the multiway layouts with several observations per cell as a composite of two-sample shift models, Lehmann developed new estimates of contrasts and thus new tests. Interpretation of tests remains the same as in least squares; only the estimates have been changed to protect the experiments.

The second approach attacks the issue more directly. Namely, replace the L_2 norm by the weighted L_1 norm, and proceed immediately to estimation by minimizing the new distance measure to the model subspaces and to testing by comparing these new distances. The inferential strategy remains the same but the norm (and hence metric) have been changed. The value of the second approach lies in the breadth of application. Models ranging from simple one-sample location through the linear model with AOV designs, regression designs and analysis of covariance designs are handled in a unified way.

The implementation of this second approach requires a fully developed asymptotic distribution theory for the estimates and tests, estimation methods for a ubiquitous scaling parameter ($\theta = \int f^2(x)\, dx$, where $f(x)$ is the density of the error distribution) and the development of algorithms to carry out the required computations.

Most of what is known about the estimation of θ has been mentioned by Draper. We would like to close this discussion with some additional comments and references on the asymptotics and on computation.

In her seminal paper, Jurečková (1971) made rather complicated assumptions about the design matrix in order to develop the asymptotic theory for her regression R-estimates. Unfortunately, in practical problems, there is no way to check whether these assumptions are reasonable. Subsequent authors who built on this work adopted the same assumptions. However, as Heiler and Willers (1979) show, the only necessary assumption on the design matrix is the same as for the asymptotic theory of least squares procedures: Huber's assumption that the diagonal elements of the least squares projection matrix (the leverages) tend to zero as n tends to infinity.

Published work on rank-based methods for linear models typically suggests doing the computations via Newton's method (using the Hessian of the quadratic approximation developed in the asymptotic theory). Osborne (1985) has derived a rather different approach which takes full advantage of the fact that the dispersion is a convex polyhedral function. This approach should be seriously considered by anyone implementing these methods on the computer.

Derivation of confidence and multiple comparison procedures through replacement of the normal theory parameter estimates and estimated error variance by their robust analogues is connected with the Wald test statistic: a quadratic form in the full model estimate of the parameter vector. To develop confidence procedures which would be tied to the drop-in-dispersion test statistic, one would have to find, for example, all values of the parameter vector which could not be rejected by the test. This presents a rather difficult computational problem which, we believe, has not been attacked as yet.

In closing, we would like to reiterate the fact that both approaches described by Draper have been implemented. Fortran routines are available from Draper for the L_1ehmann methods and from J. W. McKean at Western Michigan for the Jaeckel and Hettmansperger-McKean methods, while an experimental rank regression command will be available in Release 6 of Minitab for many computer systems. It is hoped that people will subject these methods to the ultimate test: real data.

ADDITIONAL REFERENCES

HEILER, S. and WILLERS, R. (1979). Asymptotic normality of R-estimates in the linear model. Forschungsbericht N.R. 79/6, Univ. Dortmund.

OSBORNE, M. R. (1985). *Finite Algorithms in Optimization and Data Analysis.* Wiley, New York.

Comment

Peter J. Bickel

I want to use the opportunity of discussing this excellent exposition of rank-based methods in the linear model in part to revive a suggestion I made in a review paper on robust methods generally (Bickel, 1976). This approach may have some computational advantages over the Jaeckel-McKean-Hettmansperger (I would add Jurečková-Koul to the list) approach and relates the methods more closely to classical analysis of variance. The idea is to first fit the full

Peter J. Bickel is Professor of Statistics, University of California, Berkeley, California 94720.