Sankhyā : The Indian Journal of Statistics
2023, Volume 85-A, Part 2, pp. 1667-1683

Generalized Bayes Minimax Estimators of the Variance of a Multivariate Normal Distribution

Shokofeh Zinodiny
Institute for Research in Fundamental Sciences, Tehran, Iran
Saralees Nadarajah
Howard University, Washington DC, USA

Abstract

The problem of estimating the variance of a multivariate normal distribution is considered under quadratic loss. A large class of generalized Bayes minimax estimators for the variance is found. This class include estimators obtained by Ghosh (1994). A simulation study shows superior performance of our estimators.

AMS (2000) *subject classification*. Primary 62E99.
Keywords and phrases. Generalized Bayes estimators, Quadratic loss function, Unknown variance.

1 Introduction

Let X and S be independent random variables having

$$X \sim N_p\left(\theta, \sigma^2 I_p\right), \ S \sim \sigma^2 \chi_n^2, \tag{1.1}$$

where $N_p\left(\theta, \sigma^2 I_p\right)$ denotes the p-variate normal distribution with unknown mean vector θ and covariance matrix $\sigma^2 I_p$, where I_p denotes the $p \times p$ identity matrix, and χ_n^2 denotes a chisquare variable with n degree of freedom. We consider the problem of estimating σ^2 when the loss function is

$$L\left(\delta; \sigma^2\right) = \left(\frac{\delta}{\sigma^2} - 1\right)^2, \tag{1.2}$$

where $\delta = \delta(X, S)$ is an estimator of σ^2.

The best affine equivariant estimator is $\delta_0 = (n+2)^{-1} S$ which is a minimax estimator with constant risk $2(n+2)^{-1}$. Stein (1964) showed that δ_0 can be improved by considering a class of scale equivariant estimators $\delta =$

$(n+2)^{-1}(1-\phi(F))\,S$ for $F=\frac{X'X}{S}$. He found a specific better estimator $\delta^S=(n+2)^{-1}\left(1-\phi^S(F)\right)S$, where $\phi^S(F)=\max\left\{0,\frac{p-(n+2)F}{p+n+2}\right\}$. Brewster and Zidek (1974) obtained an improved generalized Bayes estimator $\delta^{BZ}=(n+2)^{-1}\left(1-\phi^{BZ}(F)\right)S$, where

$$\phi^{BZ}(F)=1-\frac{n+2}{p+n+2}\frac{\int_0^1\lambda^{\frac{p}{2}-1}(1+\lambda F)^{-\frac{p+n}{2}-1}d\lambda}{\int_0^1\lambda^{\frac{p}{2}-1}(1+\lambda F)^{-\frac{p+n}{2}-2}d\lambda}. \tag{1.3}$$

They also gave a general sufficient condition for minimaxity, using an integral expression of the difference in risks between δ_0 and δ. Strawderman (1974) derived another sufficient condition for minimaxity. Using conditions of Brewster and Zidek (1974), Ghosh (1994) obtained a class of generalized Bayes estimators for σ^2. Maruyama and Strawderman (2006) proposed another class of improved generalized Bayes estimators.

In this paper, we derive a large class of generalized Bayes minimax estimators of σ^2 which contains estimators of Ghosh (1994) as special cases. To do so, we use techniques of Wells and Zhou (2008) and Brewster and Zidek (1974). Section 3 considers some examples of classes of generalized Bayes minimax estimators. In particular, Example 1 demonstrates that a result in Ghosh (1994) follows from our main theorem. Section 4 compares the minimax estimators of Sections 2, 3 and the equivariant estimator δ_0 by simulation.

2 A class of generalized Bayes minimax estimators

In this section, we consider the problem of estimating σ^2 in (1.1) under the loss function (1.2). Our main result is Theorem 2.2. Before stating and proving this theorem, we state a theorem due to Brewster and Zidek (1974) and Kubokawa (1994), discuss a class of priors, borrow some notations from Wells and Zhou (2008) and state and prove two technical lemmas (Lemmas 2.1 and 2.2).

Brewster and Zidek (1974) derived general sufficient conditions for minimaxity of estimators having the form $\delta=(n+2)^{-1}(1-\phi(F))\,S$, where $\phi(F)$ is a function of $F=\frac{X'X}{S}$. For the purpose of verifying the minimaxity of a generalized Bayes estimator, we use the following specialized result.

Theorem 2.1. *The estimator $\delta(X,S)$ given by*

$$\delta=(n+2)^{-1}(1-\phi(F))\,S \tag{2.1}$$

is minimax for σ^2 under the loss function (1.2) provided that the following conditions hold

I. $\phi(F)$ *is nonincreasing,*

II. $0 \leq \phi(F) \leq \phi^{BZ}(F)$*, where* $\phi^{BZ}(F)$ *is given by* (1.3).

PROOF. See Brewster and Zidek (1974) and Kubokawa (1994). □

Now, we construct generalized Bayes minimax estimators of σ^2 under the loss function (1.2). To do so, we consider the following class of prior distributions.

For $\eta = \sigma^2$, let the conditional distribution of θ given ν and η be normal with zero mean vector and covariance matrix $\nu\eta^{-1}I_p$ and let the generalized density of (ν, η) given by $h(\nu, \eta) = \eta^b g(\nu)$, $\nu > 0$, $\eta > 0$, where $b > -\frac{n+p}{2} - 1$ and $g(\nu)$ is a continuously differentiable positive function on $[0, \infty)$ such that the following conditions hold

C1. $\int_0^1 \lambda^{\frac{p}{2}-2} g\left(\frac{1-\lambda}{\lambda}\right) d\lambda < \infty$,

C2. $\lim\limits_{\nu\to\infty} \frac{g(\nu)}{(1+\nu)^{\frac{p}{2}-1}} = 0$.

In the following discussion, we obtain conditions on g and b such that the generalized Bayes estimators satisfy the conditions of Theorem 2.1, and hence are minimax. Note that the joint density function $f(\eta, x, s)$ of η, X, S is

$$\begin{aligned} f(\eta, x, s) &\propto \int_0^\infty \int_{R^p} \eta^{\frac{p}{2}} e^{-\frac{\eta\|x-\theta\|^2}{2}} \nu^{-\frac{p}{2}} \eta^{\frac{p}{2}} e^{-\frac{\eta\|\theta\|^2}{2\nu}} g(\nu)\, \eta^b \eta^{\frac{n}{2}} e^{-\frac{\eta s}{2\nu}} d\theta d\nu \\ &\propto \int_0^\infty \int_{R^p} \eta^{\frac{2p+n}{2}+b} e^{-\frac{\eta}{2}\left[\|x-\theta\|^2 + \frac{\|\theta\|^2}{\nu}\right]} \nu^{-\frac{p}{2}} g(\nu)\, e^{-\frac{\eta s}{2\nu}} d\theta d\nu \\ &\propto \int_0^\infty \eta^{\frac{p+n}{2}+b} g(\nu)\,(1+\nu)^{-\frac{p}{2}} e^{-\frac{\eta}{2}\left[s + \frac{\|x\|^2}{1+\nu}\right]} d\nu, \end{aligned}$$

where $\|.\|$ denotes the Euclidean norm. Therefore, the generalized Bayes estimator of $\eta = \sigma^{-2}$ with respect to the loss function (1.2) is

$$\begin{aligned} \delta_B &= \frac{E[\eta|\, X, S]}{E[\eta^2|\, X, S]} \\ &= \frac{\int_0^\infty \int_0^\infty \eta^{\frac{p+n}{2}+b+1} g(\nu)\,(1+\nu)^{-\frac{p}{2}} e^{-\frac{\eta S}{2}\left(1+\frac{F}{1+\nu}\right)} d\eta d\nu}{\int_0^\infty \int_0^\infty \eta^{\frac{p+n}{2}+b+2} g(\nu)\,(1+\nu)^{-\frac{p}{2}} e^{-\frac{\eta S}{2}\left(1+\frac{F}{1+\nu}\right)} d\eta d\nu} \\ &= \frac{S}{n+p+2b+4} \frac{\int_0^\infty g(\nu)\,(1+\nu)^{-\frac{p}{2}} \left(1+\frac{F}{1+\nu}\right)^{-\frac{n+p}{2}-b-2} d\nu}{\int_0^\infty g(\nu)\,(1+\nu)^{-\frac{p}{2}} \left(1+\frac{F}{1+\nu}\right)^{-\frac{n+p}{2}-b-3} d\nu}. \end{aligned}$$

Using the change of variables $\lambda = \frac{1}{1+\nu}$, we have

$$\delta_B = \frac{S}{n+p+2b+4} \frac{\int_0^1 \lambda^{\frac{p}{2}-2} g\left(\frac{1-\lambda}{\lambda}\right) (1+\lambda F)^{-\frac{n+p}{2}-b-2} \, d\lambda}{\int_0^1 \lambda^{\frac{p}{2}-2} g\left(\frac{1-\lambda}{\lambda}\right) (1+\lambda F)^{-\frac{n+p}{2}-b-3} \, d\lambda}. \tag{2.2}$$

This estimator is of the form (2.1) with

$$\phi(F) = 1 - d\,(1 + r(F)),$$

where $d = \frac{n+2}{n+p+2b+4}$ and

$$r(F) = F \frac{\int_0^1 \lambda^{\frac{p}{2}-1} g\left(\frac{1-\lambda}{\lambda}\right) (1+\lambda F)^{-A} \, d\lambda}{\int_0^1 \lambda^{\frac{p}{2}-2} g\left(\frac{1-\lambda}{\lambda}\right) (1+\lambda F)^{-A} \, d\lambda}, \tag{2.3}$$

where $A = \frac{n+p}{2} + b + 3$.

To continue discussion, we need the following notations borrowed from Wells and Zhou (2008). Define the function $I_{\alpha,A,g}(F)$ as

$$I_{\alpha,A,g}(F) = \int_0^1 \lambda^{\alpha} (1+\lambda F)^{-A} g\left(\frac{1-\lambda}{\lambda}\right) d\lambda. \tag{2.4}$$

Using integration by parts, we obtain

$$\begin{aligned} F I_{\frac{p}{2}-1,A,g}(F) &= \int_0^1 \lambda^{\frac{p}{2}-1} g\left(\frac{1-\lambda}{\lambda}\right) d\left[\frac{(1+\lambda F)^{1-A}}{1-A}\right] \\ &= g(0) \frac{(1+F)^{1-A}}{1-A} \\ &\quad + \frac{\frac{p}{2}-1}{A-1} \int_0^1 (1+\lambda F)^{-A} (1+\lambda F) \lambda^{\frac{p}{2}-2} g\left(\frac{1-\lambda}{\lambda}\right) d\lambda \\ &\quad - \frac{1}{A-1} \int_0^1 (1+\lambda F)^{-A} (1+\lambda F) \frac{1}{\lambda^2} \lambda^{\frac{p}{2}-1} g\left(\frac{1-\lambda}{\lambda}\right) d\lambda. \end{aligned} \tag{2.5}$$

Also, we define the functions $J_a\left(g\left(\frac{F-u}{u}\right)\right)$ and $J_a\left(\frac{Au}{1+u} g\left(\frac{F-u}{u}\right)\right)$ as

$$J_a\left(g\left(\frac{F-u}{u}\right)\right) = \int_0^F u^a (1+u)^{-A} g\left(\frac{F-u}{u}\right) du = F^{a+1} I_{a,A,g}(F) \tag{2.6}$$

and

$$J_a\left(\frac{Au}{1+u}g\left(\frac{F-u}{u}\right)\right) = \int_0^F u^a(1+u)^{-A}\frac{Au}{1+u}g\left(\frac{F-u}{u}\right)du$$
$$= F^{a+1}\int_0^1 \lambda^a(1+\lambda F)^{-A}g\left(\frac{1-\lambda}{\lambda}\right)\frac{A\lambda F}{1+\lambda F}d\lambda, \tag{2.7}$$

respectively. By integration by parts, we have

$$\begin{aligned}J_a\left(\frac{Au}{1+u}g\left(\frac{F-u}{u}\right)\right) &= \int_0^F u^a(1+u)^{-A}\frac{Au}{1+u}g\left(\frac{F-u}{u}\right)du\\ &= -\int_0^F u^{a+1}g\left(\frac{F-u}{u}\right)d(1+u)^{-A}\\ &= -F^{a+1}g(0)(1+F)^{-A}+(a+1)\int_0^F(1+u)^{-A}u^a g\left(\frac{F-u}{u}\right)du\\ &\quad+\int_0^F(1+u)^{-A}u^{a+1}g'\left(\frac{F-u}{u}\right)\left(-\frac{F}{u^2}\right)du.\end{aligned} \tag{2.8}$$

To show $\phi(F)$ is a decreasing function in F, it is sufficient to show that $r(F)$ is an increasing function in F. The following lemma gives conditions under which $\widetilde{r}(F)=F^c r(F)$ is an increasing function in F.

Lemma 2.1. *If $\psi(\nu)=-(1+\nu)\frac{g'(\nu)}{g(\nu)}$ can be decomposed as $l_1(\nu)+l_2(\nu)$, where $l_1(\nu)$ is increasing in ν and $0\le l_2(\nu)\le c$, a constant, then $\widetilde{r}(F)=F^c r(F)$ is nondecreasing.*

Proof. The proof is similar to the proof of Lemma 3.2 in Wells and Zhou (2008). Differentiating $\widetilde{r}(F)=F^c r(F)$ with respect to F, we have

$$\frac{\partial\widetilde{r}(F)}{\partial F}=F^c\left(c\frac{r(F)}{F}+r'(F)\right)=F^c\left((1+c)R(F)+FR'(F)\right),$$

where $R(F)=\frac{r(F)}{F}$. Therefore, $\frac{\partial\widetilde{r}(F)}{\partial F}\ge 0$ is equivalent to

$$(1+c)\frac{I_{\frac{p}{2}-1,A,g}(F)}{I_{\frac{p}{2}-2,A,g}(F)}+F\frac{\left\{I'_{\frac{p}{2}-1,A,g}(F)\,I_{\frac{p}{2}-2,A,g}(F)-I'_{\frac{p}{2}-2,A,g}(F)\,I_{\frac{p}{2}-1,A,g}(F)\right\}}{I^2_{\frac{p}{2}-2,A,g}(F)}\ge 0,$$

which in turn is equivalent to

$$\begin{aligned}-FI'_{\frac{p}{2}-1,A,g}(F)\,I_{\frac{p}{2}-2,A,g}(F) &\le (1+c)I_{\frac{p}{2}-2,A,g}(F)\,I_{\frac{p}{2}-1,A,g}(F)\\ &\quad-FI'_{\frac{p}{2}-2,A,g}(F)\,I_{\frac{p}{2}-1,A,g}(F).\end{aligned} \tag{2.9}$$

Now, we see

$$-FI'_{a,A,g}(F)=\int_0^1\lambda^a(1+\lambda F)^{-A}g\left(\frac{1-\lambda}{\lambda}\right)\frac{A\lambda F}{1+\lambda F}d\lambda.$$

Using (2.6) and (2.7), (2.9) can be written as

$$\frac{J_{\frac{p}{2}-1}\left(\frac{Au}{1+u}g\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)\right)} \leq 1+c+\frac{J_{\frac{p}{2}-2}\left(\frac{Au}{1+u}g\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)\right)}. \tag{2.10}$$

By applying (2.8), (2.10) is equivalent to

$$\begin{aligned} & \frac{-F^{\frac{p}{2}}g(0)(1+F)^{-A}}{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)\right)}+\left(\frac{p}{2}\right)+\frac{\int_0^F u^{\frac{p}{2}-1}(1+u)^{-A}g\left(\frac{F-u}{u}\right)\left[\frac{g'\left(\frac{F-u}{u}\right)}{g\left(\frac{F-u}{u}\right)}\left(-\frac{F}{u}\right)\right]du}{\int_0^F u^{\frac{p}{2}-1}(1+u)^{-A}g\left(\frac{F-u}{u}\right)du} \\ \leq \; & 1+c+\frac{-F^{\frac{p}{2}-1}g(0)(1+F)^{-A}}{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)\right)}+\frac{p}{2}-1+\frac{\int_0^F u^{\frac{p}{2}-2}(1+u)^{-A}g\left(\frac{F-u}{u}\right)\left[\frac{g'\left(\frac{F-u}{u}\right)}{g\left(\frac{F-u}{u}\right)}\left(-\frac{F}{u}\right)\right]du}{\int_0^F u^{\frac{p}{2}-2}(1+u)^{-A}g\left(\frac{F-u}{u}\right)du}, \end{aligned}$$

which in turn is equivalent to

$$\begin{aligned} & \frac{-g(0)(1+F)^{-A}}{I_{\frac{p}{2}-1,A,g}(F)}+\frac{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)l_1\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)\right)}+\frac{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)l_2\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)\right)} \\ \leq \; & c+\frac{-g(0)(1+F)^{-A}}{I_{\frac{p}{2}-2,A,g}(F)}+\frac{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)l_1\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)\right)}+\frac{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)l_2\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)\right)}. \end{aligned} \tag{2.11}$$

Since $I_{\frac{p}{2}-1,A,g}(F) \leq I_{\frac{p}{2}-2,A,g}(F)$, we have

$$\frac{-g(0)(1+F)^{-A}}{I_{\frac{p}{2}-1,A,g}(F)} \leq \frac{-g(0)(1+F)^{-A}}{I_{\frac{p}{2}-2,A,g}(F)}.$$

Note also that $l_1(\nu)$ is increasing in ν implies that for all F fixed, $l_1\left(\frac{F-u}{u}\right)$ is decreasing in u. When $t < u$, we have

$$\frac{u^{\frac{p}{2}-2}(1+u)^{-A}g\left(\frac{F-u}{u}\right)1(u\leq F)}{t^{\frac{p}{2}-2}(1+t)^{-A}g\left(\frac{F-t}{t}\right)1(t\leq F)} \leq \frac{u^{\frac{p}{2}-1}(1+u)^{-A}g\left(\frac{F-u}{u}\right)1(u\leq F)}{t^{\frac{p}{2}-1}(1+t)^{-A}g\left(\frac{F-t}{t}\right)1(t\leq F)}.$$

By a monotone likelihood argument, we have

$$\begin{aligned} & \frac{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)l_1\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)\right)}=\frac{\int_0^F u^{\frac{p}{2}-1}(1+u)^{-A}g\left(\frac{F-u}{u}\right)l_1\left(\frac{F-u}{u}\right)du}{\int_0^F u^{\frac{p}{2}-1}(1+u)^{-A}g\left(\frac{F-u}{u}\right)du} \\ \leq \; & \frac{\int_0^F u^{\frac{p}{2}-2}(1+u)^{-A}g\left(\frac{F-u}{u}\right)l_1\left(\frac{F-u}{u}\right)du}{\int_0^F u^{\frac{p}{2}-2}(1+u)^{-A}g\left(\frac{F-u}{u}\right)du}=\frac{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)l_1\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)\right)} \end{aligned}$$

and

$$0 \leq \frac{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)l_2\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-2}\left(g\left(\frac{F-u}{u}\right)\right)} \leq c,\ 0 \leq \frac{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)l_2\left(\frac{F-u}{u}\right)\right)}{J_{\frac{p}{2}-1}\left(g\left(\frac{F-u}{u}\right)\right)} \leq c.$$

Thus, we have established the inequality (2.11) and the proof is complete.

□

The next lemma gives conditions under which a lower bound of $r(F)$ can be determined.

Lemma 2.2. *With the regularity conditions C1 and C2, assume that* $\psi(\nu) = -(1+\nu)\frac{g'(\nu)}{g(\nu)} \geq M$, *where* M *is a finite real number. For the* $r(F)$ *function, we have*

$$r(F) \geq \frac{\frac{p}{2}-1+M}{A-\frac{p}{2}-M}.$$

PROOF. The proof is similar to the proof of Lemma 3.1 in Wells and Zhou (2008). According to (2.3), we have

$$r(F) = F\frac{\int_0^1 \lambda^{\frac{p}{2}-1} g\left(\frac{1-\lambda}{\lambda}\right)(1+\lambda F)^{-A}\,d\lambda}{\int_0^1 \lambda^{\frac{p}{2}-2} g\left(\frac{1-\lambda}{\lambda}\right)(1+\lambda F)^{-A}\,d\lambda} = F\frac{I_{\frac{p}{2}-1,A,g}(F)}{I_{\frac{p}{2}-2,A,g}(F)}.$$

Using (2.4) and (2.5), we obtain

$$N_1 = \frac{1}{A-1}\int_0^1 (1+\lambda F)^{-A}\left(\frac{p}{2}-1\right)\lambda^{\frac{p}{2}-2} g\left(\frac{1-\lambda}{\lambda}\right)d\lambda = \frac{\frac{p}{2}-1}{A-1} I_{\frac{p}{2}-2,A,g}(F),$$

$$\begin{aligned}
N_2 &= \frac{1}{A-1}\int_0^1 (1+\lambda F)^{-A}\lambda^{\frac{p}{2}-2} g'\left(\frac{1-\lambda}{\lambda}\right)\left(\frac{-\lambda}{\lambda^2}\right)d\lambda \\
&= \frac{1}{A-1}\int_0^1 (1+\lambda F)^{-A}\lambda^{\frac{p}{2}-2} g\left(\frac{1-\lambda}{\lambda}\right)\left[\frac{g'\left(\frac{1-\lambda}{\lambda}\right)}{g\left(\frac{1-\lambda}{\lambda}\right)}\left(-\frac{1-\lambda}{\lambda}-1\right)\right]d\lambda \\
&= \frac{I_{\frac{p}{2}-2,A,g}(F)}{A-1}\frac{\int_0^1 \lambda^{\frac{p}{2}-2}(1+\lambda F)^{-A} g\left(\frac{1-\lambda}{\lambda}\right)\psi\left(\frac{1-\lambda}{\lambda}\right)d\lambda}{\int_0^1 \lambda^{\frac{p}{2}-2}(1+\lambda F)^{-A} g\left(\frac{1-\lambda}{\lambda}\right)d\lambda} \\
&\geq \frac{M}{A-1} I_{\frac{p}{2}-2,A,g}(F),
\end{aligned}$$

$$N_3 = \frac{\frac{p}{2}-1}{A-1} F I_{\frac{p}{2}-1,A,g}(F) = \frac{\left(\frac{p}{2}-1\right)r(F)}{A-1} I_{\frac{p}{2}-2,A,g}(F)$$

and

$$
\begin{aligned}
N_4 &= \frac{I_{\frac{p}{2}-2,A,g}(F)}{A-1}\frac{F\int_0^1 \lambda^{\frac{p}{2}-1}(1+\lambda F)^{-A} g'\left(\frac{1-\lambda}{\lambda}\right)\left(\frac{-1}{\lambda}\right)d\lambda}{I_{\frac{p}{2}-2,A,g}(F)} \\
&= \frac{I_{\frac{p}{2}-2,A,g}(F)}{A-1}\frac{F\int_0^1 (1+\lambda F)^{-A}\lambda^{\frac{p}{2}-1} g\left(\frac{1-\lambda}{\lambda}\right)\left[\frac{g'\left(\frac{1-\lambda}{\lambda}\right)}{g\left(\frac{1-\lambda}{\lambda}\right)}\left(-\frac{1-\lambda}{\lambda}-1\right)\right]d\lambda}{I_{\frac{p}{2}-2,A,g}(F)} \\
&= \frac{I_{\frac{p}{2}-2,A,g}(F)}{A-1}\frac{F\int_0^1 (1+\lambda F)^{-A}\lambda^{\frac{p}{2}-1} g\left(\frac{1-\lambda}{\lambda}\right)\psi\left(\frac{1-\lambda}{\lambda}\right)d\lambda}{I_{\frac{p}{2}-2,A,g}(F)} \\
&\geq \frac{Mr(F)}{A-1}I_{\frac{p}{2}-2,A,g}(F).
\end{aligned}
$$

Combining all the terms, we obtain the following inequality

$$(A-1)r(F) \geq \left(\frac{p}{2}-1\right) + M + \left(\frac{p}{2}-1\right)r(F) + Mr(F),$$

implying

$$r(F) \geq \frac{\frac{p}{2}-1+M}{A-\frac{p}{2}-M}.$$

Thus, we have the needed bound on the $r(F)$ function.

Now we use Lemmas 2.1 and 2.2 to show minimaxity of the generalized Bayes estimator δ_B in (2.2). In fact, this is our main result.

THEOREM 2.2. *If* $\psi(\nu) = -(1+\nu)\frac{g'(\nu)}{g(\nu)}$ *is increasing in* ν *and* $\psi(\nu) = -(1+\nu)\frac{g'(\nu)}{g(\nu)} \geq M$*, where* M *is a finite real number and also* $1 \leq d \left(1+\left(\frac{\frac{p}{2}-1+M}{A-\frac{p}{2}-M}\right)\right)$*, then* δ_B *in* (2.2) *is minimax under the loss function* (1.2).

PROOF. First, assume $l_2(\nu) = 0$ and $l_1(\nu) = \psi(\nu)$. By using Lemma 2.1 for the case $c = 0$, we see that $r(F)$ is an increasing function in F, hence $\phi(F)$ is deceasing in F. By (1.2), we have

$$\varphi^{BZ}(F) = 1 - \frac{n+2}{p+n+2}r_1(F),$$

where

$$r_1(F) = \frac{\int_0^1 \lambda^{\frac{p}{2}-1}(1+\lambda F)^{-\frac{n+p}{2}-1}d\lambda}{\int_0^1 \lambda^{\frac{p}{2}-1}(1+\lambda F)^{-\frac{n+p}{2}-2}d\lambda}.$$

Using the change of variables $u = \lambda F$, we have

$$r_1(F) = \frac{\int_0^F u^{\frac{p}{2}-1}(1+u)^{-\frac{n+p}{2}-1}du}{\int_0^F u^{\frac{p}{2}-1}(1+u)^{-\frac{n+p}{2}-2}du}.$$

Since $r_1(F)$ is increasing in F, we have

$$r_1(F) \leq \frac{\int_0^\infty u^{\frac{p}{2}-1}(1+u)^{-\frac{n+p}{2}-1}\,du}{\int_0^\infty u^{\frac{p}{2}-1}(1+u)^{-\frac{n+p}{2}-2}\,du} = \frac{p+n+2}{n+2}.$$

By Lemma 2.2, we obtain

$$\varphi(F) = 1 - d\,(1+r(F)) \leq 1 - d\left[1 + \left(\frac{\frac{p}{2}-1+M}{A-\frac{p}{2}-M}\right)\right].$$

Also we have

$$\varphi^{BZ}(F) = 1 - \frac{n+2}{p+n+2} r_1(F) \geq 1 - \frac{n+2}{p+n+2}\frac{p+n+2}{n+2} = 0.$$

Therefore, if $1 - d\left[1 + \left(\frac{\frac{p}{2}-1+M}{A-\frac{p}{2}-M}\right)\right] \leq 0$ is satisfied, then condition (II) in Theorem 2.2 is satisfied and hence δ_B in (2.2) is minimax under the loss function (1.2). □

3 Examples

In this section, we give three examples to which our results can be applied. We also make some connections to existing literature (Ghosh, 1994).

Example 1. *The class of priors studied by* Ghosh (1994) *(in the setup and notation of Section 2) corresponds to*

$$h(\nu,\eta) = k\eta^b (1+\nu)^{-b-2},\ \nu > 0,\ \eta > 0,$$

where k is a positive constant. If $M = b+2$ and $-\frac{p}{2}-1 < b \leq -1$, then we can show that the class of priors of Ghosh (1994) *satisfies the conditions and hence our results include the results of* Ghosh (1994).

Example 2. *Another class of prior distributions is*

$$h(\nu,\eta) = k\eta^b e^{-\nu},\ \nu > 0,\ \eta > 0,$$

where k is a positive constant. If $M = 1$, $p \geq 2$ and $-\frac{p+n}{2}-1 < b \leq -1$ then Theorem 2.2 is satisfied and the generalized Bayes estimator will be minimax under the loss function (1.2).

Example 3. *Consider the following prior distribution*

$$h(\nu,\eta) = k\eta^b (1+\nu)^{-a-c-2}\nu^c,\ \nu > 0,\ \eta > 0,$$

where k is a positive constant. If $M = a+c$, $c \geq 0$ and $-\frac{p+n}{2}-1 < b \leq 2+a+c$, then the conditions of Theorem 2.2 are satisfied, so the corresponding generalized Bayes estimators are minimax.

4 Simulation study

In this section, we compare the performance of the affine equivariant estimator δ_0 with the generalized Bayes estimator in Theorem 2.2. The comparison uses the following simulation scheme computing bias and mean squared error:

a) set values for n, θ and σ^2;

b) simulate a random sample of size n from a seven-dimensional normal distribution with mean vector θ and covariance matrix $\sigma^2 I_p$;

c) compute

$$S = \sum_{i=1}^{n} \sum_{j=1}^{7} \left(X_{ij} - \overline{X} \right)^2,$$

where

$$\overline{X} = \frac{\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{7} X_{ij}}{7n};$$

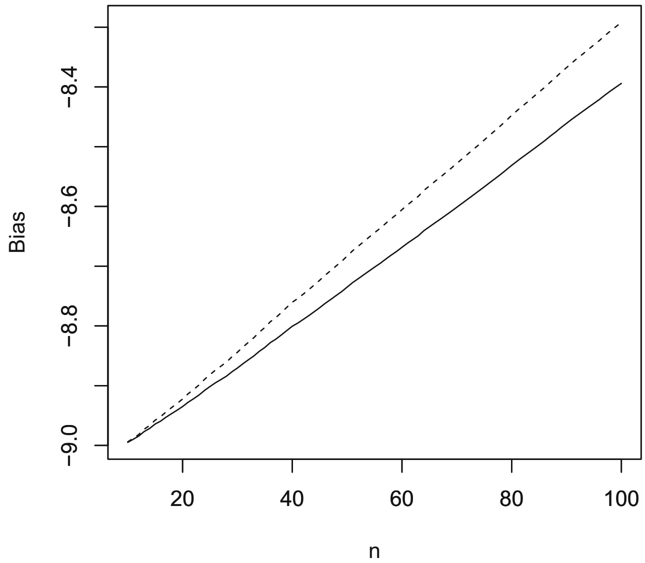

Figure 1: The biases of δ_0 (solid line) and δ_B (broken line) versus $n = 10, 11, \ldots, 100$ when $\theta = (0, 0, 0, 0, 0, 0, 0)$ and $\sigma^2 = 1$

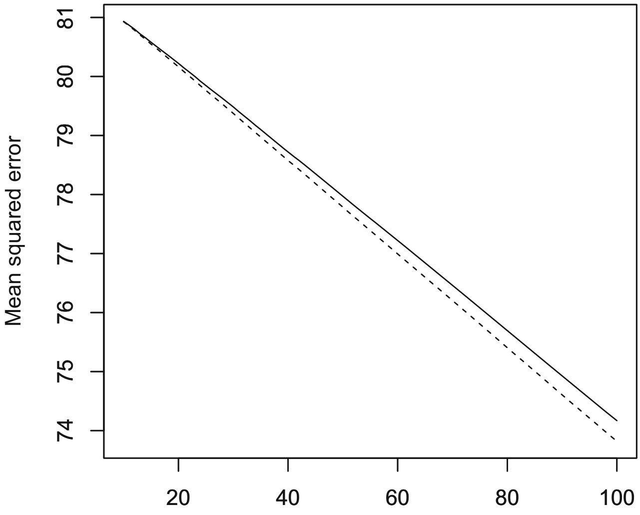

Figure 2: The mean squared errors of δ_0 (solid line) and δ_B (broken line) versus $n = 10, 11, \ldots, 100$ when $\theta = (0, 0, 0, 0, 0, 0, 0)$ and $\sigma^2 = 1$

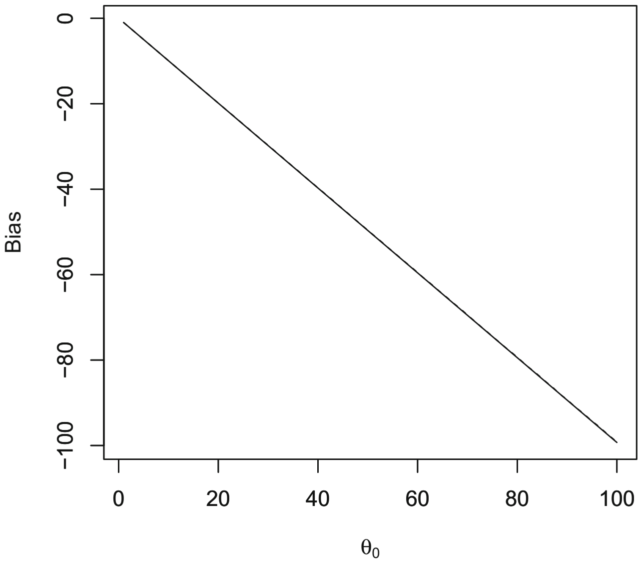

Figure 3: The biases of δ_0 (solid line) and δ_B (broken line) versus $\theta_0 = -50, 49, \ldots, 50$ when $\theta = (\theta_0, \theta_0, \theta_0, \theta_0, \theta_0, \theta_0, \theta_0)$, $n = 100$ and $\sigma^2 = 1$

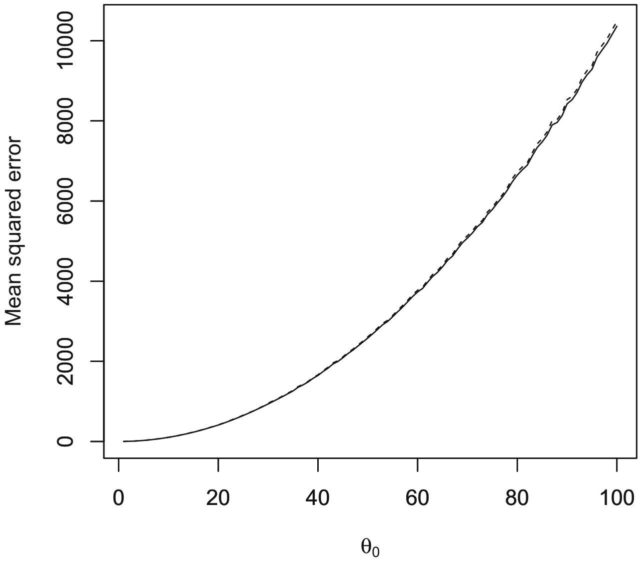

Figure 4: The mean squared errors of δ_0 (solid line) and δ_B (broken line) versus $\theta_0 = -50, 49, \ldots, 50$ when $\theta = (\theta_0, \theta_0, \theta_0, \theta_0, \theta_0, \theta_0, \theta_0)$, $n = 100$ and $\sigma^2 = 1$

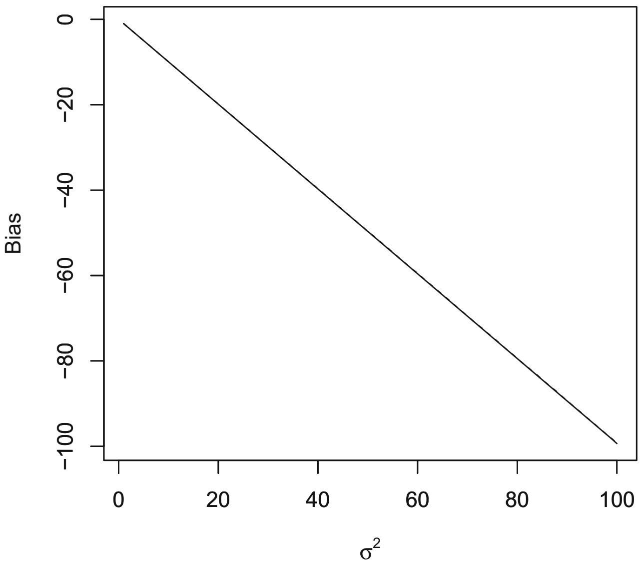

Figure 5: The biases of δ_0 (solid line) and δ_B (broken line) versus $\sigma^2 = 1, 2, \ldots, 100$ when $\theta = (0, 0, 0, 0, 0, 0, 0)$ and $n = 100$

d) compute the equivariant estimator δ_0 by $\delta_0 = \frac{S}{n+2}$;

e) compute the estimator δ_B given by Theorem 2.2 with $b = -2$ and $g(\nu) = e^{-\nu}$;

f) repeat steps b) to e) one thousand times;

g) compute the biases of the estimators as

$$\text{bias}\,(\delta_0) = \frac{1}{1000} \sum_{i=1}^{1000} \left(\delta_{0,i} - \sigma^2\right)$$

and

$$\text{bias}\,(\delta_B) = \frac{1}{1000} \sum_{i=1}^{1000} \left(\delta_{B,i} - \sigma^2\right),$$

where $\delta_{0,i}$ and $\delta_{B,i}$ denote the estimates of δ_0 and δ_B, respectively, in the ith iteration;

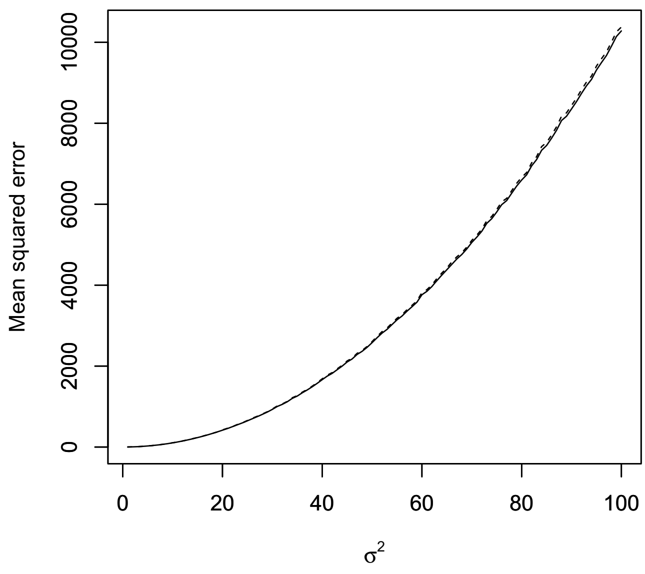

Figure 6: The mean squared errors of δ_0 (solid line) and δ_B (broken line) versus $\sigma^2 = 1, 2, \ldots, 100$ when $\theta = (0, 0, 0, 0, 0, 0, 0)$ and $n = 100$

h) compute the mean squared errors of the estimators as

$$\operatorname{mse}\left(\delta_0\right)=\frac{1}{1000}\sum_{i=1}^{1000}\left(\delta_{0,i}-\sigma^2\right)^2$$

and

$$\operatorname{mse}\left(\delta_B\right)=\frac{1}{1000}\sum_{i=1}^{1000}\left(\delta_{B,i}-\sigma^2\right)^2.$$

Plots of the biases and mean squared errors versus $n = 10, 11, \ldots, 100$ when $\theta = (0, 0, 0, 0, 0, 0, 0)$ and $\sigma^2 = 1$ are shown in Figs. 1 and 2. Plots of the biases and mean squared errors versus $\theta_0 = -50, 49, \ldots, 50$ when $\theta = (\theta_0, \theta_0, \theta_0, \theta_0, \theta_0, \theta_0, \theta_0)$, $n = 100$ and $\sigma^2 = 1$ are shown in Figs. 3 and 4. Plots of the biases and mean squared errors versus $\sigma^2 = 1, 2, \ldots, 100$ when $\theta = (0, 0, 0, 0, 0, 0, 0)$ and $n = 100$ are shown in Figs. 5 and 6.

We can observe the following from Figs. 1 to 6. The biases are generally negative for both estimators. The biases approach zero in magnitude as n increases. δ_B has smaller bias for every n. The mean squared errors approach zero as n increases. δ_B has smaller mean squared error for every n. The biases decrease from being positive to negative as θ_0 increases from -50 to 50. The biases are smallest in magnitude when $\theta_0 = 0$. The mean squared errors take a parabolic shape as θ_0 increases from -50 to 50. The mean squared errors are smallest when $\theta_0 = 0$. The biases are negative and decrease as σ^2 increases from 1 to 100. The biases are smallest in magnitude when $\sigma^2 = 1$. The mean squared errors increase as σ^2 increases from 1 to 100. The mean squared errors are smallest when $\sigma^2 = 1$.

The computations were performed using the R software (R Development Core Team, 2023) and the package mvtnorm (Genz et al., 2021). The codes used are given in the Appendix A.

Acknowledgments. The authors would like to thank the Editor, the Associate Editor and the referee for careful reading and comments which greatly improved the paper.

Code Availability. The code are in the Appendix A.

Compliance with Ethical Standards.

Conflict of Interest. Authors declare no conflicts of interest.

Consent for Publication. All authors gave explicit consent to publish this manuscript.

References

BREWSTER, J. F. and ZIDEK, J. C. (1974). Improving on equivariant estimators. *Ann. Stat.* **2**, 21–38.

GENZ, A., BRETZ, F., MIWA, T., MI, X., LEISCH, F., SCHEIPL, F., BORNKAMP, B., MAECHLER, M. and HOTHORN, T. (2021). mvtnorm: Multivariate normal and t distributions. *R package version* **1**, 1–3.

GHOSH, M. (1994). *On some Bayesian solutions of the Neyman-Scott problem,*.

KUBOKAWA, T. (1994). An unified approach to improving equivariant estimators. *Ann. Stat.* **22**, 290–299.

MARUYAMA, Y. and STRAWDERMAN, W. E. (2006). A new class of minimax generalized Bayes estimators of a normal variance. *Journal of Statistical Planning and Inference* **136**, 3822–3836.

R DEVELOPMENT CORE TEAM (2023). A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Stat. Math.* **16**, 155–160.

STRAWDERMAN, W. E. (1974). Minimax estimation of powers of the variance of a normal population under squared error loss. *Ann. Stat.* **2**, 190–198.

WELLS, T. M. and ZHOU, G. (2008). Generalized Bayes minimax estimators of the mean of multivariate normal distribution with unknown variance. *J. Multivar. Anal.* **99**, 2208–2220.

Appendix A. : R codes

```
###############################################################
# computes the bias and mean squared error with respect to n #
###############################################################

nn=seq(10,100)
bias1=nn
bias2=nn
mse1=nn
mse2=nn
nsim=1000
est1=rep(0,nsim)
est2=est1

for (n in seq(10,100))
{for (i in 1:nsim) {x=rmvnorm(n,mean=rep(0,7),sigma=diag(7))
mm=mean(x)
```

```
S=sum((x-mm)**2)
tt=0
for (i in seq(1,n)) tt=tt+sum((x[i,])**2)
F=tt/S

f1=function (x) {x**(3/2)*exp((x-1)/x)*(1+x*F)**(-(n+7)/2)}
f2=function (x) {x**(3/2)*exp((x-1)/x)*(1+x*F)**(-(n+7)/2-1)}

est1[i]=S/(n+2)
est2[i]=S*integrate(f1,lower=0,upper=1)$value/
((n+7)*integrate(f2,lower=0,upper=1)$value)}

bias1[n-9]=mean(est1-9)
bias2[n-9]=mean(est2-9)

mse1[n-9]=mean((est1-9)**2)
mse2[n-9]=mean((est2-9)**2)
}

####################################################################
# computes the bias and mean squared error with respect to sigma #
####################################################################

nn=seq(1,100)
bias1=nn
bias2=nn
mse1=nn
mse2=nn
nsim=1000
est1=rep(0,nsim)
est2=est1

n=100
for (s in seq(1,100))
{for (i in 1:nsim)
{x=rmvnorm(n,mean=rep(0,7),sigma=(s*diag(7)))
mm=mean(x)

S=sum((x-mm)**2)
tt=0 for (i in seq(1,n)) tt=tt+sum((x[i,])**2)
F=tt/S

f1=function (x) {x**(3/2)*exp((x-1)/x)*(1+x*F)**(-(n+7)/2)}
f2=function (x) {x**(3/2)*exp((x-1)/x)*(1+x*F)**(-(n+7)/2-1)}

est1[i]=S/(n+2)
est2[i]=S*integrate(f1,lower=0,upper=1)$value/
((n+7)*integrate(f2,lower=0,upper=1)$value)}

bias1[s]=mean(est1-s)
bias2[s]=mean(est2-s)

mse1[s]=mean((est1-s)**2)
mse2[s]=mean((est2-s)**2)
}
```

```
#####################################################################
#  computes the bias and mean squared error with respect to theta #
#####################################################################

nn=seq(1,101)
bias1=nn
bias2=nn
mse1=nn
mse2=nn
nsim=1000
est1=rep(0,nsim)
est2=est1

n=100
for (mu in seq(-50,50))
{for (i in 1:nsim)
{x=rmvnorm(n,mean=rep(mu,7),sigma=(diag(7)))
mm=mean(x)

S=sum((x-mm)**2)
tt=0
for (i in seq(1,n)) tt=tt+sum((x[i,])**2)
F=tt/S

f1=function (x) {x**(3/2)*exp((x-1)/x)*(1+x*F)**(-(n+7)/2)}
f2=function (x) {x**(3/2)*exp((x-1)/x)*(1+x*F)**(-(n+7)/2-1)}

est1[i]=S/(n+2)
est2[i]=S*integrate(f1,lower=0,upper=1)$value/
((n+7)*integrate(f2,lower=0,upper=1)$value)}

bias1[mu+51]=mean(est1-mu)
bias2[mu+51]=mean(est2-mu)

mse1[mu+51]=mean((est1-mu)**2)
mse2[mu+51]=mean((est2-mu)**2)
}
```

Shokofeh Zinodiny
Institute for Research in
Fundamental Sciences, Tehran, Iran
E-mail: shokofeh_zinodiny@yahoo.com

Saralees Nadarajah
Howard University, Washington DC,
USA
E-mail: mbbsssn2@manchester.ac.uk

Paper received: 23 October 2022; accepted 20 March 2023.