Journal of Applied Statistics
2012, iFirst article

Taylor & Francis
Taylor & Francis Group

BOOK REVIEW

Statistical inference: The minimum distance approach, by A. Basu, H. Shioya and C. Park, Boca Raton, Chapman & Hall/CRC Press, 2011, xix+409 pp., £57.99 or US$89.95 (hardback), ISBN 978-1-4200-9965-2 **Q1**

Estimation in statistical inference is associated with the notion of closeness between sample data and a parametric model. The required estimator is obtained via the minimisation of a suitable measure of distance (divergence). In this book, the authors focus on density-based measures by providing a comprehensive treatment of distances of the chi-square type although some other distance procedures are briefly discussed. Minimum distance methods are investigated in relation to parametric inference and multivariate tests of fit.

After the first introductory chapter, Chapter 2 deals with a brief reference to the divergence between two distribution functions followed by a thorough coverage of divergence between two probability density functions. The chapter discusses density-based distances in discrete models with special attention given to the popular family of chi-square type distances. The authors rely on the *disparity generating function* defined early in the chapter. Special attention is given to the Hellinger distance which plays the central role in the book. In Chapter 3, the authors consider continuous models and construct estimators by introducing an appropriate kernel function. The chapter is divided into two main sections which provide the methodology associated with Beran's influential work and the model smoothing method known as the *Basu-Lindsay approach*.

Chapter 4 discusses in detail various robustness indicators and ends with the investigation of the stability of distances under contamination by constructing contamination envelopes. Chapter 5 is entirely devoted to hypothesis testing problems based on divergences. The first part of the chapter deals with the Hellinger distance which is later extended to a general class of disparities.

Chapter 6 deals with the main deficiency of the minimum distance methodology, namely its applicability in small sample cases. The authors present five inlier (observations with fewer data than predicted) control techniques for improving the small sample performance of the methodology. Penalized distances (the only method appropriate exclusively for discrete data) and combined distances are the techniques that are fully discussed while three more methods are briefly mentioned in the later sections of the chapter.

In Chapter 7, the weighted likelihood estimation approach based on disparities is considered. Chapter 8 covers multinomial goodness-of-fit (gof) tests and offers a useful summary of existing results in this major research area. In addition, the authors discuss special issues that are not found elsewhere like the gof test for kappa statistics and inlier modification.

Chapter 9 is devoted to the BHHJ power divergence class of divergences which does not belong to the class of disparities. The fact that this technique does not require the use of any non-parametric smoothing (for the tuning parameter involved) is a great advantage. An adaptation of the technique useful in survival analysis is also included.

ISSN 0266-4763 print/ISSN 1360-0532 online
http://www.tandfonline.com

Applications of the density-based minimum distance methods presented in Chapter 10 cover survival models, mixture models, group data and two-sample semi-parametric models. Further applications in the context of information science and engineering are included in the last two chapters with references to f-divergences, Neural Networks, Fuzzy Set Theory and Structure Analysis with the development of an iterative algorithm for phase retrieval.

The main contribution of the book is that it brings together the available methodology for minimum distance techniques with the emphasis placed on chi-square type distances within the class of density-based techniques. The level of the book, although quite high, can be followed by graduate students with a good background in mathematical statistics. The presentation is clear, precise and easily understood by both specialists and non-specialists. The (few) examples presented are well chosen and representative but the inclusion of more examples would have been highly appreciated by readers. The authors should consider including a list of exercises at the end of each chapter in a future edition, not only for better understanding of the material but also for making the book much more attractive as a textbook.

In conclusion, the book provides a comprehensive overview of the theory of density-based minimum distance methods and it is well written and easy to read and understand. The book is well suited for graduate students, professionals and researchers not only in statistics but also in biosciences, engineering and various other fields where statistical inference plays a fundamental role.

Alex Karagrigoriou

University of Cyprus, Cyprus

alex@ucy.ac.cy

http://dx.doi.org/10.1080/02664763.2012.681565