



## Book reviews

### *Books for review*

If you would like to review a book, and thereby to retain it for your collection, please contact the Book Reviews Editor, whose details can be found by clicking on ‘book review list’ in the information on the Royal Statistical Society’s Web site:

<http://www.rss.org.uk/bookreviews>

### *Reviews in this issue*

Böhning, D., Kuhnert, R. and Rattanasiri, S. *Meta-analysis of Binary Data using Profile Likelihood*, p. 691

Clarke, B., Fokoué, E. and Zhang, H. H. *Principles and Theory for Data Mining and Machine Learning*, p. 691

Dietz, T. and Kalof, L. *Introduction to Social Statistics*, p. 692

Goupy, J. and Creighton, L. *Introduction to Design of Experiments with JMP Examples*, p. 692

Hastie, T., Tibshirani, R. and Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, p. 693

Hoff, P. D. *A First Course in Bayesian Statistical Methods*, p. 694

Horgan, J. M. *Probability with R: an Introduction with Computer Science Applications*, p. 695

National Research Council. *State and Local Government Statistics at a Crossroads*, p. 695

Panik, M. *Regression Modeling: Methods, Theory, and Computation with SAS*, p. 696

Uusipaikka, E. *Confidence Intervals in Generalized Regression Models*, p. 696

Zuur, A. F., Ieno, E. N. and Meesters, E. H. W. G. *A Beginner’s Guide to R*, p. 697

**Meta-analysis of Binary Data using Profile Likelihood**

D. BÖHNING, R. KUHNERT AND S. RATTANASIRI, 2008  
Boca Raton, Chapman and Hall–CRC  
190 pp., \$79.95  
ISBN 978-1-584-88630-3

In the usual meta-analysis of data from clinical experiments, the data on all individuals are generally available. When such information is derived from secondary sources, then the data on individuals may not be obtainable. Instead, the information is available in terms of statistical measures, e.g. odds ratios. The contents of this book assume the availability of information about the outcome of binary data in the form of a  $2 \times 2$  contingency table from a secondary source, e.g. published literature, and is not accessible on individuals. It presents the meta-analysis of data in such a situation which is termed ‘Meta-analysis with individually pooled data’ (MAIPD). Different statistical measures for MAIPD-type meta-analytic situations are presented and developed in the book.

The book is developed in 10 chapters followed by an appendix. Chapter 1 is introductory and presents several data-based examples to motivate the MAIPD-type meta-analysis and illustrates some useful concepts. The basic model and profile likelihood are introduced in Chapter 2. The application of the MAIPD approach in handling profile likelihood for homogeneity is illustrated in this chapter. The issues that are related to the handling of unobserved heterogeneity in meta-analysis are discussed in Chapter 3 and a non-parametric profile maximum likelihood estimator is characterized. Next, Chapter 4 considers the classical methods, i.e. weighted regression, logistic regression and profile likelihood methods for modelling covariate information. The approximate likelihood method and multilevel approach to estimate the treatment effect in an MAIPD are discussed in Chapter 5. Different approaches to model together the unobserved heterogeneity and observed heterogeneity in the form of covariate information are illustrated in Chapter 6. The authors have developed the freely downloadable software CAMAP (‘computer-assisted analysis of meta-analysis using the profile likelihood’) for the estimation of relative risk based on the profile likelihood models. Chapter 7 describes the uses and applications of CAMAP software. How to estimate the odds ratio by using profile likelihood under homogeneity and heterogeneity is the subject matter of Chapter 8. The issues that are related to the quantification of heterogeneity in an MAIPD are addressed in Chapter 9. Chapter 10 applies the methodology that is developed in this book to the surveillance of scrapie in

Europe. The objective and motivation behind this chapter is to illustrate that the utility and applications of the methods developed in this book are not restricted only to the meta-analysis of data on clinical trials but can be used in other situations also. This is illustrated with an example of a surveillance problem. The appendix is brief and gives details of the derivation of some of the results that are used in the book. The terminology that is used in the book is explained at the point wherever it is used. This helps a reader in smooth reading and understanding of the topic. Throughout the book, the emphasis is on the demonstration of topics and developed statistical tools via data-based examples.

The authors have succeeded in demonstrating recent developments and the utility of statistical tools for MAIPD-type meta-analysis. A reader needs to have a basic understanding of topics in meta-analysis but a strong background in mathematics is not needed. The material that is covered in this book can be a part of an advanced biostatistics course. The book should be accessible and useful to graduate students in biostatistics and biostatisticians working in theory as well as in applied areas. The book is well worth recommending for purchase by a library.

Shalabh  
Indian Institute of Technology Kanpur  
E-mail: shalab@iitk.ac.in

**Principles and Theory for Data Mining and Machine Learning**

B. CLARKE, E. FOKOUÉ AND H. H. ZHANG, 2009  
New York, Springer  
xvi + 782 pp., £58.99  
ISBN 978-0-387-98134-5

Given the advances in computational sciences and experimental high throughput technologies, high dimensional and unstructured data are generated in many fields. To deal with this situation, effective data mining (exploration, simplification, ...) and machine learning (analysis, inference, ...) methods are needed.

The present book covers all the important subjects regarding data mining and machine learning (DMML) methods:

- the context of DMML and all the difficulties that arise when working with high dimensional unstructured data, such as noise, sparsity and multicollinearity;
- the theory and application of early (line smoothers), classical (kernel, spline and nearest neighbour methods), new wave (additive mod-