IJMS, *Vol. 15, Nos. 1-2, (January-June 2016), pp. 31-37*

ISSN: 0972-754X

STEIN-RULE ESIMATOR UNDER BALANCED LOSS

Madhulika Dube, Veena Manocha

Abstract

The article studies and compares the performance properties of a weighted average estimator of Ordinary Least Squares and Stein-rule considering balanced loss function proposed by Zellner (1994). Superiority conditions have been derived assuming error distribution to be non-normal.

1. INTRODUCTION

Recently a number of research articles have appeared that utilize balanced loss, proposed by Zellner (1994), to compare the performance properties of various estimators in linear regression model [see, e.g. Wan (1994), Giles et.al. (1996) and Ohtani (1998)].The idea behind using balanced loss is to take into account both the "goodness of fit" and the precision of estimator simultaneously.

Assuming normally distributed errors, Giles et.al.(1996) studied the performance of Stein-rule estimator using balanced loss. In many practical situations, however, normality assumption is quite often unwarranted and may lead to invalid and erroneous inferences. Therefore, in this paper considering a weighted average estimator of the ordinary least squares and the Stein-rule, the conditions of dominance have been derived using small disturbance asymptotic theory and assuming error distribution to be not necessarily normal.

In Section 2, the model and the estimators have been described. Section 3 discusses the properties of the estimator and lastly in Appendix, proof of the theorem is provided.

2. THE MODEL AND THE ESTIMATORS

Let us postulate the following linear regression model

Received 28.02.04

Key words and phrases: Linear Regression Model, Ordinary Least Squares (OLS), Stein-rule, Small Disturbance asymptotic expansions, non-normality, Balanced loss Function.

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

where y is an n x 1 vector of observations on the variable to be explained, X is an n x p full column rank matrix of n observations on p explanatory variables, β is a p x 1 vector of regression coefficients associated with them and u is an n x 1 vector of disturbances, the elements of which are assumed to be independently and identically distributed having finite moments up to order four so that

$$E(u_t) = 0, \quad E(u_t^2) = \sigma^2, \quad E(u_t^3) = \sigma^3\gamma_1, \quad E(u_t^4) = \sigma^4(\gamma_2 + 3); \qquad t = 1,2,3.........n. \tag{2.2}$$

where γ_1 and γ_2 are Pearsonian measures of skewness and kurtosis respectively. For normally distributed errors both γ_1 and γ_2 are zero so that for non-normal distributions γ_1 and γ_2 provide measures of departure from normality. Using (2.2) some expectations that are useful in studying the properties of estimators can be written as follows:

$$E(u'Au) = \sigma^2(\mathrm{tr}\,A) \tag{2.3}$$

$$E(u'Au, u) = \sigma^3\gamma_1(I_n * A)e \tag{2.4}$$

$$E(u'Au.uu') = \sigma^4\,[\gamma_2(I_n * A) + (\mathrm{tr}\,A)I_n + 2A] \tag{2.5}$$

where A is any symmetric matrix with nonstochastic elements, '*' denotes the Hadamard product operator and e is an nx1 vector with all elements unity.

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

$$b = (X'X)^{-1}X'y \tag{2.6}$$

which is well known to be the best linear unbiased estimator of β, having variance-covariance matrix

$$V(b) = \sigma^2(X'X)^{-1}. \tag{2.7}$$

The Stein-rule estimator of the regression coefficient from (2.1) is given by

$$\hat{\beta} = \left[1 - k\frac{(y - Xb)'(y - Xb)}{b'X'Xb}\right]b \tag{2.8}$$

where k is any positive non-stochastic scalar characterizing the estimator.

The balanced loss function proposed by Zellner (1994) is given by

$$L(\tilde{\beta},\beta) = w_1(y - X\tilde{\beta})'(y - X\tilde{\beta}) + (1 - w_1)(\tilde{\beta} - \beta)'X'X(\tilde{\beta} - \beta) \quad ; \quad 0 \le w_1 \le 1 \tag{2.9}$$

which for $w_1 = 1$ provides goodness of fit of the estimator $\tilde{\beta}$ whence $w_1 = 0$ focuses on the precision of the estimator. Any other value of w_1 between 0 and 1 gives different weight to the goodness of fit and the precision of estimation. Following Ohtani (1998), when balanced loss function is used another estimator, which is a weighted average estimator of the OLS and the Stein-rule, should be used which takes into account both, the goodness of fit and the precision of estimation simultaneously. It is given by

$$\tilde{\beta} = w_2 b + (1 - w_2)\hat{\beta} \qquad (2.10)$$

where $w_2 \in [0,1]$ .Clearly $w_2 = 1$ yields OLS estimator and $w_2 = 0$ gives the Stein-rule estimator. If $w_1 \neq w_2$, the weight to goodness of fit in the loss function is different from the weight in the estimator, i.e., the magnitude of importance for goodness of fit is different when we evaluate loss and when we estimate parameters; the same being true for precision of estimation. Similarly $w_1 = w_2$ indicates that the same weight is used for the loss function and the estimator.

3. PROPERTIES OF ESTIMATORS

Assuming error distribution to be normal Giles (1996) derived the exact risk of the Stein-rule estimator under balanced loss and found that some well-known results under quadratic error loss are not robust to balanced loss. When the error distribution is non-normal the exact expressions of risk of Stein-rule estimator is not only intricate but it is difficult to draw clear inferences from them. Therefore, here small disturbance asymptotic approximations are used in order to obtain expression of risk of $\tilde{\beta}$ under balanced loss when errors are not necessarily normal.

THEOREM: When disturbances are small and not necessarily normal, the risk of $\tilde{\beta}$ to order $O(\sigma^4)$, is given by

$$R(\tilde{\beta}) = \sigma^2[p + (n-2p)w_1)] - 2\sigma^3 k\gamma_1 \frac{(1-w_1)(1-w_2)}{\beta' X' X\beta} \beta' X'(I * \bar{P}_X) e \qquad (3.1)$$

$$+ \frac{\sigma^4 k(1-w_2)}{\beta' X' X\beta}\Big[k(1-w_2)\{\gamma_2 \operatorname{tr}(\bar{P}_X * \bar{P}_X) + (n-p)(n-p+2)\}$$

$$- 2(1-w_1)\{\gamma_2 \operatorname{tr} M_2 (I_n * \bar{P}_X) + (n-p)(p-2)\}\Big]$$

where $P_X = X(X'X)^{-1}X'$, $\bar{P}_X = I - P_X$ and $M_2 = P_X - \dfrac{2}{\beta' X' X\beta} X\beta\beta' X'$.

Proof: See Appendix

When $w_2 = 1$, we get the expression of the risk of OLS under balanced loss while $w_2 = 0$ gives the risk of Stein-rule estimator under balanced loss. These are given by

$$R(b) = \sigma^2 \left[p + (n-2p) w_1 \right] \qquad \text{and} \tag{3.2}$$

$$R(\hat{\beta}) = \sigma^2 \left[p + (n-2p) w_1) \right] - 2\sigma^3 \frac{k\gamma_1 (1-w_1)}{\beta' X' X \beta} \beta' X' (I * \overline{P}_X)\, e \tag{3.3}$$

$$+ \frac{\sigma^4 k}{\beta' X' X \beta} \Big[k \left\{ \gamma_2 \mathrm{tr}\left(\overline{P}_X * \overline{P}_X \right) + (n-p)(n-p+2) \right\}$$

$$- 2(1-w_1) \{ \gamma_2 M_2 \left(I_n * \overline{P}_X \right) + (n-p)(p-2) \} \Big]$$

respectively. Looking at the expression (3.1) we observe that due to non-normality of the disturbances the terms of order $O(\sigma^4)$ are generally non-zero. The skewness shows its effect in terms of order $O(\sigma^3)$ while the effect of kurtosis appears in terms of order $O(\sigma^4)$. Comparison of risk functions (3.1) and (3.2)

$$R(b) - R(\tilde{\beta}) = 2 \frac{\sigma^3 k \gamma_1 (1-w_1)(1-w_2)}{\beta' X' X \beta} \beta' X' \left(I_n * \overline{P}_X \right) e \tag{3.4}$$

$$- \frac{\sigma^4 k (1-w_2)}{\beta' X' X \beta} \Big[k(1-w_2) \left\{ \gamma_2 \mathrm{tr}\left(\overline{P}_X * \overline{P}_X \right) + (n-p)(n-p+2) \right\}$$

$$- 2(1-w_1) \left\{ \gamma_2 \mathrm{tr}\, M_2 \left(I_n * \overline{P}_X \right) + (n-p)(p-2) \right\} \Big]$$

For skewed distributions of disturbances, the leading term in (3.4) depends upon γ_1, a measure of departure from symmetry. Positivity of this term implies that $\tilde{\beta}$ will be superior to b while negativity implies otherwise.

For symmetrical distributions ($\gamma_1 = 0$), the change in efficiency is determined by the terms of order $O(\sigma^4)$. Clearly $\tilde{\beta}$ dominates b under balanced loss if k is chosen to satisfy

$$0 < k < 2 \frac{(1-w_1)}{(1-w_2)} \frac{g}{q} \tag{3.5}$$

where $q = \gamma_2 \operatorname{tr}\left(\bar{P}_X * \bar{P}_X\right) + (n-p)(n-p+2)$ and

$$g = \gamma_2 \operatorname{tr}\{M_2\left(I_n * \bar{P}_X\right)\} + (n-p)(p-2)$$

Following Vinod and Ullah (1978) similar conditions for symmetrical leptokurtic $(\gamma_1 = 0, \gamma_2 > 0)$ and symmetrical platykurtic $(\gamma_1 = 0, \gamma_2 < 0)$ distributions of disturbances can be derived. For this purpose we introduce the following notations:

$$\theta = \frac{\gamma_2}{(n-p)}, \quad \phi = \operatorname{tr}(X'X)G \qquad \text{with } G = (X'X)^{-1}X'(I_n * \bar{P}_X)X(X'X)^{-1}$$

Using these, the inequality (3.5) can be rewritten as

$$0 < k < \frac{2(1-w_1)(n-p)}{(1-w_2)q}\left[(1+\theta\phi)p - 2\left(1+\theta\frac{\beta'X'GX'X\beta}{\beta'X'X\beta}\right)\right] \qquad (3.6)$$

provided the quantity in square brackets is positive. Observing that

$$0 \le \mu_* \le \frac{\beta'X'XGX'X\beta}{\beta'X'X\beta} \le \mu^* \le 1$$

$$0 \le \mu_* \le \operatorname{tr} G(X'X) \le \mu^* \le 1$$

where μ_* and μ^* are the smallest and largest characteristic root of the matrix G and using it in (3.6),we notice that the inequality is satisfied for symmetrical leptokurtic distribution of disturbances so long as

$$0 < k < 2\frac{(1-w_1)(n-p)}{(1-w_2)q}\left[(1+\theta\mu_*)p - 2\left(1+\theta\mu^*\right)\right]; \qquad p > 2\frac{(1+\theta\mu^*)}{\left(1+\theta\mu_*\right)} \qquad (3.7)$$

Similarly when the distribution of errors is symmetrical platykurtic (i.e. $\gamma_1 = 0$ and $\gamma_2 < 0$) then $\theta < 0$ holds for all (n -p)$\geq$2. If (n -p)=1 then $\theta < 0$ provided 1+γ_2>0. Therefore in this case (3.6) holds so long as

$$0 < k < \frac{2(1-w_1)(n-p)}{(1-w_2)q}(1+\theta\mu^*)\left[p - 2\frac{\left(1+\theta\mu_*\right)}{(1+\theta\mu^*)}\right], \qquad p > 2\frac{\left(1+\theta\mu_*\right)}{(1+\theta\mu^*)} \qquad (3.8)$$

For normally distributed disturbances, the condition on k reduces to

$$0<k<\frac{2(1-w_1)}{(1-w_2)}\frac{(p-2)}{(n-p-2)}\ ; \qquad p>2 \tag{3.9}$$

It is interesting to note that when $w_1 < w_2$, i.e., the weight to goodness of fit in the loss function is smaller than that given to the ordinary least squares estimator, then the range of dominance of $\tilde{\beta}$ over b in all (3.6)-(3.9) is larger than that of $\hat{\beta}$ over b and when $w_1 > w_2$ the range of dominance of $\tilde{\beta}$ over b is smaller than that of $\hat{\beta}$ over b. This is in accordance to the fact that the OLS estimator takes in to account the goodness of fit while Stein-rule estimation takes into account precision of estimation only. Interestingly, when $w_1 = w_2$, (3.9) reduces to

$$0<k<\ 2\frac{(p-2)}{(n-p-2)}\ ; \qquad p>2 \tag{3.10}$$

which is the usual condition of dominance of Stein-rule estimator over OLS. Hence in this case the range of k where a weighted average estimator dominates OLS under balanced loss is same as the range of k where Stein-rule estimator dominates the OLS estimator.

APPENDIX

To obtain the expressions small σ asymptotic approximations, we rewrite model (2.1) as

$$y = X\beta + \sigma v \quad ; \quad (u = \sigma v) \tag{A.1}$$

where the elements of v are independently and identically distributed with finite moments at least up to order $O(\sigma^4)$. We also notice from (2.10) that

$$\begin{aligned} R(\tilde{\beta}) &= E[L(\tilde{\beta},\beta)] \\ &= E[(\tilde{\beta}-\beta)'X'X(\tilde{\beta}-\beta)] - 2\sigma w_1 E\{v'X(\tilde{\beta}-\beta)\} + \sigma^2 w_1 E(v'v) \end{aligned} \tag{A.2}$$

Using (2.6), (2.8) and (A.1) we can write

$$X(\tilde{\beta}-\beta) = X(b-\beta) - k(1-w_2)\frac{(y-Xb)'(y-Xb)}{b'X'Xb}\,b \tag{A.3}$$

$$= \sigma P_X v - \sigma^2 k\frac{(1-w_2)v'\overline{P}_X v}{\beta'X'X\beta}\left[1+\frac{2\sigma\beta'X'v}{\beta'X'X\beta}+\frac{\sigma^2 v'P_X v}{\beta'X'X\beta}\right]^{-1}(X\beta+\sigma P_X v)$$

Since σ is small, we can expand the quantity in square brackets in increasing powers of σ. Retaining terms up to order $O(\sigma^3)$, we get

$$X(\tilde{\beta}-\beta)=\sigma\pi_1+\sigma^2\pi_2+\sigma^3\pi_3 \tag{A.4}$$

where

$$\pi_1 = P_X v$$

$$\pi_2 = -\frac{k(1-w_2)}{\beta' X' X\beta} v'\bar{P}_X v.X\beta$$

$$\pi_3 = -\frac{k(1-w_2)}{\beta' X' X\beta} v'\bar{P}_X v.M_2 v$$

Using the above expression in (A.1), we get the result in the theorem.

REFERENCES

[1] Dube, M. (1995), “Stein-rule Estimation in Misspecified Linear Regression Models under non-normal Disturbances”, *Journal of Scientific Research*, 44, pp. 171-189.

[2] Giles, J.A., Giles, D.E.A. and Ohtani, K, (1996), “The Exact Risks of Some Pre-test and Stein–type Regression Estimators Under Balanced Loss” Communications in Statistics–Theory & Methods, 25(12), pp. 2901-24.

[3] Ohtani, K. (1998), “The Exact Risks of Weighted Average Estimator of OLS and Stein –rule Estimator in Regression Under Balanced Loss” Statistics and Decisions,16, pp. 35-45.

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

[5] Ullah, A. Srivastava,V.K. and Chandra, R. (1983), “Properties of Shrinkage Estimators in Linear Regression when Disturbances are Not Normal” *Journal of Econometrics*, 21, pp. 389-402.

[6] Wan, A. T.K. (1994), “Risk Comparison of the Inequality Constrained Least Squares and other Related Estimators under Balanced Loss” *Economic Letters*, 46, pp. 203-10.

[7] Zellner, A. (1994), “Bayesian and Non-Bayesian Estimation using Balanced Loss Function” In Statistical Decision Theory and Related Topics eds. S.S.Gupta and J.O.Berger (Springer Verlag), pp. 377-90.

Madhulika Dube
Department of Statistics
M.D. University, Rohtak

Veena Manocha
Joint Director (Statistics)
Directorate of Human Resource Management
CCS HAU, Hisar