Ann. Inst. Statist. Math.
Vol. 43, No. 2, 369–375 (1991)

STEIN-TYPE IMPROVEMENTS OF CONFIDENCE INTERVALS FOR THE GENERALIZED VARIANCE

SANAT K. SARKAR

Department of Statistics, Temple University, Philadelphia, PA 19122, U.S.A.

(Received October 24, 1989; revised March 19, 1990)

Abstract. Based on independent random matices X: $p \times m$ and S: $p \times p$ distributed, respectively, as $N_{pm}(\mu,\ \Sigma \otimes I_m)$ and $W_p(n,\ \Sigma)$ with μ unknown and $n \geq p$, the problem of obtaining confidence interval for $|\Sigma|$ is considered. Stein's idea of improving the best affine equivariant point estimator of $|\Sigma|$ has been adapted to the interval estimation problem. It is shown that an interval estimator of the form $|S|(b^{-1},\ a^{-1})$ can be improved by $\min\{|S|,\ c|S + XX'|\}(b^{-1},\ a^{-1})$ for a certain constant c depending on $(a,\ b)$.

Key words and phrases: Generalized variance, invariant interval estimators, Stein-type improvements.

1. Introduction

Let the random matrices X: $p \times m$ and S: $p \times p$ be independently distributed as $N_{pm}(\mu,\ \Sigma \otimes I_m)$ and $W_p(n,\ \Sigma)$ $(n \geq p)$, respectively. This is the canonical form of the multivariate linear model. Assuming that Σ is positive definite, we consider the problem of obtaining confidence interval for $|\Sigma|$ based on $(X,\ S)$.

Observing that the problem is invariant under the affine group of transformations, one can check that affine invariant interval estimators with confidence coefficient $\beta \in (0,\ 1)$ are of the form

$$J = |S|(b^{-1},\ a^{-1}) \tag{1.1}$$

with $0 < a < b < \infty$ satisfying

$$\int_a^b f_{p,\,n}(x)\,dx = \beta, \tag{1.2}$$

where $f_{p,\,n}$ is the density of $|S|$ when $\Sigma = I$. Further restrictions like

$$f_{p,\,n+h}(a) = f_{p,\,n+h}(b), \tag{1.3}$$

for some particular choices of h, will provide certain types of optimum invariant confidence intervals. For example, the shortest length invariant confidence interval with confidence coefficient β is given by (1.1)–(1.3) with $h = 4$ (Sarkar (1989)); and, as it will be shown in this paper, the unbiased invariant confidence interval corresponds to $h = 2$. These optimum intervals provide multivariate extensions of the corresponding optimum intervals known in the literature for normal variance. The confidence interval (1.1) with $a < b$ giving equal probabilities (i.e., $\beta/2$), although does not have any other optimum property, is quite often used, at least in the univariate case. Also, instead of reducing the problem through invariance, if the likelihood ratio principle is used to construct an interval based on (X, S), one would get (1.1), with $a < b$ satisfying (1.2) and another condition slightly different from (1.3).

For the problem of point estimation of $|\Sigma|$, Stein (1964) first showed that, when $p = 1$, the best affine equivariant estimator, which is based on $|S|$, is inadmissible with respect to the class of estimators based on the minimal sufficient statistic (X, S). Shorrock and Zidek (1976) and Sinha (1976) have extended this result to the multivariate case. In view of these results, it seems likely that an interval estimator based on $|S|$ would also be inadmissible in the class of interval estimators based on (X, S). Considering $p = 1$, Cohen (1972) first demonstrated this inadmissibility result for the shortest length confidence interval by adapting Brown's idea (1968). An extension of this result for general p was given by Sarkar (1989). The idea of incorporating X into the estimation procedure used in this improvement is, however, only restricted to the shortest length confidence interval. Moreover, the improvement is achieved only in terms of the coverage probability. A stronger result is obtained in this paper. We adapt Stein's idea of improving the best equivariant estimator of $|\Sigma|$ to the interval estimation problem. The Stein-type improvement works for any interval of the form (1.1), where $0 < a < b < \infty$ are subject to a condition that is less restricted than (1.3). Also, the improvement is obtained in terms of both the coverage probability and the length of interval. Our result is a multivariate extension of the recent paper by Nagata (1989).

In Section 3, we prove that (1.3) with $h = 2$ is the condition of unbiasedness of an invariant confidence interval (1.1). Stein-type improvement of an interval of the form (1.1) is presented in Section 4. Section 2 contains the distributional properties of sample generalized variance that are useful in Sections 3 and 4.

2. Some necessary results

Suppose that $\Omega = \Sigma^{-1}\mu\mu'$ is of rank t ($\leq s = \min(p, m)$). Let $C_\kappa(\Omega)$ denote the zonal polynomial of Ω corresponding to the partition $\kappa = (\kappa_1, \ldots, \kappa_t)$ of the integer $k \geq 0$ in terms of the integers $\kappa_1 \geq \cdots \geq \kappa_t \geq 0$. This is a symmetric homogeneous polynomial in the characteristic roots of Ω and is known to be nonnegative because Ω is nonnegative definite. It is also known that

$$(2.1) \qquad (\mathrm{tr}\Omega)^k = \sum_{\kappa:\|\kappa\|=k} C_\kappa(\Omega),$$

where $\|\kappa\| = \sum_{i=1}^{t} \kappa_i$. See Muirhead (1982) for the aforementioned and other properties of zonal polynomial. We use the following notations: $S^* = S + XX'$ and

$F = \mathrm{Diag}(f_1, \ldots, f_s)$, where $1 \geq f_1 \geq \cdots \geq f_s \geq 0$ are the ordered characteristic roots of $S^{*-1}XX'$. Also, let χ_n^2 denote the chi-squared random variable with n degrees of freedom. We now have the following result which provides the first step in the derivation of Stein-type improvement.

LEMMA 2.1. *Conditionally given* $\kappa = (\kappa_1, \ldots, \kappa_t)$,

$$|S^*|/|\Sigma| \sim \prod_{i=1}^{p} \chi^2_{n+m+2\kappa_i-i+1}, \tag{2.2}$$

independently of F, *where* $\chi^2_{n+m+2\kappa_i-i+1}$ *are independent, and* κ *has the following probability distribution:*

$$\pi(\kappa, \Omega) = \begin{cases} \exp\left(-\frac{1}{2}\mathrm{tr}\Omega\right) C_\kappa\left(\frac{1}{2}\Omega\right) / \|\kappa\|! & if\ \kappa_1 \geq \cdots \geq \kappa_t \geq 0 \\ 0 & otherwise. \end{cases} \tag{2.3}$$

PROOF. Without any loss of generality we can assume that $\Sigma = I_p$ and replace μ by $\nu = \Sigma^{-1/2}\mu$. Then, as Shorrock and Zidek ((1976), p. 635) proved, conditionally given κ, with the marginal distribution of κ being given by (2.3), S^* and F are independent. The conditional density, $\phi_\kappa(S^*)$, of S^* is given by

$$\phi_\kappa(S^*) = \phi_0(S^*)C_\kappa(S^*)/E_0 C_\kappa(S^*), \tag{2.4}$$

where $\phi_0(S^*)$ is the density of S^* under $\kappa = 0$ (i.e., $\Omega = 0$), which is that of $W_p(n+m,\ I_p)$, and E_0 denotes the expectation with respect to this central distribution. The conditional moments of $|S^*|$ obtained from (2.4) are given by

$$E_\kappa(|S^*|^l) = \frac{2^{pl}\Gamma_p\left(\frac{n+m+2l}{2}\right)}{\Gamma_p\left(\frac{n+m}{2}\right)} \frac{EC_\kappa(A)}{EC_\kappa(B)}, \tag{2.5}$$

where $A \sim W_p(n+m+2l,\ I)$ and $B \sim W_p(n+m,\ I)$. The following formula (Muirhead (1982), p. 251)

$$\begin{aligned} E_\kappa C_\kappa(B) &= 2^k \left(\frac{n+m}{2}\right)_\kappa C_\kappa(I) \\ &= \frac{2^k \prod_{i=1}^p \Gamma\left(\frac{n+m+2\kappa_i-i+1}{2}\right)}{\prod_{i=1}^p \Gamma\left(\frac{n+m-i+1}{2}\right)} C_\kappa(I) \end{aligned} \tag{2.6}$$

simplifies (2.5) to

$$\frac{2^{pl}\prod_{i=1}^p \Gamma\left(\frac{n+m+2l+2\kappa_i-i+1}{2}\right)}{\prod_{i=1}^p \Gamma\left(\frac{n+m+2\kappa_i-i+1}{2}\right)}, \tag{2.7}$$

which are the corresponding moments of $\prod_{i=1}^{p}\chi^2_{n+m+2\kappa_i-i+1}$. Hence, the lemma follows.

The next lemma gives monotone likelihood ratio properties of the conditional distribution of $|S^*|$ obtained in Lemma 2.1. For this, let $h_p(x, \boldsymbol{n})$, with $\boldsymbol{n} = (n_1, \dots, n_p)$, denote the density of $\prod_{i=1}^{p}\chi^2_{n_i}$, where $\chi^2_{n_i}$ are independent.

LEMMA 2.2. *The ratio* $h_p(c^{-1}b, \boldsymbol{n})/h_p(c^{-1}a, ,\boldsymbol{n})$ *is* (i) *strictly increasing in each* $n_i > 0$, *for any fixed* $p \geq 1$, $c > 0$, *and* $0 < a < b < \infty$; (ii) *strictly increasing in* $c > 0$, *for any fixed* $p \geq 1$, $0 < a < b < \infty$, *and* $\boldsymbol{n} > \boldsymbol{0}$.

PROOF. In Lemma 2.2 of Sarkar (1989) it is proved that $h_p(x, ,\boldsymbol{n})$ is TP_2 in (x, n_i), for fixed $(n_1, \dots, n_{i-1}, n_{i+1}, \dots, n_p)$, for each i. A close study of this proof would reveal that the inequalities involved there are, in fact, strict; so that $h_p(x, \boldsymbol{n})$ is strictly TP_2 in (x, n_i), which is equivalent to the result stated in part (i) of the present lemma.

To prove part (ii), we need to show that $h_p(c^{-1}x, \boldsymbol{n})$ is strictly TP_2 in (c, x). This result is known to be true for $p = 1$. To see that it is also true for $p > 1$, we first note that $c^{-1}h_p(c^{-1}x, \boldsymbol{n})$, which is the density of $c\prod_{i=1}^{p}\chi^2_{n_i}$ at x, can be written as

$$c^{-1}h_p(c^{-1}x, \boldsymbol{n}) = \int_0^\infty t^{-1}h_1(t^{-1}x, n_1)c^{-1}h_{p-1}(c^{-1}t, \boldsymbol{n_2})\,dt, \tag{2.8}$$

where $h_{p-1}(x, \boldsymbol{n_2})$, with $\boldsymbol{n_2} = (n_2, \dots, n_p)$, is the density of $\prod_{i=2}^{p}\chi^2_{n_i}$. The result then follows from the basic composition theorem for TP_2 functions (Karlin (1968), p. 17) and by using induction, which is the kind of arguments used in proving Lemma 2.2 of Sarkar (1989).

3. The unbiased invariant confidence interval

In this section, we prove that the unbiased invariant confidence interval with confidence coefficient β is given by (1.1)–(1.3) with $h = 2$.

LEMMA 3.1. *An invariant confidence interval of the form* (1.1) *is unbiased iff* $0 < a < b < \infty$ *satisfy* (1.3) *with* $h = 2$.

PROOF. The probability that (1.1) contains $|\Sigma|$ when Σ' is true is given by

$$\begin{aligned}\psi(\theta) &= P_I(\theta a < |S| < \theta b) \\ &= \int_{\theta a}^{\theta b} f_{p,n}(x)\,dx,\end{aligned} \tag{3.1}$$

where $\theta = |\Sigma|/|\Sigma'|$. The condition of unbiasedness in terms of $\psi(\theta)$ is that $\psi(\theta) \leq \psi(1)$ for all θ, a necessary condition for which is $\psi'(1) = 0$. Using

$$\psi'(\theta) = bf_{p,n}(\theta b) - af_{p,n}(\theta a), \tag{3.2}$$

and Remark 2.2 of Sarkar (1989), we see that this necessary condition is same as (1.3) with $h = 2$.

If (1.3) with $h = 2$ is true, then we have using Lemma 2.2(ii) with $c = \theta^{-1}$

$$\psi'(\theta) \leq a f_{p,n}(\theta a) \left[\frac{b f_{p,n}(b)}{a f_{p,n}(a)} - 1 \right] = 0, \tag{3.3}$$

iff $\theta \geq 1$, which implies the unbiasedness of (1.1). This proves the sufficiency part of the lemma.

4. Stein-type improvement of (1.1)

We illustrate in this section how an interval estimator of the form (1.1) can be improved by incorporating X into the estimation procedure. The technique used is a multivariate analog of Nagata's univariate result (1989) and is an adaptation to the interval estimation problem of the method used by Shorrock and Zidek (1976) in improving the best affine equivariant point estimator of $|\Sigma|$.

THEOREM 4.1. *Let $J = |S|(b^{-1}, a^{-1})$ be a confidence interval for $|\Sigma|$, where $0 < a < b < \infty$. Assume that there exists a constant g_0 less than 1 satisfying*

$$f_{p,\,n+m+2}(g_0^{-1} b) = f_{p,\,n+m+2}(g_0^{-1} a). \tag{4.1}$$

If

$$J^* = \min\{|S|,\, g_0|S + XX'|\}(b^{-1}, a^{-1}), \tag{4.2}$$

then

$$\text{(i)} \quad P(J^* \text{ contains } |\Sigma|) > P(J \text{ contains } |\Sigma|), \quad \text{and} \tag{4.3}$$

$$\text{(ii)} \quad E(\text{length}\,(J^*)) < E(\text{length}\,(J)), \tag{4.4}$$

uniformly in (μ, Σ).

PROOF. All probabilities in the following are obtained when $\Sigma = I_p$. The right-hand side of (4.3) is

$$P(a < |S| < b) = P(a < |S^*| g(F) < b), \tag{4.5}$$

where $g(F) = \prod_{i=1}^{s}(1 - f_i)$. Because of Lemma 2.1, (4.5) is equal to

$$E_{(F,\,\kappa)} P\left(a < g(F) \prod_{i=1}^{p} \chi^2_{n+m+2\kappa_i - i+1} < b \right), \tag{4.6}$$

where the probability is obtained with respect to the conditional distribution of $\prod_{i=1}^{p} \chi^2_{n+m+2\kappa_i-i+1}$ given F and κ, which is then integrated with respect to the distribution of (F, κ). Using the notation $f_{p,\,n+m,\,\kappa}(x)$ for the density of

$\prod_{i=1}^{p} \chi^2_{n+m+2\kappa_i-i+1}$, we observe that the derivative of the probability in (4.6) with respect to g is

$$\begin{aligned}(4.7)\qquad &\frac{a}{g^2} f_{p,\,n+m,\,\kappa}(g^{-1}a) - \frac{b}{g^2} f_{p,\,n+m,\,\kappa}(g^{-1}b)\\ &= \frac{a}{g^2} f_{p,\,n+m,\,\kappa}(g^{-1}a)\left[1 - \frac{b f_{p,\,n+m,\,\kappa}(g^{-1}b)}{a f_{p,\,n+m,\,\kappa}(g^{-1}a)}\right].\end{aligned}$$

Hence, if g_κ is such that

$$(4.8)\qquad b f_{p,\,n+m,\,\kappa}(g_\kappa^{-1}b) = a f_{p,\,n+m,\,\kappa}(g_\kappa^{-1}a),$$

we get, using Lemma 2.2(ii), that as g increases from 0 to 1, the probability in (4.6) is strictly increasing for $g \in (0,\, g_\kappa)$ and strictly decreasing for $g \in (g_\kappa,\, 1)$. If we consider (4.8) for $\kappa = 0$, so that g_0 satisfies

$$(4.9)\qquad b f_{p,\,n+m}(g_0^{-1}b) = a f_{p,\,n+m}(g_0^{-1}a),$$

which is same as (4.1) (because of Remark 2.2 in Sarkar (1989)), we see from Lemma 2.2 that $g_\kappa < g_0 < 1$. Hence, the probability in (4.6) is less than

$$(4.10)\qquad P\left(a < \min(g_0,\, g(F)) \prod_{i=1}^{p} \chi^2_{n+m+2\kappa_i-i+1} < b\right),$$

for all $(F,\, \kappa)$. Integrating (4.10) with respect to $(F,\, \kappa)$, we then obtain that the right-hand side of (4.3) is less than

$$(4.11)\qquad P(a < \min\{|S|,\, g_0|S^*|\} < b),$$

which is the coverage probability of the left-hand side of (4.3). This proves part (i) of the theorem. The other part is obvious.

Remark. It is clear that if g_0 satisfying (4.1) is not less than 1, then $J^* \equiv J$, which means that Stein-type improvement does not work in this case. Because of Lemma 2.2, if $0 < a < b < \infty$ are obtained from Condition (1.3), then g_0 satisfying (4.1) is less than 1 iff $m + 2 > h$. Hence, the unbiased confidence interval ($h = 2$) can always be improved by Stein's idea. But, for the shortest length confidence interval ($h = 4$), we need $m \geq 3$.

Since the exact form of $f_{p,\,n}(x)$ for $p > 2$ is complicated, the exact computation of $a < b$ from (1.2) and (1.3), and hence of g_0 from (4.1), becomes difficult for these values of p. When n is large, we could, however, use the following approximation for $f_{p,\,n}$:

$$\gamma p^{-1} f_\tau(\gamma x^{1/p}) x^{1/p-1},$$

where $f_\tau \equiv f_{1,\,\tau}$ and

$$\begin{aligned}\gamma &= p\left\{1 - \frac{1}{2}(n+1)^{-1}(p-1)(p-2)\right\}^{1/p}\\ \tau &= p(n-p+1),\end{aligned}$$

which is, of course, exact for $p = 1$ and $p = 2$ (Anderson (1984), p. 265).

References

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed., Wiley, New York.

Brown, L. D. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters, *Ann. Math. Statist.*, **39**, 29–48.

Cohen, A. (1972). Improved confidence intervals for the variance of a normal distribution, *J. Amer. Statist. Assoc.*, **67**, 382–387.

Karlin, S. (1968). *Total Positivity*, Vol. 1, Stanford Press, Stanford, California.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley, New York.

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

Sarkar, S. K. (1989). On improving the shortest length confidence interval for the generalized variance, *J. Multivariate Anal.*, **31**, 136–147.

Shorrock, R. W. and Zidek, J. V. (1976). An improved estimator of the generalized variance, *Ann. Statist.*, **4**, 629–638.

Sinha, B. K. (1976). On improved estimators of the generalized variance, *J. Multivariate Anal.*, **6**, 617–626.

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.