Robustness Properties of Johns' Estimators

Peter McBurney
Mathematics Department
University of Zimbabwe
Harare
Zimbabwe
Current email: `p.j.mcburney@csc.liv.ac.uk`

December 1981

Paper presented at the Inaugural Symposium of the Southern African Mathematical Sciences Association (SAMSA), in Gaberone, Botswana, 14–18 December 1981.

Abstract

M. V. Johns [6] proposed two estimators for the symmetric location case, which, following Andrews *et al.* [1], we call JOH and JOE. This paper reports on an empirical, computer-aided evaluation of the small-sample robustness properties of these estimators.

1 Description of the Estimators

We consider estimators of a location parameter θ of a symmetric distribution $F(\ .\ |\theta)$, from which a sample of size n is drawn. We assume for conciseness that n is even. Before observation of the sample, let $T_{0,n}, T_{1,n}, \ldots, T_{k,n}$ be a partition of the set of integers $\{1, \ldots, 0.5n\}$ such that each $T_{i,n}$ consists of $t_{i,n}$ successive integers. Here $k \in \{1, \ldots, 0.5n\}$ and $\sum_{i=0}^{k} t_{i,n} \ = \ 0.5n$.

Now observe the sample and let $X_{(i)}$ be the order statistics. For $i = 1(1)k$, let

$$S_{i,n} \ = \ \sum_{j \in T_{i,n}} (X_{(j)} + X_{(n-j+1)})$$

For $i = 1(1)k$, let $c_{i,n}$ be a sequence of constants such that

$$\sum_{j=1}^{k} c_{j,n} t_{j,n} \ = \ 0.5n.$$

Consider estimators of the form:

$$\theta_{k,n} \ = \ \frac{1}{n} \sum_{j=1}^{k} c_{j,n} S_{j,n},$$

where $c_{i,n}$ may be functions of the sample values. Let $c' = (c_{1,n}, \ldots, c_{k,n})$. Johns, in introducing these estimators [6], shows that under suitable conditions, as $n \to \infty$, there is a convergence of distributions:

$$\sqrt{n}(\theta_{k,n} - \theta) \rightarrow N(0, \sigma_k^2)$$

where $\sigma_k^2 = 2c'Ac$, and A is a k-by-k matrix whose elements are functions of the quantiles ξ_i of $F(\, . \, |\theta)$.

An estimator, JOE, is obtained by replacing each ξ_i by its sample estimate $\hat{\xi}_i$ and obtaining an estimate for A. By using Lagrange multiplier methods, it is then possible to choose c which minimizes σ_k^2. The estimate $\hat{\theta}_{k,n}$ which results from this choice of c is called JOE. The other estimator, called JOH in [1], results from a simplification to the above, to avoid the matrix inversion neccessary to minimize σ_k^2. Complete details are provided by Johns [6], who also shows that JOH is asymptotically Normally distributed. For sufficiently large k, the asymptotic variance of JOH may be made arbitrarily close to that of JOE.

Although not immediately evident from the brief description above, both JOE and JOH are special cases of trimmed means. That is to say, both ignore data values which either fall below a certain point or above another point. For this reason, they are both less susceptible to the influence of "outlying" observations, or aberrant data values, than is the ordinary mean. It is for this reason that both are said to be robust estimators.

2 Robustness Properties of Johns' Estimators

Various definitions, and hence measures, of robustness estimators have been given. The measures used here were chosen for their ability to assess the presence of commonly-accepted definitions of robustness. Further information, and introductions to the concepts of robustness in general, can be obtained from the review articles [2, 3, 4, 5] and the book [1].

The following reports on the Monte Carlo study of the robustness properties of the estimators, each taken with parameter $k = 2$. This value was chosen because of Johns' concentration on it, despite the fact that he introduces JOH to avoid difficult k-by-k matrix inversion.

(a) The values and sample variances of the two estimators were compared across a class of four continuous distributions — the Gumbel, Normal, Double Exponential and Cauchy — with tails ranging from light to heavy. Various sample sizes (from $n = 6$ to $n = 42$) and various trimming properties were considered.

- The values and sample variances of the two estimators in each case differed only in the second or third decimal place, when estimating a value of $\theta = 0$.
- The estimated asymptotic variances $\hat{\sigma}_k^2$ of JOE were almost invariably smaller than those for JOE. This reflects the choice of $k = 3$.

(b) The Sensitivity Curves (SC) for the two estimators were calculated at the $N(0,1)$ distribution, with sample sizes between $n = 18$ and $n = 32$. The SC (or Empirical Influence Curve) is a measure of how well the estimator reacots to *one* aberrant data value, when a sample is drawn from a standard Normal distibution. The SC graphs obtained showed that — as expected — both JOH and JOE behaved as trimmed means, and are hardly distinguishable by their SC graphs.

(c) The Breakdown Bounds (again for $k = 2$ and $F = N(0,1)$) were calculated for various trimming proportions and sample sizes ($n = 6$ to $n = 42$). The

Breakdown Bound is a measure of how much aberrant data the estimator can withstand, being the maximum proportion of gross outliers in the data for which the estimator remains within certain predetermined bounds of the true value of θ. The results showed that JOH is much better behaved than JOE. In some cases, the latter estimator "collapsed" with only 8.3% aberrant data, while JOH was able to tolerate 25% aberrant data in the same situation.

3 Conclusion

From the Monte-Carlo comparison, it can be seen that the two estimators JOE and JOH do not behave radically differently, except in the proportion of contaminated data each can tolerate. Thus, the small sample behaviour of the estimators mirrors their asymptotic equivalence. Taking breakdown bounds into account, we are forced to conclude that JOH is the more robust of the two. Interestingly, Johns claims in [6] to have introduced JOH in order to avoid k-by-k matrix inversion, but then concentrates most of his attentions on the case $k = 2$.

Acknowledgments

The work reported here was undertaken on the Australian National University's Univac 1100/42 computer installation in Canberra, Australia, and formed part of the author's Honours degee thesis in Mathematical Statistics [7]. The author is grateful to his supervisor, Dr T. J. O'Neill, for his help and encouragement on this project.

References

[1] D. F. Andrews, P. J. Bickel, F. R. Hampel, P. J. Huber, W. H. Rogers, and J. W. Tukey. *Robust Estimates of Location*. Princeton University Press, Princeton, NJ, USA, 1972.

[2] P. J. Bickel. Another look at robustness: a review of reviews and some new developments. *Scandinavian Journal of Statistics*, 3:145–168, 1976.

[3] F. R. Hampel. Modern trends in the theory of robustness. *Math. Operationsforsch. Statist., Series Statist.*, 9:425–442, 1978.

[4] R. V. Hogg. Adaptive robust procedures: a partial review and suggestions for future applications and theory. *Journal of the American Statistical Association*, 69:909–927, 1974.

[5] P. J. Huber. Robust statistics: a review. *Annals of Mathematical Statistics*, 43:1041–1067, 1972.

[6] M. V. Johns. Nonparametric estimation of location. *Journal of the American Statistical Association*, 69:453–460, 1974.

[7] P. J. McBurney. Adaptive Estimators of Location. Honours thesis, Department of Statistics, Australian National University, Canberra, ACT, Australia, 1979.