Minimax estimators of a normal variance

Yuzo Maruyama

Abstract

In the estimation problem of unknown variance of a multivariate normal distribution, a new class of minimax estimators is obtained. It is noted that a sequence of estimators in our class converges to the Stein's truncated estimator.

1 Introduction

Let X and S be independent random variables where X has p-variate normal distribution $N_p(\theta, \sigma^2 I_p)$ and S/σ^2 has chi square distribution χ^2_n with n degrees of freedom. We deal with the problem of estimating the unknown variance σ^2 by an estimator δ relative to the quadratic loss $(\delta/\sigma^2 - 1)^2$. Stein(1964) showed that the best affine equivariant estimator is $\delta_0 = (n+2)^{-1}S$ and it can be improved by considering a class of scale equivariant estimators $\delta_\phi = \phi(Z)S$, for $Z = \|X\|^2/S$. He really found an improved estimator $\delta^{ST} = \phi^{ST}(Z)S$, where $\phi^{ST}(Z) = \min((n+2)^{-1}, (n+p+2)^{-1}(1+Z))$. Brewster and Zidek(1974) derived an improved generalized Bayes estimator $\delta^{BZ} = \phi^{BZ}(Z)S$, where

$$\phi^{BZ}(Z) = \frac{1}{n+p+2} \frac{\int_0^1 \lambda^{p/2-1}(1+\lambda Z)^{-(n+p)/2-1} d\lambda}{\int_0^1 \lambda^{p/2-1}(1+\lambda Z)^{-(n+p)/2-2} d\lambda}.$$

We note that shrinkage estimators such as Stein's procedure and Brewster-Zidek's procedure are derived by using the vague prior information that $\lambda = \|\theta\|^2/\sigma^2$ is close to 0. It goes without saying that we would like to get significant improvement of risk when the prior information is accurate. Though δ^{ST} improves on δ_0 at $\lambda = 0$, it is not analytic and thus inadmissible. On the other hand, Brewster-Zidek's estimator does not improve on δ_0 at $\lambda = 0$ though it is admissible as shown in Proskin(1985). Therefore it is desirable to get analytic improved estimators dominating δ_0 especially at $\lambda = 0$. In this paper, we give such a class of improved estimators $\delta^V_\alpha = \phi^V_\alpha(Z)S$, where

$$\phi^V_\alpha(Z) = \frac{1}{n+p+2} \frac{\int_0^1 \lambda^{\alpha p/2-1}(1+\lambda Z)^{-\alpha((n+p)/2+1)} d\lambda}{\int_0^1 \lambda^{\alpha p/2-1}(1+\lambda Z)^{-\alpha((n+p)/2+1)-1} d\lambda},$$

for $\alpha > 1$. It is noted that δ_1^V coincides with Brewster-Zidek's estimator. Further we demonstrate that δ_α^V with $\alpha > 1$ improves on δ_0 especially at $\lambda = 0$ and that δ_α^V approaches Stein's estimator when α tends to infinity.

2 Main results

We first derive the estimator δ_α^V. The problem of estimating α times the variance $\alpha\sigma^2$ relative to the loss $(\delta/(\alpha\sigma^2) - 1)^2$ is considered, which is slight different from that of the variance σ^2. Among many generalized Bayes estimators for this problem, by selecting a suitable prior distribution, we can propose the estimator δ_α^V with $\alpha > 1$ which is not suitable for minimax estimator of $\alpha\sigma^2$ but that of σ^2. So far we have not determine whether or not δ_α^V with $\alpha > 1$ is the generalize Bayes estimator of σ^2.

Calculation for deriving the estimator δ_α^V is following. For $\eta = 1/(\alpha\sigma^2)$, let the conditional distribution of θ given λ, $0 < \lambda < 1$, be normal with mean 0 and covariance matrix $\lambda^{-1}(1-\lambda)\alpha^{-1}\eta^{-1}I_p$ and density functions of λ and η are proportional to $\lambda^{(\alpha-1)p/2-1}I_{(0,1)}(\lambda)$ and $\eta^{(\alpha-1)((p+n)/2+1)-1}I_{(0,\infty)}(\eta)$, respectively . Then the joint distribution $g(\eta, x, s)$ of η, X, S is given by

$$\begin{aligned} g(\eta,x,s) &\propto \int \eta^{p/2}\exp(-\frac{\alpha\eta}{2}\|x-\theta\|^2)\left(\eta\frac{\lambda}{1-\lambda}\right)^{p/2}\exp\left(-\frac{\lambda}{1-\lambda}\frac{\alpha\eta}{2}\|\theta\|^2\right) \\ &\quad \cdot\eta^{(\alpha-1)((p+n)/2+1)-1}\lambda^{(\alpha-1)p/2-1}\eta^{n/2}\exp(-\alpha\eta s/2)d\theta d\lambda \\ &\propto \int \eta^{p/2}\left(\eta\frac{\lambda}{1-\lambda}\right)^{p/2}\exp\left(-\alpha\eta\frac{\|\theta-(1-\lambda)x\|^2}{2(1-\lambda)} - \frac{\alpha\eta\|x\|^2\lambda}{2}\right) \\ &\quad \cdot\eta^{(\alpha-1)((p+n)/2+1)-1}\lambda^{(\alpha-1)p/2-1}\eta^{n/2}\exp(-\alpha\eta s/2)d\theta d\lambda \\ &\propto \eta^{\alpha((p+n)/2+1)-2}\int_0^1 \lambda^{\alpha p/2-1}\exp\left(-\alpha\eta\frac{\|x\|^2\lambda+s}{2}\right)d\lambda. \end{aligned}$$

As the generalized Bayes estimator for the loss $(\delta/(\alpha\sigma^2)-1)^2$ is $E(\eta \mid X,S)/E(\eta^2 \mid X,S)$, we have

$$\begin{aligned} \delta_\alpha^V &= \frac{\int_0^1 \lambda^{\alpha p/2-1}\int_0^\infty \eta^{\alpha((p+n)/2+1)-1}\exp\left(-\alpha\eta(\|X\|^2\lambda+S)/2\right)d\lambda}{\int_0^1 \lambda^{\alpha p/2-1}\int_0^\infty \eta^{\alpha((p+n)/2+1)}\exp\left(-\alpha\eta(\|X\|^2\lambda+S)/2\right)d\lambda} \\ &= \frac{1}{n+p+2}\frac{\int_0^1 \lambda^{\alpha p/2-1}(1+\lambda Z)^{-\alpha((n+p)/2+1)}d\lambda}{\int_0^1 \lambda^{\alpha p/2-1}(1+\lambda Z)^{-\alpha((n+p)/2+1)-1}d\lambda}S. \end{aligned}$$

Next, in the same way as Maruyama(1998), $\phi_\alpha^V(z)$ is represented through the hypergeometric function

$$F(a,b,c,x) = 1 + \sum_{n=1}^\infty \frac{(a)_n(b)_n}{(c)_n}\frac{x^n}{n!} \qquad \text{for} \qquad (a)_n = a\cdot(a+1)\cdots(a+n-1).$$

The following facts about $F(a,b,c,x)$, from Abramowitz and Stegun(1964), are needed;

$$\int_0^x t^{a-1}(1-t)^{b-1}dt = \frac{x^a}{a}F(a,1-b,a+1,x) \quad \text{for} \quad a,b>1, \tag{2.1}$$

$$F(a,b,c,x) = (1-x)^{c-a-b}F(c-a,c-b,c,x), \tag{2.2}$$

$$(c-a-b)F(a,b,c,x) - (c-a)F(a-1,b,c,x) + b(1-x)F(a,b+1,c,x) = 0, \tag{2.3}$$

$$(b-a)(1-x)F(a,b,c,x) - (c-a)F(a-1,b,c,x) + (c-b)F(a,b-1,c,x) = 0, \tag{2.4}$$

$$F(a,b,c,1) = \infty \qquad \text{when } \ c-a-b \leq -1. \tag{2.5}$$

Making a transformation and using (2.1) and (2.2), we have

$$\begin{aligned}
&\int_0^1 \lambda^{\alpha p/2-1}(1+\lambda Z)^{-\alpha((n+p)/2+1)}d\lambda \Big/ \int_0^1 \lambda^{\alpha p/2-1}(1+\lambda Z)^{-\alpha((n+p)/2+1)-1}d\lambda \\
&= \int_0^{\frac{z}{z+1}} t^{\alpha p/2-1}(1-t)^{\alpha(n/2+1)-1}dt \Big/ \int_0^{\frac{z}{z+1}} t^{\alpha p/2-1}(1-t)^{\alpha(n/2+1)}dt \\
&= (z+1)\frac{F\left(1,\alpha(n+p+2)/2,\alpha p/2+1,z/(z+1)\right)}{F\left(1,\alpha(n+p+2)/2+1,\alpha p/2+1,z/(z+1)\right)}.
\end{aligned}$$

Moreover by (2.3) and (2.4), $\phi_\alpha^V(z)$ is expressed as

$$\begin{aligned}
\phi_\alpha^V(z) &= \frac{1}{n+2}\left[1-\frac{p}{p+n+2}\frac{z+1}{F(1,\alpha(p+n+2)/2+1,\alpha p/2+1,z/(z+1))}\right] \\
&= \frac{1}{n+2}\left[1-\frac{p}{(n+2)F(1,\alpha(p+n+2)/2,\alpha p/2+1,z/(z+1))+p}\right].
\end{aligned} \tag{2.6}$$

Making use of (2.6), we can easily prove the theorem.

Theorem 2.1. *The estimator δ_α^V with $\alpha \geq 1$ is minimax.*

Proof. We shall verify that $\phi_\alpha^V(z)$ with $\alpha \geq 1$ satisfies the condition for minimaxity proposed by Brewster and Zidek(1974): for $\delta^{IM} = \phi^{IM}(z)S$, where $\phi_{IM}(z)$ is nondecreasing and $\phi^{BZ}(z) \leq \phi^{IM}(z) \leq 1/(n+2)$. Since $F(1,\alpha(p+n+2)/2,\alpha p/2+1,z/(z+1))$ is increasing in z, $\phi_\alpha^V(z)$ is increasing in z. By (2.5), it is clear that $\lim_{z\to\infty}\phi_\alpha^V(z) = 1/(n+2)$. In order to show that $\phi_\alpha^V(z) \geq \phi_1^V(z)$ for $\alpha \geq 1$, we have only to check that $F(1,\alpha(p+n+2)/2,\alpha p/2+1,z/(z+1))$ is increasing in α, which is easily verified because the coefficient of each term of the r.h.s. of the equation

$$\begin{aligned}
&F(1,\alpha(p+n+2)/2,\alpha p/2+1,z/(z+1)) \\
&= 1+\frac{p+n+2}{p+2/\alpha}\frac{z}{1+z}+\frac{(p+n+2)(p+n+2+2/\alpha)}{(p+2/\alpha)(p+4/\alpha)}\left(\frac{z}{1+z}\right)^2+\cdots
\end{aligned} \tag{2.7}$$

is increasing in α. We have thus proved the theorem. □

Now we investigate the nature of the risk of δ_α^V with $\alpha > 1$. By using Kubokawa(1994)'s method, the risk difference between δ_0 and δ_α^V at $\lambda = 0$ is written as

$$R(0,\delta_0) - R(0,\delta_\alpha^V) = 2\int_0^\infty \frac{d}{dz}\phi_\alpha^V(z)\left(\phi_\alpha^V(z) - \phi_1^V(z)\right)\int_0^\infty t^2 F_p(zt) f_n(t)\,dt\,dz,$$

where $f_k(t)$ and $F_k(t)$ designate the density and the distribution functions of χ_k^2. Therefore we see that Brewster-Zidek's estimator (δ_α^V with $\alpha = 1$) does not improve upon the best equivariant estimator at $\lambda = 0$. See also Rukhin(1991). On the other hand, since $\phi_\alpha^V(z)$ is strictly increasing in α, δ_α^V with $\alpha > 1$ improves on the best equivariant estimator especially at $\lambda = 0$. Figure 1 gives a comparison of the respective risks of the best equivariant estimator, Stein's estimator, Brewster-Zidek's estimator, δ_α^V with $\alpha = 2, 4$ and 10 for $p = 10$ and $n = 4$. This figure reveals that the risk behavior of δ_α^V with $\alpha = 10$ is similar to that of Stein's truncated estimator. In fact, we have the following result.

Proposition 2.1. *δ_α^V approaches Stein's estimator when α tends to infinity.*

Proof. Since $F(1, \alpha(p+n+2)/2, \alpha p/2+1, z/(z+1))$ is increasing in α, by the monotone convergence theorem this function converges to $\sum_{i=0}^\infty\{(p(z+1))^{-1}(n+p+2)z\}^i$ when α tends to infinity. Considering two cases: $(n+p+2)z < (\geq) p(z+1)$, we obtain $\lim_{\alpha\to\infty}\phi_\alpha^V(z) = (1+z)/(n+p+2)$ if $z < p/(n+2)$; $= 1/(n+2)$ otherwise. This completes the proof. $\square$

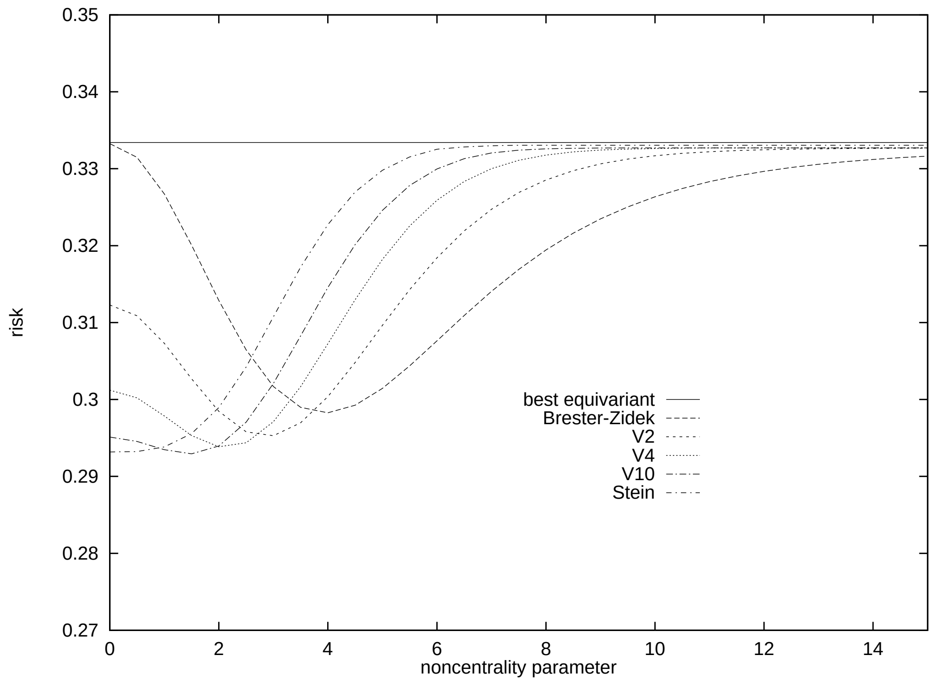

Figure 1 : Comparison of the risks of the estimators δ_0, δ^{ST}, δ^{BZ}, δ_α^V with $\alpha = 2, 4, 10$. 'noncentrality parameter' denotes $\|\theta\|/\sigma$. Vα denotes the risk of δ_α^V.

Remark 2.1 Ghosh(1994) proposed the minimax generalized Bayes estimator $\delta_k^{GH} = \phi_k^{GH}(Z)S$, where

$$\phi_k^{GH}(Z) = \frac{1}{n+p+2(k+2)} \frac{\int_0^1 \lambda^{p/2+k}(1+\lambda Z)^{-(n+p)/2-(k+2)} d\lambda}{\int_0^1 \lambda^{p/2+k}(1+\lambda Z)^{-(n+p)/2-(k+3)} d\lambda},$$

for $-1-p/2 < k \leq -1$. Clearly δ_k^{GH} with $k=-1$ coincide with Brewster-Zidek's estimator and we can see that δ_k^{GH} with $-1-p/2 < k < -1$ improves on the best equivariant estimator especially at $\lambda = 0$. It is noted that without the troublesome calculation in Ghosh(1994), the minimaxity of Ghosh's estimators is easily proved in the same way as Theorem 2.1 because ϕ_k^{GH} with $-1-p/2 < k \leq -1$ satisfies Brewster-Zidek's condition.

Remark 2.2 For the entropy loss function $\delta/\sigma^2 - \log(\delta/\sigma^2) - 1$, the discussions in this paper are directly applied. In this case, the best equivariant estimator is the unbiased estimator $\delta_0 = S/n$ and Stein's truncated estimator is $\delta^{ST} = \phi^{ST}(Z)S$, where $\phi^{ST}(Z) = \min(1/n, (n+p)^{-1}(1+Z))$. Moreover Brewster-Zidek's estimator is $\delta^{BZ} = \phi^{BZ}(Z)S$, where

$$\phi^{BZ}(Z) = \frac{1}{n+p} \frac{\int_0^1 \lambda^{p/2-1}(1+\lambda Z)^{-(n+p)/2} d\lambda}{\int_0^1 \lambda^{p/2-1}(1+\lambda Z)^{-(n+p)/2-1} d\lambda}.$$

Then the proposed minimax estimator is $\delta_\alpha^V = \phi_\alpha^V(Z)S$, where

$$\phi_\alpha^V(Z) = \frac{1}{n+p} \frac{\int_0^1 \lambda^{\alpha p/2-1}(1+\lambda Z)^{-\alpha(n+p)/2} d\lambda}{\int_0^1 \lambda^{\alpha p/2-1}(1+\lambda Z)^{-\alpha(n+p)/2-1} d\lambda},$$

for $\alpha > 1$.

References

[1] Abramowitz M, Stegun, IA (1964) Handbook of Mathematical Functions, Dover Publications, New York

[2] Brewster JF, Zidek JV (1974) Improving on equivariant estimators. Ann Statist 2: 21-38

[3] Ghosh M (1994) On some Bayesian solutions of the Neyman-Scott problem. In Statistical Decision Theory and Related Topics V, Springer-Verlag, New York pp 267-276

[4] Kubokawa T (1994) An unified approach to improving equivariant estimators. Ann Statist 22: 290-299

[5] Maruyama Y (1998) Improving on the James-Stein estimator. Unpublished manuscript

[6] Proskin HM (1985) An admissibility theorem with applications to the estimation of variance of the normal distribution. Ph.D dissertation, Dept Statist, Rutgers University

[7] Rukhin AL (1992) Asymptotic risk behavior of mean vector and variance estimators and the problem of positive mean. Ann Inst Statist Math 44: 299-311

[8] Stein C (1964) Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. Ann Inst Statist Math 16: 155-160