Ann. Inst. Statist. Math.
Vol. 53, No. 3, 436–446 (2001)

DECISION THEORETIC ESTIMATION OF THE RATIO OF VARIANCES IN A BIVARIATE NORMAL DISTRIBUTION

GEORGE ILIOPOULOS

Department of Mathematics, University of Patras 26500 Rio, Patras, Greece

(Received May 6, 1999; revised March 1, 2000)

Abstract. In this paper the problem of estimating the ratio of variances, σ, in a bivariate normal distribution with unknown mean is considered from a decision–theoretic point of view. First, the UMVU estimator of σ is derived, and then it is shown to be inadmissible under two specific loss functions, namely, the squared error loss and the entropy loss. The derivation of the results is done by conditioning on an auxiliary negative binomial random variable.

Key words and phrases: Decision theory, bivariate normal distribution, ratio of variances, squared error loss, entropy loss.

1. Introduction

In many problems we are interested in comparing the variabilities of two populations and this can usually be done by infering from their variance ratio. The typical situation is to assume that the populations are independent. For a decision theoretic approach, see Gelfand and Dey (1988), Kubokawa (1994), Madi (1995), Ghosh and Kundu (1996), Kubokawa and Srivastava (1996), and Iliopoulos and Kourouklis (1999) for the point estimation problem, and Nagata (1989), and Iliopoulos and Kourouklis (2000) for the problem of interval estimation.

However there are cases where the observations are paired and consequently we cannot suppose independence. Data of this kind arise in many cases of practical interest. For example, suppose a pair of observations to be measurements of an individual's ability before and after a certain event (such as training) or values of the same characteristic in twin brothers. Assuming normality (as it is also assumed in this paper), Pitman (1939) and Morgan (1939) derived the likelihood ratio test for the hypothesis that the ratio, σ say, of two variances is equal to a preassigned value. In contrast, the point estimation problem of σ has not been treated yet.

The main results of this paper are contained in Section 3 where the uniformly minimum variance unbiased (UMVU) estimator of σ is derived and it is shown to be inadmissible under two specific loss functions, namely, the squared error loss

$$L_s(t,\sigma) = \left(\frac{t}{\sigma} - 1\right)^2,$$

and the entropy loss

$$L_e(t,\sigma) = \frac{t}{\sigma} - \log\frac{t}{\sigma} - 1.$$

Furthermore, using Stein's (1964) technique for improving estimators of a normal variance, improved testimators are also produced. Note that although, typically, the problem is an extension of that of estimating the ratio of variances of independent populations, in fact it is different in structure. This is justified by the following. First, the UMVU estimator is a non–constant multiple of the ratio of sample variances (see Theorem 3.1). Second, the independence case allows for a unified treatment with respect to the loss based on the monotone likelihood ratio property which is valid for appropriate underlying marginal and conditional distributions. When the two populations are dependent this property does not hold and thus the problem requires separate treatment for each loss considered (see Subsections 3.1 and 3.2).

In Section 2 some distributional results of independent interest are obtained by expanding the 2–dimensional Wishart probability density function (pdf) in a power series.

2. Distributional results

Let $(X_1, Y_1), \ldots, (X_N, Y_N)$, $N \geq 6$, be a random sample from a bivariate normal distribution with mean vector $\mu = (\mu_1, \mu_2)' \in \mathrm{R}^2$ and positive definite covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix},$$

both being unknown. The complete sufficient statistic is the pair (X, A),

$$X = \begin{pmatrix} \bar{X} \\ \bar{Y} \end{pmatrix}, \qquad A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{pmatrix},$$

where $\bar{X} = N^{-1}\sum_{i=1}^{N} X_i$, $\bar{Y} = N^{-1}\sum_{i=1}^{N} Y_i$, $A_{11} = \sum_{i=1}^{N}(X_i - \bar{X})^2$, $A_{22} = \sum_{i=1}^{N}(Y_i - \bar{Y})^2$, $A_{12} = \sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})$. It is well known that X, A are independent following $N_2(\mu, N^{-1}\Sigma)$, $W_2(n, \Sigma)$ distributions respectively, with $n = N - 1$.

The problem of estimation of $\sigma = \sigma_{11}/\sigma_{22}$ remains invariant under the group of transformations $(X, A) \to (CX + b, CAC')$, with $C(2 \times 2)$ diagonal matrix, $b \in \mathrm{R}^2$, and the equivariant estimators have the form $\delta = \phi(R^2)S$, where $S = A_{11}/A_{22}$, $R = A_{12}/(A_{11}A_{22})^{1/2}$, and $\phi(\cdot)$ is a positive function. Let $\rho = \sigma_{12}/(\sigma_{11}\sigma_{22})^{1/2}$ be the population correlation coefficient. Then A_{11}, A_{22}, R have common pdf

$$\begin{aligned} f(a_{11}, a_{22}, r) &= \frac{a_{11}^{n/2-1} a_{22}^{n/2-1} (1 - r^2)^{(n-3)/2}}{2^n \Gamma(1/2)\Gamma((n-1)/2)\Gamma(n/2)\sigma_{11}^{n/2}\sigma_{22}^{n/2}(1-\rho^2)^{n/2}} \\ &\quad \times \exp\left\{ -\frac{a_{11}/\sigma_{11} + a_{22}/\sigma_{22} - 2\rho r (a_{11}/\sigma_{11})^{1/2}(a_{22}/\sigma_{22})^{1/2}}{2(1-\rho^2)} \right\} \\ &= \sum_{\kappa=0}^{\infty} \frac{a_{11}^{(n+\kappa)/2-1} a_{22}^{(n+\kappa)/2-1} (1-r^2)^{(n-3)/2} r^\kappa \rho^\kappa}{2^n \kappa! \Gamma(1/2)\Gamma((n-1)/2)\Gamma(n/2)\sigma_{11}^{(n+\kappa)/2}\sigma_{22}^{(n+\kappa)/2}(1-\rho^2)^{n/2+\kappa}} \\ &\quad \times \exp\left\{ -\frac{a_{11}/\sigma_{11} + a_{22}/\sigma_{22}}{2(1-\rho^2)} \right\}, \qquad a_{11} > 0, \quad a_{22} > 0, \ -1 < r < 1. \end{aligned}$$

The above equality is obtained by expanding $\exp\{\rho r (a_{11}/\sigma_{11})^{1/2}(a_{22}/\sigma_{22})^{1/2}/(1-\rho^2)\}$ in a power series. Set now $\xi = \rho^2$ and $T = R^2$ and observe that the terms of the sum cancel out when κ is odd. Then, A_{11}, A_{22}, T have density

$$f(a_{11}, a_{22}, t) = \sum_{\kappa=0}^{\infty} \pi(\kappa;\xi) b_\kappa(t) \frac{1}{\sigma_{11}(1-\xi)} g_{n+2\kappa}\left(\frac{a_{11}}{\sigma_{11}(1-\xi)}\right) \cdot \frac{1}{\sigma_{22}(1-\xi)} g_{n+2\kappa}\left(\frac{a_{22}}{\sigma_{22}(1-\xi)}\right), \tag{2.1}$$

where $b_\kappa(\cdot)$, $g_{n+2\kappa}(\cdot)$ are the Beta$(\kappa+1/2,(n-1)/2)$, $\chi^2_{n+2\kappa}$ pdfs respectively and

$$\pi(\kappa;\xi) = \frac{\Gamma(n/2+\kappa)}{\Gamma(n/2)}(1-\xi)^{n/2}\frac{\xi^\kappa}{\kappa!}, \qquad \kappa = 0, 1, 2, \ldots,$$

is the negative binomial probability mass function. Denote by K a random variable with probability mass $\pi(\kappa;\xi)$. Then, from (2.1), it can be seen that conditionally on $K = \kappa$, A_{11}, A_{22}, T are mutually independent with distributions

$$A_{11} \sim \sigma_{11}(1-\xi)\chi^2_{n+2\kappa}, \qquad A_{22} \sim \sigma_{22}(1-\xi)\chi^2_{n+2\kappa}, \quad T \sim \text{Beta}\left(\kappa+\frac{1}{2}, \frac{n-1}{2}\right).$$

The distribution of $S = A_{11}/A_{22}$ is given by the following theorem whose proof is straightforward.

THEOREM 2.1. (i) *Conditionally on* $K = \kappa$, S *and* T *are independent and* $S/\sigma \sim F_{n+2\kappa, n+2\kappa}$.

(ii) *The marginal pdf of* S/σ *is given by*

$$f(s) = \sum_{\kappa=0}^{\infty} \pi(\kappa;\xi) f_{n+2\kappa}(s),$$

where $f_{n+2\kappa}(\cdot)$ *denotes the pdf of* $F_{n+2\kappa,n+2\kappa}$.

Note that this particular form of the pdf of S/σ can also be derived from a result of Finney (1938).

Let $E^{(m)}K^j$ denote the j-th—moment of K when the degrees of freedom of the Wishart distribution is m. It is quite easy to verify that the following recursive formula holds.

$$\begin{aligned} E^{(n)}K &= \frac{n\xi}{2(1-\xi)}, \\ E^{(n)}K^j &= \frac{\xi}{2}E^{(n)}[(n+2K)(K+1)^{j-1}] \\ &= E^{(n)}K + \frac{\xi}{2(1-\xi)}\sum_{i=1}^{j-1}\left[n\binom{j-1}{i} + 2\binom{j-1}{i-1}\right]E^{(n)}K^i, \qquad j \ge 2. \end{aligned}$$

The above expressions can be used to get the moments of S/σ in a convenient way, as the following theorem demonstrates.

THEOREM 2.2. *For positive integer* $\nu < n/2$ *the* ν*-th moment of* S/σ *is given by*

$$E(S/\sigma)^\nu = \frac{(1-\xi)^\nu}{\prod_{i=1}^{\nu}(n-2i)} E^{(n-2\nu)}\left\{\prod_{i=1}^{\nu}[n+2K+2(i-1)]\right\}.$$

PROOF. Using the fact that $EF_{p,p}^{\nu} = \prod_{i=1}^{\nu} \frac{p+2(i-1)}{p-2i}$ we get

$$\begin{aligned} E(S/\sigma)^{\nu} &= E^{(n)}E\left[(S/\sigma)^{\nu} \mid K\right] = E^{(n)}[E(F_{n+2K,n+2K}^{\nu} \mid K)] \\ &= \sum_{\kappa=0}^{\infty}\prod_{i=1}^{\nu} \frac{n+2\kappa+2(i-1)}{n+2\kappa-2i}\frac{\Gamma(n/2+\kappa)}{\Gamma(n/2)}(1-\xi)^{n/2}\frac{\xi^{\kappa}}{\kappa!} \\ &= \frac{(1-\xi)^{\nu}}{\prod_{i=1}^{\nu}(n-2i)}\sum_{\kappa=0}^{\infty}\prod_{i=1}^{\nu}[n+2\kappa+2(i-1)]\frac{\Gamma((n-2\nu)/2+\kappa)}{\Gamma((n-2\nu)/2)}(1-\xi)^{(n-2\nu)/2}\frac{\xi^{\kappa}}{\kappa!} \\ &= \frac{(1-\xi)^{\nu}}{\prod_{i=1}^{\nu}(n-2i)}E^{(n-2\nu)}\prod_{i=1}^{\nu}[n+2K+2(i-1)], \qquad n>2\nu. \end{aligned}$$

This completes the proof.

Using Theorem 2.2 we obtain the first two moments of S as

$$ES = \frac{n-2\xi}{n-2}\sigma \quad \text{and} \tag{2.2}$$
$$ES^2 = \frac{n(n+2)-8(n+2)\xi+24\xi^2}{(n-2)(n-4)}\sigma^2.$$

3. Estimation of the ratio of variances

When $\xi = 0$, i.e. the X_i's and Y_i's are independent, the UMVU estimator of σ is known to be $\delta_{U,0} = \frac{n-2}{n}S$. However in our general case, $\delta_{U,0}$ is no more unbiased as it is obvious from (2.2). Derivation of an unbiased estimator of σ which is a function of the complete sufficient statistic can easily be done using the negative binomial random variable K defined above.

THEOREM 3.1. *The UMVU estimator of σ is given by*

$$\delta_U = \frac{n-3+2T}{n-1}S$$

with variance

$$\frac{4[(n^3-5n^2+3n+13)-(n^3+2n^2-45n+102)\xi+(7n^2-48n+89)\xi^2]}{(n-2)(n-4)(n-1)^2}\sigma^2.$$

PROOF. Recall that conditionally on $K=\kappa$, S and T are independent. Then we have

$$\begin{aligned} E\delta_U = E[E\delta_U \mid K] &= E\left\{\frac{n-3+2E[T \mid K]}{n-1}E[S \mid K]\right\} \\ &= E\left\{\frac{(n-3)(n+2K)+2(1+2K)}{(n-1)(n+2K-2)}\sigma\right\} = \sigma. \end{aligned}$$

Since δ_U is a function of the complete sufficient statistic it is the unique UMVU estimator of σ. The computation of its variance can easily be done using the expressions for the moments of K.

The maximum likelihood estimator (MLE) of σ is the same as in the case of independence, i.e. $\delta_{MLE} = S$, being the ratio of MLEs of σ_{11} and σ_{22}.

3.1 *Estimation under squared error loss*
Consider the class of estimators

$$\mathcal{C} = \{\delta_c = [T + c(1-T)]S; c \geq 0\}$$

and observe that δ_U and δ_{MLE} are members of $\mathcal{C}$ with $c = (n-3)/(n-1)$ and $c = 1$ respectively. The risk of an estimator of the above form under squared error loss is

$$E(\delta_c/\sigma - 1)^2 = c^2 E(1-T)^2 S^2/\sigma^2 - 2cE[(1-T)S/\sigma - (1-T)TS^2/\sigma^2] + E(TS/\sigma - 1)^2 ,$$

which is quadratic in c and uniquely minimized at

$$c_0 = \frac{E(1-T)S/\sigma - E(1-T)TS^2/\sigma^2}{E(1-T)^2 S^2/\sigma^2} = \frac{n-5}{n+1}.$$

The last equality can be seen to hold by substituting $E(1-T)S/\sigma$, $E(1-T)TS^2/\sigma^2$, $E(1-T)^2S^2$ by $E(n-1)/(n+2K-2)$, $E[(n-1)(1+2K)]/[(n+2K-2)(n+2K-4)]$, $E[(n-1)(n+1)]/[(n+2K-2)(n+2K-4)]$ respectively.

The following theorem establishes the inadmissibility of δ_U and δ_{MLE} and it is a consequence of the above argument.

THEOREM 3.2. *Under squared error loss δ_U and δ_{MLE} are inadmissible both being dominated by the estimator*

$$\delta_0 = \frac{n-5+6T}{n+1} S. \tag{3.1}$$

However δ_0 is inadmissible too as can be seen from the following theorem.

THEOREM 3.3. *Assume that the loss function is squared error. If $\delta_\phi = \phi(T)S$ is an estimator of σ satisfying*

$$P[\phi(T) < (n-4)/(n+2)] > 0, \tag{3.2}$$

then it is inadmissible being dominated by the estimator

$$\delta_\phi^* = \max\left\{\phi(T), \frac{n-4}{n+2}\right\} S.$$

PROOF. The method of proof is based on the ideas of Stein (1964) and Brewster and Zidek (1974) for the estimation of a normal variance. The conditional expected loss of an estimator of the form $\phi(T)S$ given $T = t$, $K = \kappa$ is quadratic in $\phi(t)$ and uniquely minimized at $\phi_\kappa(t) = E(S \mid T = t, K = \kappa)/E(S^2 \mid T = t, K = \kappa)$. Now this value does not depend on t because S, T are independent conditionally on $K = \kappa$. Hence, $\phi_\kappa(t) = \phi_\kappa = E(S \mid K = \kappa)/E(S^2 \mid K = \kappa) = (n + 2\kappa - 4)/(n + 2\kappa + 2)$. One can easily check that $\phi_\kappa > \phi_0 = (n-4)/(n+2)$, for $\kappa > 0$. Thus, by the convexity of the conditional risk, $\phi(t) < \phi_0$ implies

$$E[(\phi(t)S/\sigma - 1)^2 \mid T = t, K = \kappa] > E[(\phi_0 S/\sigma - 1)^2 \mid T = t, K = \kappa], \qquad \kappa = 0, 1, 2, \ldots.$$

Taking expectations over T, K we obtain the desired result.

Remark 3.1. Note that $\delta^* = \phi_0 S$ is the best estimator of σ of the form $\phi(T)S$, when ξ is known to be zero, i.e. X_i, Y_i are independent. Thus, δ^*_ϕ is a testimator which chooses between δ^* and δ_ϕ depending on whether or not the test for $H_0 : \xi = 0$ with critical region $\phi(T) > \phi_0$ accepts H_0.

The condition (3.2) is satisfied by the estimator δ_0 and this proves its inadmissibility. Hence, we have the following corollary.

COROLLARY 3.1. *The estimator δ_0 in (3.1) is inadmissible being dominated by the estimator*

$$\delta_0^* = \max\left\{\frac{n-5+6T}{n+1}, \frac{n-4}{n+2}\right\} S. \tag{3.3}$$

In contrast, the condition (3.2) is not satisfied by δ_U and δ_{MLE}.

For further improving on δ_0^* consider estimators of the form

$$\delta = \psi(W_2, T)S, \tag{3.4}$$

where $W_2 = N\bar{Y}^2/A_{22}$ and $\psi(\cdot,\cdot)$ is a positive function. It is well known that conditionally on $\bar{X} = \bar{x}$, $N^{1/2}\bar{Y} \sim N(N^{1/2}\mu_2^*(\bar{x}), \sigma_{22}(1-\xi))$, $\mu_2^*(\bar{x}) = \mu_2 + (\bar{x} - \mu_1)\sigma_{12}/\sigma_{11}$, and thus, by conditioning in addition on $L = \ell$, a Poisson random variable with mean $N\mu_2^*(\bar{x})^2/(2\sigma_{22}(1-\xi))$, $N\bar{Y}^2$ is distributed as $\sigma_{22}(1-\xi)\chi^2_{1+2\ell}$. Hence, by conditioning on $\bar{X} = \bar{x}$, $L = \ell$, $K = \kappa$, and recalling that $A_{22} \mid K = \kappa \sim \sigma_{22}(1-\xi)\chi^2_{n+2\kappa}$, it follows that W_2 is ancillary.

Now it can be shown that the conditional risk of an estimator of the form (3.4) given $\bar{X}$, W_2, T, L, K, is uniquely minimized at

$$\psi_{\kappa,\ell}(w_2) = \frac{E[S \mid \bar{X} = \bar{x}, W_2 = w_2, L = \ell, K = \kappa]}{E[S^2 \mid \bar{X} = \bar{x}, W_2 = w_2, L = \ell, K = \kappa]} = \frac{n + 2\kappa + 2\ell - 3}{(n + 2\kappa + 2)(1 + w_2)}$$

which attains its minimum with respect to κ, ℓ at $\kappa = \ell = 0$. Using now an argument similar to that in the proof of Theorem 3.3 we obtain the following result.

THEOREM 3.4. *Assume that the loss function is squared error. If $\delta_\psi = \psi(W_2, T)S$ is an estimator of σ satisfying*

$$P\left[\psi(W_2, T) < \frac{n-3}{(n+2)(1+W_2)}\right] > 0,$$

then it is inadmissible being dominated by the estimator

$$\delta_\psi^{**} = \max\left\{\psi(W_2, T), \frac{n-3}{(n+2)(1+W_2)}\right\} S.$$

COROLLARY 3.2. *The estimator*

$$\delta_0^{**} = \max\left\{\frac{n-5+6T}{n+1}, \frac{n-4}{n+2}, \frac{n-3}{(n+2)(1+W_2)}\right\} S$$

dominates δ_0^* in (3.3).

3.2 *Estimation under entropy loss*

Consider now estimation of σ under entropy loss. This loss function has been used by many researchers for the estimation of scale parameters, as is σ for the distribution of S (see Theorem 2.1). Furthermore one can argue that this loss is more reasonable for this kind of problem than squared error, since underestimation as well as overestimation is heavily penalized (it is $\lim_{t\to 0} L_e(t) = \lim_{t\to\infty} L_e(t) = \infty$). On the other hand, squared error loss does not penalize underestimation of the scale parameter so much as $\lim_{t\to 0} L_s(t) = 1$, while $\lim_{t\to\infty} L_s(t) = \infty$.

Let $\delta_c \in \mathcal{C}$, $c > 0$, an estimator of σ. The conditional risk of such an estimator given $K = \kappa$ is

$$E[(T + c(1-T))S/\sigma \mid K = \kappa] - E[\log(T + c(1-T)) \mid K = \kappa] - E[\log(S/\sigma) \mid K = \kappa] - 1,$$

and, by differentiation with respect to c, is uniquely minimized at $c = c_\kappa$ satisfying

$$\begin{aligned}(3.5)\qquad 0 &= E[(1-T)S/\sigma \mid K = \kappa] - E\left[\frac{1-T}{c(1-T)+T}\,\middle|\, K = \kappa\right] \\ &= \frac{n-1}{n+2\kappa-2} - \int_0^1 \frac{\Gamma(n/2+\kappa)}{\Gamma(1/2+\kappa)\Gamma((n-1)/2)} \frac{t^{\kappa-1/2}(1-t)^{(n-1)/2}}{c(1-t)+t}\,dt.\end{aligned}$$

Now, for $c \in (0,1)$ replace $[c(1-t)+t]^{-1}$ by its power series representation about 1, $\sum_{\lambda=0}^{\infty}(1-t)^\lambda(1-c)^\lambda$. Then (3.5) becomes

$$\begin{aligned}(3.6)\quad 0 &= \int_0^1 \frac{\Gamma(n/2+\kappa)t^{\kappa-1/2}(1-t)^{(n-1)/2}}{\Gamma(1/2+\kappa)\Gamma((n-1)/2)} \sum_{\lambda=0}^{\infty}(1-t)^\lambda(1-c)^\lambda dt - \frac{n-1}{n+2\kappa-2} \\ &= \sum_{\lambda=0}^{\infty} \frac{\Gamma(n/2+\kappa)(1-c)^\lambda}{\Gamma(1/2+\kappa)\Gamma((n-1)/2)} \int_0^1 t^{\kappa-1/2}(1-t)^{\lambda+(n-1)/2}dt - \frac{n-1}{n+2\kappa-2} \\ &= \sum_{\lambda=0}^{\infty} \frac{\Gamma((n+1)/2+\lambda)\Gamma(n/2+\kappa+1)}{\Gamma((n+1)/2)\Gamma(n/2+\kappa+1+\lambda)}(1-c)^\lambda - \frac{n+2\kappa}{n+2\kappa-2} \\ &= \sum_{\lambda=1}^{\infty} \frac{\Gamma((n+1)/2+\lambda)\Gamma(n/2+\kappa+1)}{\Gamma((n+1)/2)\Gamma(n/2+\kappa+1+\lambda)}(1-c)^\lambda - \frac{2}{n+2\kappa-2}.\end{aligned}$$

Notice that the minimizing value c_κ must be in the interval $(0,1)$ because the right hand side (rhs) of (3.6) is a strictly decreasing function of c which is positive near 0 and negative at $c = 1$.

We will prove now that $\min_\kappa c_\kappa = c_0 > (n-3)/(n-1)$. For this we need first an upper bound for c_κ. By (3.6) we obtain after some simplifications

$$\begin{aligned}(3.7)\qquad 0 &= \left\{\frac{n+1}{n+2\kappa+2}(1-c_\kappa) + \frac{(n+1)(n+3)}{(n+2\kappa+2)(n+2\kappa+4)}(1-c_\kappa)^2 + \cdots\right\} \\ &\quad - \frac{2}{n+2\kappa-2} \le \frac{n+1}{n+2\kappa+2}\sum_{\lambda=1}^{\infty}(1-c_\kappa)^\lambda - \frac{2}{n+2\kappa-2} \\ &= \frac{n+1}{n+2\kappa+2}\,\frac{1-c_\kappa}{c_\kappa} - \frac{2}{n+2\kappa-2}\end{aligned}$$

hence,

$$c_\kappa \le \frac{(n+1)(n+2\kappa-2)}{(n+1)(n+2\kappa-2)+2(n+2\kappa+2)} = \bar{c}_\kappa,$$

say. Since the rhs of (3.7) is strictly decreasing function of c for every κ, if it is positive at a value c then it must hold $c < c_\kappa$.

Now, substituting $\kappa = 0$ and $c = (n-3)/(n-1)$ in the rhs of (3.7) we get

$$\left\{\frac{n+1}{n+2}\frac{2}{n-1} + \frac{(n+1)(n+3)}{(n+2)(n+4)}\left(\frac{2}{n-1}\right)^2 + \cdots\right\} - \frac{2}{n-2}$$
$$> \frac{n+1}{n+2}\frac{2}{n-1} + \frac{(n+1)(n+3)}{(n+2)(n+4)}\left(\frac{2}{n-1}\right)^2 + \frac{(n+1)(n+3)(n+5)}{(n+2)(n+4)(n+6)}\left(\frac{2}{n-1}\right)^3$$
$$-\frac{2}{n-2} = \frac{4(n^4+8n^3+5n^2-86n-24)}{(n-1)^3(n-2)(n+2)(n+4)(n+6)} > 0$$

(recall that is assumed $n \ge 5$). Thus, $c_0 > (n-3)/(n-1)$ holds. In a similar way, taking $\kappa > 0$ and $c = \bar{c}_0 = (n^2-n-2)/(n^2+n+2)$ we get that

$$\left\{\frac{n+1}{n+2\kappa+2}(1-\bar{c}_0) + \frac{(n+1)(n+3)}{(n+2\kappa+2)(n+2\kappa+4)}(1-\bar{c}_0)^2 + \cdots\right\} - \frac{2}{n+2\kappa-2}$$

is greater than or equal to a fraction with positive denominator and numerator

$$\begin{aligned}&(8\kappa-4)n^7 + (32\kappa^2+112\kappa-72)n^6 + (32\kappa^3+272\kappa^2+696\kappa-548)n^5\\ &\quad+(64\kappa^3+736\kappa^2+2496\kappa-2512)n^4 + (160\kappa^3+1488\kappa^2+5936\kappa-7168)n^3\\ &\quad+(128\kappa^3+1504\kappa^2+8224\kappa-12224)n^2 + (128\kappa^3+1216\kappa^2+6336\kappa-11456)n\\ &\quad+(128\kappa^2+1408\kappa-4806).\end{aligned}$$

It is easy to verify that for $n \ge 5$ and $\kappa \ge 1$ the above expression is positive, implying $\bar{c}_0 < c_\kappa$ and hence, $c_0 < c_\kappa$. Now, observe that for every κ the conditional risk of δ_c is a convex function of c, and recall that $\delta_U \in C$ with $c = (n-3)/(n-1)$. Summarizing the above results we obtain the following theorem.

THEOREM 3.5. *Under entropy loss δ_U is inadmissible being dominated by the estimator*

(3.8) $$\delta_0 = [T + c_0(1-T)]S,$$

where $c = c_0$ is the solution to the equation

$$E\left[\frac{1-T}{c(1-T)+T}\middle| K=0\right] = \frac{n-1}{n-2}.$$

The proof of the following theorem is analogous to that of Theorem 3.3 and therefore is ommited.

THEOREM 3.6. *Assume that the loss function is the entropy loss and $\delta_\phi = \phi(T)S$ is an estimator of σ satisfying*

(3.9) $$P[\phi(T) < (n-2)/n] > 0.$$

Then the estimator δ_ϕ is inadmissible being dominated by

$$\delta_\phi^* = \max\left\{\phi(T), \frac{n-2}{n}\right\} S.$$

Remark 3.2. Analogous comments as those in Remark 3.1 hold in the present case too. Thus, δ_ϕ^* is a testimator which chooses between $\delta_{U,0} = \frac{n-2}{n}S$, which is the best estimator of σ of the form $\phi(T)S$ when it is known that $\xi = 0$, and δ_ϕ depending on whether or not the test for $H_0 : \xi = 0$ with critical region $\phi(T) > (n-2)/n$ accepts H_0.

The condition (3.9) holds for both δ_0 (since it holds for $\delta_{\bar{c}_0}$) and δ_U. Hence, we have the next result which is the analogous to that of Corollary 3.1.

COROLLARY 3.3. (i) *The estimator δ_0 in* (3.8) *is inadmissible being dominated by the estimator*

$$\delta_0^* = \max\left\{T + c_0(1-T), \frac{n-2}{n}\right\} S. \tag{3.10}$$

(ii) *The estimator*

$$\delta_U^* = \max\left\{\frac{n-3+2T}{n-1}, \frac{n-2}{n}\right\} S \tag{3.11}$$

dominates the UMVU estimator δ_U.

The estimators δ_0^*, δ_U^* can be further improved by an estimator of the form (3.4). The proof of the next theorem is analogous to that of Theorem 3.4 and therefore it is ommited.

THEOREM 3.7. *Assume that the loss function is the entropy loss and $\delta_\psi = \psi(W_2, T)S$ is an estimator of σ satisfying*

$$P\left[\psi(W_2, T) < \frac{n-1}{n(1+W_2)}\right] > 0.$$

Then the estimator δ_ψ is inadmissible being dominated by

$$\delta_\psi^{**} = \max\left\{\psi(W_2, T), \frac{n-1}{n(1+W_2)}\right\} S.$$

COROLLARY 3.4. (i) *The estimator*

$$\delta_0^{**} = \max\left\{T + c_0(1-T), \frac{n-2}{n}, \frac{n-1}{n(1+W_2)}\right\} S$$

dominates δ_0^* in (3.10).

(ii) *The estimator*

$$\delta_U^{**} = \max\left\{\frac{n-3+2T}{n-1}, \frac{n-2}{n}, \frac{n-1}{n(1+W_2)}\right\} S$$

dominates δ_U^* in (3.11).

Table 1. (Squared error loss) Percentage risk improvement of δ_0^* in (3.3) over δ_{MLE}, δ_U, δ_0 in (3.1) under squared error loss for $n = 5$ and $n = 10$.

	$n=5$			$n=10$		
ξ	δ_{MLE}	δ_U	δ_0	δ_{MLE}	δ_U	δ_0
0	89.14	69.59	23.98	61.54	36.10	0.46
.1	87.82	66.82	22.03	58.30	33.04	0.39
.2	86.18	63.57	19.86	54.68	29.86	0.30
.3	84.13	59.75	17.48	50.64	26.56	0.21
.4	81.50	55.22	14.92	46.10	23.15	0.13
.5	78.05	49.83	12.21	40.95	19.62	0.08
.6	73.35	43.37	9.40	35.09	15.97	0.04
.7	66.63	35.56	6.56	28.33	12.20	0.01
.8	56.29	26.03	3.84	20.46	8.28	0.00
.9	38.41	14.32	1.46	11.16	4.22	0.00

Table 2. (Entropy loss) Percentage risk improvement of δ_0^* in (3.10) over δ_U, δ_0 in (3.8) and of δ_U^* in (3.11) over δ_U under entropy loss for $n = 5$ and $n = 10$.

	$n=5$			$n=10$		
Estimator	δ_0^*		δ_U^*	δ_0^*		δ_U^*
ξ	δ_U	δ_0	δ_U	δ_U	δ_0	δ_U
0	0.62	0.26	1.02	0.04	0.04	0.07
.1	0.66	0.25	0.98	0.05	0.03	0.06
.2	0.69	0.23	0.91	0.04	0.03	0.05
.3	0.70	0.21	0.83	0.04	0.02	0.04
.4	0.70	0.18	0.73	0.04	0.01	0.03
.5	0.68	0.15	0.61	0.03	0.01	0.01
.6	0.65	0.12	0.48	0.03	0.00	0.01
.7	0.59	0.08	0.34	0.02	0.00	0.00
.8	0.49	0.05	0.20	0.02	0.00	0.00
.9	0.33	0.02	0.08	0.01	0.00	0.00

3.3 *Numerical results and some final remarks*

The percentage risk improvements of the estimators δ_0^* in (3.3), under squared error loss, and δ_0^* in (3.10), δ_U^* in (3.11), under entropy loss, over the standard ones have been calculated using Mathematica v.3.0 for $n = 5$, 10 and $\xi = 0(.1).9$ by taking without loss of generality $\sigma = 1$. The numerical study shows that the risk reduces substantially when squared error loss is used (see Table 1). On the contrary the reduction is very small under entropy loss (see Table 2). Observe that in the latter case δ_U^* improves on δ_U more than δ_0^* when ξ is small whereas δ_0^* behaves better for large ξ's. However δ_0^* and δ_U^* have similar risks and this suggests the use of δ_U^* rather than δ_0^* because of its easier calculation, especially when small correlation is suspected.

The improved estimators in Corollaries 3.1, 3.2, 3.3, and 3.4 have the ability of pre—testing whether the nuisance parameters ξ and μ_2 equal zero. In particular, in the case of entropy loss the improved estimator δ_U^* is the maximum between the UMVU estimator when it is known that $\xi = 0$ and the UMVU estimator in the general case. All these estimators are non—smooth as they are derived by Stein's (1964) technique. It would be nice to construct smooth improved estimators for σ using, for instance, Brewster and Zidek's (1974) idea as Madi (1995) did in the case of independence under an arbitrary strictly bowl—shaped loss function. Nevertheless, technical difficulties due to the particular structure of the problem seem to make this pursuit hard to accomplish. Notice here that the use of Brewster and Zidek's (1974) technique (by conditioning on $T \leq t$ instead of $T = t$) leads to the elimination of T from the estimation procedure. On the other hand, an improved estimator must include T, since when $\xi = 1$ it holds $P(T = 1) = P(S = \sigma) = 1$ and thus any estimator of the form $\phi(T)S$ with $\phi(1) = 1$ (like UMVU estimator and the improved estimators presented in this paper) has zero risk.

Acknowledgements

This research was motivated by a discussion with Professor T. Cacoullos. The author wishes also to thank Professor S. Kourouklis for his helpful comments and suggestions.

References

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

Finney, D. J. (1938). The distribution of the ratio of estimates of the two variances in a sample from a normal bivariate population, *Biometrika*, **30**, 190–192.

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

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

Iliopoulos, G. and Kourouklis, S. (1999). Improving on the best affine equivariant estimator of the ratio of generalized variances, *J. Multivariate Anal.*, **68**, 176–192.

Iliopoulos, G. and Kourouklis, S. (2000). Interval estimation for the ratio of scale parameters and for ordered scale parameters, *Statist. Decisions*, **18**, 169–184.

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

Kubokawa, T. and Srivastava, M. S. (1996). Double shrinkage estimators of ratio of variances, *Multidimensional Statistical Analysis and Theory of Random Matrices* (eds. A. K. Gupta and V. L. Girko), 139–154, VSP, Netherlands.

Madi, T. M. (1995). On the invariant estimation of a normal variance ratio, *J. Statist. Plann. Inference*, **44**, 349–357.

Morgan, W. A. (1939). A test for the significance of the difference between the two variances in a sample from a normal bivariate population, *Biometrika*, **31**, 13–19.

Nagata, Y. (1989). Improvements of interval estimations for the variance and the ratio of two variances, *J. Japan Statist. Soc.*, **19**, 151–161.

Pitman, E. J. G. (1939). A note on normal correlation, *Biometrika*, **31**, 9–12.

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.