Statistical Analysis of Extreme Values

by R. D. Reiss and M. Thomas

David Walshaw
University of Newcastle upon Tyne

david.walshaw@ncl.ac.uk

This book is currently in its Second Edition, released March 2001, and comes with a customized software package on CD-ROM; Xtremes 3.0. The version included with the book is the Academic Edition, which does not allow for as many rows of data as the professional version, and does not provide the server, which exports estimators, data generation routines and plotting facilities for use by other packages. The professional version of the software must be purchased separately. The book is in paperback, and runs to 462 pages. It is currently available for £39.

The textbook itself constitutes a compendium of Extreme Value Analysis as an area of Applied Statistics. It is divided into five sections, dealing with: I Modelling; II Inference for Parametric Models; III Multivariate Analysis; IV Topics in Hydrology, Insurance and Finance; and V a section comprised of five case studies. Sections I to IV contain a total of 14 chapters. There is also a substantial Appendix which provides an introduction to Xtremes. The book is comprehensive in scope. In the Preface to the First Edition (also included in the Second Edition), the authors emphasise that practitioners in many fields of modern science, engineering or insurance may profitably employ the textbook and the software. I believe this to be true, and I feel that this is the book's strength. I am less convinced, however, of its value for teaching purposes.

A stated aim of the second edition is to reinforce a characteristic of the first edition, in providing a broad statistical background to enhance the material on extremes. However the authors state in the Preface to the First Edition that for large parts of the book it is assumed the reader has some knowledge of basic statistics, and that "yet more and more statistical prerequisites are needed in the course of reading this book." This remains true of the second edition. The approach taken is to integrate the exposition of relevant ideas on extreme value analysis with bits of general statistical theory and methodology which are deemed relevant at the time. However, as the material becomes more sophisticated, it becomes less practicable to fill in the likely gaps in statistical knowledge of a less experienced reader, hence the need for more background knowledge. Thus the diversions into general statistical issues are more prevalent near the beginning of the book. As an example, in Chapter 1, in addition to the expected ideas on the limiting distributions for maxima, on return levels, and so on, we find sections on the Poisson Approximation of Binomial Distributions; on Kurtosis and a Concept of Fat-Tailedness; and even a rather philosophical comparison of viewpoints on robust statistics. In my view, this has the effect of making the organisation of material rather incoherent. In particular, issues of modelling and theory become confused.

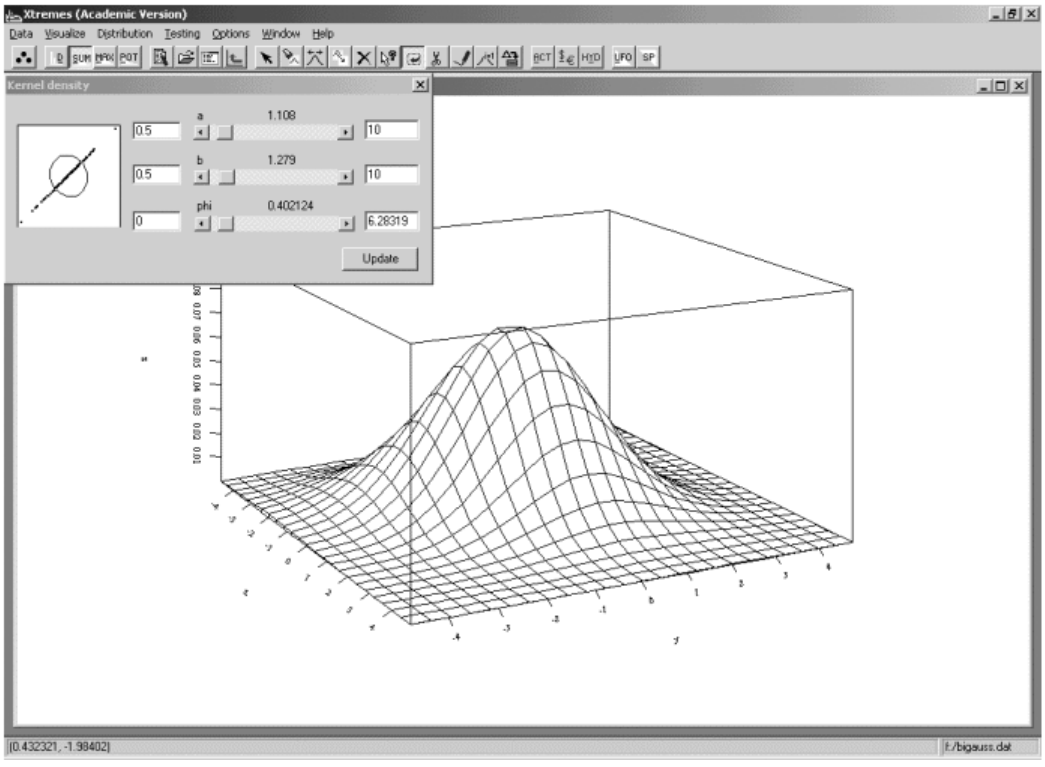
The strategy is more successful later in the book, when the necessary background material is often collected together to form an entire chapter. Section II, on Statistical Inference in Parametric Models (for extremes), begins with a chapter which gives a general introduction to parametric inference. Likewise Section III, concerning Multivariate Analysis, commences with a chapter on Basic Multivariate Concepts.

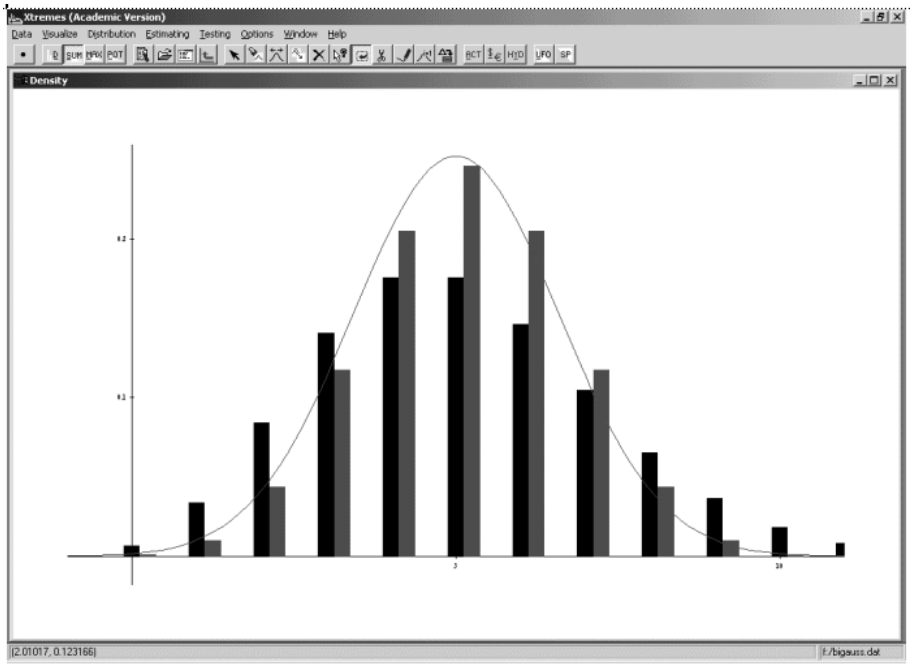
In this way the general issues are conveniently revised prior to switching attention to particular problems concerning extreme values. This is particularly useful when we bear in mind that each chapter concludes with a section dealing with the implementation of the material covered using the Xtremes package. We see how general problems of inference and exploratory analysis can be handled using Xtremes, prior to switching attention to the extreme values themselves.

In considering the value of this book for teaching, I certainly cannot imagine it as an appropriate resource for service courses, and my judgement is that the style and the organisation of material would make it a fairly daunting proposition for most undergraduate statistics students. In my own personal experience of supervising a substantial undergraduate project on extremes of financial time series, I referred the student elsewhere (Coles, 2001) to gain a general introduction and background to extreme-value theory, prior to directing them to the text under review for a more specialist handling of extremes in finance. Similarly, in planning a future course on the analysis of extremes, I would be tempted to use Coles as a basis for structuring the course, while using the text under discussion for drawing on specific examples concerning hydrology and finance. This illustrates both the strength and weakness of *Statistical Analysis of Extreme Values*. The comprehensive coverage makes it a useful resource in the analysis of extremes for experienced practitioners in a wide range of disciplines. It is not, however, designed as an introductory text for extreme value analysis, and for this reason it loses out as a front-line teaching resource.

Similar arguments apply in considering the potential of the book for postgraduate training in statistics. It is undoubtedly a useful source of material, containing many interesting and specialized examples. A postgraduate student with an applied interest in the analysis of extreme values is almost certain to find relevant examples here. However, they will probably not find the opening chapters to be the most accessible introduction to the field, and it is unlikely that a statistics postgraduate student would wish to rely too heavily on the Xtremes package before needing to work with a more general-purpose language. Having said this, it is easy to envisage that postgraduates in other disciplines, such as engineering or economics, may find the book and software indispensable. Faced with the need to carry out extreme value analyses of their data as part of their research, such a student may find that the book is almost unique in providing dedicated, easy-to-use software to fulfil their needs, which is well supported by the text.

Turning now to the software provided with the book, the Xtremes package is based on a menu-driven Graphical User Interface. In addition, the authors have integrated the StatPascal textual programming language into the package, which allows the user to extend the menu system themselves. A range of example data sets are provided with the package. The whole set-up is very easy to use, and gives the user access to a wide range of facilities, including many which are more general than the extremes-oriented functions which distinguish this software. There are separate modes for univariate and multivariate analysis, and the user switches between four domains: DIStribution; SUM; MAX and POT (for Peaks over Thresholds). While the MAX and POT domains handle the kind of functions usually associated with extreme value analysis, the SUM domain is devoted to sum-stable distributions, and the DIS domain deals with analysis involving Uniform, Poisson, Binomial and Negative Binomial Distributions. A full range of estimation and testing facilities are incorporated. There is also a UserFormula (UFO) facility, which serves as a pocket-calculator and curve-plotting





I must confess to being taken aback on first finding that some of the examples in the book involve the inputting of (supplied) data sets which are too large to be handled by the academic edition of Xtremes! Generally, however, the Xtremes package and the way in which it is integrated with the text are very positive features of the book. They make extreme value analysis very accessible to the user, and very easy to demonstrate. It is this aspect which would most commend the whole package, book and software, if it were used as a teaching resource. I would recommend it to teachers who are happy working with a dedicated package, or to those who wish to convey the *ideas* underlying extreme value analysis without the students getting their hands too dirty! This latter approach may be appropriate for students studying a more general topic, such as environmental statistics.

tool (with slider controlled parameters). This enables the execution of simple calculations without having to invoke the full power of the StatPascal language. New functions can be added to the UFO facility by the user, although the scope is necessarily more limited than that of the StatPascal language. The graphical capabilities are well integrated into the menu system, and very powerful. For example, the slider-driven 3-D visualization features are excellent, and great fun to use.

Returning to consider the scope of the package, it is very handy that Xtremes does have general non-extreme value oriented facilities built in, saving the user from the need to switch between packages. This ties in with the philosophy of the book, in giving the reader access to a general statistical background to underpin the material extreme values. Each chapter finishes with a section (usually fairly brief) in which examples, using data supplied with (or simulated by) the package, are used to illustrate the methods discussed. Given that some of the chapters do not concern themselves with extreme value analysis, we are necessarily given useful illustrations of the more general capabilities of the package. And of course, given that the user can pick and choose which aspects of the program to access, the non-extreme value aspects do not interrupt the flow of material, as they sometimes do in the text. Also tailor-designed for the book are three specialist menus: the actuarial menu, the finance menu and the hydrology menu. These correspond to the three areas of application considered specifically by the authors in Section IV, and contain some useful dedicated functions for users in these fields.

My personal preference for teaching a course specifically on *extremes* would be primarily to work with a standard language, such as S, to encourage a fuller understanding of the procedures involved in modelling. The students taking such a course will necessarily be reasonably advanced, and while I would certainly draw on *Statistical Analysis of Extreme Values* and Xtremes for motivation, I would prefer to frame most of the practical work within a general-purpose statistical language. I maintain that this book and the accompanying software would be of most use in a corporate or government research environment, where practitioners need to implement standard extreme value procedures efficiently, and convey their findings easily to non-specialists. I commend the authors on producing an extremely useful and comprehensive resource, but I think it is fair to say that the teaching of courses on extreme value analysis was not their primary aim in producing this book.

Reference

Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values*. Springer: London.