Journal of the Korean Data & Information Science Society
2011, **22**(6), 1251–1256

한국데이터정보과학회지

An approach to improving the Lindley estimator[†]

Tae Ryoung Park[1] · Hoh Yoo Baek[2]

[1]Department of Computer Engineering, Seokyeong University
[2]Division of Mathematics and Informational Statistics, Wonkwang University

Received 29 September 2011, revised 26 October 2011, accepted 6 November 2011

Abstract

Consider a p-variate $(p \geq 4)$ normal distribution with mean θ and identity covariance matrix. Using a simple property of noncentral chi square distribution, the generalized Bayes estimators dominating the Lindley estimator under quadratic loss are given based on the methods of Brown, Brewster and Zidek for estimating a normal variance. This result can be extended the cases where covariance matrix is completely unknown or $\sum = \sigma^2 \boldsymbol{I}$ for an unknown scalar σ^2.

Keywords: Generalized Bayes estimator, Lindley estimator, normal distribution, quadratic loss.

1. Introduction

Let $\boldsymbol{X} = (X_1, \cdots, X_p)'$ be a p-variate random vector normally distributed with unknown mean θ and the identity covariance matrix $\boldsymbol{I}$. Then we consider the problem of estimating θ by $\delta(\boldsymbol{X})$ relative to the quadratic loss function $\|\delta(\boldsymbol{X}) - \theta\|^2 = (\delta(\boldsymbol{X}) - \theta)'(\delta(\boldsymbol{X}) - \theta)$. Every estimator will be evaluated by the risk function $R(\theta, \delta(\boldsymbol{X})) = E\left[\|\delta(\boldsymbol{X}) - \theta\|^2\right]$.

Stein (1956) showed that the usual estimator $\boldsymbol{X}$ is inadmissible for $p \geq 3$ and James and Stein (1961) constructed the improved estimator, $\delta_1^{JS} = \left(1 - (p-2) \;/\|\boldsymbol{X}\|^2\right)\boldsymbol{X}$. Also, Lindley (1962) has proposed the another improved estimator $\delta_1^L = \overline{\mathrm{X}}\mathbf{1} + (1 - (p-3)/\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2)\left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)$ with $\overline{X} = (X_1 + X_2 + \cdots + X_p)/p$ and $\mathbf{1} = (1, \cdots, 1)'$ dominating $\boldsymbol{X}$ for $p \geq 4$. With the similar process of Baranchick (1964), we can construct the positive part estimator $\delta_1^{+L} = \overline{X}\mathbf{1}$ if $\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2 \leq p-3$; $\delta_1^{+L} = \delta_1^L$, otherwise, and we can show that δ_1^{+L} has a smaller risk than δ_1^L by Baranchik's (1964) method. This is known as an estimator eliminating undesirable properties of δ_1^L that it has singularity at $\overline{X}$ and changes the sign of each $X_i - \overline{X}$ for $\|X - \overline{X}\mathbf{1}\|^2 \leq p-3$. However δ_1^{+L} itself is unsatisfactory for θ must be estimated by $\overline{X}\mathbf{1}$ when $\|X - \overline{X}\mathbf{1}\| \leq p-3$. Of course, it is known that δ_1^{+L} is inadmissible. Kim and Baek (2005) developed a sequence of improvement over the δ_1^L with the various cases of covariance matrices.

† This work was supported by Wonkwang University Research Fund 2009.
1 Associate professor, Department of Computer Engineering, Seokyeong University, 124 Seokyeongro, Sungbuk-gu, Seoul 136-704, Korea.
2 Corresponding author: Professor, Division of Mathematics and Informational Statistics, Wonkwang University, Iksan, Jeonbuk 570-749, Korea. E-mail: hybaek@wonkwang.ac.kr

In this paper we propose a generalized Bayes estimator dominating δ_1^L based on the ideas used in Brown (1968), Brewster and Zidek (1974), and Kubokawa (1991) for estimating a normal variance. In Section 2, such a smooth estimator is derived and it is shown to be admissible. It should be noted that this admissible estimator dominating δ_1^L is just identical to the generalized Bayes estimator given by Strawderman (1971), Lindley (1962), and Berger (1976) with $a\,(=c) = 2$. Section 3 discusses the cases where the covariance matrix $\sum$ of $\boldsymbol{X}$ is fully unknown or $\sum = \sigma^2 \boldsymbol{I}$ for an unknown scalar σ^2.

2. Admissible estimator dominating δ_1^L

To improve on δ_1^L, we consider the estimator

$$\delta_1(c,r) = \begin{cases} \overline{X}\mathbf{1} + (1 - c/\|\,\boldsymbol{X} - \overline{X}\mathbf{1}\|^2)(\boldsymbol{X} - \overline{X}\mathbf{1}), & \text{if } \|\,\boldsymbol{X} - \overline{X}\mathbf{1}\|^2 \le r \\ \delta_1^L, & \text{otherwise,} \end{cases} \tag{2.1}$$

where c and r are positive constants. For a fixed r, we shall find the best $c = c(r)$ in the sense of minimizing the risk. Such an idea is due to Brown (1968) which constructed an improved estimator for a normal variance. Let $\lambda = ||\theta - \bar{\theta}\mathbf{1}||/2$ and $f_{p-1}(t;\lambda)$ denote the density of a noncentral chi square random variable with the degrees of freedom $p-1$ and the noncentrality λ. Letting

$$c_1(r,\lambda) = p - 3 - 2 f_{p-1}(r;\lambda) \Big/ \int_0^r t^{-1} f_{p-1}(t;\lambda)\, dt, \tag{2.2}$$

we can obtain the following lemma which will be proved later.

Lemma 2.1. (i) The risk function of $\delta_1(c_2, r)$ is quadratic with respect to c and is minimized at $c = c_1(r,\lambda)$.

(ii) $c_1(r,\lambda) \le c_1(r;0) = c_1(r)$, where $c_1(r)$ is expressed as

$$c_1(r) = p - 3 - 2\left[\int_0^1 t^{(p-1)/2-2} \exp\left\{\frac{1}{2}(1-t)\,r\right\} dt\right]^{-1}.$$

(iii) $c_1(r)$ is increasing in r and $0 < c_1(r) < p - 3$.

Lemma 2.1. implies that for all λ, $c_1(r)$ is closer to minimizing value of the risk $R(\theta, \delta_1(c,r))$ than $p-3$, so that we obtain the following theorem.

Theorem 2.1. The estimator $\delta_1(c_1(r), r)$ dominates $\delta_1(p-3, r)$ or δ_1^L.

Further select $0 < r' < r$. By the property (iii) of Lemma 2.1. and a similar manner, it can be seen that $\delta_1(c_1(r), r)$ is dominated by another estimator of the form

$$\delta_1'(c_1, r', r) = \begin{cases} \overline{X}\mathbf{1} + (1 - c_1(r')/\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2)(\boldsymbol{X} - \overline{X}\mathbf{1}), & \text{if } \|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2 \le r' \\ \overline{X}\mathbf{1} + (1 - c_1(r)/\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2)(\boldsymbol{X} - \overline{X}\mathbf{1}), & \text{if } r' < \|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2 \le r \\ \delta_1^L, & \text{otherwise.} \end{cases} \tag{2.3}$$

Now from the innovative idea of Brewster and Zidek (1974), we select a finite partition of $[0,\infty)$ represented by $0 = r_{i,0} < \cdots < r_{i,n_i - 1} < r_{i,n_i} = \infty$ for each $i = 1, 2 \cdots$, and a corresponding estimator

$$\delta_1^{(i)} = \overline{X}\mathbf{1} + (1 - c_1(r_{ij})/\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2)(\boldsymbol{X} - \overline{X}\mathbf{1}) \text{ if } r_{i,j-1} < \|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2 \le r_{ij}.$$

Then, providing $\max_j |r_{i,j} - r_{i,j-1}| \to 0$ and $r_{i,n_{i-1}} \to \infty$ as $i \to \infty$, the sequence $\delta_1^{(i)}$ will converge pointwise to δ_1^*, where

$$\delta_1^* = \overline{X}\mathbf{1} + (1 - c_1(\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2)/\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2)(\boldsymbol{X} - \overline{X}\mathbf{1}). \tag{2.4}$$

It should be noted that δ_1^* is the generalized Bayes estimator given by Strawderman (1971), Berger (1976) with $a(= c) = 2$, and Lindley's (1962) method against the prior density

$$\pi^*(\theta) = \int_0^1 (2\pi)^{-\frac{p-1}{2}} \lambda^{-2} (\lambda/(1-\lambda))^{\frac{p-1}{2}} \exp\left\{ -\frac{1}{2}(\lambda/(1-\lambda))\|\theta - \bar{\theta}\mathbf{1}\|^2 \right\} d\lambda.$$

Theorem 2.2. The estimator δ_1^* is an admissible estimator dominating δ_1^L.

Proof. Since $\delta_1^{(i)}$ has uniformly smaller risk than δ_1^L for each i, applying Fatou's lemma gives that δ_1^* dominates δ_1^L. The admissibility follows from the result of Brown and Hwang (1982) for the prior density $\pi^*(\theta)$ which satisfies the conditions of (ii) in page 213 of their paper. Hence we get the desired conclusion.

Proof of Lemma 2.1. Let $W = \|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2$ and $I(\cdot)$ denote the indicator function. Then for a fixed r, the risk function of $\delta_1(c, r)$ is minimized at

$$\begin{aligned} c &= \frac{E[\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^{-2}(\boldsymbol{X} - \overline{X}\mathbf{1})'(\boldsymbol{X} - \theta)I(\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2 \le r)]}{E[\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^{-2}I(\|\boldsymbol{X} - \overline{X}\mathbf{1}\||^2 \le r)]} \\ &= E\left[(1 - \frac{\theta'(\boldsymbol{X} - \overline{X}\mathbf{1})}{W})I(W \le r)\right] / E[\frac{1}{W}I(W \le r)] = c_1^*, \text{ say,} \end{aligned} \tag{2.5}$$

so that we shall demonstrate that c_1^* given by (2.5) is expressed as $c_1(r, \lambda)$ given in (2.2). Using the similar calculation by Kim *et al.* (1995) and Bock (1975) c_1^* can be represented as

$$c_1^* = \frac{E^J\left[I_r(p-1+2J) - \dfrac{2J}{p-3+2J}I_r(p-3+2J)\right]}{E^J\left[\dfrac{1}{p-3+2J}I_r(p-3+2J)\right]}, \tag{2.6}$$

where J is a random variable having a Poisson distribution with mean λ and $I_r(X) = \int_0^r f_\alpha(x)dx$ for a central chi square density $f_\alpha(x)$ with degrees of freedom α. Since $I_r(\alpha+2) = -2f_{\alpha+2}(r) + I_r(\alpha)$, we observe that $c_1^* = p-3-2E^J[f_{p-1+2J}(r)]/E^J[(p-3+2J)^{-1}I_r(p-3+2J)]$, which can be rewritten as $c_1(r; \lambda)$ given by (2.2), and we obtain part(i). For part(ii), It is sufficient to show that

$$f_{p-1}(r, \lambda)/\int_0^r t^{-1}f_{p-1}(t; \lambda)dt \ge f_{p-1}(r)/\int_0^r t^{-1}f_{p-1}(t)dt, \tag{2.7}$$

which follows from the fact that $f_{p-1}(t; \lambda)/f_{p-1}(t)$ is increasing in t. Part(iii) can be easily checked and Lemma 2.1 is proved.

3. The cases of unknown covariance matrices

In this section we extend the result derived in section 2 to the case where covariance matrix is completely unknown or $\Sigma = \sigma^2 I$ for an unknown scalar σ^2. At first, the case of $\Sigma = \sigma^2 I$ is treated.

Let $\boldsymbol{X}$ and S be a independent random variables with $\boldsymbol{X} \sim N_p\left(\theta, \sigma^2 I\right)$ and $S \sim \sigma^2\ \chi^2_n$. Here we want to estimate θ under the loss $\|\widehat{\theta} - \theta\mathbf{1}\|^2/\sigma^2$. For positive constants c and r, a corresponding estimator to (2.1) is of the form

$$\delta_2(c,r) = \begin{cases} \overline{X}\mathbf{1} + \left(1 - cS/\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2\right)\left(\boldsymbol{X} - \overline{X}\mathbf{1}\right), & \text{if } \dfrac{\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2}{S} \le r \\ \delta_2^L, & \text{otherwise,} \end{cases} \tag{3.1}$$

where, in this case, the Lindley estimator is given by

$$\delta_2^L = \overline{X}\mathbf{1}\left\{\left(1 - \frac{p-3}{n+2}\right) S/\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2\right\}\left(\boldsymbol{X} - \overline{X}\mathbf{1}\right).$$

Define $c_2(r)$ by

$$c_2(r) = \frac{p-3}{n+2} - \frac{2}{n+2}\left[\int_0^1 \frac{(1+r)^{(p+n-1)/2}}{(1+rz)^{(p+n-1)/2}} z^{\frac{p-1}{2}-2} dz\right]^{-1}. \tag{3.2}$$

Theorem 3.1. The estimator $\delta_2(c_2(r), r)$ dominates δ_2^L.

Proof. Let $\lambda = \|\theta - \bar{\theta}\mathbf{1}\|^2/\left(2\sigma^2\right)$. Note that the risk function of $\delta_2(c,r)$ is minimized at

$$c_2(r;\lambda) = \frac{\mathrm{E}\left[\left((\mathrm{S}/\sigma^2\right)\left\{1 - \left(\boldsymbol{X} - \overline{\mathrm{X}}\mathbf{1}\right)'\theta\right\}/\|\boldsymbol{X} - \overline{\mathrm{X}}\mathbf{1}\|^2) I\left(\|\boldsymbol{X} - \overline{\mathrm{X}}\mathbf{1}\|^2/\mathrm{S} \le \mathrm{r}\right)\right]}{\mathrm{E}\left[\left(\mathrm{S}/\sigma^2\right)^2\left(\sigma^2/\|\boldsymbol{X} - \overline{\mathrm{X}}\mathbf{1}\|^2\right) I\left(\|\boldsymbol{X} - \overline{\mathrm{X}}\mathbf{1}\|^2/\mathrm{S} \le \mathrm{r}\right)\right]},$$

which from (2.6), can be expressed by

$$\begin{aligned} c_2(r;\lambda) &= \frac{E^J\left[\int_0^\infty v^{\frac{n}{2}} e^{-\frac{v}{2}}\left\{I_{rv}(p-3+2J) - \dfrac{2J}{p-3+2J} I_{rv}(p-3+2J)dv\right\}\right]}{E^J\left[\int_0^\infty v^{\frac{n}{2}} e^{-\frac{v}{2}} \dfrac{1}{p-3+2J} I_{rv}(p-3+2J)dv\right]} \\ &= \frac{\int_0^\infty v^{\frac{n}{2}} e^{-\frac{v}{2}}\left\{(p-3)\int_0^{rv} w^{-1} f_{p-1}(w;\lambda)dw - 2f_{p-1}(rv;\lambda)\right\}dv}{\int_0^\infty v^{\frac{n}{2}+1} e^{-\frac{v}{2}} \int_0^{rv} w^{-1} f_{p-1}(w;\lambda)\,dwdv} \end{aligned}$$

By integration by parts,

$$\begin{aligned} &\int_0^\infty e^{-\dfrac{v}{2}}\left\{v^{\frac{n}{2}+1}\int_0^{rv} w^{-1} f_{p-1}(w;\lambda)\,dw\right\}dv \\ &\quad = (n+2)\int_0^\infty v^{\frac{n}{2}} e^{-\frac{v}{2}} \int_0^{rv} w^{-1} f_{p-1}(w;\lambda)\,dwdv + 2\int_0^\infty v^{\frac{n}{2}} e^{-\frac{v}{2}} f_{p-1}(rv;\lambda)\,dv, \end{aligned}$$

so that

$$c_2(r;\lambda) = (p-3-2H(\lambda)) / (n+2+2H(\lambda)), \tag{3.3}$$

where

$$H(\lambda) = \frac{\int_0^\infty v^{\frac{n}{2}} e^{-\frac{v}{2}} f_{p-1}(rv;\lambda)\,dv}{\int_0^\infty v^{\frac{n}{2}} e^{-\frac{v}{2}} \int_0^{rv} w^{-1} f_{p-1}(w;\lambda)\,dwdv}.$$

Let $A(\alpha) = 2^{-\frac{\alpha}{2}} \left(\Gamma\left(\frac{\alpha}{2}\right)\right)^{-1}$ and let

$$g_{p,n}(z,\lambda) = E^J \left[\frac{A(p+2J-1)}{A(n+p+2J-1)} z^{(p+2J-1)/2-1} (1+z)^{-(n+p+2J-1)/2} \right].$$

Then $H(\lambda)$ can be rewritten as $H(\lambda) = g_{p,n}(r;\lambda) / \int_0^r z^{-1} g_{p,n}(z,\lambda)\,dz$. Similar to (2.7), we can show that $H(\lambda) \geq H(0)$, so that from (3.3), $c_2(r;\lambda) \leq c_2(r;0)$. Here we can verify that $c_2(r;0)$ is equal to $c_2(r)$ given by (3.2), and that $c_2(r)$ is increasing in r and $0 < c_2(r) < (p-3)/(n+2)$. Therefore the proof of Theorem 3.1 is completed.

As a limiting form corresponding to (2.4), we can take the estimator

$$\delta_2^* = \overline{X}\mathbf{1} + \left\{1 - c_2\left(\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2/S\right) S/\|\boldsymbol{X} - \overline{X}\mathbf{1}\|^2\right\} \left(\boldsymbol{X} - \overline{X}\mathbf{1}\right), \tag{3.4}$$

which is the generalized Bayes estimator modified from Kubokwa (1991) and Lin and Tsai (1973). By the same arguments as in Section 2, we can prove the following theorem.

Theorem 3.2. The estimator δ_2^* is the generalized Bayes estimator dominating δ_2^L.

For the case where Σ is fully unknown, the above discussions are directly applied. Let X and S be independent random variables with $\boldsymbol{X} \sim N_p(\theta, \boldsymbol{\Sigma})$ and $S \sim W_p(n, \boldsymbol{\Sigma})$. Assume that we want to estimate θ under the loss $\left(\widehat{\theta} - \theta\right)' \boldsymbol{\Sigma^{-1}} \left(\widehat{\theta} - \theta\right)$. Define $c_3(r)$ by

$$c_3(r) = \frac{p-3}{n-p+2} - \frac{2}{n-p+2} \left[\int_0^1 \frac{(1+r)^{\frac{n+1}{2}}}{(1+rt)^{\frac{n+1}{2}+1}} t^{\frac{p-1}{2}-2} dt \right]^{-1}.$$

The estimator $\delta_3^* = \overline{X}\mathbf{1} + \left[1 - \frac{c_3\left\{\left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)' S^{-1} \left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)\right\}}{\left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)' S^{-1} \left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)}\right] \left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)$ is the generalized Bayes estimator modified from Kubokawa (1991) and Lin and Tsai (1973). Note that $\left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)' \Sigma^{-1} \left(\boldsymbol{X} - \overline{X}\mathbf{1}\right) / \left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)' S^{-1} \left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)$ is distributed as χ^2_{n-p+2} independent of X. Then from Theorem 3.2, it is seen the δ_3^* dominates the Lindley estimator

$$\delta_3^L = \overline{X}\mathbf{1} + \left[1 - \frac{p-3}{n-p+4} \left\{\left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)' S^{-1} \left(\boldsymbol{X} - \overline{X}\mathbf{1}\right)\right\}^{-1}\right] \left(\boldsymbol{X} - \overline{X}\mathbf{1}\right).$$

References

Baranchick, A. (1964). *Multiple regression and estimation of the mean of a multivariate normal distribution,* Technical Report. **51,** Department of Statistics, Stanford University.

Berger, J. O. (1976). Admissible minimax estimation of a multivariate normal mean with arbitrary quadratic loss. *Annals of Statistics*, **4,** 223-226.

Bock, M. E. (1975). Minimax estimation of the mean of a multivariate normal distribution. *Annals of Statistics*, **3,** 209-218.

Brewster, J. F. and Zidek, J. V. (1974). Improving on equivariant estimators, *Annals of Statistics,* **2,** 21-38.

Brown, L. D. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Annals of Mathematical Statistics,* **39,** 29-48.

Brown, L. D. and Hwang, J. T. (1982), A unified admissibility proof. In *Statistical Decision Theory and Related Topics III,* edited by S. S. Gupta and J. Berger, Academic Press, New York.

James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability,* **1,** 361-379.

Kim, B. H. and Baek, H. Y. (2005). A sequence of improvement over the Lindley type estimator with the cases of unknown covariance matrices. *The Korean Communications in Statistics,* **12**, 463-472.

Kim, B. H., Koh, T. W. and Baek, H. Y. (1995). Estimator with nondecreasing risk in a multivariate normal distribution. *Journal of the Korean Statistical Society,* **24**, 257-266.

Kubokawa, T. (1991). An approach to improving the James-Stein estimator. *Journal of Multivariate Analysis,* **36**, 121-126.

Lin, P. E. and Tsai, H. L. (1973). Generalized Bayes minimax estimators of the multivariate normal mean with unknown covariance matrix. *Annals of Statistics,* **1,** 142-145.

Lindley, D. V. (1962). Discussion of paper by C. Stein. *Journal of Royal Statistical Society,* **24,** 265-296.

Strawderman, W. E. (1971). Proper Bayes minimax estimator for the mean of multivariate normal mean. *Annals of Mathematical Statistics,* **42,** 385-388.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability,* **1,** 197-206.