95-F-33

Double Shrinkage Estimators of Ratio of Variances

by

T. Kubokawa
University of Tokyo

M.S. Srivastava
University of Toronto

December 1995

Double Shrinkage Estimators of Ratio of Variances

T. Kubokawa and M.S. Srivastava

University of Tokyo and University of Toronto

The problem of estimating the ratio of two variances $\rho = \sigma_2^2/\sigma_1^2$, of normal distributions with unknown means are treated relative to the Kullback-Leibler loss function in a decision-theoretic framework. Using Stein's truncated estimators $1/\hat{\sigma}_1^{2ST}$ and $\hat{\sigma}_2^{2ST}$ for $1/\sigma_1^2$ and σ_2^2, respectively, it is shown that the unbiased estimator of ρ is improved upon by the double shrinkage estimator $\delta^{TR} = \hat{\sigma}_2^{2ST}/\hat{\sigma}_1^{2ST}$, which is shown to be an empirical Bayes estimator. A generalized Bayes estimator which is better than the unbiased one is also obtained. Interpreting δ^{TR} as a linearly combined estimator of two single shrinkage estimators with a random weight, some other improved combined procedures with double shrinkage are proposed. The risks of the estimators are compared by Monte Carlo simulation method. A generalization to the convex loss functions and to the distributions with monotone likelihood ratio properties is given. Improvements to the usual F-statistic as an estimator of ρ are also given.

Keywords and phrases: Point estimation, ratio of variances, shrinkage estimation, inadmissibility, Stein's truncated rule, monotone likelihood ratio properties, exponential, inverse Gaussian distributions, convex loss.

1. Introduction

Let S_1, S_2, X_1 and X_2 be independent random variables where for $i = 1, 2$, S_i/σ_i^2 has chi-square distribution $\chi^2_{m_i}$ with m_i degrees of freedom, and X_i has multivariate normal distribution $\mathcal{N}_{p_i}(\mu_i, \sigma_i^2 I_{p_i})$ with unknown mean μ_i. Such a problem arises naturally in many situations. For example, suppose we wish to compare the two mean vectors when $p_1 = p_2$ in the situation considered above. If the ratio of the two variances is considerably different from one, then we must consider test-statistics designed for Behrens-Fisher problem. On the other hand, if the ratio of the two variances is not too much different from one, we may continue using the usual test-procedures designed for the equal variance situations. Also, if possible, in future experiments more observations should be taken from the population with the large variance as recomended by Carter, Khatri and Srivastava(1979). The estimation of the ratio of two variances is also needed in estimating the power of the usual test for testing the equality of two variances. Thus, it seems desirable to estimate the ratio of the variances $\rho = \sigma_2^2/\sigma_1^2$. We consider the general situation when p_1 need not be equal to p_2 and where we use the information available from sample means. We shall denote our estimator by $\delta(S_1, S_2, X_1, X_2)$ where X_i denotes the normalized sample mean so that it has variance σ_i^2. We consider Kullback-Leibler loss function

$$L_{KL}(\delta/\rho) = \delta/\rho - \log \delta/\rho - 1, \tag{1.1}$$

which is motivated from the Kullback-Leibler information loss

$$\int \left\{ \log \frac{h(x,\delta)}{h(x,\rho)} \right\} h(x,\delta) dx,$$

$h(x,\rho)$ being the density function of $F = S_2/S_1$. Every estimator will be evaluated by the risk function $R(\omega,\delta) = E_\omega[L_{KL}(\delta/\rho)]$ for unknown parameters ω.

The uniformly minimum variance unbiased estimator is given by $\delta_0 = \{(m_1-2)/m_2\}F$, $F = S_2/S_1$, which is also the best among the class of estimators cF, $c > 0$, in the sense that it has minimum risk with respect to the loss (1.1). Our interest is to improve upon δ_0 by use of the information contained in X_1 and X_2. Gelfand and Dey(1988) applied Stein's arguments to get improved estimators. Let

$$W_1 = ||X_1||^2/S_1, \qquad W_2 = ||X_2||^2/S_2,$$
$$G_1 \equiv G_1(W_1) = \min\left\{1,\ \frac{m_1-2}{m_1+p_1-2}(1+W_1)\right\},$$
$$G_2 \equiv G_2(W_2) = \min\left\{1,\ \frac{m_2}{m_2+p_2}(1+W_2)\right\}.$$

Then $1/\sigma_1^2$ and σ_2^2 can respectively be estimated by Stein estimators for the Kullback-Leibler loss functions

$$\frac{1}{\hat{\sigma}_1^{2ST}} = \frac{m_1-2}{S_1}\frac{1}{G_1} = \left[\min\left\{\frac{S_1}{m_1-2}, \frac{S_1+||X_1||^2}{m_1+p_1-2}\right\}\right]^{-1},$$
$$\hat{\sigma}_2^{2ST} = \frac{S_2}{m_2}G_2 = \min\left\{\frac{S_2}{m_2}, \frac{S_2+||X_2||^2}{m_2+p_2}\right\}.$$

Using these estimators, Gelfand and Dey (1988) proposed two estimators δ_1 and δ_2 for ρ, given by

$$\delta_1 = \delta_0/G_1, \tag{1.2}$$
$$\delta_2 = \delta_0 G_2, \tag{1.3}$$

where $\delta_0 = \{(m_1-2)/m_2\}F$. Since these estimators use shrinkage estimators for only one of the parameters in the ratio, they may be called *single shrinkage estimators.* The estimators δ_1 and δ_2 have shrinkage in the opposite directions. This causes some technical dificulty in deriving an improved *double shrinkage estimator* using both $||X_1||^2$ and $||X_2||^2$ in the shrinkage. Kubokawa(1994b) recently succeeded in proposing one such an estimator given by

$$\delta_3 = \delta_1 + \delta_2 - \delta_0 \tag{1.4}$$
$$= \delta_1 + (G_2-1)\delta_0 \tag{1.5}$$
$$= \delta_2 + \left(\frac{1}{G_1}-1\right)\delta_0, \tag{1.6}$$

which dominates both δ_1 and δ_2, and hence δ_0. The second terms in (1.5) and (1.6) may be interpreted as an adjustment for over-shrinkage in δ_1 and δ_2 respectively.

Other possible double shrinkage estimators are considered in this paper. We recall that single shrinkage rules δ_1 and δ_2 are based on two parts of the dominance results :

(R.1) the estimator $(m_1-2)/S_1$ of $1/\sigma_1^2$ is improved on by $1/\hat{\sigma}_1^{2ST}$,

(R.2) the estimator S_2/m_2 of σ_2^2 is improved on by $\hat{\sigma}_2^{2ST}$.

These results have been studied in many papers. See Kubokawa (1994a,b) for bibliography. Here, the interesting issue is: Does combining (R.1) and (R.2) give a further dominance result ? In other words, we want to investigate whether the etimator $(S_2/m_2)/(S_1/(m_1-2))$ of $\rho = \sigma_2^2/\sigma_1^2$ can be improved upon by the multiplication of the improved procedures $1/\hat{\sigma}_1^{2ST}$ and $\hat{\sigma}_2^{2ST}$ as

$$\delta^{TR} = \frac{\hat{\sigma}_2^{2ST}}{\hat{\sigma}_1^{2ST}} = \frac{G_2}{G_1}\delta_0 = \frac{\min\{S_2/m_2, (S_2+||X_2||^2)/(m_2+p_2)\}}{\min\{S_1/(m_1-2), (S_1+||X_1||^2)/(m_1+p_1-2)\}}. \tag{1.7}$$

The answer is affirmative, and the main purpose of the paper is to provide its justification.

In Section 2, we obtain conditions which ensure that the double shrinkage estimator dominates both of the single shrinkage ones relative to the loss function (1.1). This result is easily exploited to develop two kinds of procedures: One is the truncated estimator δ^{TR} and the other is a smooth estimator δ^*. Bayesian properties of these rules are demonstrated: δ^{TR} is empirical Bayes and δ^* is generalized Bayes.

For another possible choice, it will be reasonable to consider a linearly combined estimator of δ_1 and δ_2, which is also discussed in Section 2. It is interesting to point out that the estimators δ_3 and δ^{TR} can be expressed as a linear combination of δ_1 and δ_2 with random weights. Other possible combined estimators are proposed and their improvements on δ_0 are analytically verified. In order to compare the risk performances of the improved estimators, simulation results are presented in Tables 1 and 2. They reveal that δ^{TR} and δ^* have uniformly smaller risks than others, and that δ^* is the best estimator with significant improvements for $\lambda_i = ||\mu_i||^2/\sigma_i^2$ far from zero. In particular, δ^{TR} and δ^* are better than δ_3 and so we propose the use of δ^{TR} and δ^*.

Some results of Section 2 are extended in Section 3 to the general convex loss functions and to the distributions with monotone likelihood ratio properties, including normal, lognormal, exponential, pareto and inverse Gaussian distributions. As a result, we get an improved double shrinkage estimator in the general situation. In Section 4, we discuss improvements of the usual F-statistic as an estimator of ρ, which is the best among a class of estimators cF with respect to the loss function proposed by Bilodeau and Srivastava (1992). It is analytically demonstrated that for this loss, the F-statistic $\delta_0^F = (S_2/m_2)/(S_1/m_1)$ is dominated by

$$\delta^{FTR} = \delta^F(\psi_1^{TR}, \psi_2^{TR}) = \frac{\min\{S_2/m_2, (S_2 + ||X_2||^2)/(m_2 + p_2)\}}{\min\{S_1/m_1, (S_1 + ||X_1||^2)/(m_1 + p_1)\}}. \tag{1.8}$$

2. Derivation of improved double shrinkage rules

Consider a class of double shrinkage estimators given by

$$\delta(\phi_1, \phi_2) = \frac{\phi_2(W_2)}{\phi_1(W_1)} F, \tag{2.1}$$

where

$$F = S_2/S_1, \quad W_i = ||X_i||^2/S_i, \quad i = 1, 2,$$

and ϕ_1 and ϕ_2 are positive and absolutely continuous functions. Single shrinkage estimators are written by

$$\delta_1(\phi_1) = \delta(\phi_1, m_2^{-1}) = \{m_2\phi_1(W_1)\}^{-1} F, \tag{2.2}$$

and

$$\delta_2(\phi_2) = \delta((m_2-2)^{-1}, \phi_2) = (m_1-2)\phi_2(W_2)F. \tag{2.3}$$

From Kubokawa(1994b), we first note that for $i = 1, 2$,

(S.i) the estimator δ_0 is dominated by the single one $\delta_i(\phi_i)$ if the following conditions hold:

(S.i.a) $\phi_i(w_i)$ is nondecreasing and $\lim_{w_i\to\infty}\phi_i(w_i) = \{m_i + 2(i-2)\}^{-1}$,

(S.i.b) $\phi_i(w_i) \geq \phi_i^*(w_i)$, where

$$\phi_i^*(w_i) = \frac{E^{v_i}[v_i^{i-2}F_i(w_iv_i)]}{E^{v_i}[v_i^{i-1}F_i(w_iv_i)]} = \frac{1}{m_i + p_i + 2(i-2)} \frac{\int_0^{w_i} z^{\frac{p_i}{2}-1}/(1+z)^{\frac{m_i+p_i}{2}+(i-2)}dz}{\int_0^{w_i} z^{\frac{p_i}{2}-1}/(1+z)^{\frac{m_i+p_i}{2}+(i-1)}dz},$$

where $v_i = S_i/\sigma_i^2$ and $F_i(y)$ is a distribution function of the chi-square random variable $\chi^2_{p_i}$.

Our main assertion is to indicate that double shrinkage estimator $\delta(\phi_1, \phi_2)$ dominates single shrinkage ones, which is ensured by the above condtions in (S.1) and (S.2) as shown in the next theorem.

Theorem 2.1. *The estimator $\delta(\phi_1, \phi_2)$, given by (2.1), is better than $\delta_1(\phi_1)$, given by (2.2), if $\phi_2(w_2)$ satisfies the conditions in (S.2) and if $\phi_1(w_1)$ is nondecreasing and $\phi_1(w_1) \leq (m_1-2)^{-1}$. Also $\delta(\phi_1, \phi_2)$ is better than $\delta_2(\phi_2)$, given by (2.3), if $\phi_1(w_1)$ satisfies the conditions in (S.1) and if $\phi_2(w_2)$ is nondecreasing and $\phi_2(w_2) \leq m_2^{-1}$.*

Corollary. *Assume that $\phi_1(w_1)$ and $\phi_2(w_2)$ satisfy the conditions in (S.1) and (S.2). Then $\delta(\phi_1, \phi_2)$ dominates both of $\delta_1(\phi_1)$ and $\delta_2(\phi_2)$, and hence superior to δ_0.*

The proof is given later. A class of estimators given by the theorem includes two interesting types of estimators. Letting

$$\phi_1^{TR}(w_1) = G_1(w_1)/(m_1-2) \quad \text{and} \quad \phi_2^{TR}(w_2) = G_2(w_2)/m_2,$$

we can see that $\phi_1^{TR}(w_1)$ and $\phi_2^{TR}(w_2)$ satisfy the conditions in (S.1) and (S.2). Also the conditions are satisfied by $\phi_1^*(w_1)$ and $\phi_2^*(w_2)$. Hence we get two double shrinkage estimators improving on the single shrinkage ones:

$$\delta^{TR} = \frac{\phi_2^{TR}(W_2)}{\phi_1^{TR}(W_1)}F, \quad \delta^* = \frac{\phi_2^*(W_2)}{\phi_1^*(W_1)}F,$$

where δ^{TR} is proposed in (1.7).

We shall provide Bayesian properties of δ^{TR} and δ^*. First, it is demonstrated that δ^{TR} is an empirical Bayes estimator. For $i = 1, 2$, let $\eta_i = 1/\sigma_i^2$ and suppose that μ_i and η_i have the following prior distributions due to Kubokawa, Robert and Saleh(1992):

$$\mu_i \mid \eta_i, \tau_i \sim \mathcal{N}_{p_i}\left(0, \frac{1-\tau_i}{\tau_i}\frac{1}{\eta_i}I_{p_i}\right),$$
$$\eta_i \sim \eta_i^{-1} d\eta_i, \tag{2.4}$$

where τ_i, $0 < \tau_i < 1$, is an unknown hyper-parameter and $\mu_i|\eta_i$ designates the conditional distribution of μ_i given η_i. Then the posterior distribution of η_i is given by

$$\eta_i \mid S_i, X_i \ \sim \ \tau_i^{\frac{p_i}{2}} \eta_i^{\frac{m_i+p_i}{2}-1} e^{-\frac{\eta_i}{2}(S_i+\tau_i||X_i||^2)}.$$

The Bayes estimator of $\rho = \eta_1/\eta_2$ for the loss (1.1) is

$$\delta^B = \left(E[\rho^{-1}|D]\right)^{-1} = \left(E[\eta_2|D]E[\eta_1^{-1}|D]\right)^{-1} \tag{2.5}$$

where $D = (S_1, S_2, X_1, X_2)$ and $E[\cdot|D]$ denotes the posterior expectation of η_1 and η_2. Here,

$$E[\eta_2|D] = \frac{m_2 + p_2}{S_2 + \tau_2||X_2||^2},$$
$$E[\eta_1^{-1}|D] = \frac{S_1 + \tau_1||X_1||^2}{m_1 + p_1 - 2}. \tag{2.6}$$

Since τ_1 and τ_2 are unknown, they need be estimated from the marginal distributions of (S_i, X_i), $i = 1, 2$, which are given by

$$S_i^{\frac{m_i}{2}-1} \frac{\tau_i^{p_i/2}}{(S_i + \tau_i||X_i||^2)^{(m_i+p_i)/2}}, \ i = 1, 2.$$

The MLE of τ_2 is $\hat{\tau}_2 = \min\{p_2S_2/(m_2||X_2||^2), 1\}$ and the adjusted MLE of τ_1 is $\hat{\tau}_1 = \min\{p_1S_1/\{(m_1-2)||X_1||^2\}, 1\}$ where m_1 is replaced with $m_1 - 2$. Substituting $\hat{\tau}_1$ and $\hat{\tau}_2$ into (2.6), we get an empirical Bayes estimator, which is identical to δ^{TR} given by (1.7).

For the generalized Bayesness of δ^*, following Brewster and Zidek (1974), suppose the prior distributions (2.4) and that

$$\tau_i \sim \tau_i^{-1} I_{(0,1)}(\tau_i) d\tau_i, \ i = 1, 2. \tag{2.7}$$

A similar argument can show that $E[\eta_1^{-1}|D] = \phi_1^*(W_1)S_1$ and $1/E[\eta_2|D] = \phi_2^*(W_2)S_2$, which can be verified to satisfy the conditions in (S.1) and (S.2). Hence we get the double shrinkage and generalized Bayes estimator $\delta^* = \delta(\phi_1^*, \phi_2^*)$ improving upon the single ones.

Other possible candidates for double shrinkage estimators are obtained by taking linear combinations of δ_1 and δ_2. Such an estimator is of the form $\alpha\delta_1 + (1-\alpha)\delta_2$ for $0 < \alpha < 1$,and from the convexity of the loss function, it can be seen to be superior to δ_0. Since the optimal value of α depends on two noncentrality parameters, it is reasonable to take random weight $\hat{\alpha}$ based on W_1 and W_2. Here it may be interesting to note that the double shrinkage estimators δ_3 and δ^{TR} can be written as a linear combination of δ_1 and δ_2, namely,

$$\delta_3 = \hat{\alpha}\delta_1 + (1-\hat{\alpha})\delta_2 \quad \text{for} \quad \hat{\alpha} = \frac{1-G_1}{1-G_1G_2},$$

$$\delta^{TR} = \hat{\alpha}\delta_1 + (1-\hat{\alpha})\delta_2 \quad \text{for} \quad \hat{\alpha} = \frac{G_2(1-G_1)}{1-G_1G_2}$$

in the case of $G_1G_2 \neq 1$. When $G_1 = G_2 = 1$, $\hat{\alpha}$ may be defined to be any constant since $\delta_1 = \delta_2 = \delta_0$. The random weight $\hat{\alpha}$ shall take a value close to one when W_1 is quite small and W_2 is relatively large. Taking this aspect into account, we may consider the following weight functions:

$$\hat{\alpha}_1 = \frac{G_2(W_2)}{G_1(W_1) + G_2(W_2)}, \tag{2.8}$$

$$\hat{\alpha}_2 = \frac{\frac{m_2}{m_2+p_2}(1+W_2)}{\frac{m_1-2}{m_1+p_1-2}(1+W_1) + \frac{m_2}{m_2+p_2}(1+W_2)}, \tag{2.9}$$

$$\hat{\alpha}_3 = \frac{(m_2/p_2)W_2}{(m_1/p_1)W_1 + (m_2/p_2)W_2}. \tag{2.10}$$

For $i = 1, 2, 3$, let $\delta_i^C = \hat{\alpha}_i\delta_1 + (1-\hat{\alpha}_i)\delta_2$. In general, it is not ensured that the random-weighted combination estimators will have superior properties. For a weight with a specific monotonicity, however, the superiority is guranteed as shown in the following theorem.

Theorem 2.2. *If weighting function $r(w_1, w_2)$ is nonincreasing in w_1 and nondecreasing in w_2, then the estimator*

$$\delta^C(r) = r(W_1, W_2)\delta_1 + (1 - r(W_1, W_2))\delta_2 \tag{2.11}$$

dominates δ_0 relative to the loss (1.1).

The proof is given at the end of this section. Theorem 2.2 implies that the random-weight combined estimators δ_1^C, δ_2^C and δ_3^C given above have smaller risks than δ_0.

We now provide Monte Carlo simulation results for the risk function of estimators δ_0, δ_1, δ_2, δ_3, δ^{TR}, δ^*, δ_1^C, δ_2^C and δ_3^C treated in the above discussions. The simulation experiments are done in the cases of $m_1 = m_2 = 3; p_1 = p_2 = 3, 10; \sigma_1^2 = \sigma_2^2 = 1; \lambda_1 = ||\mu_1||^2/\sigma_1^2 = 0.0, 0.5, 1.0, 5.0, 10.0; \lambda_2 = ||\mu_2||^2/\sigma_2^2 = 0.0, 0.5, 1.0, 5.0, 10.0$. Tables 1 and 2 report the average values of the risks based on 50,000 replications. From the tables, it is revealed that δ^{TR} and δ^* have smaller risks than others, and that δ^* is the best estimator with significant improvements for λ_i far from zero. Among combined estimators, δ_2^C and δ_3^C are better than δ_1^C and also superior to δ_1 and δ_2. It is of interest to point out that δ^{TR} and δ^* have uniformly smaller risks than δ_3. It is also indicated that the risk gain of δ_1 is much greater than that of δ_2. This may arise from the unstableness of the denominator of δ_0 in comparison with the numerator. That is, the simulation result for δ_1 and δ_2 implies that stabilizing the denominator yields a more improvement than stabilizing the numerator. Although δ^* has a complicated form including the ratio of integrals, it can be expressed by the incomplete beta functions ratio $I_x(\cdot,\cdot)$, for the integrals are written as

$$\begin{aligned}\int_0^w z^\alpha/(1+z)^{\alpha+\beta}dz &= \int_0^{w/(1+w)} x^\alpha(1-x)^{\beta-2}dx \\ &= B(\alpha+1,\beta-1)I_{w/(1+w)}(\alpha+1,\beta-1).\end{aligned}$$

When a table of values of the incomplete beta functions ratio is available, one can employ δ^* in a practical use. In this way, δ^{TR} and δ^* derived in this paper have superior risk performances to others as in the simulation results, and so we propose the use of δ^{TR} and δ^*.

We conclude this section with giving the proofs of the theorems.

Proof of Theorem 2.1. We shall prove that $\delta(\phi_1, \phi_2)$ dominates $\delta_2(\phi_2)$. Since parts of the results in this section are extended to the general convex loss functions in Section 3, we here prove it for such a general situation. Let $L(t)$ be a positive and convex function such that the first derivative $L'(t)$ is strictly increasing and $L(1) = 0$. To evaluate the risk difference of two estimators, we exploit the *IERD(Integral Expression of Risk Difference) method* proposed by Kubokawa(1994a,b). Noting that $\phi_1(\infty) = (m_1 - 2)^{-1}$, we observe that

$$\begin{aligned}&R(\omega, \delta_2(\phi_2)) - R(\omega, \delta(\phi_1, \phi_2)) \\ &\quad = E_\omega\left[L\left(\frac{\phi_2(W_2)}{\phi_1(\infty)}\frac{F}{\rho}\right) - L\left(\frac{\phi_2(W_2)}{\phi_1(W_1)}\frac{F}{\rho}\right)\right] \\ &\quad = E_\omega\left[L\left(\frac{\phi_2(W_2)}{\phi_1(tW_1)}\frac{F}{\rho}\right)\Big|_{t=1}^{\infty}\right] \\ &\quad = E_\omega\left[\int_1^\infty \frac{d}{dt}\left\{L\left(\frac{\phi_2(W_2)}{\phi_1(tW_1)}\frac{F}{\rho}\right)\right\}dt\right]\end{aligned}$$

$$= E_\omega \left[\int_1^\infty L'\left(\frac{\phi_2(W_2)}{\phi_1(tW_1)} \frac{F}{\rho} \right) \left\{ -\frac{\phi_2(W_2)}{\phi_1(tW_1)^2} \phi_1'(tW_1) W_1 \frac{F}{\rho} \right\} dt \right]$$
$$\geq -E_\omega \left[\int_1^\infty L'\left(\frac{F/\rho}{m_2\phi_1(tW_1)} \right) \frac{F}{\rho} W_1 \frac{\phi_2(W_2)}{\phi_1(tW_1)^2} \phi_1'(tW_1) dt \right], \quad (2.12)$$

since L' and ϕ_1 are nondecreasing and $\phi_2 \leq m_2^{-1}$. We here consider the following conditional expectation given $||X_2||^2/\sigma_2^2 = c$:

$$I_0(c) = E_\omega \left[\int_1^\infty L'\left(\frac{F/\rho}{m_2\phi_1(tW_1)} \right) \frac{F}{\rho} W_1 \frac{\phi_2(c\sigma_2^2/S_2)}{\phi_1(tW_1)^2} \phi_1'(tW_1) dt \Big| \frac{||X_2||^2}{\sigma_2^2} = c \right].$$

Then it is sufficient to show that $I_0(c) \leq 0$ for any positive c.

Let $v_i = S_i/\sigma_i^2$ and $u_i = ||X_i||^2/\sigma_i^2$ for $i = 1, 2$. Then v_i and u_i have respectively central chi-square $\chi^2_{m_i}$ and non-central chi-square $\chi^2_{p_i}(\lambda_i)$ distributions with noncentrality parameters $\lambda_i = ||\mu_i||^2/\sigma_i^2$, whose densities are designated by $g_i(v_i)$ and $f_i(u_i; \lambda_i)$, respectively. Based on these random variables,

$$I_0(c) = E^{v_2} \left[\int \int \int_1^\infty L'\left(\frac{v_2/v_1}{m_2\phi_1(tu_1/v_1)} \right) \frac{v_2}{v_1} \frac{u_1}{v_1} \frac{\phi_1'(tu_1/v_1)}{\phi_1(tu_1/v_1)^2} dt \right.$$
$$\left. \times f_1(u; \lambda_1) g_1(v_1) du_1 dv_1 \phi_2\left(\frac{c}{v_2}\right) \right]. \quad (2.13)$$

Making the transformation $w_1 = (t/v_1)u_1$ with $dw_1 = (t/v_1)du_1$ gives

$$I_0(c) = E^{v_2} \left[\int \int \int_1^\infty L'\left(\frac{v_2/v_1}{m_2\phi_1(w_1)} \right) \frac{v_2}{v_1} \frac{w_1 v_1}{t^2} \frac{\phi_1'(w_1)}{\phi_1(w_1)^2} \right.$$
$$\left. \times f_1(\frac{w_1 v_1}{t}; \lambda_1) g_1(v_1) dt dw_1 dv_1 \phi_2\left(\frac{c}{v_2}\right) \right]. \quad (2.14)$$

Making again the transformation $x = (w_1 v_1)/t$ with $dx = (w_1 v_1/t^2)dt$ yields

$$I_0(c) = E^{v_2} \left[\int \int \int_0^{w_1 v_1} L'\left(\frac{v_2/v_1}{m_2\phi_1(w_1)} \right) \frac{V_2}{v_1} \frac{\phi_1'(w_1)}{\phi_1(w_1)^2} \right.$$
$$\left. \times f_1(x; \lambda_1) g_1(v_1) dx dv_1 dw_1 \phi_2\left(\frac{c}{v_2}\right) \right]$$
$$= \int \frac{\phi_1'(w_1)}{\phi_1(w_1)^2} E^{v_1, v_2} \left[L'\left(\frac{v_2/v_1}{m_2\phi_1(w_1)} \right) \frac{v_2}{v_1} F_1(w_1 v_1; \lambda_1) \phi_2\left(\frac{c}{v_2}\right) \right] dw_1, \quad (2.15)$$

where $F_1(y; \lambda_1) = \int_0^y f_1(x; \lambda_1) dx$.

Letting

$$G_1(v_1) = E^{v_2} \left[L'\left(\frac{v_2}{v_1 m_2 \phi_1(w_1)} \right) v_2 \phi_2\left(\frac{c}{v_2}\right) \Big| v_1 \right],$$

we see that $G_1(v_1)$ has one sign change, that is, there exists a point v_{10} such that $G_1(v_1) > 0$ for $v_1 < v_{10}$ and $G_1(v_1) < 0$ for $v_1 > v_{10}$. Thereby, $G_1(v_1)v_1^{-1}F_1(w_1v_1 : 0)$ has one sign change at v_{10}. Combining this fact and the monotonicity of $F_1(x;\lambda_1)/F(x;0)$ guarantees that the following inequality holds:

$$E^{v_1}\left[G_1(v_1)v_1^{-1}F_1(w_1v_1;0)\frac{F_1(w_1v_1;\lambda_1)}{F_1(w_1v_1;0)}\right] \leq E^{v_1}\left[G_1(v_1)v_1^{-1}F_1(w_1v_1;0)\right]\cdot\left\{\frac{F_1(w_1v_{10};\lambda_1)}{F_1(w_1v_{10};0)}\right\}. \tag{2.16}$$

From (2.15) and (2.16), it is seen that $I_0(c) \leq 0$ if we can show that

$$E^{v_1,v_2}\left[L'\left(\frac{v_2/v_1}{m_2\phi_1(w_1)}\right)\frac{v_2}{v_1}F_1(w_1v_1)\phi_2\left(\frac{c}{v_2}\right)\right] \leq 0 \tag{2.17}$$

for $F_1(y) = F_1(y;0)$. Letting

$$G_2(v_2) = E^{v_1}\left[L'\left(\frac{v_2}{v_1m_2\phi_1(w_1)}\right)v_1^{-1}F_1(w_1v_1)\Big|v_2\right],$$

we see that $G_2(v_2)$ has one sign change at some point v_{20}. Since $\phi_2(c/v_2)$ is nonincreasing in v_2, the same argument as in (2.17) establishes the inequality

$$E^{v_2}\left[G_2(v_2)v_2\phi_2\left(\frac{c}{v_2}\right)\right] \leq E^{v_2}\left[G_2(v_2)v_2\right]\cdot\phi_2\left(\frac{c}{v_{20}}\right), \tag{2.18}$$

which yields a sufficient condition

$$E^{v_1,v_2}\left[L'\left(\frac{v_2/v_1}{m_2\phi_1(w_1)}\right)\frac{v_2}{v_1}F_1(w_1v_1)\right] \leq 0. \tag{2.19}$$

Since $L'_{KL}(x) = 1 - 1/x$, the condition (2.19) is expressed by

$$m_2\phi_1(w_1) \geq E^{v_1,v_2}\left[\frac{v_2}{v_1}F_1(w_1v_1)\right]/E^{v_1}\left[F_1(w_1v_1)\right],$$

or

$$\phi_1(w_1) \geq E^{v_1}\left[\frac{1}{v_1}F_1(w_1v_1)\right]/E^{v_1}\left[F_1(w_1v_1)\right], \tag{2.20}$$

the r.h.s. of which is equal to $\phi_1^*(w_1)$ given in the condition (S.1.b). Hence the first assertion is proved.

Similarly, we can show that $\delta(\phi_1,\phi_2)$ dominates $\delta_1(\phi_1)$ where in this case, we need to restrict the loss function to (1.1). For the general convex loss, we could not utilize an inequality similar to (2.18) so as to approximate the risk difference. In any case, Theorem 2.1 is established for the loss (1.1).

Proof of Theorem 2.2. From the convexity of the loss function,

$$\begin{aligned} &E[L_{KL}(\hat{\alpha}\delta_1 + (1-\hat{\alpha})\delta_2, \rho)] \\ &\quad \le E[\hat{\alpha}L_{KL}(\delta_1,\rho) + (1-\hat{\alpha})L_{KL}(\delta_2,\rho)] \\ &\quad = E[\hat{\alpha}\{L_{KL}(\delta_1,\rho) - L_{KL}(\delta_0,\rho)\}] \\ &\qquad + E[(1-\hat{\alpha})\{L_{KL}(\delta_2,\rho) - L_{KL}(\delta_0,\rho)\}] + E[L_{KL}(\delta_0,\rho)] \\ &\quad = I_1 + I_2 + I_3, \quad \text{say.} \end{aligned} \tag{2.21}$$

It suffices to show that $I_1 \le 0$ and $I_2 \le 0$. To show $I_1 \le 0$, we write

$$\begin{aligned} L_{KL}(\delta_1,\rho) - L_{KL}(\delta_0,\rho) &= \frac{\delta_0}{\rho G_1} - \log\frac{\delta_0}{\rho G_1} - \frac{\delta_0}{\rho} + \log\frac{\delta_0}{\rho} \\ &= \frac{\delta_0}{\rho}\left(\frac{1}{G_1} - 1\right) + \log G_1. \end{aligned} \tag{2.22}$$

From the monotonicity of $r(w_1, w_2)$,

$$E\left[r\left(w_1, \frac{||X_2||^2}{S_2}\right) S_2 \Big| w_1\right] \le E\left[r\left(w_1, \frac{||X_2||^2}{S_2}\right)\Big| w_1\right] E\left[S_2 \Big| w_1\right] \tag{2.23}$$

which implies that

$$\begin{aligned} &E\left[r(w_1, W_2)\left\{\frac{\delta_0}{\rho}\left(\frac{1}{G_1} - 1\right) + \log G_1\right\}\Big| w_1\right] \\ &\quad \le E\left[r(w_1, W_2)\Big| w_1\right] E\left[\left\{\frac{\delta_0}{\rho}\left(\frac{1}{G_1} - 1\right) + \log G_1\right\}\Big| w_1\right] \\ &\quad = E\left[r(w_1, W_2)\Big| w_1\right] \\ &\qquad \times E\left[\left\{L_{KL}\left(\frac{\sigma_1^2}{\hat{\sigma}_1^{2ST}}\right) - L_{KL}\left(\frac{\sigma_1^2}{S_1/(m_1-2)}\right)\right\}\Big| w_1\right], \end{aligned} \tag{2.24}$$

which is not positive as verified by recalling Stein's original method (1964) for the domination.

For I_2, note that

$$L_{KL}(\delta_2,\rho) - L_{KL}(\delta_0,\rho) = \frac{\delta_0}{\rho}(G_2 - 1) - \log G_2, \tag{2.25}$$

and that

$$E\left[\left\{1-r\left(\frac{||X_1||^2}{S_1},w_2\right)\right\}\frac{1}{S_1}\Big|w_2\right]\geq E\left[1-r\left(\frac{||X_1||^2}{S_1},w_2\right)\Big|w_2\right]E\left[\frac{1}{S_1}\Big|w_2\right]. \quad (2.26)$$

The remainder can be shown similarly to the case $I_1 \leq 0$, and therefore the theorem is proved.

3. Generalization of distributions and loss functions

The results of Section 2 are extended to distributions with monotone likelihood ratio properties and to the general convex loss functions.

Let S_1, S_2, T_1 and T_2 be independent random variables where for $i = 1, 2$, $v_i = S_i/\sigma_i$ and $u_i = T_i/\sigma_i$ have densities

$$g_i(v_i)I_{[v_i>0]} \quad \text{and} \quad h_i(u_i;\lambda_i)I_{[u_i>k_i(\lambda_i)]} \quad (3.1)$$

for unknown real parameter λ_i, real function $k_i(\lambda_i)$, $k_i(0) = 0$, and the indicator function $I_{[\cdot]}$. Then we want to estimate the ratio of the scales $\rho = \sigma_2/\sigma_1$ by an estimator δ relative to the convex loss function $L(\delta/\rho)$ where $L'(t)$ is strictly increasing and $L(1) = 0$.

The best multipliers c_1 and c_2 of the estimators $c_2S_2/(c_1S_1)$ are defined by solutions of the equations

$$E^{v_1,v_2}\left[L'\left(\frac{c_2v_2}{c_1v_1}\right)\frac{v_2}{v_1}\right]=0, \quad (3.2)$$

$$E^{v_1,v_2}\left[L'\left(c_2v_2\right)v_2\right]=0. \quad (3.3)$$

For improving on $\delta_0 = c_2S_2/(c_1S_1)$, consider a class of estimators

$$\delta(\phi_1,\phi_2)=\begin{cases}\{\phi_2(W_2)/\phi_1(W_1)\}S_2/S_1 & \text{if } W_1>0, W_2>0;\\ \{c_2/\phi_1(W_1)\}S_2/S_1 & \text{if } W_1>0, W_2\leq 0;\\ \{\phi_2(W_2)/c_1\}S_2/S_1 & \text{if } W_1\leq 0, W_2>0;\\ (c_2/c_1)S_2/S_1 & \text{if } W_1\leq 0, W_2\leq 0,\end{cases} \quad (3.4)$$

where $W_i = T_i/S_i$, $i = 1, 2$, and ϕ_1, ϕ_2 are positive and absolutely continuous. To establish the dominance, we assume that

(A.1) $H_i(x;\lambda_i)/H_i(x)$ is nondecreasing in $x > 0$ for $i = 1, 2$,

where $H_i(x;\lambda_i) = \int_0^x h_i(u;\lambda_i)I_{[u\geq k_i(\lambda_i)]}du$ and $H_i(x) = \int_0^x h_i(u)du$ for $h_i(u) = h_i(u;0)$. Note that (A.1) is guaranteed if

(A.1′) $h_i(x;\lambda_i)/h_i(x)$ is nondecreasing in $x > \max(0, k_i(\lambda_i))$.

The following dominance results due to Kubokawa(1994b) hold under the assumption (A.1):

(SG.1) The estimator δ_0 is dominated by the single shrinkage estimator $\delta_1(\phi_1) = \delta(\phi_1, c_2)$ if

(SG.1.a) $\phi_1(w_1)$ is nondecreasing and $\lim_{w_1\to\infty}\phi_1(w_1) = c_1$,

(SG.1.b) $E^{v_1,v_2}[L'(c_2v_2/\phi_1(w_1)v_1)(v_2/v_1)H_1(w_1v_1)] \leq 0$.

(SG.2) δ_0 is dominated by $\delta_2(\phi_2) = \delta(c_1, \phi_2)$ if

(SG.2.a) $\phi_2(w_2)$ is nondecreasing and $\lim_{w_2\to\infty}\phi_2(w_2) = c_2$,

(SG.2.b) $E^{v_1,v_2}[L'(\phi_2(w_2)v_2/c_1v_1)(v_2/v_1)H_2(w_2v_2)] \geq 0$.

The same arguments as in Kubokawa(1994a) and the proof of Theorem 2.1 can be exploited to show

Theorem 3.1. *Assume (A.1). Then,*

(1) The double shrinkage estimator $\delta(\phi_1, \phi_2)$, given by (3.4), dominates $\delta_2(\phi_2)$ if the conditions in (SG.1) are satisfied and if $\phi_2(w_2)$ is nondecreasing and $\phi_2(w_2) \leq c_2$.

(2) If the conditions in (SG.1) and (SG.2) hold, then $\delta(\phi_1, \phi_2)$ is better than δ_0.

(3) If $L(x)$ is specified by $L(x) = x - \log x - 1$, then $\delta(\phi_1, \phi_2)$ dominates $\delta_1(\phi_1)$ under the conditions in (SG.2) and if $\phi_1(w_1)$ is nondecreasing and $\phi_1(w_1) \leq c_1$.

Note. It is not easy to prove part (3) of Theorem 3.1 for the general convex loss functions.

The proof is omitted. In the case where $L(x) = x - \log x - 1$, the conditions are simplied as follows: Let

$$\phi_1^{G*}(w_1) = E^{v_1}[v_1^{-1}H_1(w_1v_1)]/E^{v_1}[H_1(w_1v_1)]$$
$$\phi_2^{G*}(w_2) = E^{v_2}[H_2(w_2v_2)]/E^{v_2}[v_2H_2(w_2v_2)].$$

Then (SG.1.b) and (SG.2.b) are replaced with $\phi_1(w_1) \geq \phi_1^{G*}(w_1)$ and $\phi_2(w_2) \geq \phi_2^{G*}(w_2)$, respectively. Denote

$$\phi_1^{GTR}(w_1) = \min\{c_1, E^{v_1}[h_1(w_1v_1)]/E^{v_1}[v_1h_1(w_1v_1)]\}$$
$$\phi_2^{GTR}(w_2) = \min\{c_2, E^{v_2}[v_2h_2(w_2v_2)]/E^{v_2}[h_2(w_2v_2)]\}$$

for $c_1 = E[v_1^{-1}]$ and $c_2 = 1/E[v_2]$. It is easily checked that $\phi_1^{G*}(w_1)$, $\phi_2^{G*}(w_2)$, $\phi_1^{GTR}(w_1)$ and $\phi_2^{GTR}(w_2)$ satisfy the conditions in (SG.1) and (SG.2) if the following assumption holds:

(A.2) $H_i(d_1x)/H_i(d_2x)$ is nondecreasing in x for $0 < d_1 < d_2$ and $i = 1, 2$.

This assumption is guaranteed if

(A.2$'$) $h_i(d_1x)/h_i(d_2x)$ is nondecreasing in x for $0 < d_1 < d_2$ and $i = 1, 2$,

which also implies that $xh_i(x)/H_i(x)$ is decreasing. So under the distributional assumptions (A.1) and (A.2), we get two kinds of double shrinkage estimators $\delta(\phi_1^{G*}, \phi_2^{G*})$ and $\delta(\phi_1^{GTR}, \phi_2^{GTR})$ improving both of single shrinkage ones for the loss (1.1).

The assumptions (A.1) and (A.2) are satisfied for normal, lognormal, exponential, pareto and inverse Gaussian distributions. For the exponential distributions, see Kubokawa (1994b) and Madi and Tsui(1990). For the inverse Gaussian distribution, Kourouklis(1995) recently proved that it satisfies the assumptions (A.1) and (A.2), and so we get improved double shrinkage estimators.

4. Improvement on the F-statistic as an estimator

In the previous sections, the improvements on the unbiased estimator $\delta_0 = \{(m_1 - 2)/m_2\}F$, $F = S_2/S_1$, are dealt with. This is not a usual F-statistic, which is given by

$$\delta_0^F = \frac{m_1}{m_2}F \tag{4.1}$$

and is of the natural form such that the numerator and the denomenator are divided by their degrees of freedom.

In this section, we consider improving on δ_0^F relative to the following loss function due to Bilodeau and Srivastava(1992):

$$L_{BS}(\delta, \rho; F) = \frac{\delta + F}{\rho + F} - \log\frac{\delta + F}{\rho + F} - 1, \tag{4.2}$$

which satisfies the convexity and $L_{BS}(\rho, \rho; F) = 0$, but depends on the data through F. It is easy to see that δ_0^F is the best of estimators cF for the loss $L_{BS}(\delta, \rho; F)$. For improving on δ_0^F, consider a class of the estimators

$$\delta^F(\psi_1, \psi_2) = \frac{\psi_2(W_2)}{\psi_1(W_1)}F, \quad W_i = ||X_i||^2/S_i. \tag{4.3}$$

The single shrinkage estimators are defined by $\delta_1^F(\psi_1) = \{m_2\psi_1(W_1)\}^{-1}F$ and $\delta_2^F(\psi_2) = m_1\psi_2(W_2)F$.

The conditions for dominance **(SF.i)** are described for $i = 1, 2$ as

(SF.i.a) $\psi_i(w_i)$ is nondecreasing and $\lim_{w_i\to\infty}\psi_i(w_i) = m_i^{-1}$,

(SF.i.b) $\psi_i(w_i) \geq \psi_i^*(w_i)$, where

$$\psi_i^*(w_i) = \frac{1}{m_i + p_i}\frac{\int_0^{w_i} z^{\frac{p_i}{2}-1}/(1+z)^{\frac{m_i+p_i}{2}}dz}{\int_0^{w_i} z^{\frac{p_i}{2}-1}/(1+z)^{\frac{m_i+p_i}{2}+1}dz}.$$

Then the following results are obtained for the loss (4.2).

Theorem 4.1.

(1) For $i = 1, 2$, δ_0 is dominated by the single shrinkage estimator $\delta_i^F(\psi_i)$ under the conditions in (SF.i).

(2) The double shrinkage estimator $\delta^F(\psi_1, \psi_2)$ dominates $\delta_2^F(\psi_2)$ if the conditions in (SF.1) are satisfied and if $\psi_2(w_2)$ is nondecreasing and $\psi_2 \le m_2^{-1}$.

(3) If all the conditions in (SF.1) and (SF.2) hold, then $\delta^F(\psi_1, \psi_2)$ dominates δ_0^F.

Note. Similar to Theorem 3.1, we could not establish that $\delta^F(\psi_1, \psi_2)$ dominates $\delta_1^F(\psi_1)$ for the loss (4.2).

The conditions in (SF.i) are satisfied by $\psi_i^*(w_i)$ and $\psi_i^{TR}(w_i) = \min\{m_i^{-1}, (1 + w_i)(m_i + p_i)^{-1}\}$, which yield superior double shrinkage estimators $\delta^F(\psi_1^*, \psi_2^*)$ and

$$\delta^{FTR} = \delta^F(\psi_1^{TR}, \psi_2^{TR}) = \frac{\min\{S_2/m_2, (S_2 + ||X_2||^2)/(m_2 + p_2)\}}{\min\{S_1/m_1, (S_1 + ||X_1||^2)/(m_1 + p_1)\}}. \tag{4.4}$$

Proof of Theorem 4.1. We first prove part (2). From the IERD method,

$$\begin{aligned} &R(\omega, \delta_1^F(\psi_1) - R(\omega, \delta^F(\psi_1, \psi_2)) \\ &= E_\omega\left[\int_1^\infty \frac{d}{dt}\left\{L\left(\frac{\psi_2(W_2)}{\psi_1(tW_1)}F, \rho; F\right)\right\} dt\right] \\ &= E_\omega\left[\int_1^\infty \left\{\frac{F}{\rho + F} - \frac{1}{\psi_2(W_2)/\psi_1(tW_1) + 1}\right\}\left\{-\frac{\psi_2(W_2)}{\psi_1(tW_1)^2}W_1\psi_1'(tW_1)\right\} dt\right] \\ &\ge E_\omega\left[\int_1^\infty \left\{\frac{1}{\{m_2\psi_1(tW_1)\}^{-1} + 1} - \frac{F}{\rho + F}\right\}\frac{\psi_2(W_2)}{\psi_1(tW_1)^2}\psi_1'(tW_1)W_1 dt\right], \end{aligned} \tag{4.5}$$

since $\psi_2 \le m_2^{-1}$. Noting that $-F/(\rho + F)$ and $\psi_2(||X_2||^2/S_2)$ are monotone in the same direction with respect to S_2, we see that the r.h.s. of the inequality on (4.5) is greater than or equal to

$$E_\omega[\psi_2(W_2)]\, E_\omega\left[\int_1^\infty \left\{\frac{1}{\{m_2\psi_1(tW_1)\}^{-1} + 1} - \frac{F}{\rho + F}\right\}\frac{\psi_1'(tW_1)}{\psi_1(tW_1)^2}W_1 dt\right]. \tag{4.6}$$

By making transformations similar to as in the proof of Theorem 2.1, the second expectation in (4.6) can be expressed as

$$\int_0^\infty \frac{\psi_1'(w_1)}{\psi_1^2(w_1)}E^{v_1, v_2}\left[\left(A(w_1) - \frac{v_2}{v_1 + v_2}\right)F_1(w_1v_1; \lambda_1)\right] dw_1, \tag{4.7}$$

where $A(w_1) = 1/\{(m_2\psi_1(w_1))^{-1} + 1\}$ and $F_1(y; \lambda_1)$ is a distribution function of $||X_1||^2/\sigma_1^2$. Let $G(v_1) = E^{v_2}[A(w_1) - v_2/(v_1 + v_2)|v_1]$, then it is increasing in v_1, so that

$$E^{v_1}\left[G(v_1)F_1(w_1v_1)\frac{F_1(w_1v_1; \lambda_1)}{F_1(w_1v_1)}\right] \ge E^{v_1}[G(v_1)F_1(w_1v_1)]\, E^{v_1}\left[\frac{F_1(w_1v_1; \lambda_1)}{F_1(w_1v_1)}\right].$$

This implies that it is sufficient to show that

$$m_2\psi_1(w_1) \geq \frac{E^{v_1,v_2}[v_2/(v_1+v_2)F_1(w_1v_1)]}{E^{v_1,v_2}[v_1/(v_1+v_2)F_1(w_1v_1)]}. \tag{4.8}$$

To approximate the r.h.s. of (4.8), let $V = v_1 + v_2$ and $Z = v_1/(v_1+v_2)$, which are independent, and we shall prove that

$$\frac{E[(1-Z)F_1(w_1ZV)]}{E[ZF(w_1ZV)]} \leq \frac{E[(1-Z)VF_1(w_1ZV)]}{E[ZVF(w_1ZV)]}, \tag{4.9}$$

which is equivalently written by

$$E^*\left[\frac{1-Z}{Z}\right]E^*[V] \leq E^*\left[\frac{1-Z}{Z}\cdot V\right], \tag{4.10}$$

where $E^*[\cdot]$ is taken with respect to the probability measure

$$P^*(A) = E[I_A Z F_1(w_1ZV)]/E[ZF_1(w_1ZV)].$$

Note that given V, the conditional expectation of $(1-Z)/Z$ is represented by

$$E^*\left[\frac{1-Z}{Z}\middle|V=v\right] = \frac{E[(1-Z)F_1(w_1Zv)|V=v]}{E[ZF_1(w_1Zv)|V=v]}. \tag{4.11}$$

Differentiating the r.h.s. of (4.11) with respect to v, we can verify that $E^*[(1-Z)/Z|V=v]$ is increasing in v if $xf_1(x)/F_1(x)$ is decreasing in x, which is also guaranteed for chi-square distributions. So, for the l.h.s. of (4.11), we have that

$$\begin{aligned} E^*\left[\frac{1-Z}{Z}\right]E^*[V] &= E^{*V}\left[E^{*Z|V}\left[\frac{1-Z}{Z}\middle|V\right]\right]E^*[V] \\ &\leq E^{*V}\left[E^{*Z|V}\left[\frac{1-Z}{Z}\middle|V\right]\cdot V\right] \\ &= E^{*V}\left[E^{*Z|V}\left[\frac{1-Z}{Z}\cdot V\middle|V\right]\right], \end{aligned} \tag{4.12}$$

which shows (4.10) or (4.9). Combining (4.8) and (4.9) gives the sufficient condition

$$m_2\psi_1(w_1) \geq \frac{E^{v_1,v_2}[v_2F_1(w_1v_1)]}{E^{v_1,v_2}[v_1F_1(w_1v_1)]} = m_2\psi_1^*(w_1), \tag{4.13}$$

which proves the first assertion.

For part (1), we only need to verify that $\delta_2^F(\psi_2)$ dominates δ_0^F. Observe that

$$
\begin{aligned}
&R(\omega, \delta_0^F) - R(\omega, \delta_2^F(\psi_2)) \\
&= E_\omega \left[\int_1^\infty \frac{d}{dt} \left\{ L(m_1\psi_2(tW_2)F, \rho; F) \right\} dt \right] \\
&= E_\omega \left[\int_1^\infty \left\{ \frac{F}{\rho + F} - \frac{1}{m_1\psi_2(tW_2)+1} \right\} m_1\psi_2'(tW_2)W_2 dt \right], \qquad (4.14)
\end{aligned}
$$

which is nonnegative if

$$
E^{v_1,v_2}\left[\left\{ \frac{v_2}{v_1+v_2} - \frac{1}{m_1\psi_2(w_2)+1} \right\} F_2(w_2v_2; \lambda_2) \right] \geq 0.
$$

The remainder of the proof is quite similar to the above arguments, and, therefore we get Theorem 4.1.

Acknowledgments. This work was done while the first author was visiting Department of Statistics, University of Toronto. This research was partially supported by National Sciences and Engineering Research Council of Canada and Mitsubishi Seimeikai in Japan. We are grateful to Mr. David Emig and Mr. Dai Ushijima for assistance with the numerical computations. Also we are grateful to Professor S. Kourouklis, University of Patras, for his valuable comments.

REFERENCES

Bilodeau, M. and Srivastava, M.S. (1992). Estimation of the eigenvalues of $\Sigma_1\Sigma_2^{-1}$. *J. Multivariate Anal.*, **41**, 1-13.

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

Carter, E.M., Khatri, C.G. and Srivastava, M.S. (1979). The effect of inequality of variances on the t-test. *Sankhya, Series B*, **41**, 216-225.

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

Kourouklis, S. (1995). A new property of the inverse Gaussian distribution with applications. Unpublished manuscript, University of Patras, Greece.

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

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

Kubokawa, T., Robert, C. and Saleh, A.K.Md.E. (1992). Empirical Bayes estimation of the covariance matrix of a normal distribution with unknown mean under an entropy loss. *Sankhya, Series A*, **54**, 402-410.

Madi, M. and Tsui, K.-W. (1990). Estimation of the ratio of the scale parameters of two exponential distributions with unknown location parameters. *Ann. Inst. Statist. Math.* **42**, 77-87.

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.

Table 1. MSEs of δ_0, δ_1, δ_2, δ_3, δ^{TR}, δ^*, δ_1^C, δ_2^C and δ_3^C for $m_1 = m_2 = 3$ and $p_1 = p_2 = 3$.

λ_1	λ_2	δ_0	δ_1	δ_2	δ_3	δ^{TR}	δ^*	δ_1^C	δ_2^C	δ_3^C
	0.0	1.075	0.851	1.040	0.740	0.729	0.743	0.823	0.756	0.757
	0.5	1.075	0.851	1.039	0.748	0.737	0.750	0.827	0.761	0.760
0.0	1.0	1.075	0.851	1.039	0.757	0.745	0.757	0.832	0.767	0.764
	5.0	1.075	0.851	1.050	0.807	0.798	0.811	0.864	0.802	0.794
	10.0	1.075	0.851	1.066	0.837	0.833	0.866	0.885	0.824	0.819
	0.0	1.075	0.851	1.040	0.745	0.740	0.737	0.837	0.765	0.774
	0.5	1.075	0.851	1.039	0.753	0.747	0.743	0.841	0.770	0.776
0.5	1.0	1.075	0.851	1.039	0.761	0.754	0.750	0.845	0.776	0.778
	5.0	1.075	0.851	1.050	0.808	0.802	0.799	0.875	0.808	0.802
	10.0	1.075	0.851	1.066	0.837	0.834	0.849	0.895	0.829	0.824
	0.0	1.075	0.855	1.040	0.755	0.754	0.733	0.851	0.777	0.793
	0.5	1.075	0.855	1.039	0.761	0.760	0.739	0.854	0.781	0.793
1.0	1.0	1.075	0.855	1.039	0.769	0.766	0.744	0.858	0.787	0.795
	5.0	1.075	0.855	1.050	0.814	0.809	0.789	0.886	0.818	0.815
	10.0	1.075	0.855	1.066	0.842	0.839	0.836	0.906	0.838	0.834
	0.0	1.075	0.921	1.040	0.849	0.861	0.734	0.938	0.868	0.904
	0.5	1.075	0.921	1.039	0.852	0.862	0.735	0.939	0.871	0.902
5.0	1.0	1.075	0.921	1.039	0.857	0.864	0.737	0.942	0.875	0.901
	5.0	1.075	0.921	1.050	0.887	0.888	0.759	0.961	0.901	0.908
	10.0	1.075	0.921	1.066	0.909	0.909	0.789	0.978	0.918	0.920
	0.0	1.075	0.989	1.040	0.936	0.946	0.760	0.993	0.940	0.974
	0.5	1.075	0.989	1.039	0.937	0.945	0.759	0.993	0.942	0.972
10.0	1.0	1.075	0.989	1.039	0.939	0.946	0.758	0.994	0.945	0.971
	5.0	1.075	0.989	1.050	0.959	0.961	0.767	1.010	0.970	0.978
	10.0	1.075	0.989	1.066	0.979	0.979	0.787	1.025	0.987	0.991

Table 2. MSEs of δ_0, δ_1, δ_2, δ_3, δ^{TR}, δ^*, δ_1^C, δ_2^C and δ_3^C for $m_1 = m_2 = 3$ and $p_1 = p_2 = 10$.

λ_1	λ_2	δ_0	δ_1	δ_2	δ_3	δ^{TR}	δ^*	δ_1^C	δ_2^C	δ_3^C
	0.0	1.075	0.749	1.016	0.593	0.566	0.530	0.673	0.602	0.605
	0.5	1.075	0.749	1.016	0.600	0.572	0.536	0.678	0.608	0.609
0.0	1.0	1.075	0.749	1.016	0.606	0.578	0.541	0.682	0.613	0.612
	5.0	1.075	0.749	1.025	0.651	0.624	0.584	0.714	0.650	0.641
	10.0	1.075	0.749	1.039	0.689	0.669	0.635	0.743	0.682	0.669
	0.0	1.075	0.750	1.016	0.596	0.574	0.530	0.683	0.608	0.616
	0.5	1.075	0.750	1.016	0.603	0.580	0.535	0.687	0.613	0.619
0.5	1.0	1.075	0.750	1.016	0.609	0.585	0.540	0.691	0.618	0.622
	5.0	1.075	0.750	1.025	0.653	0.629	0.581	0.723	0.655	0.649
	10.0	1.075	0.750	1.039	0.691	0.672	0.629	0.751	0.686	0.676
	0.0	1.075	0.751	1.016	0.600	0.582	0.530	0.693	0.614	0.626
	0.5	1.075	0.751	1.016	0.606	0.587	0.535	0.697	0.620	0.629
1.0	1.0	1.075	0.751	1.016	0.612	0.593	0.539	0.701	0.625	0.632
	5.0	1.075	0.751	1.025	0.655	0.634	0.579	0.731	0.660	0.658
	10.0	1.075	0.751	1.039	0.692	0.676	0.625	0.759	0.690	0.683
	0.0	1.075	0.778	1.016	0.645	0.650	0.541	0.765	0.672	0.706
	0.5	1.075	0.778	1.016	0.650	0.654	0.544	0.768	0.676	0.707
5.0	1.0	1.075	0.778	1.016	0.654	0.657	0.547	0.771	0.680	0.709
	5.0	1.075	0.778	1.025	0.691	0.686	0.574	0.795	0.711	0.726
	10.0	1.075	0.778	1.039	0.724	0.717	0.608	0.819	0.737	0.745
	0.0	1.075	0.829	1.016	0.713	0.730	0.566	0.835	0.741	0.786
	0.5	1.075	0.829	1.016	0.717	0.732	0.567	0.837	0.745	0.787
10.0	1.0	1.075	0.829	1.016	0.721	0.733	0.569	0.839	0.749	0.788
	5.0	1.075	0.829	1.025	0.751	0.754	0.586	0.859	0.775	0.800
	10.0	1.075	0.829	1.039	0.779	0.779	0.610	0.879	0.799	0.815