J. Chosun Natural Sci.
Vol. 6, No. 4 (2013) pp. 251-255
http://dx.doi.org/10.13160/ricns.2013.6.4.251

Improvement of the Modified James-Stein Estimator with Shrinkage Point and Constraints on the Norm

Jae Hyun Kim[1†] and Hoh Yoo Baek[2†]

Abstract

For the mean vector of a p-variate normal distribution $(p \geq 4)$, the optimal estimation within the class of modified James-Stein type decision rules under the quadratic loss is given when the underlying distribution is that of a variance mixture of normals and when the norm $\|\theta - \bar{\theta}1\|$ it known.

Key words : Modified James-Stein Type Decision Rule, Mean Vector, Quadratic Loss, Underlying Distribution

1. Introduction

The problem considered that of estimating with quadratic loss function the mean vector of a compound multinormal distribution when the norm $\|\theta - \bar{\theta}1\|$ is known. The class of estimation rules will consist of modified James-Stein type estimators only. Such a class was introduced by James-Stein(1961) and Lindley(1962) in order to prove that some of members dominate the sample mean in the multinormal case.

Strawderman(1974) also derived a similar result for the more general case considered in this paper of a compound multinormal distribution.

The problem of estimation of a mean vector under constraint has an old origin and recently focussed again in the context of curved model in the works of Efron(1978), Hinckley(1977), Amari(1982), Kariya(1989), Perron and Giri(1989), Marchand and Giri(1993), Park and Baek(2011) among others. A study of compound multinormal distributions and the estimation of their location vectors was carried out by Berger(1976) and Kubokawa(1991).

In section 2, the general setting of this problem and necessary notations are presented. In section 3, the best modified James-Stein type estimator of a mean vector in the viewpoint of Branchick's method in derived when it's norm is known and an example is given. Concluding remarks are presented in section 4.

2. Notation and Preliminaries

Let $x = (x_1, \cdots, x_p)'$, $p \geq 4$, be an observation from a compound multinormal distribution with unknown location parameter $\theta(p \times 1)$ and mixture parameter $H(\cdot)$, where $H(\cdot)$ represents a known c.d.f. defined on the interval $(0, \infty)$.

In other words, it is assumed that the random variable X generating the observation x admits the representation,

$$\mathcal{L}(X \mid Z = z) = N_p(\theta, zI_p), \forall z > 0 \tag{2.1}$$

Z being the positive random variable with c.d.f. $H(\cdot)$.

Then we consider the problem of estimating θ by $\delta(X)$ relative to the quadratic loss function

$$L(\theta, \delta(x)) = (\delta_1(x) - \theta)^2 + \cdots + (\delta_p(x) - \theta)^2$$
$$= (\delta(x) - \theta)'(\delta(x) - \theta) = \| \delta(x) - \theta \|^2$$

with $\theta \in \Theta_\lambda = \{\theta \in R^p \mid \|\theta - \bar{\theta}1\| = \lambda,\ 0 \leq \lambda < \infty\}$,

where $\bar{\theta} = \frac{1}{p}\sum_{i=1}^{p}\theta_i, 1 = (1, \cdots, 1)'$ and the decision rule δ^c, $\delta^c(\cdot) : R^p \rightarrow R^p$, is of the form

$$\delta^c(X) = \bar{X}1 + \left(1 - \frac{c}{\| X - \bar{X}1 \|^2}\right)(X - \bar{X}1), \quad c \in R$$

[1]Department of Computer Engineering, Seokyeong University, Seoul 136-704, Korea
[2]Division of Mathematics and Informational Statistics, Wonkwang University, Iksan, Jeonbuk 570-749, Korea
†Corresponding author : statsr@skuniv.ac.kr, hybaek@wonkwang.ac.kr
(Received : November 28, 2013, Revised : December 13, 2013, Accepted : December 20, 2013)

Let another decision rule δ^{c_1}, is of the truncated form

$$\delta^{c_1}(X) = \begin{cases} \overline{X}1 + \left(1 - \dfrac{c_1}{\| X - \overline{X}1 \|^2}\right)(X - \overline{X}1), & \text{if } \|X - \overline{X}1\| \le r, \\ \delta^{c}(X), & otherwise, \end{cases}$$

where c_1 and r are positive constants. For a fixed r, it will be found that $c_1 = c_1(r)$ in the sense of which constructed an improved estimator for a normal variance. Restated in terms of the family of probability functions of X, the distributional assumption given by expression (2.1) and the restriction on the location parameter θ indicate that the p.d.f. of X is

$$p_\theta(x) = \int_0^\infty (2\pi z)^{-p/2} \exp\left(\frac{-\|x - \theta\|^2}{2z}\right) dH(z) \quad (2.2)$$

$x \in R^p$ and $\theta \in \Theta_\lambda$. It will be also assumed that $E(Z) < \infty$ which will guarantee the existence of the covariance matrix $\sum = Cov(X) = E(Z) I_p$ and the mean vector $E(X) = \theta$. The performance of the estimator δ_1 will be measured by its risk function

$$R(\theta, \delta^{c_1}) = E_\theta[L(\theta, \delta^{c_1}(X)] = E_\theta[(\delta^{c_1}(X) - \theta)'(\delta^{c_1}(X) - \theta)], \ \theta \in \Theta_\lambda.$$

3. Optimal Estimation with Truncated Type When the Norm $\|\theta - \overline{\theta}1\|$ is Known.

In this section, the best estimation is derived within $\mathscr{D}_1 = \{ \delta : R^p \to R^p | \delta(X) = \delta^{c_1}(X) \}$, where

$$\delta^{c_1}(X) = \begin{cases} \overline{X}1 + \left(1 - \dfrac{c_1}{\| X - \overline{X}1 \|^2}\right)(X - \overline{X}1), & \text{if } \| X - \overline{X}1 \|^2 \le r, \\ \delta^{c}(X), & otherwise \end{cases}$$

where c and c_1 are in R and the parameter space is of the form

$$\Theta_\lambda = \left\{ \Theta_\lambda = \theta \in R^p \mid \|\theta - \overline{\theta}1\| = \lambda \right\}, \ \lambda \ge 0.$$

If $c_1 = c$, the class $\mathscr{D}_1$ is coincided with the type of Baek(2000) and Baek and Lee(2005) as a special case.

The following lemmas will prove useful in the evaluation of the risk function of the decision rule δ^{c_1}, $c_1 \in R$.

Lemma 3.1. Let X be a random multinormal vector $N_p(\theta, I_p)$, $p \ge 4$ and $\theta \in R^p$ and $I(\cdot)$ denote the indicator function. Then

(i) $$E_\theta\left[\frac{1}{(X - \overline{X}1)'(X - \overline{X}1)} I(\|X - \overline{X}1\|^2 \le r)\right] = E^L\left[\frac{1}{p + 2L - 3} I_r(p + 2L - 3)\right]$$

(ii) $$E_\theta\left[\frac{(X - \theta)'(X - \overline{X}1)}{(X - \overline{X}1)'(X - \overline{X}1)} I(\|X - \overline{X}1\|^2 \le r)\right] = E^L\left[\frac{p - 3}{p + 2L - 3} I_r(p + 2L - 3)\right],$$

where L is Poisson random variable with mean $\dfrac{\lambda^2}{2}$ and $I_r(\alpha) = \int_0^r g_\alpha(x)dx$ for a central chi square density $g_\alpha(x)$ with degrees of freedom α.

Proof. Using the similar calculation by Bock(1975) and Kim et al.(1995), this lemma can be proved.

Lemma 3.2. Let X be a compound multinormal vector with location parameter θ, $p \ge 4$ and $\theta \in R^p$, and known mixture parameter $H(\cdot)$ with p.d.f. of the form given in (2.2). Then with $\|\theta - \overline{\theta}1\| = \lambda$,

(i) $$E_\theta\left[\frac{1}{(X - \overline{X}1)'(X - \overline{X}1)} I(\|X - \overline{X}1\|^2 \le r)\right] = \int_0^\infty E^L\left[\frac{1}{p + 2L - 3} I_{\frac{r}{z}}(p + 2L - 3)\right] \frac{dH(z)}{z}$$

(ii) $$E_\theta\left[\frac{(X - \theta)'(X - \overline{X}1)}{(X - \overline{X}1)'(X - \overline{X}1)} I(\|X - \overline{X}1\|^2 \le r)\right] = \int_0^\infty E^L[I_{\frac{r}{z}}(p - 1 + 2L) - \frac{2L}{p + 2L - 3} I_{\frac{r}{z}}(p + 2L - 3)] dH(z)$$

where L is Poisson random variable with mean $\lambda^2/2z$ and $I_{\frac{r}{z}}(\alpha) = \int_0^{\frac{r}{z}} g_\alpha(x)dx$ for a central chi square density $g_\alpha(x)$ with degrees of freedom α.

Proof. (i) Using both the representation given in (2.1) and part (i) of Lemma 3.1, the following expression can be obtained;

$$E_{\boldsymbol\theta}\left[\frac{1}{(\boldsymbol X-\overline X\mathbf 1)'(\boldsymbol X-\overline X\mathbf 1)}I(\|\boldsymbol X-\overline X\mathbf 1\|^2\le r)\right]$$
$$=E^Z\left\{Z^{-1}E_{\boldsymbol\theta}^{X|Z}\left[\frac{Z}{(\boldsymbol X-\overline X\mathbf 1)'(\boldsymbol X-\overline X\mathbf 1)}I\left(\frac{\|\boldsymbol X-\overline X\mathbf 1\|^2}{Z}\le\frac{r}{Z}\right)\right]\right\}$$
$$=\int_0^\infty\sum_{l=0}^\infty\frac{e^{-\frac{\lambda^2}{2z}}\left(\frac{\lambda^2}{2z}\right)^l}{l!(p+2l-3)}\left(\int_0^{\frac{r}{z}}\frac{v^{\frac{p+2l-3}{2}-1}e^{-\frac{v}{2}}}{\Gamma\left(\frac{p+2l-3}{2}\right)2^{\frac{p+2l-3}{2}}}dv\frac{dH(z)}{z}\right)$$
$$=\int_0^\infty E^L\left[\frac{1}{p+2L-3}I_{\frac{r}{z}}(p+2L-3)\right]\frac{dH(z)}{z}$$

(ii) Again, combining the representation given in (2.1) and part (ii) of Lemma 3.2, the following expression can be obtained;

$$E_{\boldsymbol\theta}\left[\frac{(\boldsymbol X-\boldsymbol\theta)'(\boldsymbol X-\overline X\mathbf 1)}{(\boldsymbol X-\overline X\mathbf 1)'(\boldsymbol X-\overline X\mathbf 1)}I(\|\boldsymbol X-\overline X\mathbf 1\|^2\le r)\right]$$
$$=E^Z\left\{E_{\boldsymbol\theta}^{X|Z}\left[\frac{\left(\frac{\boldsymbol X-\boldsymbol\theta}{\sqrt Z}\right)'\left(\frac{\boldsymbol X-\overline X\mathbf 1}{\sqrt Z}\right)}{\left(\frac{\boldsymbol X-\overline X\mathbf 1}{\sqrt Z}\right)'\left(\frac{\boldsymbol X-\overline X\mathbf 1}{\sqrt Z}\right)}I\left(\frac{\|\boldsymbol X-\overline X\mathbf 1\|^2}{Z}\le\frac{r}{Z}\right)\right]\right\}$$
$$=E^Z\left\{E^{X|Z}\left[\left(1-\frac{\left(\frac{\boldsymbol\theta}{\sqrt Z}\right)'\left(\frac{\boldsymbol X-\overline X\mathbf 1}{\sqrt Z}\right)}{\left(\frac{\boldsymbol X-\overline X\mathbf 1}{\sqrt Z}\right)'\left(\frac{\boldsymbol X-\overline X\mathbf 1}{\sqrt Z}\right)}\right)I\left(\frac{\|\boldsymbol X-\overline X\mathbf 1\|^2}{Z}\le\frac{r}{Z}\right)\right]\right\}$$
$$=\int_0^\infty E^L\left[I_{\frac{r}{z}}(p-1+2L)-\frac{2L}{p+2L-3}I_{\frac{r}{z}}(p+2L-3)\right]dH(z)$$

Assume that the function $f_p(\cdot,\cdot):[0,\infty)\to(0,\infty)$ is defined by the relation $f_p(\lambda,z)$

$$=\sum_{l=0}^\infty\frac{e^{-\frac{\lambda^2}{2z}}\left(\frac{\lambda^2}{2z}\right)^l}{l!(p+2l-3)}\int_0^{\frac{r}{z}}\frac{v^{\frac{p+2l-3}{2}}e^{-\frac{v}{2}}}{\Gamma\left(\frac{p+2l-3}{2}\right)2^{\frac{p+2l-3}{2}}}dv.$$

Then the main result of this section now follows.

Theorem 3.3. Let x be a single observation from a p-dimensional location parameter with p.d.f. of the form given by (2.2). Under the assumptions $\theta\in\Theta_\lambda,\ p\ge4$ and $E[Z]<\infty$, the unique best estimator within the class $\mathscr{D}_1$ is given by $\delta^{c_1^*}$ where

$$c_1^*=(p-3)\frac{\int_0^\infty f_p(\lambda,z)dH(z)}{\int_0^\infty f_p(\lambda,z)\frac{dH(z)}{z}}-2\frac{\int_0^\infty E^L\left(g_{p+2L-1}\left(\frac{r}{z}\right)\right)dH(z)}{\int_0^\infty f_p(\lambda,z)\frac{dH(z)}{z}}\tag{3.1}$$

Proof. It can be shown that for a fixed r, the risk function of $\delta_1(\boldsymbol X)$ is minimized at

$$c_1=c_1^*=\frac{E\left[\frac{(\boldsymbol X-\boldsymbol\theta)'(\boldsymbol X-\overline X\mathbf 1)}{(\boldsymbol X-\overline X\mathbf 1)'(\boldsymbol X-\overline X\mathbf 1)}I(\|\boldsymbol X-\overline X\mathbf 1\|^2\le r)\right]}{E\left[\frac{1}{(\boldsymbol X-\overline X\mathbf 1)'(\boldsymbol X-\overline X\mathbf 1)}I(\|\boldsymbol X-\overline X\mathbf 1\|^2\le r)\right]}$$

By Lemma 3.2

$$c_1^*=\frac{\int_0^\infty E^L\left[I_{\frac{r}{z}}(p-1+2L)-\frac{2L}{p+2L-3}I_{\frac{r}{z}}(p+2L-3)\right]dH(z)}{\int_0^\infty E^L\left[\frac{1}{p+2L-3}I_{\frac{r}{z}}(p+2L-3)\right]\frac{dH(z)}{z}}$$

Since $I_r(\alpha+2)=-2g_{\alpha+2}(r)+I_r(\alpha)$, this theorem can be proved.

Example 3.1. For $\boldsymbol X\sim N_p(\boldsymbol\theta,\sigma^2 I_p)$, $p\ge4$, ($i.e.$ $H(z)=1_{(\sigma^2,\infty)}(z)$ with $1_A(\cdot)$ being the indicator function of the set A); it can be deduced from Theorem 3.1 that

$$c_1^*=(p-3)\sigma^2-\frac{2E^L\left[g_{p+2L-1)}\left(\frac{r}{\sigma^2}\right)\right]}{f_p(\lambda,\sigma^2)/\sigma^2}$$
$$=\sigma^2\left\{(p-3)-\frac{2E^L\left[g_{p+2L-1}\left(\frac{r}{\sigma^2}\right)\right]}{f_p(\lambda,\sigma^2)}\right\}$$

we can show that and the best estimator $\delta^{(p-3)\sigma^2}(X) = \left(1 - \frac{(p-3)\sigma^2}{\|X - \overline{X}1\|^2}\right)(X - \overline{X}1)$ within $\mathscr{D}$ (See Merchand and Giri(1993)) is improved by $\delta_1^{c_1^*}(X)$ in $\mathscr{D}_1$, which is shown in Kubokawa(1991).

4. Conclusions

The main result of this paper can be extended to the case where the mean θ in restricted to a known interval $[\lambda_1, \lambda_2]$ case;

$$\Theta_{\lambda_2}^{\lambda_1} = \left\{\theta \in R^p \mid \|\theta - \bar{\theta}1\| \in [\lambda_1, \lambda_2], 0 \le \lambda_1 \le \lambda_2 < \infty\right\}.$$

Also, It can be considered another estimator

$$\delta_1'(c_1, r', r) = \begin{cases} \overline{X}1 + \left(1 - \frac{c_1(r')}{\| X - \overline{X}1 \|^2}\right)(X - \overline{X}1), & \text{if } \| X - \overline{X}1 \|^2 \le r' \\ \overline{X}1 + \left(1 - \frac{c_1(r)}{\| X - \overline{X}1 \|^2}\right)(X - \overline{X}1), & r' \le \| X - \overline{X}1 \|^2 \le r \\ \delta^c, & otherwise \end{cases}$$

We can show that $\delta_1'(c_1, r', r)$ dominate δ^{c_1} by Kubokawa(1991) and Park and Baek(2011). Futhermore, it is possible to be derived the convergent form, which is left a further research.

Acknowledgements

This study was supported by Wonkwang University Research Fund 2013.

References

[1] S. Amari, "Differential geometry of curved exponential families, curvature and information loss", Ann. Stat., Vol. 10, pp. 357- 385, 1982.

[2] Baranchick, A. "Multiple regression and estimation of the mean of a multivariate normal distribution", Technical Report, 51, Department of Statistics, Stanford university, 1984.

[3] H. Y. Baek, "Lindley type estimators with the known norm", Journal of the Korean Data & Information Science Society, Vol. 11, pp. 37-45, 2000.

[4] H. Y. Baek and J. M. Lee, "Lindley type estimators when the norm is restricted to an interval", Journal of the Korean Data & Information Science Society, Vol. 16, pp. 1027-1039, 2005.

[5] J. 0. Berger, "Admissible minimax estimation of a multivariate normal mean with arbitrary quadratic loss", Ann. Stat., Vol. 4, pp. 223-226, 1976.

[6] M. E. Bock, "Minimax estimation of the mean of a multivariate normal distribution", Ann. Stat., Vol. 3, No. 1, pp. 209-218, 1975.

[7] L. D. Brown, "Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters", Annals of Mathematical Statistics, Vol. 39, pp. 29-48, 1968.

[8] B. Efron, "The geometry of exponential families", Ann. Stat., Vol. 6, pp. 362-376, 1978.

[9] D. V. Hinkley, "Conditional inference about a normal mean with known coefficient of variation", Biometrika, Vol. 64, pp. 105-108, 1977.

[10] W. James and C. Stein, "Estimation with quadratic loss", In proceeding Fourth Berkeley Symp. Math. Statis. Probability, 1, University of California Press, Berkeley, pp. 361-380, 1961.

[11] T. Kariya, "Equivariant estimation in a model with ancillary statistics", Ann. Stat., Vol. 17, pp. 920-928, 1989.

[12] B. H. Kim, T. W. Koh, and H. Y. Baek, "Estimators with nondecreasing risk in a multivariate normal distribution", J. Korean Stat. Soc., Vol. 24, pp. 257-266, 1995.

[13] B. H. Kim and H. Y. Baek "A sequence of improvement over the Lindley type estimator with the cases of unknown covariance matrices", Korean Communications in Statistics, Vol. 12, pp. 463-472, 2005.

[14] T. Kubokawa, "An approach to improving the James-Stein Estimator", J. Multivar. Anal., Vol. 36, pp. 121-126, 1991.

[15] D. V. Lindley, "Discussion of paper by C. Stein", J. Roy. Statist. Soc. Ser. B, Vol. 24, pp. 265-296, 1962.

[16] E. Marchand and N. C. Giri, "James- Stein estimation with constraints on the norm", Commun. Stat.-Theory Methods, Vol. 22, pp. 2903-2924, 1993.

[17] T. R. Park and H. Y. Baek, "An approach to

improving the Lindley estimator", Journal of the Korean Data & Information Science Society, Vol. 22, pp. 1251-1256, 2011.

[18] F. Perron and N. Giri, "On the best equivariant estimator of mean of a multivariate normal population", J. Multivar. Anal., Vol. 32, pp. 1-16, 1989.

[19] W. E. Strawderman, "Minimax estimation of location parameters for certain spherically symmetric distributions", J. Multivar. Anal., Vol. 4, pp. 255-264, 1974.