数理解析研究所講究録 1334 巻 2003 年 175-182

Another Minimax Generalized Bayes Estimators of a Normal Variance

丸山 祐造
YUZO MARUYAMA
東京大学・空間情報科学研究センター
CENTER FOR SPATIAL INFORMATION SCIENCE, THE UNIVERSITY OF TOKYO
ビル・ストローダーマン
WILLIAM, E. STRAWDERMAN
ラトガース大学・統計学部
DEPARTMENT OF STATISTICS, RUTGERS UNIVERSITY

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_n^2 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 function $L_2(\delta,\sigma^2)=(\delta/\sigma^2-1)^2$.

Stein (1964) showed that the best affine equivariant minimax estimator is $\delta_0=S/(n+2)$ and it can be improved by considering a class of scale equivariant estimators

$$\delta_\phi=\frac{1}{n+2}(1-\phi(W))S, \tag{1}$$

for $W=\|X\|^2/S$. Explicitly, he found an improved estimator $\delta^S=\min\{S/(n+2),(\|X\|^2+S)/(p+n+2)\}=(n+2)^{-1}(1-\phi^S(W))S$, where $\phi^S(w)=\max\{0,(p-(n+2)w)/(p+n+2)\}$. Brewster and Zidek (1974) derived an improved generalized Bayes estimator $\delta^{BZ}=(n+2)^{-1}(1-\phi^{BZ}(W))S$, where

$$\phi^{BZ}(w)=\frac{2(1+w)^{-n/2-1}}{n+p+2}\left(\int_0^1 t^{p/2-1}\{1-wt/(w+1)\}^{n/2+1}dt\right)^{-1}. \tag{2}$$

They also gave the general sufficient condition for minimaxity, which we denote by the BZ-condition in this paper. They showed that δ_ϕ is minimax if

(BZ1) $\phi(w)$ is nonincreasing,

(BZ2) $0\le\phi(w)\le\phi^{BZ}(w)$.

Clearly both $\phi^S(w)$ and $\phi^{BZ}(w)$ satisfy the BZ-condition.

On the other hand, Strawderman (1974) derived another sufficient condition for minimaxity, which we denote by the ST-condition in this paper. He showed that δ_ϕ is minimax if

(ST1) $(1+w)^\epsilon \phi(w)$ is nonincreasing,

(ST2) $0 \le \phi(w) \le B_2(p,n,\epsilon)$.

The upper bound $B_2(p,n,\epsilon)$ will be discussed detail in Section 2. Strawderman (1974) claimed that δ^{BZ} satisfies the ST-condition. But his claim is incorrect as pointed out in Ghosh (1994) and Pal *et al.* (1998). Ghosh (1994) proposed a generalized Bayes estimator $\delta^G = (n+2)^{-1}(1-\phi^G(W))S$ where

$$\phi^G(w) = \frac{2(1+w)^{-n/2-1}}{n+p+2(a+2)} \left(\int_0^1 t^{p/2+a}\{1-wt/(w+1)\}^{n/2+1} dt \right)^{-1} \tag{3}$$

and showed that $\phi^G(w)$ for $-p/2-1 < a \le -1$ satisfies the BZ-condition for minimaxity. Pal *et al.* (1998) pointed out that $\phi^G(w)$ for some $a(< -1)$ also satisfies the ST-condition. As far as we know, however, generalized Bayes estimators which satisfy the ST-condition but do not satisfy the BZ-condition have not been found to date.

In this paper, we propose such estimators. We consider a generalized Bayes estimator $\delta^{GB} = (n+2)^{-1}(1-\phi^{GB}(W))S$ where

$$\phi^{GB}(w) = \frac{2b(w+1)^{-1}}{p+n+2(a+2)} \frac{\int_0^1 t^{p/2+a+1}(1-t)^{b-1}\{1-wt/(w+1)\}^{n/2-b} dt}{\int_0^1 t^{p/2+a}(1-t)^{b}\{1-wt/(w+1)\}^{n/2-b+1} dt} \tag{4}$$

which, for $b > 0$ does not satisfy the BZ-condition. We show that $\phi^{GB}(w)$ for some a and $b > 0$ satisfies the ST-condition. We make two main contributions. The first is to enrich the class of minimax generalized Bayes estimators under L_2. The second is to find within our class, a subclass of estimators of a particularly simple form $(n+2)^{-1}(1+\alpha/(W+1))^{-1}S$. This second contribution is interesting since most known generalized Bayes minimax procedures such as (2) or (3) seem quite complicated while the empirical Bayes estimator δ^S is quite simple. Hence we produce generalized Bayes estimators with a form as simple as δ^S.

2 Strawderman's type sufficient condition for minimaxity

First we review Strawderman (1974)'s sufficient condition for minimaxity. Under L_2, Strawderman (1974) proposed as a upper bound of ϕ

$$B_2^{ST}(p,n,\epsilon) = \min\left(\frac{2}{1+\epsilon}, 2\frac{\Gamma(p/2+n/2+2\epsilon+2)\Gamma(n/2+\epsilon+1)}{\Gamma(p/2+n/2+\epsilon+2)\Gamma(n/2+2\epsilon+2)} \frac{p\epsilon}{p+n+2} \right). \tag{5}$$

He claimed that $B \le 2/(1+\epsilon)$ is required to guarantee that the function $g(u)u^{\epsilon+1}$, where $g(u) = 2u - Bu^{\epsilon+1} - 2(n+2)/(p+n+2)$, changes sign only once from negative to positive on $u \in [0,1]$. Pal *et al.* (1998) pointed out that $g(1) > 0$, that is, $B < 2p/(p+n+2)$, should be also required and hence proposed $\min(B_2^{ST}, 2p/(p+n+2))$ as the upper bound. Here noting that the equation $g(u) = 0$ has at most two solutions on $[0, \infty]$, $g(0) < 0$ and $g(u)$ is negative for sufficiently large u, we see that $g(u)$ changes sign only once from negative to positive on $[0, 1]$ if and only if $g(1) > 0$. Therefore we propose the modified version of the ST-condition.

Theorem 2.1. *Estimators of the form (1) are minimax under L_2 provided $\epsilon > 0$, $(1+w)^\epsilon \phi(w)$ is nonincreasing and $0 \le \phi < B_2(p, n, \epsilon)$ where*

$$B_2(p,n,\epsilon) = \min\left(\frac{2p}{p+n+2}, 2\frac{\Gamma(p/2+n/2+2\epsilon+2)\Gamma(n/2+\epsilon+1)}{\Gamma(p/2+n/2+\epsilon+2)\Gamma(n/2+2\epsilon+2)}\frac{p\epsilon}{p+n+2}\right). \tag{6}$$

3 Minimax generalized Bayes estimators

In this section, we derive minimax generalized Bayes estimators satisfying the sufficient condition proposed in Theorem 2.1 for the loss L_2. We consider a class of generalized Bayes estimators with respect to the following prior distribution. For $\eta = \sigma^{-2}$, let the conditional distribution of θ given λ and η, for $0 < \lambda < 1$, be normal with mean vector 0 and covariance matrix $\lambda^{-1}(1-\lambda)\eta^{-1}I_p$ and let the density functions of λ and η be proportional to $\lambda^a(1-\lambda)^b I_{(0,1)}(\lambda)$ and $\eta^c I_{(0,\infty)}(\eta)$, respectively. Then the joint distribution $g(\eta, x, s)$ of η, X, S is given by

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

As the generalized Bayes estimator under L_2 loss is written as $E(\eta \mid X,S)/E(\eta^2 \mid X,S) = \int \eta g(\eta,x,s)d\eta / \int \eta^2 g(\eta,x,s)d\eta$, we have the estimator

$$\begin{aligned}
\delta^{GB} &= \frac{\int_0^1 \lambda^{p/2+a}(1-\lambda)^b\int_0^\infty \eta^{(p+n)/2+c+1}\exp\left(-\eta\frac{\|X\|^2\lambda+S}{2}\right)d\eta d\lambda}{\int_0^1 \lambda^{p/2+a}(1-\lambda)^b\int_0^\infty \eta^{(p+n)/2+c+2}\exp\left(-\eta\frac{\|X\|^2\lambda+S}{2}\right)d\eta d\lambda}\\
&= \frac{1}{p+n+2(c+2)}\frac{\int_0^1 \lambda^{p/2+a}(1-\lambda)^b(1+\lambda W)^{-(n+p)/2-c-2}d\lambda}{\int_0^1 \lambda^{p/2+a}(1-\lambda)^b(1+\lambda W)^{-(n+p)/2-c-3}d\lambda}S\\
&= \frac{1}{p+n+2(c+2)}\frac{\int_0^1 t^{p/2+a}(1-t)^b\{1-tW/(W+1)\}^{n/2-a-b+c}dt}{\int_0^1 t^{p/2+a}(1-t)^b\{1-tW/(W+1)\}^{n/2-a-b+c+1}dt}S \qquad (7)\\
&= \varphi^{GB}(W)S,\ (say)
\end{aligned}$$

which is well-defined if $a > -p/2-1$ and $b > -1$. In the following, as we have $\lim_{w\to\infty}\varphi^{GB}(w) = 1/\{n+2+2(c-a)\}$ from (7), we only consider the case $a = c$, that is,

$$\varphi^{GB}(w) = \frac{1}{p+n+2(a+2)} \frac{\int_0^1 t^{p/2+a}(1-t)^b\{1-tW/(W+1)\}^{n/2-b}dt}{\int_0^1 t^{p/2+a}(1-t)^b\{1-tW/(W+1)\}^{n/2-b+1}dt}. \tag{8}$$

In particular, we have a simple form

$$\varphi^{GB}(w) = \left(n+2+\frac{p+2a+2}{w+1}\right)^{-1} \tag{9}$$

by letting $b = n/2$ in (8). This is unexpected because generalized Bayes minimax shrinkage estimators such as (2) and (3) typically have a complicated form. The estimator (9) has a form which is comparable to Stein's estimator δ^S in its simplicity. We will show, in Proposition 3.5 below, that $\psi^{GB}(W)S$ given by (9) is minimax for certain a.

Some basic properties of behavior of φ^{GB} given by (9) are given in the following result.

Lemma 3.1. *1. $\varphi^{GB}(w)$ is increasing in w for $b \geq 0$.*

2. φ^{GB} is decreasing in a for fixed $b \geq 0$ and w.

Since δ^{GB} for $a = -1$ and $b = 0$ corresponds to δ^{BZ}, the BZ-condition together with Lemma 3.1 implies that δ^{GB} for $-p/2-1 < a \leq -1$ and $b = 0$, which is equal to Ghosh's (1994) estimator, is minimax.

Proof. By the change of variables in (8), we have

$$\varphi^{GB}(w) = \frac{1}{p+n+2(a+2)} \frac{\int_0^v t^{p/2+a}(v-t)^b(1-t)^{n/2-b}dt}{\int_0^v t^{p/2+a}(v-t)^b(1-t)^{n/2-b+1}dt},$$

where $v = w/(w+1)$. For $v_1 > v_2$ and $b \geq 0$,

$$\begin{aligned}\frac{\int_0^{v_1} t^{p/2+a}(v_1-t)^b(1-t)^{n/2-b}dt}{\int_0^{v_1} t^{p/2+a}(v_1-t)^b(1-t)^{n/2-b+1}dt} &\geq \frac{\int_0^{v_2} t^{p/2+a}(v_1-t)^b(1-t)^{n/2-b}dt}{\int_0^{v_2} t^{p/2+a}(v_1-t)^b(1-t)^{n/2-b+1}dt} \\ &\geq \frac{\int_0^{v_2} t^{p/2+a}(v_2-t)^b(1-t)^{n/2-b}dt}{\int_0^{v_2} t^{p/2+a}(v_2-t)^b(1-t)^{n/2-b+1}dt}.\end{aligned}$$

The first inequality follows from the fact that the ratio of integrands of the numerator and the denominator is increasing, the second inequality from the fact that $\{(v_1-t)/(v_2-t)\}^b$ is increasing. This completes the proof of (i).

By the change of variables $(u = \eta s,\ t = u\lambda)$, in the first equality in (7), we have

$$\varphi^{GB}(w) = \frac{\int_0^\infty u^{n/2}\exp(-u/2)h_w(u)du}{\int_0^\infty u^{n/2+1}\exp(-u/2)h_w(u)du},$$

where $h_w(u,a) = \int_0^u t^{p/2+a}(1-t/u)^b\exp(-wt/2)dt$. Hence, to prove (ii), it is sufficient to show that $h_w(u,a_1)/h_w(u,a_2)$ is increasing in u for $a_1 > a_2$. As in the proof of (i), we see that for

$u_1 > u_2$ and $b \geq 0$

$$\begin{aligned}\frac{\int_0^{u_1} t^{p/2+a_1}(u_1-t)^b \exp(-wt/2)dt}{\int_0^{u_1} t^{p/2+a_2}(u_1-t)^b \exp(-wt/2)dt} &\geq \frac{\int_0^{u_2} t^{p/2+a_1}(u_1-t)^b \exp(-wt/2)dt}{\int_0^{u_2} t^{p/2+a_2}(u_1-t)^b \exp(-wt/2)dt} \\ &\geq \frac{\int_0^{u_2} t^{p/2+a_1}(u_2-t)^b \exp(-wt/2)dt}{\int_0^{u_2} t^{p/2+a_2}(u_2-t)^b \exp(-wt/2)dt},\end{aligned}$$

which completes the proof of (ii). □

Note: Since the derivative of $\{(1-vt)/t\}^{n/2+1}$ is $(-n/2-1)\{(1-vt)/t\}^{n/2}t^{-2}$, an integration by parts in (8) gives $\varphi^{GB}(w) = (n+2)^{-1}(1-\phi^{GB}(w))$ where

$$\begin{aligned}\phi^{GB}(w) &= \frac{2b(1-v)}{p+n+2(a+2)} \frac{\int_0^1 t^{p/2+a+1}(1-t)^{b-1}(1-vt)^{n/2-b}dt}{\int_0^1 t^{p/2+a}(1-t)^b(1-vt)^{n/2-b+1}dt} \text{ for } b>0 \\ &= \frac{2(1-v)^{n/2+1}}{p+n+2(a+2)} \frac{1}{\int_0^1 t^{p/2+a}(1-vt)^{n/2+1}dt} \text{ for } b=0.\end{aligned}$$

Since $v = w/(1+w)$ and all integrals above approach constant values, we have

$$\phi^{GB}(w) = \begin{cases} O\{(w+1)^{-n/2-1}\} & \text{for } b=0 \\ O\{(w+1)^{-1}\} & \text{for } b>0. \end{cases}$$

Since $\phi^{BZ}(w)$ is $\phi^{GB}(w)$ for $a=-1$ and $b=0$, $\phi^{BZ}(w) = O\{(w+1)^{-n/2-1}\}$. This implies that $\phi^{GB}(w)$ for $b>0$ is greater than $\phi^{BZ}(w)$ for sufficiently large w. Thus we have the following result.

Theorem 3.2. *$\phi^{GB}(w)$ for $b>0$ does not satisfy (BZ2) of the BZ-condition.*

Next we investigate the properties of ϕ^{GB} in order to apply the ST-condition proposed in Theorem 2.1.

Theorem 3.3. *1. $\phi^{GB}(w) \leq (p+2a+2)/(p+n+2a+4)$.*

2. $(1+w)^\epsilon \phi^{GB}(w)$ is monotone nonincreasing if

(a) $b=0$ and $a < -p/2-2+(n/2+1)/\epsilon$ or

(b) $0 < b \leq n/2+1$, $\epsilon \leq 1$ and $a < -p/2-b-2+(n/2+1)/\epsilon$.

(c) $b > n/2+1$ and $a < -p/2-b-2+b(b-n/2)/(\epsilon+b-n/2-1)$.

Proof. By Theorem 3.1 $\phi^{GB}(w)$ is decreasing in w and hence $\phi^{GB}(w) \leq \phi^{GB}(0) = (p+2a+2)/(p+n+2a+4)$, which completes the proof of (i).

For $b=0$, The derivative of $(1+w)^\epsilon\phi(w)$ with respect to $v = w/(w+1)$ is written as

$$(1-v)^{-\epsilon-1}\phi(w)\left[-n/2-1+\epsilon+(n/2+1)(1-v)\frac{\int_0^1 t^{p/2+a+1}(1-vt)^{n/2}dt}{\int_0^1 t^{p/2+a}(1-vt)^{n/2+1}dt}\right]. \qquad (10)$$

Using the relation

$$\int_0^1 t^\alpha(1-t)^\beta(1-vt)^\gamma dt = (1-v)^{\beta+\gamma+1}\int_0^1 t^\beta(1-t)^\alpha(1-vt)^{-\alpha-\beta-\gamma-2}dt \tag{11}$$

for $\alpha > -1$ and $\beta > -1$, the term in bracket in (10) is written as

$$-n/2-1+\epsilon+(n/2+1)\frac{\int_0^1(1-t)^{p/2+a+1}(1-vt)^{-p/2-n/2-a-3}dt}{\int_0^1(1-t)^{p/2+a}(1-vt)^{-p/2-n/2-a-3}dt}$$

which is less than $-n/2-1+\epsilon+(n/2+1)(p/2+a+1)/(p/2+a+2) = \epsilon-(n/2+1)/(p/2+a+2)$ because $(1-vt)^{-1}$ is increasing in t. This completes the proof in the case $b=0$.

For $b>0$, the derivative of $(1+w)^\epsilon\phi(w)$ with respect to $v=w/(w+1)$, together with the relation (11) is written as

$$\begin{aligned}(1-v)^{-\epsilon-1}\phi(w)\Bigg[\epsilon-1+(b-n/2)&\frac{\int_0^1 t^{b-1}(1-t)^{p/2+a+2}(1-vt)^{-p/2-n/2-a-2}dt}{\int_0^1 t^{b-1}(1-t)^{p/2+a+1}(1-vt)^{-p/2-n/2-a-2}dt}\\ &+(n/2+1-b)\frac{\int_0^1 t^{b}(1-t)^{p/2+a+1}(1-vt)^{-p/2-n/2-a-3}dt}{\int_0^1 t^{b}(1-t)^{p/2+a}(1-vt)^{-p/2-n/2-a-3}dt}\Bigg].\end{aligned} \tag{12}$$

By applying a Maclaurin expansion to the integrals in (12), the term in bracket in (12) is written as

$$\begin{aligned}\epsilon-1+&\frac{(b-n/2)(p/2+a+2)}{p/2+a+b+2}\frac{F(p/2+n/2+a+2,b,p/2+a+b+3,v)}{F(p/2+n/2+a+2,b,p/2+a+b+2,v)}\\ &-\frac{(b-n/2-1)(p/2+a+1)}{p/2+a+b+2}\frac{F(p/2+n/2+a+3,b+1,p/2+a+b+3,v)}{F(p/2+n/2+a+3,b+1,p/2+a+b+2,v)},\end{aligned} \tag{13}$$

where $F(a,b,c,x)$ is the hypergeometric function

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

From the inequality

$$\frac{F(p/2+n/2+a+2,b,p/2+a+b+3,v)}{F(p/2+n/2+a+2,b,p/2+a+b+2,v)} \geq \frac{F(p/2+n/2+a+3,b+1,p/2+a+b+3,v)}{F(p/2+n/2+a+3,b+1,p/2+a+b+2,v)},$$

(13) for $b \le n/2+1$ is less than

$$\begin{aligned}&\epsilon-1+\frac{p/2+a+1-n/2+b}{p/2+a+b+2}\frac{F(p/2+n/2+a+2,b,p/2+a+b+3,v)}{F(p/2+n/2+a+2,b,p/2+a+b+2,v)}\\ &\le \epsilon-\min\left(1,\frac{n/2+1}{p/2+a+b+2}\right),\end{aligned}$$

which is nonpositive when $\epsilon \le 1$ and $a < -p/2-b-2+(n/2+1)/\epsilon$. This completes the proof in the case $0<b\le n/2+1$.

For $b > n/2+1$, (13) is less than

$$\epsilon-1+\frac{(b-n/2)(p/2+a+2)}{p/2+a+b+2},$$

which is nonpositive when $a < -p/2-b-2+b(b-n/2)/(\epsilon+b-n/2-1)$. This completes the proof. □

Combining Theorem 3.3 and Theorem 2.1, we have the following result.

Theorem 3.4. *The generalized Bayes estimator δ^{GB} given by (4) is minimax under L_2 loss if*

$$-p/2 - 1 < a < -p/2 - 1 + \max_{\epsilon} \min \left((n/2+1)\frac{B_2(p,n,\epsilon)}{1 - B_2(p,n,\epsilon)}, C_2(n,\epsilon,b) \right) \quad (14)$$

where

$$C_2(n,\epsilon,b) = \begin{cases} (n/2+1)/\epsilon - 1 & b = 0 \\ -b - 1 + \max(n/2+1, (n/2+1)/\epsilon) & 0 < b \le n/2+1 \\ -b - 1 + b(b - n/2)/(\epsilon + b - n/2 - 1) & b > n/2+1. \end{cases}$$

Figure 1 reveals the upper bounds for minimaxity in the case $p = 10$. Note that the upper bound given by (14) is not always continuous in $b = 0$.

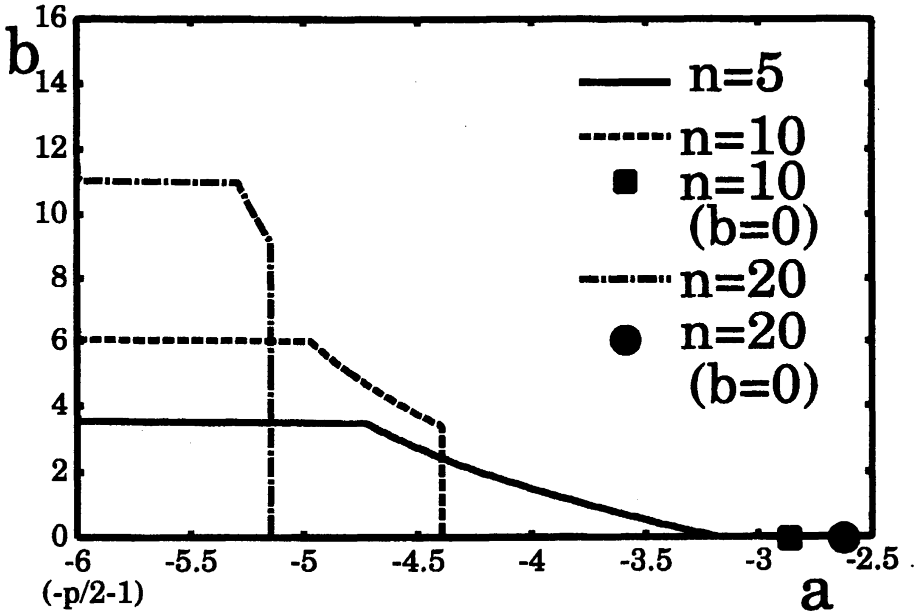

Figure 1 : Ranges of values for minimaxity in the case $p = 10$ under L_2 loss

As noted earlier, when $b = 0$ ϕ^{GB} satisfies the BZ-condition for $-p/2 - 1 < a \le -1$ but does not satisfy the ST-condition when $a = -1$ (i.e. $\delta^{GB} = \delta^{BZ}$). Our contribution when $b = 0$ is to add an explicit upper bound on a so that the ST-condition is satisfied. This upper bound is of course less than -1 so that when $b = 0$. The class of estimators satisfying the BZ-condition contains the class satisfying the ST-condition. However when $b > 0$ the class of estimators δ^{GB} satisfying the BZ-condition is empty while the class satisfying the ST-condition is non-empty and is in fact quite rich.

Finally we highlight the following very simple case.

Proposition 3.5. *the generalized Bayes estimator*

$$\frac{1}{n+2}\left(1 + \frac{\alpha}{W+1}\right)^{-1} S$$

is minimax under L_2 if $0 < \alpha \le \max_\epsilon \min \left(B_2(p,n,\epsilon)(1 - B_2(p,n,\epsilon))^{-1}, 1/\epsilon - 1\right)$.

Proof. Apply Theorem 3.4 to the estimator given in (9). □

参 考 文 献

[1] Brewster, J.F. and Zidek, J.V. (1974). Improving on equivariant estimators. *Ann. Statist.*, **2**, 21–38.

[2] Ghosh, M. (1994). On some Bayesian solutions of the Neyman-Scott problem. In *Statistical Decision Theory and Related Topics V*, 267–276.

[3] Pal, N., Ling, C., and Lin, J. (1998). Estimation of a normal variance - a critical review. *Statist. Papers*, **4**, 389–404.

[4] 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.

[5] Strawderman, W.E. (1974). Minimax estimation of powers of the variance of a normal population under squared error loss. *Ann. Statist*, **2**, 190–198.