J. Ind. Soc. Agril. Statist.
61(3), 2007 : 341-343

A Note on Inadmissibility of the Iterative Stein-rule Estimator of the Disturbance Variance

M. Dube and S. Chandra[1]
M.D. University, Rohtak
(Received: November 2004)

SUMMARY

The article investigates the effect on the disturbance variance estimation in linear regression model when Stein-rule instead of least squares estimation is used. Using small sigma asymptotics, it is demonstrated that the iterative Stein-rule estimator is not only asymptotically biased but is also dominated by its counterpart stemming from ordinary least squares.

Key words: Linear regression model, Ordinary least squares (OLS), Stein-rule (SR), Disturbance variance estimation, small sigma asymptotics.

1. INTRODUCTION

Following Stein's elegant proof of the inadmissibility of the usual estimator of variance, Maatta and Casella (1990) and Wan (1999) reviewed and examined the developments in variance estimation under decision theoretic set up. Earlier, in the context of linear regression model, Ohtani (1987, 2001) demonstrated that the iterative Stein-rule estimator of the disturbance variance is dominated by the usual estimator of the disturbance variance based on OLS under squared error loss criterion but the pre-test variance estimator dominates if the number of regressors is greater than or equal to five. In this article, using small sigma asymptotics, it is demonstrated that the iterative Stein-rule estimator of disturbance variance is not only asymptotically biased but is also dominated by the usual estimator of disturbance variance.

The plan of the paper is as follows. Section 2 describes the model and the estimators. In Section 3 the properties of the disturbance variance estimators are studied and a comparison is made. Lastly, in Appendix, the proof of the theorem is outlined.

[1] *Department of Mathematics and Statistics, Banasthali Vidyapeeth, Rajasthan*

2. THE MODEL AND THE ESTIMATORS

Let the true model be

$$y = X\beta + u \tag{2.1}$$

where y is an n × 1 vector of observations on the variable to be explained, X is an n × p full column rank matrix of n observations on p explanatory variables, β is a p × 1 vector of regression coefficients and u is an n × 1 vector of disturbances. The elements of the disturbance vector u are assumed to be independently and identically distributed each following normal distribution with mean 0 and variance σ^2, (σ^2 being unknown) so that

$$E(u) = 0 \tag{2.2}$$
$$E(uu') = \sigma^2 I_n$$

Application of least squares to (2.1) yields the ordinary least squares (OLS) estimator of β given by

$$b = (X'X)^{-1}X'y \tag{2.3}$$

Using it, the estimator of the disturbance variance is constructed as

$$s^2 = \frac{1}{m}(y - Xb)'(y - Xb) \tag{2.4}$$

where m is any arbitrary scalar. The Stein- rule estimator of β proposed by James and Stein (1961) is given by

$$\hat{\beta} = \left[1-k\frac{ms^2}{b'X'Xb}\right]b \quad (2.5)$$

where k is any positive characterizing scalar. The iterative Stein-rule estimator is constructed with the help of residuals obtained by SR estimation of the regression coefficients and is given by

$$\hat{\sigma}^2 = \frac{1}{m}\left(y-X\hat{\beta}\right)'\left(y-X\hat{\beta}\right) \quad (2.6)$$

where m is any arbitrary scalar and the properties of s^2 and that of $\hat{\sigma}^2$ can be studied for various values of m.

3. PROPERTIES OF DISTURBANCE VARIANCE ESTIMATORS

We notice from (2.3) that

$$s^2 = \frac{1}{m}y'\overline{P}_X y \text{ and } \overline{P}_X = I-X(X'X)^{-1}X' \quad (3.1)$$

so that the bias and the mean squared error of s^2 are given by

$$B(s^2) = \frac{\sigma^2}{m}(n-p-m) \quad (3.2)$$

$$\text{and } M(s^2) = \frac{\sigma^4}{m^2}\left[(n-p-m)^2+2(n-p)\right] \quad (3.3)$$

respectively. To study the properties of the iterative Stein-rule estimator of disturbance, we notice that

$$y-X\hat{\beta} = y-Xb+\frac{kms^2}{b'X'Xb}Xb \quad (3.4)$$

from which it is easy to see that

$$\hat{\sigma}^2-\sigma^2 = s^2-\sigma^2+\frac{k^2ms^4}{b'X'Xb} \quad (3.5)$$

$$\text{so that } B(\hat{\sigma}^2) = E(\hat{\sigma}^2-\sigma^2) \geq B(s^2) \quad (3.6)$$

clearly indicating that the Stein-rule based estimator is rather more biased than its counterpart stemming from least squares. In order to derive the magnitude of bias and mean squared error of $\hat{\sigma}^2$, small sigma asymptotics is used which simply requires σ to be small so that the sampling error of the estimator may be expanded in higher orders of σ and then taking term by term expectation of each term in the expansion. Assuming σ to be small is justifiable from the fact that if it is large, the model (2.1) will not be well explained by the explanatory variables in X. This technique was first suggested by Kadane (1971) and later used by many, to cite a few see; e.g. Vinod and *et al.* Ullah (1980), Dube *et al.* (1991), Srivastava and Dube (1993), Srivastava *et al.* (1996) and the references therein. In the following theorem, small sigma asymptotics is used to find the magnitude of bias and MSE of $\hat{\sigma}^2$.

Theorem. When disturbances are small, the bias to order $o(\sigma^4)$ and the mean squared error to order $o(\sigma^4)$ of the estimator $\hat{\sigma}^2$ are given by

$$B(\hat{\sigma}^2) = \frac{\sigma^2}{m}(n-p-m)+\frac{\sigma^4}{m}\frac{k^2(n-p)(n-p+2)}{\beta'X'X\beta} \quad (3.7)$$

$$M(\hat{\sigma}^2) = \frac{\sigma^4}{m^2}\left[(n-p-m)^2+2(n-p)\right] + \frac{2\sigma^6}{m^2}\frac{k^2(n-p)(n-p+2)(n-p+4)}{\beta'X'X\beta} \quad (3.8)$$

Proof. See Appendix

Let us now compare the performance of the usual (2.4) and the SR based (2.6) estimator of disturbance for various values of scalar m. It can be easily verified that s^2 is an unbiased estimator of σ^2 if m = (n – p), has the smallest mean squared error if m =(n – p + 2) and the choice m = n yields the maximum likelihood estimate of σ^2 when errors are normally distributed. Interestingly, for all the above choices of m yield consistent estimates of σ^2. However, from (3.7) it is clearly evident that $\hat{\sigma}^2$ is a biased estimator even if we choose m = (n – p). In fact, this bias does not vanish even if the number of observations is very large, clearly indicating the iterative SR estimator of the disturbance variance to be asymptotically biased. Also, the comparison of the expressions (3.3) with (3.8) establishes that the OLS based estimator of disturbance variance estimator uniformly dominates the Stein-rule based estimator under mean squared error criterion at least in small sigma sense.

4. CONCLUSION

The present article demonstrated that the iterative Stein-rule estimator of disturbance variance is not only asymptotically biased but is also dominated by the usual estimator of disturbance variance. As we know in experimental designs of agricultural studies, minimization and estimation of disturbance variance is an important issue. The result derived here will prove to be helpful in this direction with the advantage of minimum risk attached to this approach as opposed to the classical approach in which failure of distributional assumption of disturbances can seriously affect the quality of the inferences.

ACKNOWLEDGEMENT

The authors wish to acknowledge the referees helpful comments leading to the revision of the article.

REFERENCES

Dube, M., Srivastava, V.K., Toutenburg, H. and Wijekoon, P. (1991). Stein rule estimator under inclusion of superfluous variables in linear regression models. *Comm. Statist. — Theory Methods*, **20(7),** 2009-2022.

James, W. and Stein, C. (1961). Estimation with quadratic loss. *Proc. of fourth Berkley Symposium,* **1** (Univ. of California Press, Berkley, CA), 361-379.

Kadane, J.B. (1971) Comparison of K- class estimators when. the disturbances are small. *Econometrica*, **39**,723-738.

Maatta, J.M. and Casella, G. (1990). Developments in decision-theoretic variance estimation. *Statist. Sci.,* **5(1),** 90-120.

Ohtani, K. (1987). Inadmissibility of the iterative Stein-rule estimator of the disturbance. Variance in linear regression *Eco. Lett.*, **24,** 51-55.

Ohtani, K. (2001). MSE dominance of the pretest iterative variance estimator over the iterative variance estimator in regression. *Statist. Probab. Lett.*, **54**, 331-340.

Srivastava, V.K. and Dube M. (1993). Properties of the OLS and Stein-rule predictions in linear regression models with proxy variables. *Statistical Papers,* **34**, 27-41.

Srivastava,V.K., Dube, M. and Singh, V. (1996). Ordinary least squares and Stein rule predictions in regression models under inclusion of some superfluous variables. *Statistical Papers*, **37**, 253-265.

Vinod, H.D. and Ullah, A. (1981). *Recent Advances in Regression Analysis.* Marcel Dekker.

Wan, A.T.K. and Kurumai, H. (1999). An iterative feasible minimum mean squared error estimator of the disturbance variance in linear regression under asymmetric loss. *Statist. Probab. Lett*, **45**, 253-259.

APPENDIX

For application of small sigma asymptotic approximations, let us write the model (2.1) as

$$y = X\beta + u; \qquad u = w \tag{A.1}$$

so that w follows a multivariate normal distribution having mean vector 0 and variance-covariance matrix I_n. Thus we have

$$E(w'Aw) = (\text{tr}\,A)$$

$$E(w'Aw)^2 = (\text{tr}\,A)[(\text{tr}\,A)+2]$$

$$E(w'Aw)^3 = (\text{tr}\,A)[(\text{tr}\,A)+2][(\text{tr}\,A)+4]$$

where A is any $n \times n$ symmetric matrix with non-stochastic elements. Using (2.3) and (A.1) we get

$$b'X'Xb = \beta'X'X\beta + 2\sigma\beta'X' + \sigma^2 w'P_X w \text{ giving}$$

$$(b'X'Xb)^{-1} = \frac{1}{\beta'X'X\beta}\left[1 + 2\sigma\frac{\beta'X'w}{\beta'X'X\beta} + \sigma^2\frac{w'P_X w}{\beta'X'X\beta}\right]^{-1}$$

Expanding and retaining terms to order $o(\sigma)$ gives

$$(b'X'Xb)^{-1} = \frac{1}{\beta'X'X\beta}\left[1 - 2\sigma\frac{\beta'X'w}{\beta'X'X\beta}\right] \tag{A.2}$$

From (3.1) and (A.1) we can also write

$$s^2 = \frac{\sigma^2}{m} w'\bar{P}_X w \tag{A.3}$$

so that using (A.2) and (A.3) gives

$$\frac{s^4}{b'X'Xb} = \frac{\sigma^4}{m^2}\frac{(w'P_X w)^2}{\beta'X'X\beta} \tag{A.4}$$

to order $o(\sigma^4)$. Now using (A.4) in (3.5) gives the bias of $\hat{\sigma}^2$. For mean squared error of $\hat{\sigma}^2$, we notice that

$$M(\hat{\sigma}^2) = E(\hat{\sigma}^2 - \sigma^2)$$

$$= M(s^2) + 2k^2 mE\left(\frac{s^4(s^2-\sigma^2)}{b'X'Xb}\right) + E\left(\frac{k^4 m^2 s^8}{(b'X'Xb)^2}\right) \tag{A.5}$$

Using (A.3) along with (A.4), we get to order $o(\sigma^6)$

$$E\left(\frac{s^4(s^2-\sigma^2)}{b'X'Xb}\right) = \frac{\sigma^6}{m^2\beta'X'X\beta}\left[\frac{1}{m}E(w'\bar{P}_X w)^3 - E(w'\bar{P}_X w)^2\right]$$

while the contribution of the last term on the right hand side of (A.5) up to order $o(\sigma^6)$ is 0.