Journal of Applied Statistics, Vol. 23, No. 6, 1996, 677–680

Book Reviews

Kernel Smoothing
M. P. WAND & M. C. JONES, 1995
London, Chapman and Hall
pp. xxi + 212, £25.00
ISBN 0 412 55270 1

Kernel smoothing and associated methods (such as local polynomial smoothing) are fast becoming the method of choice for handling smoothing problems in data analysis. This book serves an important need in clearly explaining the mathematical background to kernel smoothing methods. Both Matt Wand and Chris Jones have played important roles in the development, refinement and analysis of kernel smoothing methods and this book is a useful guide giving their perspective of this class of techniques.

In the preface, the authors state (p. xi) their goals "are to develop the reader's intuition and mathematical skills required for a comprehensive understanding of kernel smoothing, and hence smoothing problems in general". It is not a book about how to use kernel methods. Rather it focuses on how to analyze kernel methods, and introduces the necessary mathematical tools to do this.

The book concentrates on the two most common applications of kernel smoothing, non-parametric regression and density estimation. Each of these problems is motivated in Chapter 1. The density estimation problem is covered in the next three chapters. Chapter 2 covers a large amount of material on univariate kernel density estimation including error criteria, types of kernels, asymptotic behaviour of the estimator and a few modifications to the simple kernel estimator. Bandwidth selection of univariate density estimation is the topic of Chapter 3, starting with the simple 'rules-of-thumb' and going on to plug-in rules and various cross-validation schemes. The ideas of Chapters 2 and 3 are extended to multivariate density estimation in Chapter 4. The non-parametric regression problem is discussed in Chapter 5, focusing on local polynomial kernel estimators. Most of the chapter deals with the asymptotic behaviour of the estimator with some brief discussion of bandwidth selection and multivariate regression. The final chapter deals with kernel smoothing in various other settings including density estimation of dependent, length-biased, censored and measured-with-error data, hazard function estimation, spectral density estimation, likelihood-based regression models and intensity function estimation.

Each chapter concludes with excellent bibliographical notes providing some historical references and references to recent literature containing further details of topics covered in this book. The bibliography is extensive, reflecting the enormous research interest in kernel smoothing over the last few years. There are also a large number of exercises at the end of each chapter. These fill in a few derivations that are omitted in the text and apply the general theory to particular densities and kernels. In keeping with the focus of the book, the exercises are designed to hone mathematical skills needed in the analysis of kernel smoothing methods; they do not help in understanding the application of kernel smoothing methods.

There are four appendices. The first on notation is useful for keeping track of some of the notation and abbreviations used. The fourth appendix is a very useful outline on computational algorithms for kernel estimation. This appendix advises that S-PLUS functions for the major estimators described in the book are available from the authors. Since publication, these functions have been made available on statlib.

The authors concentrate heavily on the asymptotic properties of the methods. This reflects the preoccupation of the research literature with asymptotics, but does not necessarily provide much guidance in understanding or using kernel smoothing with small to moderate data sets. Sometimes the presentation is too brief, with many details being left to the reader and discussion being sacrificed for the sake of broader coverage. Although there are a few applications of the methods to real data, they are given as very brief examples of the application of the methods rather than examples of genuine data analysis. The book is not intended to provide much guidance on the application of the methods or to warn readers of potential data problems or difficulties of interpretation.

The authors state (p. xi): "We have aimed this book at newcomers to the field. These may include students and researchers from the statistical sciences and interface disciplines." The book is indeed an excellent reference for newcomers to statistical research in kernel smoothing. I have found it useful in this regard, as has one of my PhD students. However, I am glad it was not the first book we read on kernel smoothing as the cryptic coverage of some topics and the lack of applications would probably

0266-4763/96/060677-04 $6.00

have put us off. Silverman (1986) and Scott (1992) provide better introductions to kernel density estimation and Härdle (1990) provides a nice introduction to kernel regression. This book is better for readers with some prior understanding of kernel smoothing and who want to extend their understanding of its mathematical foundations or learn about some of the recent research directions in the field.

The book would be a good supplementary text for a graduate class on kernel smoothing; the exercises would be particularly helpful for such a class. However, the course would be incomplete without additional material on data analysis using kernel smoothing. The authors claim that the book would be "appropriate for most first or second year statistics graduate students in the North American system [and] honours level students in the Commonwealth system". It is too difficult for the honours class on smoothing methods which I teach, although I did find it useful when preparing material for the class.

The book is systematically organized and well indexed. One unfortunate feature of the book is the very narrow (and variable) outside and bottom margins. On some pages, the text and figures go to within a millimetre or two of the edge. The narrow margins also make the edge of the text noticeably non-parallel with the edge of the page. It mars what is otherwise a very useful book, albeit with a relatively small potential readership.

ROB J. HYNDMAN
Monash University

REFERENCES

HÄRDLE, W. (1990) *Applied Nonparametric Regression* (Cambridge, Cambridge University Press).

SCOTT, D. W. (1992) *Multivariate Density Estimation: Theory, Practice and Visualization* (New York, Wiley).

SILVERMAN, B. W. (1986) *Density Estimation for Statistics and Data Analysis* (London, Chapman and Hall:)

Achieving Quality Improvement: A Practical Guide

R. CAULCUTT, 1995
London, Chapman and Hall
pp. x + 233, £35.00
ISBN 0 412 55930 7

"This has been a very useful discussion, Roland. Very useful indeed. You should put all this wisdom into a book." Fortunately, Roland Caulcutt heeded this advice from Bill Johnson, one of his clients, and has produced one of the most readable books on statistics I have ever come across. The book covers process capability, control charts, CUSUM charts and factorial designs for quality improvement, but it is no ordinary textbook. It is written as a series of conversations between the author and his clients which take place on visits the consultant makes to a number of fictional industrial companies. As well as discussing the above techniques and the philosophy behind them, there is a considerable amount of discussion on the human side of consulting and on such issues and companies' training programmes.

This book can be, and perhaps can only be, read from cover to cover like a novel. It is based on the author's experiences as a consultant and the reader gets a fly-on-the-wall view of an expert at work. The conversations feel realistic, even if everything runs slightly more smoothly than in real life — for example, coffee always arrives at a convenient time. This is not to say that the book lacks substance. It is written "for those who need help to achieve quality improvement" and guides them through the various stages of data collection, analysis and interpretation.

Even applied statisticians familiar with all the techniques described will be informed by the discussions of psychology and the human side of quality improvement, and entertained by the author's humane and humorous approach to consulting. I found myself laughing aloud at the chapter on designed experiments. On the other hand, the consultant's style is not one that everyone could copy. For example, he postpones a discussion of factor interactions until after lunch: ' "The next steps in the analysis will blow your mind", I promise my client in a slow deep voice. "I think you should raise your blood sugar level in order to maximise your enjoyment of the experience." ' While wholeheartedly agreeing with the sentiments, I am not sure I could get away with talking to a client like that.

Everybody concerned with implementing quality improvement should read this book, as should everybody involved in teaching statistical methods for quality management. Statisticians working in other fields would also find it worthwhile reading as an overview of the quality area. The danger in the way the book is written is that many people will read the book and enjoy it, but then never look at it again and slowly forget all the good advice contained in it. The final chapter summarizes the contents of the book and is arranged by the techniques used, which is intended to make the book more useful as a reference work. At a quick glance, I did not find this chapter helpful, but it would be of more use to someone who was prepared to read the whole chapter.

Having received Bill Johnson's advice, the consultant reflects: "Perhaps I should write a book. But it might be impossible to produce a book that would satisfy the needs of people like Bill Johnson and the