Sankhyā: The Indian Journal of Statistics
2000, Volume 62, Series A, Pt. 1, pp. 157–160

Efficient and Adaptive Estimation for Semiparametric Models
P.J. Bickel, C.A.J. Klaassen, Y. Ritov and J.A. Wellner
Springer Verlag

This book is a reprint of the book that appeared with Johns Hopkins University Press in 1993. Springer Verlag does the statistical community a great favour by making this paperback version available, for a much lower price, which brings the book within easy reach of all mathematical statisticians working on the theory of semiparametric models. From the authors I have understood that the book is unchanged apart from the correction of minor errors.

In their preface the authors recall how the project of writing this book started from lectures given in 1983 by the first and last author, but that it took another seven years for the four authors together to complete the manuscript. The reason for this is simply that only a small part of semiparametric theory existed in the early eighties. In fact the papers in the Annals of Statistics by Bickel in 1982 and Begun, Hall, Huang and Wellner in 1983 are often viewed as the beginning of semiparametric theory. The book grew as the theory developed and many parts of the book were the subject of research papers by the four authors.

This review is written seven years after the publication of the original book. Thus it should not be surprising that in the mean time semiparametric theory has evolved further again. The reader who wishes to be completely up-to-date should complement the reading of the present volume with additional material. At this time there is no single volume that could serve this purpose. A lot is still buried in journal papers. If I may, I would recommend the last chapter of the book by van der Vaart, 1998 for an overview of a part of the likelihood theory and applications of empirical processes to semiparametric models, developed since the publication of the book under review. Another area that is not covered is the statisticial inference for parameters that are not in the realm of "asymptotically normal theory", sometimes referred to as "inverse problems". There are many interesting inverse problems connected to semiparametric models. They are not covered here both by choice of the authors and because the theory on inverse problems is still in its first stages.

Anyway, the book under discussion contains a wealth of material, much of it in final form, and will remain an important reference for this area. The only competing book devoted to semiparametrics is the early work by Pfanzagl (Pfanzagl and Wefelmeyer, 1982). In my opinion Pfanzagl's exposition of tangent spaces and information bounds, a very important subject of the present work, is elegant and still deserves to be read. However, he does not develop a calculus of scores as in the book under review, has little to say concerning the construction of estimators, and contains a lot less detail.

One obtains a good overview of the wealth of detail of the book under review by looking at its front matter. Besides the usual pages of contents there is a six page list of examples, which are throughout the book. I think many of these examples were or could have been subject of full research papers. Thus one should not be surprised that an imaginary reader who would read the book from front to cover will find himself frustrated by his necessarily low speed.

The book consists of seven chapters and a long appendix. The appendix contains expository material on functional analysis, in particular Hilbert spaces, as well as statistically oriented material such as stochastic convergence theory. I would say the last part is superceded by the first chapter of van der Vaart and Wellner (1996). The part on functional analysis is useful as a summary and also because it contains some material that is not easily found in introductory books on functional analysis. In principle, all background that is needed can be found here, but I doubt that someone without prior exposure to this subject would find the book pleasant reading.

This is because functional analysis plays a major role in semiparametric theory, albeit that a good familiarity with the basic concepts suffices for most purposes. The center of the functional analytic influx is simply the score function. A score function in a smooth parametric model has a finite second moment, the Fisher information, and hence belongs to the Hilbert space of square-integrable variables. A semiparametric model could be considered as a union of infinitely many parametric models and hence yields infinitely many score functions, the "tangent space" of the model. This name was introduced by Pfanzagl in Pfanzagl and Wefelmeyer, 1982 and stems from a functional analytic view of a statistical model, in which a root of a density belongs to a Hilbert space and hence a model becomes a manifold in a Hilbert space, with a geometric tangent space if it is smooth.

Score functions are of extreme importance in statistics, because they define the Fisher information and hence how good estimators can be. The book under review starts in Chapter 2 by discussing this for ordinary, finite-dimensional models, putting the optimality in the asymptotic framework of the Hájek convolution theorem. Readers not familiar with Hájek's result may find this a pleasant introduction. Next, this chapter prepares for semiparametric models by considering a component of a vector of parameters as the parameter of interest and the rest as a nuisance parameter. The optimality bounds turn out to result from orthogonal projections of scores and hence the importance of inner products and Hilbert spaces.

Chapters 3, 5 and 6 develop information bounds, and generalizations of Hájek's convolution theorem, for semiparametric models. An information bound roughly determines the smallest possible variance of an estimator. The authors argue convincingly that even if one is not interested in efficient estimators (having the smallest possible variance) the information bounds are still a good yard stick to measure the performance or malperformance of any estimator. An alternative to the convolution theorem, the local asymptotic minimax theorem, is not covered.

Further functional analysis comes in because in semiparametric models scores often take the form of an operator acting on some class of functions. Projections can be calculated in terms of these “score operators”. The book contains a wealth of examples of such calculations, which are often impressive.

Chapter 4 explains what a semiparametric model is and gives many examples, under the general categories of group models, regression models, biased sampling models, mixture models, missing data models and transformation models. It is clear here that the models known before the eighties (mainly the Cox model and the symmetric location model) are only drops within the ocean of semiparametric possibilities. And each class of models leads to its own calculus of scores, which may range from straightforward algebra to the solution of differential and integral equations.

In summary: Chapters 2 to 6 concern models and the properties of their tangent spaces. Chapter 7 concerns the discussion of methods of estimation. Generalized minimum contrast estimation is presented as a unifying method. This includes variations extending ordinary contrasts to smoothed, sieved or penalized contrasts. The book covers the usual theory on M-estimation, as for instance given in Pollard (1984, 1985). When applied to semiparametric problems the theorems in this chapter are generally master theorems, giving results at an abstract level, leaving much to do for particular examples. The point of view is that also methods of estimation need to be tuned to particular examples. This is regrettable, although perhaps unavoidable. Whereas the lower bound theory developed in the earlier chapters has a great deal of consistency and is mathematically satisfying, the theory on estimation is closer to a collection of intriguing ideas and clever arguments. Personally I hope that in this domain semiparametric theory will develop further in the coming years, as it already has in the past decade. We are clearly not at a point, where simple theorems on methods such as maximum likelihood can be stated under general conditions. In that sense there is a discrepancy between the lower bounds in Chapters 2-6 and the results on estimation in Chapter 7. This is not a criticism of the book. Rather it is a statement on the present state of affairs and an invitation to contribute to this interesting new area in statistics.

To conclude: this is a very impressive book, very densely packed with material on semiparametric models. I believe the book is too full to be read from front to back, but by selective reading one can introduce oneself to the general theory of semiparametric information. The last chapter contains many interesting ideas about estimation and solves a great number of problems. It is a book (the book at the present time) to consult before engaging in semiparametric research.

Bickel, P.J. (1982). On adaptive estimation. *Annals of Statistics* 10, 647–

671.

BEGUN, J.M., HALL, W.J., HUANG, W.-M., WELLNER, J.A (1983). Information and asymptotic efficiency in parametric–nonparametric models. *Annals of Statistics* 11, 432–452.

PFANZAGL, J., WEFELMEYER, W. (1982). *Contributions to a General Asymptotic Statistical Theory*. Springer Verlag, New York. Lecture Notes in Statistics 13.

POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer-Verlag, New York.

−−−− (1985). New ways to prove central limit theorems. *Econometric Theory* 1, 295–314.

VAN DER VAART, A.W. (1998). *Asymptotic Statistics*. Cambridge University Press.

VAN DER VAART, A.W., WELLNER, J.A., (1996). *Weak Convergence and Empirical Processes*. Springer, New York.

Vrije Universiteit *A.W. van der Vaart*