathematica macros (SIP). The accompanying disc contains this package, as well as the Mathematica platform that is necessary for running it.

The book suffers from numerous misprints and misleading errors. There are illustrations with no or hardly any explanations, strange references appearing from nowhere, notational misprints etc. In conclusion, the book cannot be recommended as a stand-alone textbook. Its best feature is the accompanying disc which provides a useful tool for analysing quite complex models in an informative manner, providing also theoretical formulae for likelihood functions, score functions etc., and therefore making it a useful self-learning tool for students of theoretical statistics. The book cannot be recommended for applied researchers who do not have a sound statistical education.

L. T. Skovgaard
University of Copenhagen
E-mail: lts@biostat.ku.dk

A Beginner's Guide to R

R is an open source language and environment for statistical computing and graphics that is becoming increasingly popular among statisticians and scientists. This popularity has been responsible for the publication of an increasing number of books on R in the last 5 years. At the time of writing this, more than 100 books are listed on the R archive and an impressive 20% of them were published in 2009 (http://www.r-project.org/doc/bib/R-publications.html). Most of these books deal with more or less advanced uses of statistics and few are targeted at beginners. *A Beginner's Guide to R* takes a different approach because it is the experience of its authors that

'teaching statistics and R at the same time means two steep learning curves, one for the statistical methodology and one for the R code'.

For this reason, to quote the authors again,

'this book provides basic instruction in R and does not deal with statistics'.

Chapter 1 teaches how to download and install R and its packages. It also introduces the reader to the 'art of programming' by showing the importance of

writing readable and reusable code and using the R help system right from the start. Chapter 2 goes on to explain how to enter and import data into R. The main methods of organizing data in R are reviewed and compared through application to some ecological data sets. Although most users will import data from external software packages—something which can be done following the indications in the last part of the chapter—it is very important to become familiar with the internal organization of data to exploit the full potential of R. Chapter 3 shows how to select and manage subsets of data. This is probably the most difficult aspect of R for a beginner, but the presentation is very clear and the suggestions of the authors on how *not* to do things (here and in other chapters) will prove invaluable to most readers. Chapter 4 is a short introduction to functions and teaches how to calculate some descriptive statistics for the whole data set or subsets of it. Chapter 5 is a simple but effective introduction to the basic plotting tools of R. Chapter 6 goes to the heart of every programming language: loops and functions. This chapter is a wonderful example of the art of programming in action. The importance of a piecemeal approach to coding is clearly demonstrated by developing R code to draw multiple graphs and personalized functions. Chapter 7 is an extension of Chapter 5 and shows how to draw pie charts and bar charts (with some reluctance owing to their well-known limitations) together with boxplots, dot plots, pair plots and conditioning plots. Chapter 8 is a further exploration of the graphic arsenal of R and introduces the reader to the `lattice` package and to its high level functions which are put to work by showing how to draw multipanel graphs. Lastly, Chapter 9 provides a discussion of common mistakes in R that will benefit both novice users and R instructors.

In my opinion, *A Beginner's Guide to R* is a well-structured and highly didactic book that fulfils its promise of giving a simple and effective introduction to R. This may be the first book to read when deciding to enter the immensely powerful but intimidating and sometimes messy world of R. Although this book was written mostly for beginners, instructors also may want to read it because it is built on extensive teaching experience.

Giorgio Bedogni
Liver Research Centre
Trieste
E-mail: giorgiobedogni@gmail.com