http://dx.doi.org/10.1111/anzs.12060

Clarke, B.R. (2014) Book review: Methodology in Robust and Nonparametric Statistics. By J. Jurečková, P.K. Sen, and J. Picek. Australian & New Zealand Journal of Statistics, 55 (4). pp. 497-499.

Methodology in Robust and Nonparametric Statistics. By J. Jurečková, P.K. Sen, and J. Picek. Boca Raton, Florida: CRC Press. 2012. 410 pages. £63.99 (hardback). ISBN 978-1-4398-4068-9.

Brenton R. Clarke

Murdoch University

It is a pleasure to have been invited to review the book 'Methodology in Robust and Nonparametric Statistics'. I briefly met both the authors Jurečková and Sen together in 1987 at Charles University in Prague, where they were working on a joint article. The current book is a testament to their collaboration having been recently published. There are over three pages of references to articles by Jurečková and colleagues and four pages to articles by Sen and colleagues. With this book the authors cement their lifetime collaboration with the aim of setting a foundation for other research workers to learn from and develop their work even further. They are joined by Jan Picek in this noble endeavour.

This book is a substantial revision of an earlier work by Jurečková and Sen (1996) published by Wiley. They justify it by noting the 'phenomenal growth' in the research literature in robust and nonparametric statistics. The new book includes a considerable section on convergence theory, which to my mind would form a basis for a course on probability theory at the postgraduate level. The results are said by the authors (JSP) to provide a pool of basic mathematical tools to be used in later chapters, though I feel that the cross referencing to the tools is not prescriptive.

So why are all these results needed, when other authors of robustness books skirt the boundaries of the theory of M-, L-, and R-estimation? We soon see the answer to this question when we consider the broad coverage of many forms of estimator. The authors study Pitman and Pitman Type Estimators (which trace their origin to Australia's own Pitman) as well as briefly covering minimum distance estimators. The study of robustness was in part initiated by the publication of a paper by Huber in the 1964 issue of the Annals of Mathematical Statistics. This was complemented by Hampel's introduction of the influence function which plays a prominent role in the expansions of statistical estimators represented as statistical functionals. JSP relate all of the above using a functional approach which makes use of the Hadamard or compact derivative. The use of these weaker derivatives is necessary in order to encompass wide classes of estimator. The Fréchet derivative of the functionals corresponding to a number of these estimators does not always exist.

JSP include estimators such as the Huber skipped mean and Huber skipped median and L-estimators that have a weighting J-function that extends outside a confined region $[\alpha, 1-\alpha]$, $0 < \alpha < 1$. They also include R-estimators such as the normal scores estimator (which has an unbounded score function) and the log rank scores estimator. Subsequently there is a need for JSP to consider a variety of approaches to establishing asymptotic criteria. The authors emphasize that there are special considerations when carrying out regression as opposed to location estimation. JSP also consider studentised M-estimators of location, one-step M-estimators, Bayes Type estimators of a general parameter, and general M-estimation. In respect of their treatment of general M-estimation in Chapter 5, the authors say that they avoid the use of the compact derivative as it does not exist for maximum likelihood estimators with unbounded score functions. This is incorrect, as can be verified by studying the paper of Heesterman and Gill (1992). The alternative arguments of JSP which include a section on Hadamard approaches for M-estimators in linear models are still valid nevertheless. On a somewhat related note there is an omission of a reference to the paper by Huber (1967) in regard to asymptotic normality of M-estimators.

Because their net is cast so wide in order to capture essential expansions of both first or second order, the authors find a need to treat estimators within a class differently, subject to appropriate conditions on the score function. They also expand on optimality and breakdown points throughout the text. M-estimators of location and regression are treated in Chapter 3. The idea of using M- and L- estimators simultaneously is noted and some effort is devoted to elucidating the concept of quantile in linear models. There are some nice approaches to establishing asymptotic results, for instance making use of Bahadur representations in the discussion of L-estimators – see Chapter 4. R-estimates are based on ranks rather than observations which ‘guarantees’ that they are less sensitive to gross errors as well as to the effects of most heavy tailed distributions. JSP state that their main aim is to go to second-order representations and relax the conditions of other currently used methods. This they do in Chapter 6 while Chapter 7 highlights interrelations of estimators.

Chapter 8 relates to multivariate data. The authors explain the importance of affine equivariance, which plays a pivotal role in the discussion of multivariate parametric estimation theory in 'the domain of elliptically symmetric distributions'. JSP look at norms other than the Mahalanobis distance, which then leads to a more elaborate discussion of multivariate theory. They examine the diverse pattern of multivariate symmetry and also note that there is no universal optimality criterion in the multivariate data model; even the notion of robustness is imprecise. They explore multivariate estimation noting that without the assumption of elliptic symmetry the Mahalanobis distance cannot be prevailed upon to gain affine equivariance and its dual, affine invariance. JSP examine diagonal symmetry, spherical symmetry, marginal symmetry and symmetry interchangeability and how they relate to different subclasses of symmetric multivariate distributions where affine equivariance/invariance may not be an appropriate assumption. See Oja (2010) and Serfling (2010) in regard to the notion of multivariate symmetry.

Robust testing is frequently an afterthought to the usual considerations of robust estimation, but is no less important; JSP say that by no means does their final chapter constitute a complete treatise but they do nevertheless illustrate the intricacies of testing. The theory of Choquet capacities does not fully cover the problems of testing a composite null hypothesis against a composite alternative. JSP outline options: use robust estimation which is defined implicitly as a solution of a system of equations, and use this system as a test criterion, or use the Wald-type test based on robust estimators. Robust tests defined this way involve confidence set estimation. Robustness considerations are then even more important for confidence sets and intervals. JSP consider M- and R-tests for location and for use in the case of the linear model, respectively. JSP introduce two classes of confidence interval Type I and Type II and describe confidence intervals based on ranks and extensions based on M-statistics simply as an illustration of what can be obtained. Chapter 9.6 considers Affine-Equivariant Tests and Confidence Sets.

In summary, this book is very detailed and offers many ingenious ways to set up expansions for robust estimators leading to asymptotic properties of statistics. In view of the broadness of the study undertaken over a number of years, there is something for everyone. I dare say that few readers will find it easy to read the whole book without having had some experience in the areas of robust estimation and asymptotic theory. To help the reader assimilate the ideas there are ample problems at the end of each chapter.

References

Heesterman, C.C. & Gill, R.D. (1992). A central limit theorem for M-estimators by the von Mises method. *Stat. Neerl.* **46**, 165–177.

Huber, P.J. (1967). The behaviour of maximum likelihood estimates under nonstandard conditions. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1**, pp. 73–101.

Jurečková, J. & Sen, P.K. (1996). *Robust Statistical Procedures: Asymptotics and Interrelations*. New York: Wiley.

Oja, H. (2010). *Multivariate Nonparametric Methods with R. An approach based on Spatial Signs and Ranks. Lecture Notes in Statistics*. **199**, New York: Springer.

Serfling, R.J. (2010). Equivariance and invariance properties of multivariate quantile and related functions, and the role of standardization. *J. Nonpar. Statist.* **22**, 915–936.