meaningful to perform an asymptotic analysis at places where there are hardly any data. It may seem a bit ironical that Chu and Marron make the same assumption "bounded from below" in the same paper (assumption A.4 of Section 3). In an interesting paper, Fan (1990) concludes independently about the Nadaraya-Watson estimator (remark 2, Section 3) "... hence its asymptotic minimax efficiency is arbitrary small."

CONCLUSIONS

Our conclusion is that the convolution weights are clearly superior to evaluation weights for fixed design, since we have the same variance for both methods but a nasty bias for evaluation weights. For random design, the problem seems to us more open: There is a minimax argument, and we would like to repeat a general argument, which is not well quoted by Chu and Marron (Section 3): "The latter authors [Gasser and colleagues] in particular seem to feel that variability is not a major issue, apparently basing their feelings on the premise that it is always easy to gather simply more data." What we said when discussing the structural bias of the evaluation weights was the following (Gasser and Engel, 1990): "These bias problems are particularly accentuated in the scientific process of many empirical sciences: studies are usually replicated by sticking to the design of the previously published study. In this way, qualitatively misleading phenomena as obtained by the Nadaraya-Watson estimator will be attributed even more confidence."

OUTLOOK

One way out of this problem has been opened by Fan (1990), who showed that for random design local polynomials have the same bias as convolution weights and the same variance as evaluation weights (the equivalence of local polynomials to convolution type kernel estimators for fixed design had been shown by Müller, 1987). A further possibility for improving the variance properties of convolution weights has been described by Chu and Marron in Section 6.

Comment

Birgit Grund and Wolfgang Härdle

1. OBJECTIVES OF SMOOTHING

Smoothing has become a standard data analytic tool. A good indicator of this is the increased offer of smoothing procedures in a variety of standard statistical software packages. It is therefore high time to provide background information that enables statisticians and users to critically evaluate the—in the meantime—rich basket of smoothing tools. The paper by Chu and Marron meets this demand for information and compares two different kernel regression estimators on an easy, understandable level. The authors combine successfully careful mathematical discussion with heuristic arguments in a well-done exposition. Cleverly chosen striking examples provide an easy access to not immediately apparent problems in smoothing for data analysis. We congratulate the authors to this valuable contribution.

Among the many objectives of smoothing, there are certainly the two perhaps most discussed. These are P1: to find structure; and P2: to construct estimators from a probability distribution.

We agree that the interplay of these two objectives is vital for an honest parameter-free data

Birgit Grund is Assistant Professor, School of Statistics, University of Minnesota, 352 Classroom-Office Building, 1994 Bufford Avenue, St. Paul, Minnesota 55108, and CORE Research Fellow, Université Catholique de Louvain, Belgium. Wolfgang Härdle is Professor of Statistics, CORE, Université Catholique de Louvain, 34 Voie du Roman Pays, B01348 Louvain-la-Neuve, Belgium. He is currently visiting CentER, Tilburg University, The Netherlands.