

## Book review

**Statistical Data Analysis for Ocean and Atmospheric Sciences** by H. Jean Thiébaux (Academic Press, 1994). ISBN 0 12 686925 1 and 0 12 686926 X (diskette). 248 pp., \$179.95.

While this book is a useful introduction to statistics, and particularly to statistical inference, it does not deliver the promised '... uncommon introduction to methods of statistical analysis'. It certainly does not fill the need for a balanced introductory statistics text, with a practical meteorological (or oceanic) flavour. On the contrary, large parts of the book, especially from Chapter 4 onwards, are practically indistinguishable from any standard textbook on mathematical statistics. In these later chapters the text has little specific meteorological or oceanic emphasis, with most of the data analysis using the supplied datasets left to the exercises.

Before this stage is reached, however, the first few chapters are extremely tedious, especially Chapter 2. The heavy, and exclusive, emphasis on the MINITAB statistical software package is a major drawback of the book. While this package is apparently widely available in universities, especially in the United States, not all readers will have access to it. Furthermore, with its cumbersome command line structure, many users would prefer some of the more modern, user-friendly PC-based packages now available. The author also has a somewhat optimistic attitude regarding the availability of data in computer-compatible form, and adequate computing power and software. This is certainly not true for climate research in large parts of the developing world.

The emphasis on MINITAB made most of Chapter 2 and large parts of Chapter 3 irrelevant. Chapter 2 is little more than an introduction to MINITAB. Chapter 3 introduces some basic descriptive statistics, however this introduction appears to be driven by the output of the MINITAB DESCRIBE procedure and is little more than a description of the output statistics. This same information is presented in most statistical packages manuals, indeed these two chapters could easily be part of such a manual.

Beginning with basic probability models in Chapter 4, the major part of the book then works its way through a fairly standard statistical devel-

opment. The chapter concludes with a statement and an example application of Bayes Theorem. The next three chapters cover probability distribution functions, specially the Binomial and Poisson distributions in Chapter 5 and the Uniform and Exponential in Chapter 6. The more general Gamma distribution is not covered. The Normal distribution is extensively described in Chapter 7, including the multivariate Normal distribution. At this point the level of mathematics increases with the introduction of matrix and vector algebra.

The next three long chapters, on analysis of variability, hypothesis testing and linear regression, form the core of the text. These chapters handle the material competently. Chapter 9 in particular is a succinct introduction to hypothesis testing. However, any mention of meteorology or oceanography has all but disappeared from the text! One hypothesis regarding different mean temperatures in two periods is proposed near the beginning of the chapter but not referred to again. In Chapter 10 two examples are given and in this case one followed to some extent throughout the chapter. This structure may work when the book is used in conjunction with a lecture course, however, trying to work through the exercises on one's own based on the practical guidance in the text could be rather daunting.

In these later chapters new concepts and terminology are introduced at a fairly rapid pace, and not always with adequate definition or explanation. Thus in Chapter 7 marginal probability distributions are introduced with Assertion 7.6 with no prior explanation. Similarly in Chapter 8 the concept of 'degrees of freedom', a crucial concept in hypothesis testing suddenly appears in a discussion of the chi-squared distribution. The meaning of this term, which is then used extensively in Chapter 9 and 10, is never adequately explained.

The final chapter on Bootstrapping is totally out of character with the preceding chapters, and indeed most of the rest of the book. Only here is the most crucial aspect of most meteorological and oceanographic data, its spatial and temporal autocorrelation, mentioned. As the author notes, these effects tend to make most of the discussion in the text invalid, or at least incomplete. While a proper treatment of these topics is beyond the scope of an introductory text, other elementary aspects of time series analysis could have been

included, especially some discussion on the seasonality found in most meteorological or oceanographic data. Many of the exercises in Chapters 7–10 analyse data for a single month or season only.

The supplied datasets, which are promoted as a major feature of the book, have a strong North American bias, and an unfortunate mixture of non-standard units, e.g., precipitation in inches and temperatures in Fahrenheit. These should have been converted to SI units as is required by virtually all meteorological and oceanographic journals.

Overall, it is difficult to recommend this book as it has little to offer beyond that available in any standard introductory mathematical statistics

textbook, but at a much greater price at around A\$180 for the hardcover edition. It is not clear just who the intended target readers are. The preface states it has been written as a guide to students, although it does not specify at what level. Certainly the level of mathematics, in particular the matrix algebra in the later chapters, would suggest at least second or third year undergraduates. However the price would put it well beyond the reach of this group.

**Wasył Drosdowsky**

*Wasył Drosdowsky is a meteorologist in the Climate Group of the Bureau of Meteorology Research Centre.*