Ann. Inst. Statist. Math.
31 (1979), Part A, 177-183

STEIN'S POSITIVE PART ESTIMATOR AND BAYES ESTIMATOR

YOSHIKAZU TAKADA

(Received Mar. 31, 1978; revised Apr. 6, 1979)

Summary

Stein's positive part estimator for p normal means is known to dominate the M.L.E. if $p \geqq 3$. In this article by introducing some priors we show that Stein's positive part estimator is posterior mode. We also consider the Bayes estimators (posterior mean) with respect to the same priors and show that some of them dominate M.L.E. and are admissible.

1. Introduction

Let X have the p-variate normal distribution with unknown mean vector θ and covariance matrix I, and let the loss be quadratic, given by

$$L(\theta, \delta) = \|\delta - \theta\|^2 , \tag{1.1}$$

where δ is the vector of estimate. Stein [5] showed that the estimator $\delta(X) = X$ is inadmissible when $p \geqq 3$. James and Stein [2] showed that the estimator

$$\delta(X) = \left(1 - \frac{t(p-2)}{\|X\|^2}\right) X , \qquad 0 < t < 2 \tag{1.2}$$

dominates X and the uniformly best value of t is the James-Stein choice $t = 1$. But Stein [6] showed that the estimator

$$\delta(X) = \left(1 - \min\left(1, \frac{t(p-2)}{\|X\|^2}\right)\right) X , \qquad 0 < t < 2 \tag{1.3}$$

dominates the above estimator. This estimator is called Stein's positive part estimator. Efron and Morris [1] gave its justification by Empirical Bayes approach. In this article we show that Stein's positive part estimators are posterior mode with respect to properly selected priors on θ. We also consider Bayes estimates (posterior mean) with respect to the priors and show the condition under which they dominate X and are admissible.

2. Posterior mode and mean

A generalized prior distribution $\pi_\lambda(\theta)$ of θ, conditional on λ, is given by the density

$$(2.1)\qquad \pi_\lambda(\theta)=\left[\frac{\lambda}{2\pi(1-\lambda)}\right]^{p/2}\exp\left\{-\frac{\lambda}{2(1-\lambda)}\|\theta\|^2\right\}, \qquad 0<\lambda<1,$$

and λ has the density

$$(2.2)\qquad h(\lambda)\propto(1-\lambda)^{p/2}\lambda^{-pa/2}, \qquad a=1-t(p-2)/p.$$

Then it follows that the posterior density of (θ, λ) with respect to the generalized prior with the density

$$(2.3)\qquad \pi(\theta, \lambda)=\pi_\lambda(\theta)h(\lambda),$$

is

$$(2.4)\qquad p_X(\theta, \lambda)=\text{const.}\times\lambda^{t(p-2)/2}\exp\left\{-\frac{1}{2}\left(\frac{\lambda}{1-\lambda}\|\theta-(1-\lambda)X\|^2+\lambda\|X\|^2\right)\right\}.$$

We have the following result.

THEOREM 1. *Stein's positive part estimator is a posterior mode of θ with respect to the above prior.*

PROOF. From (2.4) we have

$$(2.5)\qquad 2\log p_X(\theta, \lambda)=\text{const.}+t(p-2)\log\lambda-\frac{\lambda}{1-\lambda}\|\theta-(1-\lambda)X\|^2-\lambda\|X\|^2.$$

If we denote the posterior mode of (θ, λ) by (θ^*, λ^*), then from (2.5) it follows that

$$(2.6)\qquad \theta^*=(1-\lambda^*)X$$

and λ^* is the value which maximizes the following function

$$(2.7)\qquad g(\lambda)=t(p-2)\log\lambda-\lambda\|X\|^2.$$

It is easily shown that λ^* is given by

$$(2.8)\qquad \lambda^*=\min\left(1, \frac{t(p-2)}{\|X\|^2}\right).$$

Then from (2.6) and (2.8) we have the conclusion.

From (2.4) Bayes estimator with respect to the above prior is

$$\delta(X, t)=[1-\mathrm{E}(\lambda | X, t)]X \, , \tag{2.9}$$

where

$$\mathrm{E}(\lambda | X, t)=\frac{\int_0^1 \lambda^{t(p-2)/2+1}(1-\lambda)^{p/2}\exp(-\lambda\|X\|^2/2)d\lambda}{\int_0^1 \lambda^{t(p-2)/2}(1-\lambda)^{p/2}\exp(-\lambda\|X\|^2/2)d\lambda} \, . \tag{2.10}$$

Let $M(a, b, z)$ denote the confluent hypergeometric function defined by

$$M(a, b, z)=1+\frac{az}{b}+\cdots+\frac{(a)_n z^n}{(b)_n n!}+\cdots \, ,$$

where $(a)_n=a(a+1)\cdots(a+n-1)$, $(a)_0=1$. Using the relation (see equation 13.2.1 and 13.1.27 in [4], p. 505)

$$\int_0^1 e^{z\lambda}\lambda^{a-1}(1-\lambda)^{b-a-1}d\lambda=\frac{\Gamma(b-a)\Gamma(a)}{\Gamma(b)}M(a, b, z) \, , \tag{2.11}$$

and

$$M(a, b, z)=e^z M(b-a, b, -z) \, , \tag{2.12}$$

we obtain that

$$\begin{aligned} &\int_0^1 \lambda^{t(p-2)/2+1}(1-\lambda)^{p/2}\exp(-\lambda\|X\|^2/2)d\lambda \\ &=\frac{\Gamma(p/2+1)\Gamma[t(p-2)/2+2]}{\Gamma[t(p-2)/2+3+p/2]}M\left(\frac{t(p-2)}{2}+2, \frac{t(p-2)}{2}\right. \\ &\qquad \left. +3+\frac{p}{2}, -\frac{\|X\|^2}{2}\right) \\ &=e^{-\|X\|^2/2}\frac{\Gamma(p/2+1)\Gamma[t(p-2)/2+2]}{\Gamma[t(p-2)/2+3+p/2]}M\left(\frac{p}{2}+1, \frac{t(p-2)}{2}\right. \\ &\qquad \left. +3+\frac{p}{2}, \frac{\|X\|^2}{2}\right) \end{aligned} \tag{2.13}$$

and similary

$$\begin{aligned} &\int_0^1 \lambda^{t(p-2)/2}(1-\lambda)^{p/2}\exp(-\lambda\|X\|^2/2)d\lambda \\ &=e^{-\|X\|^2/2}\frac{\Gamma(p/2+1)\Gamma[t(p-2)/2+1]}{\Gamma[t(p-2)/2+2+p/2]}M\left(\frac{p}{2}+1, \frac{t(p-2)}{2}\right. \\ &\qquad \left. +2+\frac{p}{2}, \frac{\|X\|^2}{2}\right) \, . \end{aligned} \tag{2.14}$$

Then from (2.10),

$$(2.15)\qquad \mathrm{E}(\lambda\,|\,X,t)=\left[\frac{t(p-2)+2}{t(p-2)+p+4}\right]\cdot\frac{M(p/2+1,\, t(p-2)/2+3+p/2,\, \|X\|^2/2)}{M(p/2+1,\, t(p-2)/2+2+p/2,\, \|X\|^2/2)}\,.$$

Using the relation (see equation 13.4.4 in [4], p.506)

$$(2.16)\qquad zM(a,b+1,z)=bM(a,b,z)-bM(a-1,b,z)\,,$$

and (2.15), we have

$$(2.17)\qquad \mathrm{E}(\lambda\,|\,X,t)=\tau(\|X\|^2)/\|X\|^2$$

where

$$(2.18)\qquad \tau(\|X\|^2)=[t(p-2)+2]\left[1-\frac{M(p/2,\, t(p-2)/2+2+p/2,\, \|X\|^2/2)}{M(p/2+1,\, t(p-2)/2+2+p/2,\, \|X\|^2/2)}\right]\,.$$

Lemma 1 is used to prove Lemma 2 which says that $\tau(\|X\|^2)$ is non-decreasing in $\|X\|^2$.

LEMMA 1. *Let* $h(y)=\left(\sum_{n=0}^{\infty} d_n y^n\right)\Big/\left(\sum_{n=0}^{\infty} c_n y^n\right)$ *where* d_i, c_i *are nonnegative, and* $\sum_{n=0}^{\infty} c_n y^n$ *and* $\sum_{n=0}^{\infty} d_n y^n$ *converge for all* $y>0$. *If the sequence* $\{d_n/c_n\}$ *is non-decreasing (non-increasing), then* $h(y)$ *is non-decreasing (non-increasing) in* y.

PROOF. See the problem 4 (i) in Lehmann [3], p. 312.

LEMMA 2. $\tau(\|X\|^2)$ *is non-decreasing in* $\|X\|^2$.

PROOF. We show that $M(p/2,\, t(p-2)/2+2+p/2,\, \|X\|^2/2)/M(p/2+1,\, t(p-2)/2+2+p/2,\, \|X\|^2/2)$ is non-increasing in $\|X\|^2$. Put

$$d_n=\left(\frac{p}{2}\right)_n\Big/\left\{n!\left[\frac{t(p-2)}{2}+2+\frac{p}{2}\right]_n\right\}\qquad\text{and}$$

$$c_n=\left(\frac{p}{2}+1\right)_n\Big/\left\{n!\left[\frac{t(p-2)}{2}+2+\frac{p}{2}\right]_n\right\}$$

in Lemma 1. Then we get that

$$(2.19)\qquad d_n/c_n=p/(p+2n)\,,$$

which is non-increasing in n. From Lemma 1 we have the conclusion.

LEMMA 3. $\lim_{\|X\|\to\infty}\tau(\|X\|^2)=t(p-2)+2.$

PROOF. As $|z|\to\infty$ we have (see equation 13.1.4 in [4], p. 504)

$$M(a, b, z)=\frac{\Gamma(b)}{\Gamma(a)}e^z z^{a-b}[1+O(|z|^{-1})] . \tag{2.20}$$

Then we have

$$\tau(\|X\|^2)=[t(p-2)+2][1-p\,O(\|X\|^{-2})] , \tag{2.21}$$

which shows the result.

LEMMA 4. *Let $\delta(X)$ be the estimator of the form*

$$\delta(X)=[1-\tau(\|X\|^2)/\|X\|^2]X , \tag{2.22}$$

where $\tau(\|X\|^2)$ is non-decreasing in $\|X\|^2$ and non-negative. Then $\delta(X)$ is a minimax estimator of θ subject to the loss (1.1) *if and only if* $\lim_{\|X\|\to\infty} \tau(\|X\|^2)\leqq 2(p-2)$.

PROOF. See the lemma in Strawderman [8], p. 385 and the remark in Efron and Morris [1], p. 121.

Then from Lemmas 2, 3 and 4 we have the following theorem.

THEOREM 2. *The estimator given by* (2.9) *is a minimax estimator of θ subject to the loss* (1.1) *if and only if* $t\leqq 2(1-1/(p-2))$.

The following lemma is used to show whether the estimator given by (2.9) is admissible subject to the loss (1.1).

LEMMA 5. *Let $\delta(X)$ be a bounded risk generalized Bayes estimator of the form $\delta(X)=h(\|X\|^2)X$. The following results hold:*

(a) *If there exists an M such that $y>M$ implies $h(y)\leqq 1-(p-2)/y$, then $\delta(X)$ is admissible, and*

(b) *If there exists an M such that $y>M$ implies $h(y)\geqq 1-b/y$ for some $b<p-2$, then $\delta(X)$ is inadmissible.*

PROOF. See Theorem 6.1.1 in Strawderman and Cohen [7], p. 292.

From Lemmas 2, 3 and 5 we have the following theorem.

THEOREM 3. *The estimator given by* (2.9) *is admissible if* $t>1-2/(p-2)$ *and inadmissible if* $t<1-2/(p-2)$.

Remark. If $t=1-2/(p-2)$, the estimator satisfies conditions neither (a) nor (b) in Lemma 5, so we do not know whether it is admissible.

For $t=1$, from Theorems 2 and 3, Bayes estimator is minimax and admissible if $p\geqq 4$. The risks of James-Stein's positive part estimator and Bayes estimator are graphed in Fig. 1, when $p=4$. This graph shows that when true parameter is near to zero, Stein's estimator is superior to Bayes estimator.

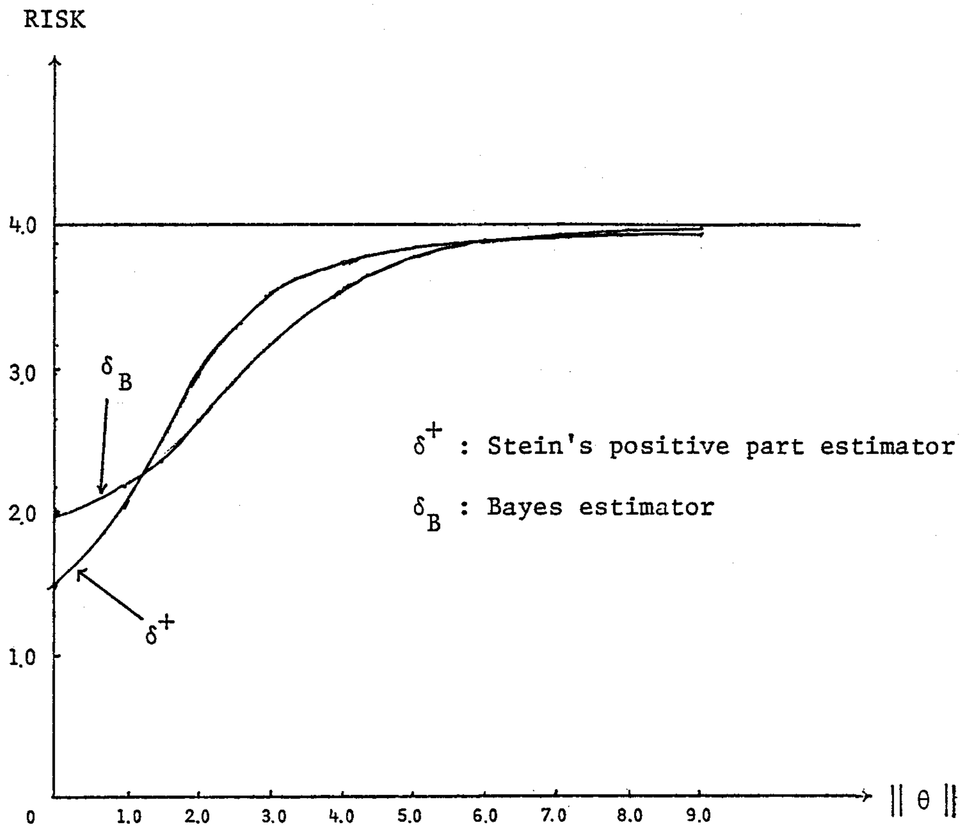

Fig. 1

Acknowledgment

The author wishes to thank Professor N. Sugiura and the referee for their valuable comments and suggestions.

UNIVERSITY OF TSUKUBA

REFERENCES

[1] Efron, B. and Morris, C. (1973). Stein's estimation rules and its competitors—An empirical Bayes approach, *J. Amer. Statist. Ass.*, 68, 117-130.

[2] James, W. and Stein, C. M. (1961). Estimation with quadratic loss function, *Proc. 4th Berkeley Symp. Math. Statist. Prob.*, 1, 361-379.

[3] Lehmann, E. L. (1959). *Testing Statistical Hypotheses*, Wiley, New York.

[4] National Bureau of Standards, Applied Mathematics Series 55 (1964). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.*

[5] Stein, C. M. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, *Proc. 3rd Berkeley Symp. Math. Staitst. Prob.*, 1. 197-206.

[6] Stein, C. M. (1966). An approach to the recovery of inter-block information in balanced incomplete block designs, In Festschrift for J. Neyman: *Research Papers in Statistics*

(ed. F. N. David), Wiley, New York, 351-366.

[7] Strawderman, W. E. and Cohen, A. (1971). Admissibility of estimators of the mean vector of a multivariate normal distribution with quardratic loss, *Ann. Math. Statist.*, **42**, 270-296.

[8] Strawderman, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean, *Ann. Math. Statist.*, **42**, 385-388.