Hacettepe Journal of Mathematics and Statistics
Volume 38 (3) (2009), 351 – 356

COMPARISON OF THE ITERATIVE STEIN-RULE AND THE USUAL ESTIMATORS OF THE ERROR VARIANCE UNDER THE PITMAN NEARNESS CRITERION

Deniz Ünal* and Güzin Yüksel*†

Received 11 : 11 : 2008 : Accepted 10 : 08 : 2009

Abstract

The iterative Stein-rule estimator and the usual estimator of the error variance are compared under the Pitman Nearness Criterion. An exact expression of Pitman's Nearness probability is derived and numerically evaluated.

Keywords: Iterative Stein-rule estimator, Pitman Nearness Criterion, Stein-rule estimator.

2000 AMS Classification: 62 J 05, 62 F 10, 62 H 12.

1. Introduction

The mean squared error (MSE) has often been used to measure the performance of estimators. However, the justification for MSE as a criterion has often be argued. Therefore Pitman [8] introduced a measure of the closeness of an estimator compared to the true parameter value, and defined Pitman nearness (PN) criteria. PN and MSE were investigated by Rao [9] in order to compare these two criteria and he claims that PN is a more intrinsic criterion than the MSE. His examples show that estimators with the minimum MSE property can have very poor performance in terms of PN. Peddada [7] characterized PN in terms of MSE to overcome the complexity of the PN computation. The least squares estimator and the James-Stein estimator of the vector-valued parameter in a multiple linear regression model are compared in the sense of PN by Keating and Czitrom [2]. The reader who wishes to find more out about PN may consult Keating, Mason and Sen [4]. The Stein-rule (SR) and ordinary least squares (OLS) were compared for regression error variance under the PN criterion when variables are omitted by Ohtani

*Department of Statistics, Faculty of Arts and Sciences, Çukurova University, 01330 Adana, Turkey.
E-mail: (D. Ünal) dunal@cu.edu.tr, denizozp@gmail.com (G. Yüksel) yguzin@cu.edu.tr
†Corresponding Author.

and Wan [6]. On the other hand, Ohtani [5] defined a new estimator of the disturbance variance, called the iterative Stein-rule estimator (ISRE) and showed the dominance of ISRE with respect to the PN criteria. Unal [10] compared this estimator with the OLS estimator with respect to the MSE criterion in a regression model with proxy variables.

The aim of this paper is to use the PN criterion to compare the ISRE of the disturbance variance to the usual estimator of the disturbance variance theoretically. In Section (2), the model and the estimators will be given. The exact formula of the PN probability for the comparison of the ISRE and the usual estimators of the disturbance variance will be obtained. In Section (3), the PN probabilities for a range of parameter values will be computed.

2. The model and the estimators

Let us first consider the classical linear regression model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim N(0, \sigma^2 I_n), \tag{2.1}$$

where $\boldsymbol{y}$ is the $n \times 1$ observation vector of a dependent variable, $\boldsymbol{X}$ the $n \times k$ matrix of observations of non-stochastic independent variables with column rank, $\boldsymbol{\beta}$ the $k \times 1$ vector of parameters and $\boldsymbol{\epsilon}$ the $n \times 1$ vector of normal disturbance terms. The usual estimator for the parameter vector $\boldsymbol{\beta}$ is

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y},$$

and the SR estimator for the parameter vector $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = \left[1 - \frac{a\boldsymbol{e}'\boldsymbol{e}}{\boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}\right]\boldsymbol{b},$$

where $\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}$ and a is a constant in $0 \leq a \leq 2(k-2)/(n-k+2)$ (e.g. [1, 2]). The usual estimator of the disturbance variance is

$$s^2 = \frac{(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})}{n-k} = \frac{\boldsymbol{e}'\boldsymbol{e}}{n-k}. \tag{2.2}$$

Using $\hat{\boldsymbol{\beta}}$ instead of $\boldsymbol{b}$ in the formula for the usual estimator of the disturbance variance, the ISRE of the disturbance variance [5]

$$\hat{\sigma}_s^2 = \frac{(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})}{n-k} \tag{2.3}$$

is obtained.

Setting $u_1 = \dfrac{\boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\sigma^2}$ and $u_2 = \dfrac{\boldsymbol{e}'\boldsymbol{e}}{\sigma^2}$, the ISRE of the disturbance variance can be written in the form (e.g. [5, 3]),

$$\hat{\sigma}_s^2 = \frac{1}{n-k}\sigma^2\left[u_2 + \frac{a^2{u_2}^2}{u_1}\right], \tag{2.4}$$

where $u_1 \sim \chi^2(k, \lambda)$ with non-centrality parameter $\lambda = \dfrac{\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}}{\sigma^2}$ and $u_2 \sim \chi^2_{(n-k)}$. Also, u_1 and u_2 are independent.

If the squared error loss is considered as the loss function, then the PN to compare $\hat{\sigma}_s^2$ and s^2 is defined as the following probability,

$$\begin{aligned} PC(\hat{\sigma}_s^2, s^2) &= Pr\left(\left[\frac{\hat{\sigma}_s^2 - \sigma^2}{\sigma^2}\right]^2 < \left[\frac{s^2 - \sigma^2}{\sigma^2}\right]^2\right) \\ &= Pr\left(\frac{a_1 u_2}{u_1}\, u_2^2\left(2 - \frac{a_2}{u_2} + \frac{a_3 u_2}{u_1}\right) < 0\right), \end{aligned} \tag{2.5}$$

where $\nu = n - k$, $a_1 = a^2/\nu^2$, $a_2 = 2\nu$ and $a_3 = a^2$. The estimator $\hat{\sigma}_s^2$ dominates s^2 whenever $PC\big(\hat{\sigma}_s^2, s^2\big) > 0.5$, and vice versa.

The probability given in (2.5) can be written as

$$\iint_R f_1(u_1) f_2(u_2) du_1 du_2, \tag{2.6}$$

where R is the region $\{(u_1, u_2) : a_1 u_2/u_1 > 0,\ 2 + a_3 u_2/u_1 - a_2/u_2 < 0\}$,

$$f_1(u_1) = e^{-\lambda/2} \sum_{i=0}^{\infty} \frac{(\lambda/2)^i}{i!} \frac{u_1^{k/2+i-1} e^{-u_1/2}}{2^{k/2+i}\Gamma(\frac{k}{2}+i)} \tag{2.7}$$

and

$$f_2(u_2) = \frac{u_2^{\nu/2-1} e^{-u_2/2}}{2^{\nu/2}\Gamma(\frac{\nu}{2})}. \tag{2.8}$$

Using (2.7) and (2.8) in (2.6) we obtain

$$\iint_R \sum_{i=0}^{\infty} K_i u_1^{k/2+i-1} u_2^{\nu/2-1} \exp\left[-(u_1+u_2)/2\right] du_1\, du_2, \tag{2.9}$$

where

$$K_i = \frac{(\lambda/2)^i}{i!} \frac{\mathrm{e}^{-\lambda/2}}{2^{(k+\nu)/2+i}\Gamma(\frac{k}{2}+i)\Gamma(\frac{\nu}{2})}$$

Making the change of variables $t_1 = a_1 \dfrac{u_2}{u_1}$ and $t_2 = a_3 \dfrac{u_2}{u_1} - \dfrac{a_2}{u_2}$ we obtain,

$$\begin{aligned} &\sum_{i=0}^{\infty} \int_0^{\infty}\int_{-\infty}^{-2} K_i \frac{a_1^{k+v/2+2i+1} a_2^{(k+\nu)/2+i}}{t_1^2 (a_3 t_1^2 - a_1 t_1 t_2)^{k/2+i-1} (a_3 t_1 - a_1 t_2)^{\nu/2+2}} \\ &\qquad\qquad \times \exp\left[-\frac{a_1 a_2 (a_1+t_1)}{2(a_3 t_1^2 - a_1 t_1 t_2)}\right] dt_2\, dt_1, \end{aligned} \tag{2.10}$$

where $J = \dfrac{-a_1^4 a_2^2}{t_1^2 (a_3 t_1 - a_1 t_2)^3}$. Again, making the change of variables, $z = \dfrac{a_1 a_2 (a_1+t_1)}{2(a_3 t_1^2 - a_1 t_1 t_2)}$, we obtain

$$\sum_{i=0}^{\infty} K_i \int_0^{\infty}\int_0^{\frac{a_1 a_2 (a_1+t_1)}{2(a_3 t_1^2 + 2a_1 t_1)}} \frac{a_1^{k/2+i} 2^{(k+\nu)/2+i} z^{(k+\nu)/2+i-1} \exp(-z)\ t_1^{\nu/2-1}}{(a_1+t_1)^{(k+\nu)/2+i}} dz\, dt_1. \tag{2.11}$$

Making the change of variable $x = \frac{t_1}{a_1+t_1}$, equation (2.11) becomes

$$\begin{aligned} &\sum_{i=0}^{\infty} K_i 2^{(k+\nu)/2+i} \\ &\quad \times \int_0^1 x^{v/2-1}(1-x)^{k/2+i-1} \int_0^{\frac{a_2(1-x)}{2((a_3-2)x^2+2x)}} z^{(k+\nu)/2+i-1} \exp(-z)\, dz\, dx. \end{aligned} \tag{2.12}$$

Using the incomplete gamma function

$$P(a, x) = \frac{1}{\Gamma(a)} \int_0^x t^{a-1} \exp(-t)\, dt,$$

one obtains

$$\begin{aligned} &\sum_{i=0}^{\infty} K_i 2^{(k+\nu)/2+i} \Gamma\left(\frac{k+v}{2} + i\right) \\ &\quad \times \int_0^1 x^{v/2-1}(1-x)^{k/2+i-1} P\left(\frac{k+v}{2} + i, \frac{a_2(1-x)}{2((a_3-2)x^2+2x)}\right) dx. \end{aligned} \tag{2.13}$$

Finally using K_i in equation (2.13), it becomes,

$$\sum_{i=0}^{\infty} \frac{e^{-\lambda/2}(\lambda/2)^i}{i!\,B(v/2,k/2+i)} \int_0^1 x^{v/2-1}(1-x)^{k/2+i-1} P\left(\frac{k+v}{2}+i, \frac{a_2(1-x)}{2[(a_3-2)x^2+2x]}\right) dx \tag{2.14}$$

where $B(x,y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the complete Beta function.

3. Numerical Analysis

The equation given in (2.14) is too complex to give an idea of its magnitude, so it requires a numerical evaluation. This equation has been evaluated for $k = 3, 5, 8$, $n = 20, 40$, and various values of λ. For the evaluation of the integral in (2.14), Simpon's 3/8 rule with 200 subdivisions was used as in [6]. The numerical evaluations were executed on a personal computer using Mathematica code. The infinite series in (2.14) converged rapidly with a convergence tolerance level set to 10^{-12}. Some original results appear in the tables.

Table 1

k	n	λ	PC	k	n	λ	PC	k	n	λ	PC
3	10	0.5	0.118692	5	10	0.5	0.099454	9	10	0.5	0.128703
3	10	1	0.211751	5	10	1	0.178262	9	10	1	0.229644
3	10	5	**0.504243**	5	10	5	0.441461	9	10	5	**0.550588**
3	10	6	**0.524083**	5	10	6	0.463092	9	10	6	**0.573933**
3	10	7	**0.536780**	5	10	7	0.478466	9	10	7	**0.589743**
3	10	8	**0.545034**	5	10	8	0.489793	9	10	8	**0.600878**
3	10	9	**0.550506**	5	10	9	0.498450	9	10	9	**0.609082**
3	10	10	**0.554221**	5	10	10	**0.505302**	9	10	10	**0.615423**
3	10	20	**0.563637**	5	10	20	**0.537072**	9	10	20	**0.646846**
3	10	23	**0.561180**	5	10	23	**0.538735**	9	10	23	**0.649073**
3	10	24	**0.559099**	5	10	24	**0.537869**	9	10	24	**0.648192**
3	10	25	**0.556080**	5	10	25	**0.535999**	9	10	25	**0.646129**
3	10	30	**0.519650**	5	10	30	**0.504638**	9	10	30	**0.609391**
3	10	40	0.317237	5	10	40	0.310383	9	10	40	0.375891
3	10	50	0.105303	5	10	50	0.103327	9	10	50	0.125325
3	20	0.5	0.113769	5	20	0.5	0.090102	9	20	0.5	0.057034
3	20	1	0.202957	5	20	1	0.161834	9	20	1	0.102882
3	20	5	0.483013	5	20	5	0.406633	9	20	5	0.270397
3	20	6	**0.501935**	5	20	6	0.427752	9	20	6	0.288112
3	20	7	**0.514013**	5	20	7	0.443034	9	20	7	0.302280
3	20	8	**0.521837**	5	20	8	0.454480	9	20	8	0.314078
3	20	9	**0.527001**	5	20	9	0.463350	9	20	9	0.324239
3	20	10	**0.530487**	5	20	10	0.470445	9	20	10	0.333223
3	20	20	**0.539056**	5	20	20	**0.503555**	9	20	20	0.392352
3	20	23	**0.536637**	5	20	23	**0.505475**	9	20	23	0.401815
3	20	24	**0.534629**	5	20	24	**0.504751**	9	20	24	0.403611
3	20	25	**0.531723**	5	20	25	**0.503074**	9	20	25	0.404492
3	20	30	**0.496828**	5	20	30	0.473883	9	20	30	0.389547
3	20	40	0.303269	5	20	40	0.291581	9	20	40	0.245413
3	20	50	0.100662	5	20	50	0.097079	9	20	50	0.082501

Table 2

k	n	λ	PC	k	n	λ	PC	k	n	λ	PC
3	40	0.5	0.111936	5	40	0.5	0.090938	9	40	0.5	0.054679
3	40	1	0.199589	5	40	1	0.163222	9	40	1	0.098874
3	40	5	0.473349	5	40	5	0.407494	9	40	5	0.264533
3	40	6	0.491550	5	40	6	0.427956	9	40	6	0.282854
3	40	7	**0.503066**	5	40	7	0.442541	9	40	7	0.297673
3	40	8	**0.510445**	5	40	8	0.453287	9	40	8	0.310104
3	40	9	**0.515251**	5	40	9	0.461470	9	40	9	0.320845
3	40	10	**0.518443**	5	40	10	0.467905	9	40	10	0.330344
3	40	20	**0.525681**	5	40	20	0.496240	9	40	20	0.391331
3	40	23	**0.523170**	5	40	23	0.497304	9	40	23	0.400646
3	40	24	**0.521170**	5	40	24	0.496358	9	40	24	0.402364
3	40	25	**0.518302**	5	40	25	0.494492	9	40	25	0.403161
3	40	30	0.484159	5	40	30	0.465009	9	40	30	0.387821
3	40	40	0.295461	5	40	40	0.285629	9	40	40	0.243892
3	40	50	0.098061	5	40	50	0.095034	9	40	50	0.081916
3	60	0.5	0.111434	5	60	0.5	0.092835	9	60	0.5	0.057982
3	60	1	0.198630	5	60	1	0.166471	9	60	1	0.104780
3	60	5	0.470003	5	60	5	0.412540	9	60	5	0.278513
3	60	6	0.487858	5	60	6	0.432535	9	60	6	0.297256
3	60	7	0.499092	5	60	7	0.446588	9	60	7	0.312250
3	60	8	**0.506240**	5	60	8	0.456783	9	60	8	0.324692
3	60	9	**0.510854**	5	60	9	0.464428	9	60	9	0.335333
3	60	10	**0.513888**	5	60	10	0.470351	9	60	10	0.344653
3	60	20	**0.520387**	5	60	20	0.495276	9	60	20	0.402512
3	60	23	**0.517814**	5	60	23	0.495767	9	60	23	0.410819
3	60	24	**0.515811**	5	60	24	0.494664	9	60	24	0.412203
3	60	25	**0.512951**	5	60	25	0.492660	9	60	25	0.412664
3	60	30	0.479087	5	60	30	0.462763	9	60	30	0.395588
3	60	40	0.292324	5	60	40	0.283934	9	60	40	0.247849
3	60	50	0.097015	5	60	50	0.094430	9	60	50	0.083117

4. Conclusion

The ISRE of the disturbance variance is not better than s^2, except for some particular λ values. For example for $k = 3$ and for $n = 10$, $n = 20$, $n = 40$ and for $n = 60$ the ISRE of the disturbance variance is superior (i.e. $PC\big(\hat{\sigma}_s^2, s^2\big) \geq 0.5$) to the usual estimator of the disturbance variance just for $5 \leq \lambda \leq 30$, $6 \leq \lambda \leq 25$, $7 \leq \lambda \leq 25$ and for $8 \leq \lambda \leq 25$, respectively. Here, as n increases, the intervals for the dominance of $\hat{\sigma}_s^2$ decreases.

When the non-centrality parameter λ is small the regression is not significant unless the sample size is very large. Conversely, if λ is large the regression is significant and the error variance is quite small. The results will be of greatest practical importance for intermediate values of λ. So, especially in such cases, the researcher is concerned with which estimator is better. And here, from the tables, it can be seen that the strength of the ISRE of the disturbance variance over the usual estimator of the disturbance variance increases for the intermediate values of λ.

For $k = 5$, these intervals for λ are getting narrower. For example for $n = 10$ and $n = 20$, $PC(\hat{\sigma}_s^2, s^2) \geq 0.5$ for $10 \leq \lambda \leq 30$ and $20 \leq \lambda \leq 25$, respectively.

For $k = 9$ and $n = 10$, $PC(\hat{\sigma}_s^2, s^2) \geq 0.5$ for $5 \leq \lambda \leq 30$, much as for $k = 3$ and $n = 10$. But here the probabilities are greater than the probabilities for $k = 3$ and $n = 10$.

For $k > 9$ there is no situation that makes $PC(\hat{\sigma}_s^2, s^2) \geq 0.5$. This means that the usual estimator of the disturbance variance is superior to the ISRE of the disturbance variance for all values of $k > 9$.

Although Ohtani [5] showed that s^2 is superior to the ISRE of the disturbance variance for $k > 5$ using the MSE criterion, we have found the same result for $k > 9$ using the PN criterion. So it can be said that this result is compatible with Ohtani's result [5], which compared the ISRE of the disturbance with the usual estimator of the disturbance variance using the MSE criterion.

Acknowledgement We thank the referee and Prof. Dr. Olcay Arslan for providing constructive comments that have helped improve the contents of this paper.

References

[1] James, W. and Stein, C. *Estimation with quadratic loss*, Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics and Probability, 361–379, 1961.

[2] Keating, J. P. and Czitrom, V. *A comparison of James-Stein regression with least squares in the Pitman nearness sense*, Journal of Statistical Computation and Simulation **34**, 1–9, 1989.

[3] Keating, J. P. and Mason,R. L. *Closeness comparison of classical and inverse regression estimators*, Computational Statistics and Data Analysis **12**, 4–11, 1991.

[4] Keating, J. P., Mason, R. L. and Sen, P. K. *Pitman's measure of closeness: a comparison of statistical estimators* (Society for Industrial and Applied Mathematics, Philadelphia, 1993).

[5] Ohtani, K. *Inadmisibility of the Iterative Stein-rule estimator of the disturbance variance in a linear regression.* Economics Letters **24**, 51–55, 1987.

[6] Ohtani, K. and Wan, T. K. A. *Comparison of the Stein and the usual estimators for the regression error variance under the Pitman nearness Criterion when variables are omitted.* Statistical Papers, 10.1007/s00362-007-0047-6, 2007.

[7] Peddada, S. D. *A short note on Pitman's measure of nearness*, The American Statistician **39** (4), 298–299, 1985.

[8] Pitman, E. J. G. *The closest estimates of statistical parameters*, Proceedings of the Cambridge Philosophical Society **33**, 212–222, 1937.

[9] Rao, C. R. *Some comments on the minimum mean square error as a criterion of estimation.* In: Csorge M, Dawson, D. A., Rao, J. N. K. and Saleh, A. K. MdE (eds) *Statistics and related topics* (North Holland, Amsterdam, 1981), 123–143.

[10] Unal, D. *The effects of the proxy information on the iterative Stein-rule estimator of the disturbance variance*, Statistical Papers, 10.1007/s00362-009-0220-1, 2009