

- [40] YAKOWITZ, S. and FISHER, L. (1975). Experiments and developments on the method of potential functions. *Proc. of Computer Science and Statistics: 8th Annual Symposium on the Interface* 419-423. Health Science Computer Facility, UCLA.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024

DISCUSSION

PETER J. BICKEL

University of California at Berkeley

As Professor Stone has pointed out, over the years a large variety of methods have been proposed for the estimation of various features of the conditional distributions of Y given X on the basis of a sample $(X_1, Y_1), \dots, (X_n, Y_n)$. The asymptotic consistency of these methods has always been subject to a load of regularity conditions. In this elegant paper, Professor Stone has given a unified treatment of consistency under what seem to be natural necessary as well as sufficient conditions.

His work really reveals the essentials of the problem. He has been able to do this by defining the notion of consistency properly from a mathematical point of view in terms of L_r convergence. However, the notions of convergence that would seem most interesting practically are pointwise notions. An example is uniform convergence on (x, y) compacts of the conditional density of Y given $X = x$. The study of this convergence necessarily involves more regularity conditions. At the very least there must be a natural, unique choice of the conditional density. However, such a study and its successors, studies of speed of asymptotic convergence, asymptotic normality of the estimates of the density at a point, asymptotic behavior of the maximum deviation of the estimated density from its limit (see [1] for the marginal case), etc., would seem necessary to me and to Professor Stone too! (He informed me, when I raised this question at a lecture he recently gave in Berkeley, that a student of his had started work on such questions.)

One important question that could be approached by such a study is, how much is lost by using a nonparametric method over an efficient parametric one? If density estimation is a guide, the efficiency would be 0 at the parametric model for any of the nonparametric methods surveyed by Professor Stone. However, even if this is the case, it seems clear that one can construct methods which are asymptotically efficient under any given parametric model and are generally consistent in Stone's sense. This could be done by forming a convex combination of the best parametric and a nonparametric estimate, with weights