General Dominance Properties of Double Shrinkage Estimators for Ratio of Positive Parameters

Tatsuya Kubokawa*
University of Tokyo

September 1, 2013

Abstract

In estimation of ratio of variances in two normal distributions with unknown means, it has been shown in the literature that simple and crude ratio estimators based on two sample variances are dominated by shrinkage estimators using information contained in sample means. Of these, a natural double shrinkage estimator is the ratio of shrinkage estimators of variances, but its improvement over the crude ratio estimator depends on loss functions, namely, the improvement has not been established except the Stein loss function.

In this paper, this dominance property is shown for some convex loss functions including the Stein and quadratic loss functions in the general framework of distributions with positive parameters and shrinkage estimators. The resulting new finding is that the generalized Bayes estimator of the ratio of variances dominates the crude ratio estimator relative to the quadratic loss. The paper also shows that the dominance property of the double shrinkage estimator holds for estimation of the difference of variances, but it does not hold for estimation of the product and sum of variances. Finally, it is demonstrated that the double shrinkage estimators for the ratio, product, sum and differences of variances are connected to estimation of linear combinations of the normal positive means, and the dominance and non-dominance results of the double shrinkage estimators coincide with the corresponding dominance results in estimation of linear combinations of means.

Key words and phrases: Decision theory, generalized Bayes estimator, improved estimation, minimaxity, quadratic loss, ratio, Stein estimator, Stein loss, variance.

1 Introduction

The estimation of a scale parameter in the presence of another nuisance parameters has been studied in the literature since Stein (1964) established the surprising result that in a normal population with unknown means, the estimator of the variance based on the sample

*Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, JAPAN. E-Mail: tatsuya@e.u-tokyo.ac.jp

variance with the optimal multiple, which is the best affine equivariant, is inadmissible and improved on by the Stein truncated estimator using information contained in the sample mean. Of these, Brown (1968), Brewster and Zidek (1974), Strawderman (1974) and Shinozaki (1995) developed other types of improved estimators. Since most improved estimators are smaller than or equal to the best affine equivariant, we call here them shrinkage estimators, and a class of improved and shrinkage estimators was derived by Kubokawa (1994a).

An inherited problem is the estimation of ratio $\rho = \theta_2/\theta_1$ for two scale parameters θ_1 and θ_2. A possible improvement is the single shrinkage estimators $\hat{\theta}_2^*/\hat{\theta}_1$ and $\hat{\theta}_2/\hat{\theta}_1^*$, where $\hat{\theta}_2^*$ and $1/\hat{\theta}_1^*$ are shrinkage estimators of θ_2 and $1/\theta_1$, respectively, improving on the crude estimators $\hat{\theta}_2$ and $1/\hat{\theta}_1$. The dominance results of such single shrinkage estimators were shown by Gelfand and Dey (1988) and Ghosh and Kundu (1996). An interesting issue is whether the single shrinkage estimators can be further improved on by a double shrinkage estimator. For the quadratic loss function $L_q(\hat{\rho}/\rho) = (\hat{\rho}/\rho - 1)^2$ for estimator $\hat{\rho}$ of ρ, Kubokawa(1994b) demonstrated that the single shrinkage estimators can be improved on by a double shrinkage estimator of the form

$$\hat{\rho}^{**} = \hat{\theta}_2^*/\hat{\theta}_1 + \hat{\theta}_2/\hat{\theta}_1^* - \hat{\theta}_2/\hat{\theta}_1.$$

For the Stein loss function $L_s(\hat{\rho}/\rho) = \hat{\rho}/\rho - \log(\hat{\rho}/\rho) - 1$, Kubokawa and Srivastava (1996) and Iliopoulos and Kourouklis (1999) showed that the single shrinkage estimators can be improved on by another type of a double shrinkage estimator

$$\hat{\rho}^* = \hat{\theta}_2^*/\hat{\theta}_1^*.$$

Bobotas, Iliopoulos and Kourouklis (2012) developed a very nice unified theory, and clarified conditions on loss functions under which the single shrinkage estimators can be dominated by $\hat{\rho}^*$ and/or $\hat{\rho}^{**}$. We are inspired from these dominance results to raise the following queries about the double shrinkage estimators.

(I) The double shrinkage estimator $\hat{\rho}^*$ has a natural form, but it could not be shown that $\hat{\rho}^*$ dominates the single shrinkage estimators relative to the quadratic loss. Does this suggest that $\hat{\rho}^*$ cannot dominate the crude ratio estimator $\hat{\theta}_2/\hat{\theta}_1$? That is, we want to investigate whether the dominance property of $\hat{\rho}^*$ over $\hat{\theta}_2/\hat{\theta}_1$ holds for the quadratic loss.

(II) As the related problems, we can consider estimation of the product $\theta_1\theta_2$, the difference $\theta_1 - \theta_2$ and the sum $\theta_1 + \theta_2$. Can we extend the dominance property of double shrinkage ratio estimators to the estimation of such parameters? That is, we want to investigate whether their double shrinkage estimators dominate the corresponding crude estimators.

The objective of this paper is to reply to these queries. In Section 2, we show the dominance of $\hat{\rho}^*$ over the crude ratio estimator relative to some convex loss functions including the Stein and quadratic loss functions. It is noted that the dominance results hold in quite general setups as given in (A1), (A2), (A3) and (A4), namely, we do not assume any distributional assumptions except that $\hat{\theta}_i^* \leq \hat{\theta}_i$ and $\hat{\theta}_i^*$ dominates $\hat{\theta}_i$ in estimation of θ_i for $i = 1, 2$. The dominance results will be applied in Section 3 to two sample problems of normal populations.

The query (II) is studied in Section 4. For estimation of the difference $\theta_1 - \theta_2$, we can get a similar dominance result as in the case of the ratio estimation. For estimation of the product $\theta_1\theta_2$ and the sum $\theta_1 + \theta_2$, however, the corresponding double shrinkage estimators cannot necessarily dominate the crude estimators. Especially, the generalized Bayes estimator $\hat{\theta}_1^{GB}\hat{\theta}_2^{GB}$ of the product never dominates the crude estimator $\hat{\theta}_1\hat{\theta}_2$ although the generalized Bayes estimator $\hat{\theta}_i^{GB}$ can dominate $\hat{\theta}_i$ in the framework of estimation of the individual parameter θ_i.

The above explanations mean that the estimation of the ratio and difference of two positive parameters has a different dominance story from the estimation of the product and sum. In Section 5, using the same arguments as in Rukhin (1992), we show that the double shrinkage estimators of the ratio, product, sum and difference are connected to estimation of the sum and difference of positive normal means. It is confirmed that the dominance and non-dominance results derived in this paper coincide with the decision-theoretic properties given by Kubokawa (2012) in the framework of estimation of the sum and difference of positive normal means.

2 General Dominance Results in Estimation of Ratio of Positive Parameters

Let θ_1 and θ_2 be positive unknown parameters. For $i = 1, 2$, let $\hat{\theta}_i$ and $\hat{\theta}_i^*$ be positive estimators of θ_i satisfying the following assumptions:

(A1) $(\hat{\theta}_1, \hat{\theta}_1^*)$ is independent of $(\hat{\theta}_2, \hat{\theta}_2^*)$.

(A2) $\hat{\theta}_1^*$ and $\hat{\theta}_2^*$ are shrinkage estimators of $\hat{\theta}_1$ and $\hat{\theta}_2$, respectively, satisfying that $\hat{\theta}_1^* \leq \hat{\theta}_1$ and $\hat{\theta}_2^* \leq \hat{\theta}_2$, where the strict inequalities hold with positive probabilities.

Consider the estimation of ratio of the positive parameters $\rho = \theta_2/\theta_1$. To evaluate an estimator $\tilde{\rho}$ of ρ, we begin by treating the risk function relative to the Stein loss function $L_s(\tilde{\rho}/\rho)$ for $L_s(t) = t - \log(t) - 1$, namely the risk function is given by

$$R_s(\omega, \tilde{\rho}) = E_\omega[\tilde{\rho}/\rho - \log(\tilde{\rho}/\rho) - 1] = E_\omega[\tilde{\rho}\theta_1/\theta_2 - \log(\tilde{\rho}\theta_1/\theta_2) - 1],$$

where ω is a collection of unknown parameters. When $\hat{\theta}_2$ and $1/\hat{\theta}_1$ are improved on by $\hat{\theta}_2^*$ and $1/\hat{\theta}_1^*$, respectively, as estimators of θ_2 and $1/\theta_1$, we want to investigate whether $\hat{\rho} = \hat{\theta}_2/\hat{\theta}_1$ can be improved on by the double shrinkage estimator $\hat{\rho}^* = \hat{\theta}_2^*/\hat{\theta}_1^*$ relative to the Stein loss. To establish the dominance property for the Stein loss function, we assume the following conditions for $\hat{\theta}_i$ and $\hat{\theta}_i^*$:

(A3) $E_\omega[\hat{\theta}_2/\theta_2] = 1$, $E_\omega[\theta_1/\hat{\theta}_1] = 1$ and

$$E_\omega[L_s(\hat{\theta}_2/\theta_2)] \geq E_\omega[L_s(\hat{\theta}_2^*/\theta_2)], \quad E_\omega[L_s(\theta_1/\hat{\theta}_1)] \geq E_\omega[L_s(\theta_1/\hat{\theta}_1^*)]$$

for any ω.

Theorem 2.1 *Assume conditions* (A1), (A2) *and* (A3). *Then, the double shrinkage estimator $\hat{\rho}^* = \hat{\theta}_2^*/\hat{\theta}_2^*$ dominates the estimator $\hat{\rho} = \hat{\theta}_2/\hat{\theta}_1$ relative to the Stein loss, namely,*

$$R_s(\omega, \hat{\rho}) \geq R_s(\omega, \hat{\rho}^*) \tag{2.1}$$

for any ω.

Proof. For notational simplicity, let $X_i = \hat{\theta}_i/\theta_i$ and $Y_i = \hat{\theta}_i^*/\theta_i$ for $i = 1, 2$. Assumption (A3) is expressed as $E_\omega[X_2] = 1$, $E_\omega[1/X_1] = 1$ and

$$\begin{aligned} E_\omega[X_2 - \log(X_2) - 1] \geq & E_\omega[Y_2 - \log(Y_2) - 1], \\ E_\omega[1/X_1 + \log(X_1) - 1] \geq & E_\omega[1/Y_1 + \log(Y_1) - 1], \end{aligned} \tag{2.2}$$

for any ω. Since $E_\omega[X_2] = 1$ and $E_\omega[1/X_1] = 1$, the inequalities in (2.2) imply that

$$\begin{aligned} E_\omega[\log(Y_2)] \geq & E_\omega[Y_2 + \log(X_2) - 1], \\ -E_\omega[\log(Y_1)] \geq & E_\omega[1/Y_1 - \log(X_1) - 1]. \end{aligned} \tag{2.3}$$

The difference of the risk functions of $\hat{\rho}$ and $\hat{\rho}^*$ is written as

$$\begin{aligned} \Delta_s(\omega) = & R_s(\omega, \hat{\rho}) - R_s(\omega, \hat{\rho}^*) \\ = & E_\omega[X_2/X_1 - \log(X_2/X_1) - 1] - E_\omega[Y_2/Y_1 - \log(Y_2/Y_1) - 1] \\ = & E_\omega[1 - \log(X_2/X_1) - Y_2/Y_1 + \log(Y_2) - \log(Y_1)], \end{aligned}$$

since $E_\omega[X_2/X_1] = E_\omega[X_2]E_\omega[1/X_1] = 1$ from (A1) and (A3). Applying the inequalities in (2.3) for $E_\omega[\log(Y_2)]$ and $-E_\omega[\log(Y_1)]$, we can evaluate the risk difference $\Delta_s(\omega)$ as

$$\begin{aligned} \Delta_s(\omega) \geq & E_\omega[1 - \log(X_2/X_1) - Y_2/Y_1 + Y_2 + \log(X_2) - 1 + 1/Y_1 - \log(X_1) - 1] \\ = & E_\omega[-Y_2/Y_1 + 1/Y_1 + Y_2 - 1] \\ = & E_\omega[1 - Y_2]E_\omega[1/Y_1 - 1]. \end{aligned}$$

From (A2), it follows that $Y_1 \leq X_1$ and $Y_2 \leq X_2$, which imply that $E_\omega[1 - Y_2] \geq E_\omega[1 - X_2] = 0$ and $E_\omega[1/Y_1 - 1] \geq E_\omega[1/X_1 - 1] = 0$. Hence, it is concluded that $\Delta_s(\omega) \geq 0$ for any ω. □

The dominance property given in Theorem 2.1 can be provided as a simple conclusion of Kubokawa and Srivastava (1996), Iliopoulos and Kourouklis (1999) and Bobotas, *et al.* (2012). However, Theorem 2.1 proves (2.1) without any distributional assumptions as long as (A1), (A2) and (A3) are assumed.

We next treat the quadratic loss function $L_q(\tilde{\rho}/\rho)$ for $L_q(t) = (t - 1)^2$. In this case, instead of (A3), we assume the following condition:

(A4) $E_\omega[(\hat{\theta}_2/\theta_2)^2] = E_\omega[\hat{\theta}_2/\theta_2]$, $E_\omega[(\theta_1/\hat{\theta}_1)^2] = E_\omega[\theta_1/\hat{\theta}_1]$ and

$$E_\omega[L_q(\hat{\theta}_2/\theta_2)] \geq E_\omega[L_q(\hat{\theta}_2^*/\theta_2)], \quad E_\omega[L_q(\theta_1/\hat{\theta}_1)] \geq E_\omega[L_q(\theta_1/\hat{\theta}_1^*)]$$

for any ω.

Theorem 2.2 *Assume conditions* (A1), (A2) *and* (A4). *Then, the double shrinkage estimator $\hat{\rho}^* = \hat{\theta}_2^*/\hat{\theta}_2^*$ dominates the estimator $\hat{\rho} = \hat{\theta}_2/\hat{\theta}_1$ relative to the quadratic loss, namely,*

$$R_q(\omega, \hat{\rho}) \geq R_q(\omega, \hat{\rho}^*) \tag{2.4}$$

for any ω.

Proof. For $X_i = \hat{\theta}_i/\theta_i$ and $Y_i = \hat{\theta}_i^*/\theta_i$, $i = 1, 2$, assumption (A4) is expressed as $E_\omega[X_2^2] = E_\omega[X_2]$, $E_\omega[1/X_1^2] = E_\omega[1/X_1]$ and

$$\begin{aligned} E_\omega[(X_2 - 1)^2] \geq & E_\omega[(Y_2 - 1)^2], \\ E_\omega[(1/X_1 - 1)^2] \geq & E_\omega[(1/Y_1 - 1)^2], \end{aligned} \tag{2.5}$$

for any ω. Since $E_\omega[X_2^2] = E_\omega[X_2]$ and $E_\omega[1/X_1^2] = E_\omega[1/X_1]$, the inequalities in (2.5) imply that

$$\begin{aligned} E_\omega[Y_2^2] \leq & E_\omega[2Y_2 - X_2], \\ E_\omega[1/Y_1^2] \leq & E_\omega[2/Y_1 - 1/X_1]. \end{aligned} \tag{2.6}$$

The difference of the risk functions of $\hat{\rho}$ and $\hat{\rho}^*$ is written as

$$\begin{aligned} \Delta_q(\omega) = & R_q(\omega, \hat{\rho}) - R_q(\omega, \hat{\rho}^*) \\ = & E_\omega[X_2^2/X_1^2 - 2X_2/X_1 + 1] - E_\omega[Y_2^2/Y_1^2 - 2Y_2/Y_1 + 1] \\ = & E_\omega[-X_2/X_1 - Y_2^2/Y_1^2 + 2Y_2/Y_1], \end{aligned}$$

since $E_\omega[X_2^2/X_1^2] = E_\omega[X_2^2]E_\omega[1/X_1^2] = E_\omega[X_2]E_\omega[1/X_1]$ from (A1) and (A4). Applying the inequalities in (2.6) for $E_\omega[Y_2^2]$ and $E_\omega[Y_1^2]$, we get that

$$\begin{aligned} E_\omega[Y_2^2/Y_1^2] = & E_\omega[Y_2^2]E_\omega[1/Y_1^2] \\ \leq & E_\omega[2Y_2 - X_2]E_\omega[2/Y_1 - 1/X_1] \\ = & E_\omega[(2Y_2 - X_2)(2/Y_1 - 1/X_1)] \end{aligned}$$

from the independence between (X_1, Y_1) and (X_2, Y_2). Thus, the risk difference $\Delta_q(\omega)$ is evaluated as

$$\begin{aligned} \Delta_q(\omega) \geq & E_\omega[-X_2/X_1 + 2Y_2/Y_1 - (2Y_2 - X_2)(2/Y_1 - 1/X_1)] \\ = & 2E_\omega[-Y_2/Y_1 + X_2/Y_1 + Y_2/X_1 - X_2/X_1] \\ = & 2E_\omega[X_2 - Y_2]E_\omega[1/Y_1 - 1/X_1]. \end{aligned}$$

From (A2), it follows that $Y_1 \leq X_1$ and $Y_2 \leq X_2$, which imply that $E_\omega[X_2 - Y_2] \geq 0$ and $E_\omega[1/Y_1 - 1/X_1] \geq 0$. Hence, it is concluded that $\Delta_q(\omega) \geq 0$ for any ω. □

The dominance property (2.4) of the double shrinkage estimator $\hat{\rho}^*$ relative to the quadratic loss is a new finding which we want to show. This gives an answer to the query (I) raised in Section 1. It is interesting to note that Theorem 2.2 proves (2.4) without any distributional assumptions as long as (A1), (A2) and (A4) are assumed.

Remark 2.1 Theorems 2.1 and 2.2 imply that the crude ratio estimator $\hat{\rho} = \hat{\theta}_2/\hat{\theta}_1$ can be improved on by the single shrinkage estimators $\hat{\rho}_1^* = \hat{\theta}_2/\hat{\theta}_1^*$ and $\hat{\rho}_2^* = \hat{\theta}_2^*/\hat{\theta}_1$ relative to the Stein and the quadratic loss functions. An interesting query is whether the single shrinkage estimators can be further improved on by the double shrinkage estimator $\hat{\rho}^* = \hat{\theta}_2^*/\hat{\theta}_1^*$. This dominance property is correct for the Stein loss. In fact, the same arguments as in the proof of Theorem 2.1 can be used to verify that the single shrinkage and improved estimators are dominated by the double shrinkage estimator. This is an extension of Kubokawa and Srivastava (1996) and Iliopoulos and Kourouklis (1999) to the general framework.

For the quadratic loss, however, the same arguments as in the proof of Theorem 2.2 do not work to show the stronger dominance property. Thus, we could not answer whether the stronger dominance property holds or not for the quadratic loss. In stead of $\hat{\rho}^* = \hat{\theta}_2^*/\hat{\theta}_1^*$, Kubokawa (1994b) and Bobotas, *et al*. (2012) showed that the double shrinkage estimator of the form

$$\hat{\rho}^{**} = \hat{\theta}_2^*/\hat{\theta}_1 + \hat{\theta}_2/\hat{\theta}_1^* - \hat{\theta}_2/\hat{\theta}_1$$

dominates the single shrinkage and improved estimators $\hat{\theta}_2^*/\hat{\theta}_1$ and $\hat{\theta}_2/\hat{\theta}_1^*$ relative to the quadratic loss. □

Remark 2.2 The dominance property of the double shrinkage estimator can be confirmed for other loss functions. For example, consider the log-transformed quadratic loss function $L_{LTQ}(\tilde{\rho}/\rho)$ for $L_{LTQ}(t) = (\log t)^2$. In this case, instead of (A3) and (A4), it is assumed that for $i = 1, 2$, $E_\omega[\log(\hat{\theta}_i/\theta_i)] = 0$ and

$$E_\omega[L_{LTQ}(\hat{\theta}_i/\theta_i)] \geq E_\omega[L_{LTQ}(\hat{\theta}_i^*/\theta_i)],$$

for any ω. Then, under assumptions (A1) and (A2), it can be shown that the estimator $\hat{\rho} = \hat{\theta}_2/\hat{\theta}_1$ is dominated by $\hat{\rho}^* = \hat{\theta}_2^*/\hat{\theta}_1^*$ relative to the log-transformed quadratic loss function.

When we consider the dual Stein loss function $L_{DS}(\tilde{\rho}/\rho)$ for $L_{DS}(t) = t + t^{-1} - 2 = L_s(t) + L_s(t^{-1})$, instead of (A3) and (A4), we assume that for $i = 1, 2$, $E_\omega[\hat{\theta}_i/\theta_i] = E_\omega[\theta_i/\hat{\theta}_i]$ and

$$E_\omega[L_{DS}(\hat{\theta}_i/\theta_i)] \geq E_\omega[L_{DS}(\hat{\theta}_i^*/\theta_i)],$$

for any ω. Then, under assumptions (A1) and (A2), it can be shown that the estimator $\hat{\rho} = \hat{\theta}_2/\hat{\theta}_1$ is dominated by $\hat{\rho}^* = \hat{\theta}_2^*/\hat{\theta}_1^*$ relative to the dual Stein loss function.

As shown above, the dominance property of the double shrinkage estimator holds for the Stein, quadratic, log-transformed quadratic and dual Stein loss functions. These results suggest the interesting conjecture that the dominance property would be established relative to convex loss functions. This will be studied as a future work. □

3 Applications to Estimation of Ratio of Normal Variances

We now apply the results given in the previous section to the estimation of ratio of the variances in two normal distributions. Let $\boldsymbol{X}_i$ and V_i, $i = 1, 2$, be mutually independent random variables such that $\boldsymbol{X}_i \sim \mathcal{N}_{p_i}(\boldsymbol{\mu}_i, \sigma_i^2 \boldsymbol{I}_{p_i})$ and $V_i/\sigma_i^2 \sim \chi^2_{n_i}$. In the framework of estimating each σ_i^2, $i = 1, 2$, a crude estimator is of the form $\hat{\sigma}_i^{2M} = c_i V_i$ for the optimal constant c_i, which is given in terms of minimizing an estimation error. The estimator $\hat{\sigma}_i^{2M}$ can be improved on by shrinkage estimators $\hat{\sigma}_i^{2*}$ using information contained in

$\boldsymbol{X}_i$. Typical shrinkage and improved estimators have been provided by Stein (1964), Brown (1968), Brewster and Zidek (1974) and Strawderman (1974). These procedures for improvement can be used to give double shrinkage estimators in estimation of the variance ratio $\rho = \sigma_2^2/\sigma_1^2$. In fact, Kubokawa and Srivastava (1996) and Iliopoulos and Kourouklis (1999) established that the double shrinkage estimator $\hat{\rho}^* = \hat{\sigma}_2^{2*}/\hat{\sigma}_1^{2*}$ dominates $\hat{\rho}^M = \hat{\sigma}_2^{2M}/\hat{\sigma}_1^{2M}$ relative to the Stein loss function. This dominance result is also an application of Theorem 2.1. For the quadratic loss function, on the other hand, the double shrinkage estimator suggested by Kubokawa (1994b) and Bobotas, *et al.* (2012) is of the form

$$\hat{\rho}^{**} = \hat{\sigma}_2^{2*}/\hat{\sigma}_1^2 + \hat{\sigma}_2^2/\hat{\sigma}_1^{2*} - \hat{\sigma}_2^2/\hat{\sigma}_1^2, \tag{3.1}$$

which is different from $\hat{\rho}^* = \hat{\sigma}_2^{2*}/\hat{\sigma}_1^{2*}$. Theorem 2.2 shows that the ratio of shrinkage estimators produces the improvement even for the quadratic loss. Thus, we shall provide some double shrinkage estimators for $\rho = \sigma_2^2/\sigma_1^2$ relative to the quadratic loss function $L_q(\hat{\rho}/\hat{\rho})$ for $L_q(t) = (t-1)^2$.

We begin by explaining dominance results in estimation of the variance σ_2^2 and the reciprocal of variance $1/\sigma_1^2$ relative to the quadratic loss $L_q(\hat{\sigma}_2^2/\sigma_2^2)$ and $L_q(\sigma_1^2/\hat{\sigma}_1^2)$. The minimax estimators of σ_2^2 and $1/\sigma_1^2$ are $\hat{\sigma}_2^{2M} = (n_2+2)^{-1}V_2$ and $1/\hat{\sigma}_1^{2M} = (n_1-4)/V_1$. Stein (1964) discovered the surprising inadmissibility result of $\hat{\sigma}_2^{2M}$, which can be improved on by the truncated estimator

$$\hat{\sigma}_2^{2ST} = \min\Big\{\frac{V_2}{n_2+2}, \frac{V_2 + \|\boldsymbol{X}_2\|^2}{n_2+p_2+2}\Big\},$$

where $\|\boldsymbol{u}\|^2 = \sum_{i=1}^p u_i^2$ for $\boldsymbol{u} = (u_1, \ldots, u_p)^t \in \mathbb{R}^p$. Similarly, $1/\hat{\sigma}_1^{2M}$ can be improved on by

$$1/\hat{\sigma}_1^{2ST} = 1/\min\Big\{\frac{V_1}{n_1-4}, \frac{V_1 + \|\boldsymbol{X}_1\|^2}{n_1+p_1-4}\Big\}.$$

Brewster and Zidek (1974) succeeded in deriving the generalized Bayes estimator improving on $\hat{\sigma}_2^{2M}$ and $1/\hat{\sigma}_1^{2M}$. Let $\eta_i = 1/\sigma_i^2$, $i = 1, 2$. The hierarchical prior distribution which they suggested is

$$\begin{aligned}\boldsymbol{\mu}_i|\eta_i, \lambda_i \sim& \mathcal{N}_{p_i}(\boldsymbol{0}, \frac{1-\lambda_i}{\lambda_i}\eta_i^{-1}\boldsymbol{I}_{p_i}),\\ \lambda_i \sim& \lambda_i^{-1}\mathrm{d}\lambda_i, \quad 0 < \lambda_i < 1,\\ \eta_i \sim& \eta_i^{-1}\mathrm{d}\eta_i,\end{aligned}$$

for $i = 1, 2$. Let $E[\cdot|V_2, \boldsymbol{X}_2]$ and $E[\cdot|V_1, \boldsymbol{X}_1]$ be posterior expectations. The generalized Bayes estimators of σ_2^2 and $1/\sigma_1^2$ against the hierarchical prior are given by $\hat{\sigma}_2^{2GB} = E[\eta_2|V_2, \boldsymbol{X}_2]/E[\eta_2^2|V_2, \boldsymbol{X}_2] = V_2\phi_2^{GB}(W_2)$ and $\hat{\sigma}_1^{1GB} = E[\eta_1^{-2}|V_1, \boldsymbol{X}_1]/E[\eta_1^{-1}|V_1, \boldsymbol{X}_1] = V_1\phi_1^{GB}(W_1)$, where

$$\begin{aligned}\phi_2^{GB}(W_2) =& \frac{1}{n_2+p_2+2}\frac{\int_0^{W_2}\lambda^{p_2/2-1}/(1+\lambda)^{(n_2+p_2)/2+1}d\lambda}{\int_0^{W_2}\lambda^{p_2/2-1}/(1+\lambda)^{(n_2+p_2)/2+2}d\lambda},\\ \phi_1^{GB}(W_1) =& \frac{1}{n_1+p_1-4}\frac{\int_0^{W_1}\lambda^{p_1/2-1}/(1+\lambda)^{(n_1+p_1)/2-2}d\lambda}{\int_0^{W_1}\lambda^{p_1/2-1}/(1+\lambda)^{(n_1+p_1)/2-1}d\lambda},\end{aligned}$$

for $W_i = \|\boldsymbol{X}_i\|^2/V_i$.

Kubokawa (1994a) constructed unified classes of improved estimators of the form $\hat{\sigma}_i^2(\phi_i) = V_i\phi_i(W_i)$ for absolutely continuous function $\phi_i(\cdot)$. In fact, $\hat{\sigma}_i^2(\phi_i)$ improves on $\hat{\sigma}_i^{2M}$ if $\phi_i(w)$ satisfies the conditions:

(C1) $\phi_i(w)$ is non-decreasing in w.

(C2) $\phi_i(w) \geq \phi_i^{GB}(w)$ for $w > 0$.

Since $\phi_i^{TR}(w)$ and $\phi_i^{GB}(w)$ satisfy these conditions, the Stein truncated and the generalized Bayes estimators belong to the classes.

The class of improved estimators can be derived by using the method of *Integral expression of Risk Difference*. Following Kubokawa (1994a, 98, 99), the risk difference between the estimators $\hat{\sigma}_2^{2M}$ and $\hat{\sigma}_2^2(\phi_2)$ can be expressed as

$$\begin{aligned}\Delta =&E[\{(n_2+2)^{-1}V_i/\sigma_2^2-1\}^2] - E[\{V_2\phi_2(W_2)/\sigma_2^2-1\}^2]\\ =&2\int_0^\infty \phi_2'(w_2)\int_0^\infty \{\phi_2(w_2)x-1\}xF_{p_2}(w_2x;\lambda_2)f_{n_2}(x)\mathrm{d}x\mathrm{d}w_2\\ =&2\int_0^\infty \phi_2'(w_2)\int_0^\infty x^2F_{p_2}(w_2x;\lambda_2)f_{n_2}(x)\mathrm{d}x\{\phi_2(w_2)-\phi_2^*(w_2;\lambda_2)\}\mathrm{d}w_2,\end{aligned}$$

where $f_{n_2}(x)$ is the density function of a central chi-square distribution $\chi^2_{n_2}$ with n_2 degrees of freedom, $F_{p_2}(x;\lambda_2)$ is the cumulative distribution function of a non-central chi-square distribution $\chi^2_{p_2}(\lambda_2)$ with p_2 degrees of freedom and noncentrality λ_2, and $\phi_2^*(w_2;\lambda_2)$ is given by

$$\phi_2^*(w_2;\lambda_2) = \int_0^\infty xF_{p_2}(w_2x;\lambda_2)f_{n_2}(x)\mathrm{d}x\Big/\int_0^\infty x^2F_{p_2}(w_2x;\lambda_2)f_{n_2}(x)\mathrm{d}x.$$

It can be verified that $\phi^*(w_2;\lambda_2) \leq \phi_2^*(w_2;0)$ and that $\phi_2^*(w_2;0)$ is identical to $\phi_2^{GB}(w_2)$. This fact implies not only that the improvement can be established under conditions (C1) and (C2), but also that the risk of the generalized Bayes estimator is equal to that of $(n_2+2)^{-1}V_2$ at $\lambda_2 = 0$, namely,

$$E_{\lambda_2=0}[\{(n_2+2)^{-1}V_2/\sigma_2^2-1\}^2] = E_{\lambda_2=0}[\{V_2\phi_2^{GB}(W_2)/\sigma_2^2-1\}^2] \tag{3.2}$$

when $\lambda_2 = 0$. This property is useful for investigating the improvement of the generalized Bayes estimator of product of variances. A similar equation to (3.2) holds for estimation of $1/\sigma_1^2$, namely,

$$E_{\lambda_1=0}[\{\sigma_1^2/\{V_1/(n_1-4)\}-1\}^2] = E_{\lambda_1=0}[\{\sigma_1^2/\{V_1\phi_1^{GB}(W_1)\}-1\}^2] \tag{3.3}$$

Another improved estimators are the Strawderman-type estimators suggested by Strawderman (1974), Maruyama and Strawderman (2006) and Bobotas and Kourouklis (2010), and are given by

$$\begin{aligned}\hat{\sigma}_2^{2SR} =&(1+W_2)V_2/\{(n_2+2)(r_2+1+W_2)\},\\ \hat{\sigma}_1^{2SR} =&(1+W_1)V_1/\{(n_1-4)(r_1+1+W_1)\},\end{aligned} \tag{3.4}$$

for $0 < r_i < r_{0i}$, $i = 1, 2$, where r_{01} and r_{02} are specified constants .

Theorem 2.2 guarantees that the shrinkage and improved estimators of σ_2^2 and $1/\sigma_1^2$ can produce the natural double shrinkage estimators which improve on the crude ratio estimator $\hat{\rho}^M = \hat{\sigma}_2^{2M}/\hat{\sigma}_1^{2M}$. The double shrinkage and improved estimators with natural forms are

$$\hat{\rho}^{*ST} = \hat{\sigma}_2^{2ST}/\hat{\sigma}_1^{2ST}, \quad \hat{\rho}^{*GB} = \hat{\sigma}_2^{2GB}/\hat{\sigma}_1^{2GB}, \quad \hat{\rho}^{*SR} = \hat{\sigma}_2^{2SR}/\hat{\sigma}_1^{2SR}.$$

It is noted that $\hat{\rho}^{*GB}$ is the generalized Bayes estimator of ρ relative to the quadratic loss. As given in (3.1), we have another types of double shrinkage and improved estimators, which are denoted by $\hat{\rho}^{**ST}$, $\hat{\rho}^{**GB}$ and $\hat{\rho}^{**SR}$.

Figure 1 illustrates the risk functions of $\hat{\rho}^{*M} = \hat{\sigma}_2^{2M}/\hat{\sigma}_1^{2M}$, $\hat{\rho}^{*ST}$, $\hat{\rho}^{*GB}$, $\hat{\rho}^{**ST}$ and $\hat{\rho}^{**GB}$ relative to the quadratic loss, where $n_1 = n_2 = 8$, $p_1 = p_2 = 10$, $\sigma_1^2 = \sigma_2^2 = 1$, $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \mu(1, \dots, 1)^t$ for μ taking values form 0 to 3. From this figure, it is seen that the double shrinkage estimators $\hat{\rho}^{*ST}$ and $\hat{\rho}^{*GB}$ are better than $\hat{\rho}^{**ST}$ and $\hat{\rho}^{**GB}$, respectively. The generalized Bayes estimator $\hat{\rho}^{*GB}$ has the smallest risks of these estimators.

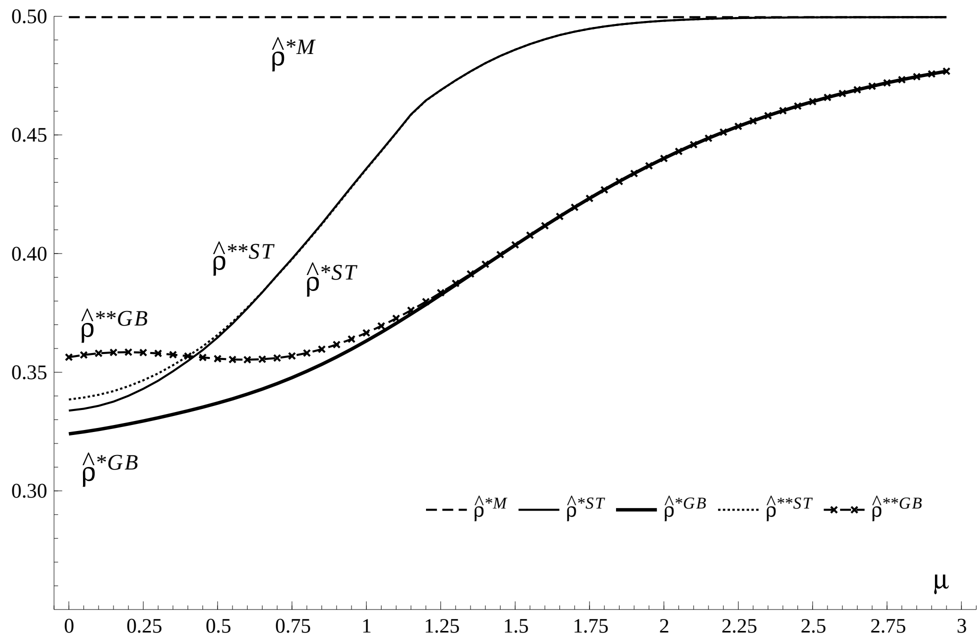

Figure 1: Plots of risk functions of $\hat{\rho}^{*M}$, $\hat{\rho}^{*ST}$, $\hat{\rho}^{*GB}$, $\hat{\rho}^{**ST}$ and $\hat{\rho}^{**GB}$

4 Non-dominance and Dominance Results in Estimation of Some Functions of Positive Parameters

In the previous sections, the dominance results of the double shrinkage estimator have been shown for estimation of ratio of the positive parameters $\rho = \theta_2/\theta_1$. In this section, we investigate whether similar dominance results hold in estimation of the product, sum and difference of positive parameters.

4.1 Non-dominance in estimation of product $\tau = \theta_1\theta_2$

We first treat estimation of the product of the positive parameters $\tau = \theta_1\theta_2$. Contrary to our expectation, we shall show that similar dominance results do not hold.

Let $\hat{\tau} = \hat{\theta}_1\hat{\theta}_2$ and $\hat{\tau}^* = \hat{\theta}_1^*\hat{\theta}_2^*$ be, respectively, a crude estimator and a double shrinkage estimator of τ. An estimator $\tilde{\tau}$ is evaluated relative to the Stein loss $L_s(\tilde{\tau}/\tau)$.

Theorem 4.1 *Assume conditions* (A1) *and* (A2) *with* $E_\omega[\hat{\theta}_i/\theta_i] = 1$ *for* $i = 1, 2$*. Assume that there exists a point* ω_0 *such that*

$$E_{\omega_0}[L_s(\hat{\tau}_i/\tau_i)] = E_{\omega_0}[L_s(\hat{\tau}_i^*/\tau_i)], \quad i = 1, 2. \tag{4.1}$$

Then, at the point ω_0*,*

$$R_s(\omega_0, \hat{\tau}) < R_s(\omega_0, \hat{\tau}^*),$$

namely, the double shrinkage estimator $\hat{\tau}^* = \hat{\theta}_1^*\hat{\theta}_2^*$ *does not dominate the estimator* $\hat{\tau} = \hat{\theta}_1\hat{\theta}_2$ *relative to the Stein loss.*

Proof. The same notations as in the proof of Theorem 2.1 are used. Condition (4.1) implies that

$$E_{\omega_0}[\log(Y_i)] = E_{\omega_0}[Y_i + \log(X_i) - 1], \tag{4.2}$$

for $i = 1, 2$. Using these equalities, we can write the difference of the risk functions of $\hat{\tau}$ and $\hat{\tau}^*$ as

$$\begin{aligned}\Delta_s(\omega_0) =& E_{\omega_0}[X_1X_2 - \log(X_1X_2) - 1] - E_{\omega_0}[Y_1Y_2 - \log(Y_1Y_2) - 1] \\ =& E_{\omega_0}[-Y_1Y_2 + Y_1 + Y_2 - 1] \\ =& - E_{\omega_0}[Y_1 - 1]E_{\omega_0}[Y_2 - 1].\end{aligned}$$

From (A2), it follows that $E_\omega[Y_i - 1] \le E_\omega[X_i - 1] = 0$ for $i = 1, 2$. This implies that $\Delta_s(\omega_0) < 0$ at ω_0. □

As described in (3.2), the generalized Bayes estimator $\hat{\sigma}_2^{2GB}$ of σ_2^2 has the same risk to that of $\hat{\sigma}_2^{2M}$ when the noncentrality parameter λ_2 is zero. Thus, condition (4.1) is satisfied for $\hat{\sigma}_2^{2GB}$. For σ_1^2, we can provide the generalized Bayes estimator $\tilde{\sigma}_1^{2GB}$ which is given by replacing n_2, p_2, V_2 and W_2 in $\hat{\sigma}_2^{2GB}$ with n_1, p_1, V_1 and W_1, and condition (4.1) is satisfied for $\tilde{\sigma}_1^{2GB}$. Hence, from Theorem 4.1 it follows that the generalized Bayes estimator $\tilde{\sigma}_1^{2GB}\hat{\sigma}_2^{2GB}$ cannot dominate $\tilde{\sigma}_1^{2M}\hat{\sigma}_2^{2M}$ for $\tilde{\sigma}_1^{2M} = V_1/(n_1 + 2)$, namely,

$$R_s(\omega, \tilde{\sigma}_1^{2M}\hat{\sigma}_2^{2M}) < R_s(\omega, \tilde{\sigma}_1^{2GB}\hat{\sigma}_2^{2GB}), \tag{4.3}$$

when $\lambda_1 = 0$ and $\lambda_2 = 0$.

The non-dominance result under the Stein loss can be shown to hold for the quadratic loss function $L_q(\tilde{\tau}/\tau)$ for $L_q(t) = (t - 1)^2$.

Theorem 4.2 *Assume conditions* (A1) *and* (A2) *with* $E_\omega[(\hat{\theta}_i/\theta_i)^2] = E_\omega[\hat{\theta}_i/\theta_i]$ *for* $i = 1, 2$*. Assume that there exists a point* ω_0 *such that*

$$E_{\omega_0}[L_q(\hat{\theta}_i/\theta_i)] = E_{\omega_0}[L_q(\hat{\theta}_i^*/\theta_i)], \quad i = 1, 2. \tag{4.4}$$

Then, at the point ω_0,

$$R_q(\omega_0, \hat{\tau}) < R_q(\omega_0, \hat{\tau}^*),$$

namely, the double shrinkage estimator $\hat{\tau}^* = \hat{\theta}_1^*\hat{\theta}_2^*$ does not dominate the estimator $\hat{\tau} = \hat{\theta}_1\hat{\theta}_2$ relative to the quadratic loss.

Proof. The same notations as in the proof of Theorem 2.2 are used. Condition (4.4) implies that

$$E_{\omega_0}[Y_i^2] = E_{\omega_0}[2Y_i - X_i], \quad i = 1, 2, \tag{4.5}$$

at ω_0. Then, the difference of the risk functions of $\hat{\tau}$ and $\hat{\tau}^*$ is written as

$$\begin{aligned}\Delta_q(\omega_0) =& E_{\omega_0}[(X_1X_2 - 1)^2 - (Y_1Y_2 - 1)^2]\\ =& E_{\omega_0}[-X_1X_2 - Y_1^2Y_2^2 + 2Y_1Y_2]\\ =& E_{\omega_0}[-X_1X_2 - (2_1 - X_1)(2Y_2 - X_2) + 2Y_1Y_2]\\ =& -2E_{\omega_0}[Y_1 - X_1]E_{\omega}[Y_2 - X_2].\end{aligned}$$

From (A2), it follows that $\Delta_q(\omega_0) < 0$ at ω_0. □

4.2 Dominance in estimation of difference $\xi = \theta_1 - \theta_2$

We next treat estimation of the difference of the positive parameters $\xi = \theta_1 - \theta_2$. Let $\xi = \hat{\theta}_1 - \hat{\theta}_2$ and $\hat{\xi}^* = \hat{\theta}_1^* - \hat{\theta}_2^*$ be, respectively, a crude estimator and a double shrinkage estimator of ξ. An estimator $\hat{\xi}$ is evaluated in terms of the mean squared error (MSE) given by $MSE(\omega, \hat{\xi}) = E[(\hat{\xi} - \xi)^2]$. Assume the following condition:

(A5) For $i = 1, 2$, $E_\omega[(\hat{\theta}_i/\theta_i)^2] = E_\omega[\hat{\theta}_i/\theta_i]$ and $E_\omega[\{\hat{\theta}_i/\theta_i - 1\}^2] \geq E_\omega[\{\hat{\theta}_i^*/\theta_i - 1\}^2]$ for any ω.

This condition implies that $E_\omega[\hat{\theta}_i/\theta_i] \leq 1$. In fact, it is seen that $0 < \mathrm{Var}(\hat{\theta}_i/\theta_i) = E[(\hat{\theta}_i/\theta_i)^2] - \{E[\hat{\theta}_i/\theta_i]\}^2 = E[\hat{\theta}_i/\theta_i] - \{E[\hat{\theta}_i/\theta_i]\}^2 = E[\hat{\theta}_i/\theta_i](1 - E[\hat{\theta}_i/\theta_i])$, which leads to the inequality $E_\omega[\hat{\theta}_i/\theta_i] < 1$. Combining this inequality and condition (A2) gives the condition

$$E_\omega[\hat{\theta}_i^*/\theta_i - 1] < E_\omega[\hat{\theta}_i/\theta_i - 1] < 0, \tag{4.6}$$

which will be used for proving the following theorem.

Theorem 4.3 *Assume conditions* (A1), (A2) *and* (A5). *In the estimation of the difference $\xi = \theta_1 - \theta_2$, the double shrinkage estimator $\hat{\xi}^* = \hat{\theta}_1^* - \hat{\theta}_2^*$ dominates the estimator $\hat{\xi} = \hat{\theta}_1 - \hat{\theta}_2$ in terms of MSE, namely,*

$$MSE(\omega, \hat{\xi}) \geq MSE(\omega, \hat{\xi}^*)$$

for any ω.

Proof. For $X_i = \hat{\theta}_i/\theta_i$ and $Y_i = \hat{\theta}_i^*/\theta_i$, $i = 1, 2$, it is noted that $\hat{\xi} - \xi$ and $\hat{\xi}^* - \xi$ are written as $\hat{\xi} - \xi = \theta_1(X_1 - 1) - \theta_2(X_2 - 1)$ and $\hat{\xi}^* - \xi = \theta_1(Y_1 - 1) - \theta_2(Y_2 - 1)$. Thus, the difference of MSEs of the estimators $\hat{\xi}$ and $\hat{\xi}^*$ is expressed as

$$\begin{aligned}\Delta(\omega) =& MSE(\omega, \hat{\xi}) - MSE(\omega, \hat{\xi}^*)\\ =& \theta_1^2\{E_\omega[(X_1 - 1)^2] - E_\omega[(Y_1 - 1)^2]\} + \theta_2^2\{E_\omega[(X_2 - 1)^2] - E_\omega[(Y_2 - 1)^2]\}\\ &+ 2\theta_1\theta_2\{E_\omega[Y_1 - 1]E_\omega[Y_2 - 1] - E_\omega[X_1 - 1]E_\omega[X_2 - 1]\}.\end{aligned}$$

The first two terms are nonnegative from (A5). Also from (4.6), it follows that

$$E_\omega[Y_1 - 1]E_\omega[Y_2 - 1] \geq E_\omega[X_1 - 1]E_\omega[X_2 - 1] > 0.$$

These facts show that $\Delta(\omega) > 0$ for any ω. $\square$

For the model treated in Section 3, we can consider the double shrinkage estimators $\hat{\xi}^{*ST} = \tilde{\sigma}_1^{2ST} - \hat{\sigma}_2^{2ST}$, $\hat{\xi}^{*GB} = \tilde{\sigma}_1^{2GB} - \hat{\sigma}_2^{2GB}$ and $\hat{\xi}^{*SR} = \tilde{\sigma}_1^{2SR} - \hat{\sigma}_2^{2SR}$ where $\tilde{\sigma}_1^{2ST}$, $\tilde{\sigma}_1^{2GB}$ and $\tilde{\sigma}_1^{2SR}$ are provided by replacing n_2, p_2, V_2 and W_2 in $\hat{\sigma}_2^{2ST}$, $\hat{\sigma}_2^{2GB}$ and $\hat{\sigma}_2^{2SR}$ with n_1, p_1, V_1 and W_1. Then, the improvements of the double shrinkage estimators can be guaranteed by Theorem 4.3.

4.3 Non-dominance in estimation of the sum $\eta = \theta_1 + \theta_2$

Finally, we consider estimation of the sum of the positive parameters $\eta = \theta_1 + \theta_2$, and we shall show that dominance results do not hold.

Let $\widehat{\eta} = \hat{\theta}_1 + \hat{\theta}_2$ and $\widehat{\eta}^* = \hat{\theta}_1^* + \hat{\theta}_2^*$ be, respectively, a crude estimator and a double shrinkage estimator of η. An estimator $\widehat{\eta}$ is evaluated in terms of the mean squared error $MSE(\omega, \widehat{\eta}) = E[(\widehat{\eta} - \eta)^2]$.

Theorem 4.4 *Assume conditions* (A1) *and* (A2) *with* $E_\omega[(\hat{\theta}_i/\theta_i)^2] = E_\omega[\hat{\theta}_i/\theta_i]$ *for* $i = 1, 2$. *Assume that there exists a point* ω_0 *such that the equalities* (4.4) *holds. Then, at the point* ω_0,

$$MSE(\omega_0, \widehat{\eta}) < MSE(\omega_0, \widehat{\eta}^*),$$

namely, the double shrinkage estimator $\widehat{\eta}^* = \hat{\theta}_1^* + \hat{\theta}_2^*$ *does not dominate the estimator* $\widehat{\eta} = \hat{\theta}_1 + \hat{\theta}_2$ *in terms of MSE.*

Proof. Using the same arguments as in the proof of Theorem 4.3, we can write the difference of MSEs of the estimators $\widehat{\eta}$ and $\widehat{\eta}^*$ as

$$\begin{aligned}\Delta(\omega) =& MSE(\omega, \widehat{\eta}) - MSE(\omega, \widehat{\eta}^*)\\ =&\theta_1^2\{E_\omega[(X_1 - 1)^2] - E_\omega[(Y_1 - 1)^2]\} + \theta_2^2\{E_\omega[(X_2 - 1)^2] - E_\omega[(Y_2 - 1)^2]\}\\ &- 2\theta_1\theta_2\{E_\omega[Y_1 - 1]E_\omega[Y_2 - 1] - E_\omega[X_1 - 1]E_\omega[X_2 - 1]\}.\end{aligned}$$

The first two terms are zero at ω_0 from condition (4.4). Also from (4.6), it follows that $E_\omega[Y_1 - 1]E_\omega[Y_2 - 1] \geq E_\omega[X_1 - 1]E_\omega[X_2 - 1] > 0$, which implies that $\Delta(\omega_0) < 0$ at ω_0. $\square$

A similar argument given around (4.3) is used to give an example of the non-dominance result in Theorem 4.4. That is, a double shrinkage estimator $\widehat{\eta}^* = \tilde{\sigma}_1^{2GB} + \hat{\sigma}_2^{2GB}$ cannot dominate $\widehat{\eta} = (n_1 + 2)^{-1}V_1 + (n_2 + 2)^{-1}V_2$, where $\tilde{\sigma}_1^{2GB}$ and $\hat{\sigma}_2^{2GB}$ are the generalized Bayes estimators of σ_1^2 and σ_2^2 relative to ${}_q(\hat{\sigma}_i^2/\sigma_i^2)$.

5 A Connection to Estimation of Positive Normal Means

In this section, we explain that the double shrinkage estimators for normal variances are connected to estimation of restricted means in normal distributions. This fact was established by Rukhin (1992). We here use the same arguments to explain how the dominance and non-dominance results treated in the previous sections link to the corresponding results derived by Kubokawa (2012) for estimation of positive normal means.

We first illustrate a one sample model with the following canonical form: $\boldsymbol{X}$ and V are mutually independent random variables such that $\boldsymbol{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \sigma^2 \boldsymbol{I})$ and $V/\sigma^2 \sim \chi^2_n$. The Stein truncated estimator of σ^2 relative to the Stein loss is $\hat{\sigma}^{2ST} = \min\{\hat{\sigma}^{2M}, (V + \|\boldsymbol{X}\|^2)/(n+p)\}$ improving on $\hat{\sigma}^{2M} = V/n$. The generalized Bayes and improved estimator is given by

$$\hat{\sigma}^{2GB} = \frac{V}{n+p} \frac{\int_0^W \lambda^{p/2-1}/(1+\lambda)^{(n+p)/2} d\lambda}{\int_0^W \lambda^{p/2-1}/(1+\lambda)^{(n+p)/2+1} d\lambda}. \tag{5.1}$$

To approximate these estimators, we assume the following conditions:

(R1) Both p and n tend to infinity under the condition that $n = O(p^\delta)$ for $0 < \delta < 1$.

(R2) $\sqrt{n}\|\boldsymbol{\mu}\|^2/(p\sigma^2) \to \sqrt{2}\xi$ as $(n, p) \to \infty$, where ξ is a positive constant.

Theorem 5.1 *Assume conditions* (R1) *and* (R2). *Then,* $\sqrt{n}(\hat{\sigma}^{2M} - \sigma^2)/\sigma^2 \to -\sqrt{2}(Y - \xi)$ *in distribution, where* Y *is a random variable having* $\mathcal{N}(\xi, 1)$ *for* $\xi > 0$. *Also,*

$$\begin{aligned}
\sqrt{n}(\hat{\sigma}^{2ST} - \sigma^2)/\sigma^2 \to& -\sqrt{2}(\max\{Y, 0\} - \xi),\\
\sqrt{n}(\hat{\sigma}^{2GB} - \sigma^2)/\sigma^2 \to& -\sqrt{2}(\hat{\xi}^{GB} - \xi),
\end{aligned}$$

where $\hat{\xi}^{GB} = \int_0^\infty \xi \exp\{-(Y-\xi)^2/2\} \mathrm{d}\xi / \int_0^\infty \exp\{-(Y-\xi)^2/2\} \mathrm{d}\xi$, *which is the generalized Bayes estimator of* ξ *against the uniform prior over the half real line* $\xi > 0$.

Proof. This theorem was established by Rukhin (1992) in the asymptotics of making $n \to \infty$ after making $p \to \infty$, which is slightly different from (R1). Thus, we provide the proof under (R1) and (R2) instructively. Without any loss of generality, we can assume that $\sigma^2 = 1$. Let $Z = (V - n)/\sqrt{2n}$. Since $Z \to \mathcal{N}(0, 1)$, it is seen that $\sqrt{n}(V/n - 1) = \sqrt{2}Z = -\sqrt{2}\{(-Z + \xi) - \xi\} \to -\sqrt{2}(Y - \xi)$. We next note that

$$\begin{aligned}
U \equiv& \sqrt{n}\Big(\frac{V + \|\boldsymbol{X}\|^2}{n+p} - \frac{V}{n}\Big) \qquad (5.2)\\
=& -\sqrt{n}\frac{p(\sqrt{2n}Z + n)}{(n+p)n} + \sqrt{n}\frac{\|\boldsymbol{X} - \boldsymbol{\mu}\|^2 + 2\boldsymbol{\mu}^t(\boldsymbol{X} - \boldsymbol{\mu}) + \|\boldsymbol{\mu}\|^2}{n+p}\\
=& -\sqrt{2}\frac{pZ}{n+p} + \frac{\sqrt{np}A + 2\sqrt{\sqrt{2np}\xi_p}B + \sqrt{2}p\xi_p}{n+p},
\end{aligned}$$

where $A = (\|\boldsymbol{X} - \boldsymbol{\mu}\|^2 - p)/\sqrt{p}$, $B = \boldsymbol{\mu}^t(\boldsymbol{X} - \boldsymbol{\mu})/\|\boldsymbol{\mu}\|$ and $\xi_p = \sqrt{n}\|\boldsymbol{\mu}\|^2/(\sqrt{2}p)$. Since $B \sim \mathcal{N}(0, 1)$, $A \to \mathcal{N}(0, 1)$ and $\xi_p \to \xi$, it is seen from conditions (R1) and (R2) that

$$U \to \sqrt{2}(-Z + \xi) = \sqrt{2}Y. \tag{5.3}$$

Thus, $\sqrt{n}(\hat{\sigma}^{2ST}-1)=\sqrt{n}(\hat{\sigma}^{2M}-1)+\sqrt{n}\min\{0,U\}\to-\sqrt{2}(Y-\xi)+\min(0,\sqrt{2}Y)=-\sqrt{2}\{\max(Y,0)-\xi\}$.

To evaluate the generalized Bayes estimator given in (5.1), make the transformation

$$\frac{1}{1+\lambda}=\frac{V}{n+p}\Big(1-\frac{t}{\sqrt{n}}\Big)\quad\text{with}\quad\frac{\mathrm{d}\lambda}{(1+\lambda)^2}=\frac{V\mathrm{d}t}{(n+p)\sqrt{n}}.$$

Then, the range of λ in the integrals is changed from $0<\lambda<W$ to $T_1<t<T_2$, where

$$T_1=\sqrt{n}\Big(1-\frac{n+p}{V}\Big),\quad T_2=\sqrt{n}\Big(1-\frac{n+p}{V(1+W)}\Big).$$

Thus, the generalized Bayes estimator $\hat{\sigma}^{2GB}$ in (5.1) can be rewritten as

$$\hat{\sigma}^{2GB}=\frac{\int_{T_1}^{T_2}\Big\{1-(n+p)^{-1}V\big(1-t/\sqrt{n}\big)\Big\}^{p/2-1}\big(1-t/\sqrt{n}\big)^{n/2-1}\mathrm{d}t}{\int_{T_1}^{T_2}\Big\{1-(n+p)^{-1}V\big(1-t/\sqrt{n}\big)\Big\}^{p/2-1}\big(1-t/\sqrt{n}\big)^{n/2}\mathrm{d}t},$$

so that $\sqrt{n}(\hat{\sigma}^{2GB}-1)$ is expressed as

$$\sqrt{n}(\hat{\sigma}^{2GB}-1)=\frac{\int_{T_1}^{T_2}\Big\{1-(n+p)^{-1}V\big(1-t/\sqrt{n}\big)\Big\}^{p/2-1}t\big(1-t/\sqrt{n}\big)^{n/2-1}\mathrm{d}t}{\int_{T_1}^{T_2}\Big\{1-(n+p)^{-1}V\big(1-t/\sqrt{n}\big)\Big\}^{p/2-1}\big(1-t/\sqrt{n}\big)^{n/2}\mathrm{d}t}. \tag{5.4}$$

We here investigate limiting values of the end points T_1 and T_2. It is observed that

$$T_1=\sqrt{n}\Big(1-\frac{n+p}{\sqrt{2n}Z+n}\Big)=\frac{\sqrt{2n}Z-p}{\sqrt{2}Z+\sqrt{n}}\to-\infty. \tag{5.5}$$

For T_2, from (5.3), W is expressed as

$$W=\|\boldsymbol{X}\|^2/V=(n+p)U/(\sqrt{n}V)+p/n,$$

which is used to rewrite T_2 as

$$\begin{aligned}T_2=&\frac{\sqrt{n}V(1+W)-\sqrt{n}(n+p)}{V(1+W)}=\frac{\sqrt{n}V+(n+p)U+\sqrt{n}(p/n)V-\sqrt{n}(n+p)}{V+(n+p)U/\sqrt{n}+(p/n)V}\\=&\frac{V\sqrt{n}+U-\sqrt{n}}{U/\sqrt{n}+V/n}=\frac{\sqrt{2}Z+U}{U/\sqrt{n}+\sqrt{2/n}Z+1}\\\to&\sqrt{2}Z+\sqrt{2}(-Z+\xi)=\sqrt{2}\xi.\end{aligned} \tag{5.6}$$

We now evaluate the integrant in (5.4). Note that

$$\begin{aligned}1-\frac{V}{n+p}(1-\frac{t}{\sqrt{n}})=&\frac{p-\sqrt{2n}Z}{n+p}+\frac{\sqrt{2}Z+\sqrt{n}}{n+p}t\\=&\frac{p-\sqrt{2n}Z}{n+p}\Big(1+\frac{\sqrt{2}Z+\sqrt{n}}{p-\sqrt{2n}Z}t\Big)\\=&\frac{p-\sqrt{2n}Z}{n+p}\Big(1+\frac{\sqrt{2}Z+\sqrt{n}}{p}t+O_p(n/p^2)\Big).\end{aligned}$$

Then the integrals in (5.4) can be rewritten as

$$\sqrt{n}(\hat\sigma^{2GB}-1)=\frac{\int_{T_1}^{T_2}\Big\{1+(\sqrt{2}Z+\sqrt{n})t/p+O_p(n/p^2)\Big\}^{p/2-1}t\big(1-t/\sqrt{n}\big)^{n/2-1}\mathrm{d}t}{\int_{T_1}^{T_2}\Big\{1+(\sqrt{2}Z+\sqrt{n})t/p+O_p(n/p^2)\Big\}^{p/2-1}\big(1-t/\sqrt{n}\big)^{n/2}\mathrm{d}t}. \quad (5.7)$$

It is here demonstrated that

$$\begin{aligned}&\Big(\frac{p}{2}-1\Big)\log\Big\{1+\frac{\sqrt{2}Z+\sqrt{n}}{p}t+O_p(n/p^2)\Big\}+\Big(\frac{n}{2}-1\Big)\log\Big\{1-\frac{t}{\sqrt{n}}\Big\}\\
=&\frac{1}{2}\Big\{(\sqrt{2}Z+\sqrt{n})t+O_p(n/p)-\sqrt{n}t-\frac{t^2}{2}+O(n^{-1/2})\Big\}\\
=&-\frac{1}{4}(t-\sqrt{2}Z)^2+\frac{Z^2}{2}+O_p(n/p)+O(n^{-1/2}).\end{aligned} \quad (5.8)$$

Combining (5.5), (5.6), (5.7) and (5.8), we can show that

$$\sqrt{n}(\hat\sigma^{2GB}-1)\to\frac{\int_{-\infty}^{\sqrt{2}\xi}t\exp\{-(t-\sqrt{2}Z)^2/4\}\mathrm{d}t}{\int_{-\infty}^{\sqrt{2}\xi}\exp\{-(t-\sqrt{2}Z)^2/4\}\mathrm{d}t}\equiv D.$$

Making the transformation $\sqrt{2}\xi-t=\sqrt{2}\mu$ with $-\mathrm{d}t=\sqrt{2}\mathrm{d}\mu$ gives

$$\begin{aligned}D=&\frac{\int_0^\infty(\sqrt{2}\xi-\sqrt{2}\mu)\exp\{-(\xi-\mu-Z)^2/2\}\mathrm{d}t}{\int_0^\infty\exp\{-(\xi-\mu-Z)^2/2\}\mathrm{d}t}\\
=&-\sqrt{2}\Big\{\frac{\int_0^\infty\mu\exp\{-(Y-\mu)^2/2\}\mathrm{d}t}{\int_0^\infty\exp\{-(Y-\mu)^2/2\}\mathrm{d}t}-\xi\Big\},\end{aligned}$$

which proves Theorem 5.1. □

We now apply Theorem 5.1 to a two sample model with the canonical forms: for $i=1,2$, $\boldsymbol{X}_i$ and V_i are mutually independent random variables such that $\boldsymbol{X}_i\sim\mathcal{N}_p(\boldsymbol{\mu}_i,\sigma_i^2\boldsymbol{I}_p)$ and $V_i/\sigma_i^2\sim\chi_n^2$. Similarly to (R2), assume the following condition for $i=1,2$:

(R2′) $\sqrt{n}\|\boldsymbol{\mu}_i\|^2/(p\sigma_i^2)\to\sqrt{2}\xi_i$ as $(n,p)\to\infty$, where ξ_i is a positive constant.

Theorem 5.2 *Assume conditions* (R1) *and* (R2′). *Then,*

$$\begin{aligned}\sqrt{n}\Big(\frac{\hat\sigma_2^{2GB}}{\hat\sigma_1^{2GB}}-\frac{\sigma_2^2}{\sigma_1^2}\Big)\frac{\sigma_1^2}{\sigma_2^2}&\to-\sqrt{2}\Big\{(\hat\xi_2^{GB}-\hat\xi_1^{GB})-(\xi_2-\xi_1)\Big\},\\
\sqrt{n}\Big(\hat\sigma_1^{2GB}\hat\sigma_2^{2GB}-\sigma_1^2\sigma_2^2\Big)/(\sigma_1^2\sigma_2^2)&\to-\sqrt{2}\Big\{(\hat\xi_1^{GB}+\hat\xi_2^{GB})-(\xi_1+\xi_2)\Big\},\\
\sqrt{n}\Big\{(\hat\sigma_1^{2GB}-\hat\sigma_2^{2GB})-(\sigma_1^2-\sigma_2^2)\Big\}&\to-\sqrt{2}\Big\{(\sigma_1^2\hat\xi_1^{GB}-\sigma_2^2\hat\xi_2^{GB})-(\sigma_1^2\xi_1-\sigma_2^2\xi_2)\Big\},\\
\sqrt{n}\Big\{(\hat\sigma_1^{2GB}+\hat\sigma_2^{2GB})-(\sigma_1^2+\sigma_2^2)\Big\}&\to-\sqrt{2}\Big\{(\sigma_1^2\hat\xi_1^{GB}+\sigma_2^2\hat\xi_2^{GB})-(\sigma_1^2\xi_1+\sigma_2^2\xi_2)\Big\},\end{aligned}$$

where $\hat\xi_i^{GB}=\int_0^\infty\xi_i\exp\{-(Y_i-\xi_i)^2/2\}\mathrm{d}\xi_i/\int_0^\infty\exp\{-(Y_i-\xi_i)^2/2\}\mathrm{d}\xi_i$ *for a random variable* Y_i *having* $\mathcal{N}(\xi_i,1)$, $i=1,2$.

Proof. It is first demonstrated that

$$\sqrt{n}\Big(\frac{\hat{\sigma}_2^{2GB}/\sigma_2^2 - 1 + 1}{\hat{\sigma}_1^{2GB}/\sigma_1^2 - 1 + 1} - 1\Big) = \sqrt{n}(\hat{\sigma}_2^{2GB}/\sigma_2^2 - 1) - \sqrt{n}(\hat{\sigma}_1^{2GB}/\sigma_1^2 - 1) + o_p(1).$$

Here from Theorem 5.1, it is noted that $\sqrt{n}(\hat{\sigma}_i^{2GB} - \sigma_i^2)/\sigma_i^2 \to -\sqrt{2}(\hat{\xi}_i^{GB} - \xi_i)$. Then, the first approximation in Theorem 5.2 can be derived. Next, it is observed that

$$\sqrt{n}\Big\{(\hat{\sigma}_1^{2GB}/\sigma_1^2 - 1 + 1)(\hat{\sigma}_1^{2GB}/\sigma_2^2 - 1 + 1)\Big\} = \sqrt{n}(\hat{\sigma}_1^{2GB}/\sigma_1^2 - 1) + \sqrt{n}(\hat{\sigma}_2^{2GB}/\sigma_2^2 - 1) + o_p(1),$$

which can yield the second approximation. Also,

$$\sqrt{n}\Big\{(\hat{\sigma}_1^{2GB} - \hat{\sigma}_1^{2GB}) - (\sigma_1^2 - \sigma_2^2)\Big\} = \sigma_1^2\sqrt{n}(\hat{\sigma}_1^{2GB}/\sigma_1^2 - 1) - \sigma_2^2\sqrt{n}(\hat{\sigma}_2^{2GB}/\sigma_2^2 - 1),$$

which can be approximated as $-\sqrt{2}\{(\sigma_1^2\hat{\xi}_1^{GB} - \sigma_2^2\hat{\xi}_2^{GB}) - (\sigma_1^2\xi_1 - \sigma_2^2\xi_2)\}$. The fourth approximation can be confirmed similarly. □

Theorem 5.2 implies that the generalized Bayes estimators of the ratio and product of normal variances can be approximated as the generalized Bayes estimators of the difference and sum of the normal positive means. Kubokawa (2012) considered the estimation of the linear combination of ξ_1 and ξ_2, namely,

$$a_1\xi_1 + a_2\xi_2,$$

where a_1 and a_2 are known and non-zero constants. Then Kubokawa (2012) derived the condition for the improvement of the generalized Bayes estimator, given in the following proposition:

Proposition 5.1 *In estimation of $a_1\xi_1 + a_2\xi_2$, the generalized Bayes estimator $a_1\hat{\xi}_1^{GB} + a_2\hat{\xi}_2^{GB}$ dominates $a_1Y_1 + a_2Y_2$ in terms of MSE if and only if $a_1a_2 < 0$.*

Hence from Proposition 5.1, it follows that the generalized Bayes estimators of the difference $\xi_1 - \xi_2$ and $\sigma_1^2\xi_1 - \sigma_2\xi_2$ for known σ_1^2 and σ_2^2 can dominate the corresponding estimators $Y_1 - Y_2$ and $\sigma_1^2Y_1 - \sigma_2^2Y_2$, respectively. However, the generalized Bayes estimators of the sum $\xi_1 + \xi_2$ and $\sigma_1^2\xi_1 + \sigma_2\xi_2$ for known σ_1^2 and σ_2^2 cannot dominate the corresponding estimators $Y_1 + Y_2$ and $\sigma_1^2Y_1 + \sigma_2^2Y_2$. These facts coincide with the results derived in the previous sections.

Acknowledgments.
Research of the author was supported in part by Grant-in-Aid for Scientific Research (21540114 and 23243039) from Japan Society for the Promotion of Science.

References

[1] Bobotas, P., Iliopoulos, G. and Kourouklis, S. (2012). Estimating the ratio of two scale parameters: a simple approach. *Ann. Inst. Statist. Math.*, **64**, 343-357.

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

[3] Gelfand, A.E. and Dey, D.K. (1988). On the estimation of a variance ratio. *J. Statist. Plann. Inference*, **19**, 121-131.

[4] Ghosh, M. and Kundu, S. (1996). Decision theoretic estimation of the variance ratio. *Statist. Decisions*, **14**, 161-175.

[5] Iliopoulos, G. and Kourouklis, S. (1999). Improving on the best affine equivariant estimator of the ratio of the generalized variances. *J. Multivariate Analysis*, **68**, 176-192.

[6] Kubokawa, T. (1994a). A unified approach to improving equivariant estimators. *Ann. Statist.*, **22**, 290-299.

[7] Kubokawa, T. (1994b). Double shrinkage estimation of ratio of scale parameters. *Ann. Inst. Statist. Math.*, **46**, 95-116.

[8] Kubokawa, T. (1998). The Stein phenomenon in simultaneous estimation: A review. *Applied Statistical Science III* (eds. S.E. Ahmed, M. Ahsanullah and B.K. Sinha), NOVA Science Publishers, Inc., New York, to appear.

[9] Kubokawa, T. (1999). Shrinkage and modification techniques in estimation of variance and the related problems: A review. *Comm. Statist. - Theory and Methods*, **28**, 613-650. A special issue on "Statistical Inference and Data Analysis" in honor of retirement of Professor Nariaki Sugiura.

[10] Kubokawa, T. (2012). Minimax estimation of linear combinations of restricted location parameters. *Contemporary Developments in Bayesian Analysis and Statistical Decision Theory: A Festschrift for William E. Strawderman*, IMS Collections, Vol. **8**, 24-41, Institute of Mathematical Statistics.

[11] Kubokawa, T. and Srivastava, M.S. (1996). Double shrinkage estimators of ratio of variances. In *Multidimensional Statistical Analysis and Theory of Random Matrices* (Proceedings of the Sixth Eugene Lukacs Symposium, Bowling Green, Editors: A.K. Gupta and V.L. Girko), 139-154, VSP, Netherlands.

[12] Maruyama, Y. and Strawderman, W.E. (2006). A new class of minimax generalized Bayes estimators of a normal variance. *J. Statist. Plann. Inf.*, **136**, 3822-3836.

[13] Rukhin, A.L. (1992). Asymptotic risk behavior of mean vector and variance estimators and the problem of positive normal mean. *Ann. Inst. Statist. Math.*, **44**, 299-311.

[14] Shinozaki, N. (1995). Some modifications of improved estimators of a normal variance. *Ann. Inst. Statist. Math.*, **47**, 273-286.

[15] 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.

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