Ann. Inst. Statist. Math.
37 (1985), Part A, 487-498

IMPROVED CONFIDENCE SET ESTIMATORS OF A MULTIVARIATE NORMAL MEAN AND GENERALIZATIONS

FANNY KI AND KAM-WAH TSUI

(Received May 16, 1984; revised Sept. 10, 1984)

Summary

Let X be the observed vector of the p-variate ($p \geqq 3$) normal distribution with mean θ and covariance matrix equal to the identity matrix. Denote $y_+ = \max\{0, y\}$ for any real number y. We consider the confidence set estimator of θ of the form $C_{\delta^{a,\phi}} = \{\theta : |\theta - \delta^{a,\phi}(X)| \leqq c\}$, where $\delta^{a,\phi} = [1 - a\phi(|X|)/|X|^2]_+ X$ is the positive part of the Baranchik (1970, *Ann. Math. Statist.*, **41**, 642–645) estimator. We provide conditions on $\phi(\cdot)$ and a which guarantee that $C_{\delta^{a,\phi}}$ has higher coverage probability than the usual one, $\{\theta : |\theta - X| \leqq c\}$. This dominance result will be shown to hold for spherically symmetric distributions, which include the normal distribution, t-distribution and double exponential distribution. The latter result generalizes that of Hwang and Chen (1983, *Technical Report*, Dept. of Math., Cornell University).

1. Introduction

The problem of constructing p-variate normal mean point estimators that improve over (dominate) the usual one (the maximum likelihood estimator) has generated considerable research since James and Stein [8] first provided such an estimator. The point estimator improves over the usual one in the sense that given a loss function, the expected loss using the point estimator is less than that using the usual estimator, regardless of the true value of the multivariate normal mean.

While point estimation is important, the associated problem of constructing confidence set estimators of the multivariate normal mean is also of interest. It has applications, for example, in multiple comparison of means. However, there has been little progress in this related confidence set estimation problem. Loosely speaking, this problem involves finding multivariate normal mean confidence set estimators whose volume is less than or equal to the volume of the usual one, and whose

Key words: Confidence set, dominate, normal, spherically symmetric distribution.

coverage probability (probability of covering the true mean) is greater than or equal to the coverage probability of the usual one. The lack of progress in this area is primarily because of difficulty in calculating the coverage probability of a confidence set estimator, even when the covariance matrix is known.

For simplicity, the focus here will be on the known covariance matrix case. Without loss of generality, the covariance matrix can be assumed to be the identity matrix, since a suitable linear transformation will reduce the original problem to one with an identity covariance matrix. Accordingly, let $X=(X_1, X_2, \cdots, X_p)$ be p-variate normal with mean $\theta=(\theta_1, \theta_2, \cdots, \theta_p)$ and covariance matrix equal to the identity matrix. The maximum likelihood estimator of θ is $\delta^0(X)=X$. The usual confidence set estimator of θ is the ball

$$C_{\delta^0}=\{\theta: |\theta-x|\leqq c\} \tag{1.1}$$

centered at $\delta^0(x)=x$ with radius c determined by the confidence coefficient $(1-\alpha)$, namely, $1-\alpha=P(\chi_p^2\leqq c^2)$ where χ_p^2 is the chi-squared random variable with p degrees of freedom, and $|\cdot|$ is the usual Euclidean norm. A confidence set estimator C^* improves over (dominates) another set estimator C if

$$\begin{aligned} &\text{(i)}\quad P_\theta(\theta \in C^*)\geqq P_\theta(\theta \in C) \qquad \text{and} \\ &\text{(ii)}\quad \mathrm{Vol}\,(C^*)\leqq \mathrm{Vol}\,(C)\,, \end{aligned} \tag{1.2}$$

with strict inequality in either (i) or (ii) for a set of θ or X with positive Lebesgue measure.

For any real number y, let $y_+=\max\{0, y\}$. The James-Stein plus rule,

$$\delta^a(X)=(1-a/|X|^2)_+X\,, \qquad \text{where } 0<a\leqq 2(p-2)\,, \tag{1.3}$$

dominates $\delta^0(X)=X$ when $p\geqq 3$ under the squared loss function $L(\theta, \delta)=|\delta-\theta|^2$. The confidence set estimator C_{δ^a} which results from using $\delta^a(X)$ as the center instead of δ^0 in the usual confidence set estimator dominates C_{δ^0}. This was proved by Hwang and Casella [5], [6] for $p\geqq 3$ and with $0<a\leqq a_0$. The proof in Hwang and Casella [6] can be generalized to obtain similar results for spherically symmetric distributions (see Hwang and Chen [7]). For a detailed history of the development of improved confidence set estimators, see Hwang and Casella [6].

We shall focus on confidence set estimators $C_{\delta^{a,\phi}}$ with centers at point estimators of the form

$$\delta^{a,\phi}=[1-a\phi(|X|)/|X|^2]_+X\,. \tag{1.4}$$

Equation (1.4) is the positive part of the Baranchik [1] estimator. The

Baranchik estimator and its positive part dominate $\delta^0(X)$ if $\phi(|X|)$ is increasing and $0 \leqq a \leqq 2(p-2)$ (Efron and Morris [4]). We will show in Section 2 that confidence set estimators $C_{\delta^{a,\phi}}$ with centers at estimators of the form (1.4) dominate the usual one in the sense of (1.2) under appropriate conditions on a and $\phi(|X|)$. In Section 3, the result is generalized to spherically symmetric distributions which include the normal distributions, t-distributions and double exponential distributions. Section 4 concludes with examples and further discussion.

2. The sufficient conditions for dominance of $C_{\delta^{a,\phi}}$ over C_{δ^0}

Both C_{δ^0} and $C_{\delta^{a,\phi}}$ have the same radius, so they have the same volume. In order to establish the domination of $C_{\delta^{a,\phi}}$ over C_{δ^0}, we must show that $C_{\delta^{a,\phi}}$ has higher coverage probability than C_{δ^0}. In evaluating the coverage probability of a confidence set $C_\delta(X)=\{\theta: |\theta-\delta(X)|\leqq c\}$, it is easier to consider the θ-section $C_\theta(\delta)=\{X: |\theta-\delta(X)|\leqq c\}$. Observe that $P_\theta[C_\theta(\delta)]=P_\theta[\{\theta \in C_\delta(X)\}]$ and

$$P_\theta[C_\theta(\delta^{a,\phi})]=\int_{|\theta-\delta^{a,\phi}(X)|\leqq c} f(X-\theta)dX\,, \tag{2.1}$$

where $f(X-\theta)$ is the density of a p-variate normal distribution with mean θ and identity covariance matrix. Another expression for the coverage probability in (2.1) is derived by using a spherical transformation. Let r be $|X|$ and β be the angle between X and θ. Define

$$u(r)=[1-a\phi(r)/r^2]_+\,. \tag{2.2}$$

Then $|\theta-\delta^0(X)|\leqq c$ and $|\theta-\delta^{a,\phi}(X)|\leqq c$ are equivalent to, respectively,

$$r^2-2r|\theta|\cos\beta+|\theta|^2\leqq c^2 \quad \text{and} \tag{2.3}$$

$$r^2(u(r))^2-2ru(r)|\theta|\cos\beta+|\theta|^2\leqq c^2\,. \tag{2.4}$$

Theorem 2.1 of Hwang and Casella ([5], p. 870) shows that $P_\theta[C_\theta(\delta^{a,\phi})] \geqq P_\theta[C_\theta(\delta^0)]$ for all $|\theta|\leqq c$. This result facilitates the proofs of Theorems 2.1 and 3.1 by permitting us to focus on finding sufficient conditions for dominance of $C_{\delta^{a,\phi}}$ over C_{δ^0} only when $|\theta|>c$. For $\phi(|x|)\equiv 1$, Hwang and Casella [6] provided one such sufficient condition. Their condition on a $(0<a\leqq a_0)$ guarantees that $\partial P_\theta[C_\theta(\delta^{a,1})]/\partial a>0$. This condition is sufficient since $\lim_{a\to 0+} P_\theta[C(\delta^{a,1})]=P_\theta[C(\delta^0)]$. In this special case, the solutions of r in the equality of (2.4) could be found explicitly. For a more general $\phi(|X|)$ such solutions cannot be solved for explicitly. Our method does not depend on these explicit solutions.

Conditions on $\phi(|X|)$ will be given so that $\partial P_\theta[C_\theta(\delta^{a,\phi})]/\partial a>0$ for $0< a\leqq a_0^*$. The dominance result is then established. To this end, a useful expression for $h_\theta(a)=P_\theta[C_\theta(\delta^{a,\phi})]$ obtained by using a spherical transformation is given in (2.7) below. If $|\theta|>c$, then the set of X satisfying (2.4) is $\{X: r_-\leqq r\leqq r_+$ and $0\leqq\beta\leqq\beta_0\}$ where $r_\pm$ are solutions to

$$r_\pm u(r_\pm)=|\theta|\cos\beta\pm\sqrt{c^2-|\theta|^2\sin^2\beta}\overset{\text{def}}{=}r_\pm^0 , \tag{2.5}$$

and $\beta_0=\sin^{-1}(c/|\theta|)$. When $|\theta|>c$, $r_\pm^0>0$ for any $0\leqq\beta\leqq\beta_0$. If condition (ii) of (2.9) is satisfied, then $u(r)$ and $ru(r)$ are increasing, and $u(r)>0$ for $r_-\leqq r\leqq r_+$. Hence, r_+ and r_-, the solutions to (2.5), are unique and positive. Let

$$\alpha(r)=r^2-2r|\theta|\cos\beta+|\theta|^2\equiv|x-\theta|^2 . \tag{2.6}$$

The coverage probability of $C_{\delta^{a,\phi}}$ is equal to

$$h_\theta(a)=K\int_0^{\beta_0}\int_{r_-}^{r_+} r^{p-1}\sin^{p-2}\beta f^*(\alpha(r))drd\beta , \tag{2.7}$$

where $K=2\prod_{i=0}^{p-3}\int_0^\pi \sin^i t dt$ and $f^*(\alpha(r))=\frac{1}{\sqrt{2\pi}^p}\exp\{-\alpha(r)/2\}$. Using (2.7) and interchanging the integration and differentiation signs, we obtain

$$h_\theta'(a)=\partial P_\theta[C_\theta(\delta^{a,\phi})]/\partial a=K\int_0^{\beta_0} m(a,\theta,\beta)d\beta , \tag{2.8}$$

where

$$m(a,\theta,\beta)=\sin^{p-2}\beta\left\{r_+^{p-2}f^*(\alpha(r_+))\phi(r_+)\Big/\frac{\partial ru(r)}{\partial r}\Big|_{r=r_+}\right.$$
$$\left.-r_-^{p-2}f^*(\alpha(r_-))\phi(r_-)\Big/\frac{\partial ru(r)}{\partial r}\Big|_{r=r_-}\right\} .$$

A sufficient condition for $\partial P_\theta[C_\theta(\delta^{a,\phi})]/\partial a>0$ is $m(a,\theta,\beta)>0$.

THEOREM 2.1. *Let X be a p-variate normal random variable with mean θ and the identity matrix as covariance matrix. Let $\delta^{a,\phi}(X)$ be as given in* (1.4) *with $0<a\leqq p-2$, and $C_{\delta^{a,\phi}}=\{\theta:|\theta-\delta^{a,\phi}(X)|\leqq c\}$, where $P\{\chi_p^2\leqq c^2\}=1-\alpha$, and ϕ is a real valued function satisfying the following conditions*:

(2.9)

- (i) $0<\phi(r)\leqq 1$, *nondecreasing, and when $u(r)$ in* (2.2) *is positive*
- (ii) $\phi(r)/r$ *is decreasing*
- (iii) $r\partial ru(r)/\partial r$ *is nondecreasing*
- (iv) $\frac{\partial ru(r)}{\partial r}\Big/\phi(r)$ *is nonincreasing.*

Then $C_{\delta^{a,\phi}}$ has higher coverage probability then the usual confidence set *estimator*, C_{δ^0}, *for all* θ, $p\geqq 3$ *if* $0<a\leqq a_0^*\leqq p-2$, *where* a_0^* *is the unique solution to the equation*

$$S_{p-2}^*(a)=t^{*p-2}e^{-a(t^*-1/t^*)/2}=1 , \tag{2.10}$$

where

$$t^*=\max_{\substack{|\theta|>c \\ 0\leqq\beta\leqq\beta_0}} t=\frac{r_+}{r_-}\bigg|_{\substack{\beta=0 \\ |\theta|=c}} .$$

PROOF. It suffices to show that if $|\theta|>c$, then $m(a,\theta,\beta)>0$, or equivalently,

$$R=\frac{r_+^{p-2}f^*(\alpha(r_+))\phi(r_+)(\partial ru(r)/\partial r|_{r=r_-})}{r_-^{p-2}f^*(\alpha(r_-))\phi(r_-)(\partial ru(r)/\partial r|_{r=r_+})}>1 . \tag{2.11}$$

Since $\partial ru(r)/\partial r[\phi(r)]^{-1}$ is nonincreasing, $R\geqq\left(\frac{r_+}{r_-}\right)^{p-2}\frac{f^*(\alpha(r_+))}{f^*(\alpha(r_-))}$ by condition (iv) of (2.9). The right hand side equals $t^{p-2}\exp\left\{-\frac{1}{2}a\left[\phi(r_+)\left(1-\frac{1}{t}\right)+\phi(r_-)(t-1)\right]\right\}$, where $t=r_+/r_-$. Since $0<\phi(r)\leqq 1$ and $t\geqq 1$,

$$R\geqq t^{p-2}\exp\left\{-\frac{1}{2}a\left[t-\frac{1}{t}\right]\right\}\stackrel{\text{def}}{=}S_{p-2}(t) . \tag{2.12}$$

Observe that $S_{p-2}(t)$ either decreases, or increases to a certain maximum point and then decreases. Since $S_{p-2}(1)=1$, it suffices to show that $S_{p-2}(t)\geqq 1$ at $t^*=\max_{\substack{|\theta|>c \\ 0\leqq\beta\leqq\beta_0}} t$.

For fixed $|\theta|$, we first maximize t w.r.t. β. Condition (ii) of (2.9) implies that $\partial ru(r)/\partial r>0$. By differentiating (2.5), we obtain $\partial r_+/\partial\beta\leqq 0$ and $\partial r_-/\partial\beta\geqq 0$, therefore t decreases in β. Thus, $\max_{0\leqq\beta\leqq\beta_0} t=t|_{\beta=0}$.

When $\beta=0$, we have $r_+u(r_+)=|\theta|+c$, $r_-u(r_-)=|\theta|-c$. Using condition (iii) of (2.9), it can be shown that $\partial t/\partial|\theta|\leqq 0$. Therefore, $t^*=\max_{\substack{|\theta|>c \\ 0\leqq\beta\leqq\beta_0}} t=t|_{\substack{|\theta|=c \\ \beta=0}}>1$ and the inequality in (2.12) is strict.

Using conditions (ii) and (iii) in (2.9), Lemma A.1 in the appendix shows that $S_{p-2}^*(a)$ is strictly decreasing in a. Since $S_{p-2}^*(a)>1$ when $a=0$, there must exist a unique $a_0^*>0$ such that $S_{p-2}^*(a_0^*)=1$. Consequently, (2.11) holds for all $0<a\leqq a_0^*$ and the proof is complete.

3. Improved confidence set estimators in the spherically symmetric distribution case

In this section, we show how the proof of Theorem 2.1 can be adapted to extend the dominance result in the normal case to the case where the underlying distribution is spherically symmetric. Let $X=(X_1,\cdots,X_p)$ have a spherically symmetric probability density, $f(|x-\theta|^2)$, with location vector $\theta=(\theta_1,\cdots,\theta_p)$. Intuitively, the flatter the tails are, the higher the probability that the shrinkage estimator $\delta^{a,\phi}$ pulls the observations in the correct direction. The relative increasing rate, RIR $=f'(s)/f(s)$, was used by Hwang and Chen [7] as a measure of the flatness of tails, since the density has flat tails if the RIR is large and the density has sharp tails if the RIR is small. The dominance of $C_{\delta^{a,\phi}}$ over C_{δ^0} when the underlying distribution is normal was established in Section 2. The same result should hold for spherically symmetric distribution with RIR at least equal to that of the normal distribution. The improvement of $C_{\delta^{a,\phi}}$ over C_{δ^0} generally increases with the RIR. Hwang and Chen [7] established such dominance results for spherically symmetric distribution for the case $\phi(|x|)\equiv 1$. The theorem below shows that $C_{\delta^{a,\phi}}$ dominates C_{δ^0} for virtually any spherically symmetric distribution.

THEOREM 3.1. *Let X have a spherically symmetric distribution with density $f(|x-\theta|^2)$. Let $\alpha(r)$ be as given in* (2.6), *and $r_\pm$ be the root of* (2.5). *Suppose the RIR of $f(s)$, $f'(s)/f(s)$, is defined for every $\alpha_0\leqq s\leqq\alpha_1$ and is nonpositive, where*

$$(3.1)\qquad \alpha_0=\begin{cases}\alpha(r_-)|_{\substack{\beta=0\\|\theta|=c}} & a<c^2\\ 0 & a\geqq c^2\end{cases}\qquad \textit{and}$$

$$(3.2)\qquad \alpha_1=\alpha(r_+)|_{\substack{\beta=0\\|\theta|=c}}\ .$$

If $a>0$ is such that

$$(3.3)\qquad \inf_{\alpha_0\leqq s\leqq\alpha_1}\frac{f'(s)}{f(s)}\geqq\frac{-(p-2)}{a}\,\frac{\ln t^*}{(t^*-1/t^*)}\qquad \textit{where } t^*=\frac{r_+}{r_-}\Big|_{\substack{\beta=0\\|\theta|=c}}\ ,$$

then the coverage probability of $C_{\delta^{a,\phi}}$ is higher than C_{δ^0} for all θ provided $\phi(r)$ satisfies conditions (2.9). *Hence $C_{\delta^{a,\phi}}$ dominates C_{δ^0}.*

PROOF. It suffices to show that $\dfrac{\partial}{\partial a}P_\theta[C_\theta(\delta^{a,\phi})]>0$ for $|\theta|>c$. For fixed $|\theta|$, the expression for $h_\theta'(a)$ is similar to that in (2.8), except that f^* is replaced by f. Showing $m(a,\theta,\beta)>0$ is equivalent to showing $\dfrac{r_+^{p-2}f(\alpha(r_+))}{r_-^{p-2}f(\alpha(r_-))}>\dfrac{(\partial ru(r)/\partial r|_{r=r_+})\phi(r_-)}{(\partial ru(r)/\partial r|_{r=r_-})\phi(r_+)}$. Since $\dfrac{\partial ru(r)}{\partial r}\Big/\phi(r)$ is nonincreasing, we need only to show that the left hand side is greater than one.

$$(3.4)\qquad (p-2)\ln t+\ln f(\alpha(r_+))-\ln f(\alpha(r_-))>0\ ,\qquad \text{where } t=r_+/r_-\ .$$

Define $g(s)=\ln f(s)$. By the mean value theorem, there exists an

s between $\alpha(r_-)$ and $\alpha(r_+)$ such that $g(\alpha(r_+))-g(\alpha(r_-))=g'(s)(\alpha(r_+)-\alpha(r_-))$. Observe that $\alpha(r_+)-\alpha(r_-)=a[\phi(r_+)(1-1/t)+\phi(r_-)(t-1)]<a(t-1/t)$. A sufficient condition for (3.4) is $ag'(s)(t-1/t)>-(p-2)\ln t$, which holds if $\inf_{\alpha(r_-)\le s\le\alpha(r_+)} g'(s)>\frac{-(p-2)\ln t}{a(t-1/t)}$. According to Lemma A.2 in the appendix, $\frac{\partial\alpha(r_+)}{\partial\beta}<0$ and $\frac{\partial\alpha(r_-)}{\partial\beta}>0$. For fixed $|\theta|$, $\alpha(r_+)$ and $\alpha(r_-)$ attain their maximum and minimum, respectively at $\beta=0$. By Lemma A.3, if $\beta=0$ then (i) $\partial\alpha(r_+)/\partial|\theta|<0$, and (ii) $\partial\alpha(r_-)/\partial|\theta|>0$ if $a<c^2$. Therefore, α_0 is the minimum of $\alpha(r_-)$ and α_1 is the maximum of $\alpha(r_+)$. As $(\ln t)/(t-1/t)$ is decreasing in t for $t\geqq 1$, (3.3) implies that (3.4) is true for all $|\theta|$ and the result follows.

COROLLARY 3.1. *Suppose $p\geqq 3$ and let $a_1^*>0$ be the unique solution to* (3.3) *with equality. If the left hand side of* (3.3) *is continuous in a, then any a in* $(0, a_1^*)$ *is a solution to the inequality* (3.3).

The existence of a_1^* in Corollary 3.1 will be discussed after the proof of Corollary 3.1.

PROOF. If the left hand side of (3.3) is decreasing in a and the right hand side of (3.3) is increasing in a, then the solution is either an interval or an empty set. Since a_1^* is assumed to exist, the solution to (3.3) is an interval. Accordingly, it will first be established that the left hand side of (3.3) is decreasing in a, and then that the right hand side is increasing in a. From Lemma A.3, if $\beta=0$, then $\partial\alpha(r_+)/\partial a>0$ and $\partial\alpha(r_-)/\partial a<0$ for all $|\theta|>c$. Thus as a increases, the interval $[\alpha_0, \alpha_1]$ becomes larger and hence the left hand side of (3.3) decreases.

To establish that the right hand side of (3.3) is increasing, it suffices to show that $a(t^*-1/t^*)/\ln t^*$ is increasing in a. To evaluate $\partial[a(t-1/t)/\ln t]/\partial a$, the following quantities are needed:

(1) $\partial[(t-1/t)/\ln t]/\partial t=[(1+1/t^2)\ln t-(1/t)(t-1/t)]/(\ln t)^2>0$ for $t>1$ (see Lemma A.1 in Hwang and Chen [7])

(2) $\partial t/\partial a=t[r_+^{-2}(\partial ru(r)/\partial r|_{r=r_+})^{-1}\phi(r_+)-r_-^{-2}(\partial ru(r)/\partial r|_{r=r_-})^{-1}\phi(r_-)]$ $\geqq -t/a$ since $\phi(r)/r$ is decreasing. Thus,

$$\frac{\partial}{\partial a}\left[\frac{a(t-1/t)}{\ln t}\right]=\frac{t-1/t}{\ln t}+a\frac{\partial}{\partial t}\left[\frac{t-1/t}{\ln t}\right]\frac{\partial t}{\partial a}$$
$$\geqq\frac{1}{(\ln t)^2 t}(-2\ln t+t^2-1)\ ,$$

which is nonnegative because $-2\ln t+t^2-1$ is zero at $t=1$ and its derivative is positive when $t>1$. Therefore, $a(t^*-1/t^*)/\ln t^*$ is increasing in a.

Under certain conditions, the existence of $a_1^*>0$ which is the unique solution to (3.3) with equality can be established. For example, $a_1^*>0$ exists if $\inf_{a_0\le s\le a_1} \frac{f'(s)}{f(s)}=d(a)>-\infty$. In that case, when $a=0$, (3.3) holds with strict inequality. From Corollary 3.1, the difference of the left hand side and the right hand side of (3.3) is decreasing in a. Therefore, there must exist a unique $a_1^*>0$ satisfying (3.4) with equality.

4. Examples and further discussion

The class Φ of functions $\phi(\cdot)$ satisfying the conditions in (2.9) is not trivial. As remarked before, $\phi(\cdot)\equiv 1$ is in Φ. Moreover, it is easy to check that Φ also contains $\phi(r)$ of the form $r^b/(r^b+w)$, when $0<b\le 1$ and $w\ge 0$. A $\phi(\cdot)$ of this form moderates the shrinkage rate of $\delta^{a,\phi}(X)$.

Once the function $\phi(\cdot)$ is specified, an explicit solution of the maximum value of t, t^*, can be obtained when $|\theta|>c$. It is then possible to obtain explicit expressions for condition (3.3) for commonly encountered spherically symmetric distributions. This, in turn, enables us to solve numerically, if necessary, for the value of a_1^* which produces equality in (3.3). The value of a_1^* depends on the underlying distribution of X and the function $\phi(\cdot)$ which determines the center of the confidence set.

As a first illustration, consider X having a p-variate normal distribution with $p\ge 3$ and the identity matrix as the covariance matrix. Then the RIR is $-1/2$ and (3.3) becomes

$$(4.1)\qquad -1/2\ge\frac{-(p-2)\ln t^*}{a(t^*-1/t^*)}\,,\qquad \text{where } t^*=r_+/r_-\big|_{\substack{|\theta|=c\\ \beta=0}}\,.$$

The inequality (4.1) is equivalent to $S_{p-2}^*(a)$, defined in (2.10), being greater than or equal to 1. Let $\phi(r)=r/(r+2d)$ with $d>0$ $\delta^{a,\phi}(X)=(1-a/[r(r+2d)])X$, where $r=|X|$. The corresponding confidence set estimator is

$$(4.2)\qquad C_{\delta^{a,\phi}}=\{\theta:|\theta-\delta^{a,\phi}(X)|\le c\}\,.$$

The constant c is $100(1-\alpha)$ percentile of the chi-square distribution with p degrees of freedom. There $1-\alpha$ is the confidence level of C_{δ^0}. The maximum value of t is

$$(4.3)\qquad t^*=[\sqrt{(d+c)^2+a}-(d-c)]/[\sqrt{d^2+a}-d]\,.$$

The value $a=a_1^*$ which makes (4.1) an equality can be solved for.

The values of a_1^* for some particular values of d (0, 0.01, and 0.1) and the corresponding coverage probabilities of $C_{\delta^{a_1^*,\phi}}$ when $1-\alpha=0.90$ are given in Table 1 for different values of $|\theta|$. The overall improve-

Table 1. Results for X having a p-variate normal distribution with mean θ and identity matrix as covariance matrix. $(1-\alpha)=0.90$.

A. The value of $a=a_1^*$ that makes (4.1) an equality.

p	$d=0$	$d=0.01$	$d=0.10$
5	2.1325	2.1248	2.0534
9	5.4131	5.4028	5.3082
15	10.4344	10.4215	10.3033

B. Coverage probability of the confidence set estimator, $C_{\delta^{a_1^*},\phi}$, in (4.2).

	$\|\theta\|=0$				$\|\theta\|=2$		
p	$d=0$	$d=0.01$	$d=0.1$	p	$d=0$	$d=0.01$	$d=0.1$
5	0.97803	0.97774	0.97498	5	0.96927	0.96895	0.96604
9	0.99616	0.99608	0.99535	9	0.99354	0.99344	0.99244
15	0.99962	0.99961	0.99951	15	0.99928	0.99926	0.99910

	$\|\theta\|=4$				$\|\theta\|=15$		
p	$d=0$	$d=0.01$	$d=0.1$	p	$d=0$	$d=0.01$	$d=0.1$
5	0.93165	0.93162	0.93108	5	0.90247	0.90246	0.90242
9	0.97818	0.97811	0.97722	9	0.90968	0.90968	0.90959
15	0.99710	0.99706	0.99661	15	0.92375	0.92374	0.92359

ment in coverage probability is very high. When $|\theta|$ is small and p is moderate, the coverage probability is as high as 0.99.

To illustrate the results for a nonnormal distribution, suppose the underlying distribution of X is a p-variate t-distribution with location vector θ and n degrees of freedom. The density of X is proportional to $(1+|x-\theta|^2/n)^{-(n+p)/2}$. The RIR$=f'(s)/f(s)$ is $-(n+p)/[2(n+s)]$ and (3.3) becomes

$$(4.4) \qquad \frac{-(n+p)}{2[n+(c-\sqrt{a})_+^2]} \geq \frac{-(p-2)\ln t^*}{a(t^*-1/t^*)} \qquad \text{where } t^* = \frac{r_+}{r_-}\bigg|_{\substack{|\theta|=c\\ \beta=0}} .$$

If the same $\phi(r)$ and constant c are used, t^* here is the same as in (4.3). The values of $a=a_1^*$ which make (4.4) an equality are given in Table 2A. The corresponding coverage probabilities of $C_{\delta^{a_1^*},\phi}$ are given in Table 2B. Table 2A indicates that the coverage probability of usual confidence set estimator C_{δ^0} can be far less than .9. In contrast, the coverage probabilities of $C_{\delta^{a_1^*},\phi}$ in (4.2) are close to .9 when $|\theta|$ is not very large and p and n are moderate. Thus, the coverage probabilities of $C_{\delta^{a_1^*},\phi}$ in (4.2) appears to be more robust than that of the usual confidence set estimator, C_{δ^0}, with respect to the distribution's departure from normality, provided that the underlying distribution is the p-variate t-distribution.

Table 2. Results for X having a p-variate t-distribution with mean θ and n degrees of freedom. $(1-\alpha)=.90$.

A. The value of $a=a_1^*$ that makes (4.4) an equality.

p	n	$d=0$	$d=0.01$	$d=0.10$	The coverage probability of the usual confidence set estimator, $C_{\delta 0}$ (i.e. 4.2 with $a=0$).
5	1	1.4365	1.4415	1.4831	0.49491
5	5	1.6361	1.6372	1.6455	0.74154
5	15	1.8465	1.8437	1.8172	0.83601
9	1	2.7363	2.7457	2.8267	0.45326
9	5	3.2198	3.2253	3.2721	0.69256
9	15	3.9282	3.9279	3.9246	0.80576
15	1	4.4113	4.4247	4.5421	0.42504
15	5	5.1055	5.1155	5.2029	0.65061
15	15	6.3607	6.3644	6.3956	0.77439

B. Coverage probability of the confidence set estimator, $C_{\delta a_1^*, \phi}$ in (4.2).

$|\theta|=0$ / $|\theta|=2$

p	n	$d=0$	$d=0.01$	$d=0.1$	p	n	$d=0$	$d=0.01$	$d=0.1$
5	1	0.54602	0.54593	0.54502	5	1	0.53641	0.53633	0.53559
5	5	0.82696	0.82665	0.82393	5	5	0.81242	0.81215	0.80979
5	15	0.92527	0.92489	0.92148	5	15	0.91204	0.91169	0.90852
9	1	0.51652	0.51646	0.51594	9	1	0.50802	0.50802	0.50762
9	5	0.81225	0.81204	0.81009	9	5	0.79856	0.79836	0.79662
9	15	0.93740	0.93713	0.93466	9	15	0.92623	0.92595	0.92348
15	1	0.49380	0.49377	0.49346	15	1	0.48731	0.48729	0.48705
15	5	0.79052	0.79037	0.78903	15	5	0.77934	0.77921	0.77800
15	15	0.93685	0.93664	0.93479	15	15	0.92806	0.92786	0.92598

$|\theta|=4$ / $|\theta|=15$

p	n	$d=0$	$d=0.01$	$d=0.1$	p	n	$d=0$	$d=0.01$	$d=0.1$
5	1	0.51242	0.51243	0.51247	5	1	0.49626	0.49626	0.49629
5	5	0.77216	0.77212	0.77170	5	5	0.74390	0.74390	0.74390
5	15	0.86972	0.86964	0.86892	5	15	0.83863	0.83863	0.83859
9	1	0.48820	0.48820	0.48812	9	1	0.45679	0.45680	0.45686
9	5	0.76342	0.76331	0.76234	9	5	0.69991	0.69991	0.69994
9	15	0.89253	0.89232	0.89047	9	15	0.81520	0.81519	0.81513
15	1	0.47155	0.47155	0.47149	15	1	0.43149	0.43150	0.43160
15	5	0.75046	0.75037	0.74954	15	5	0.66521	0.66522	0.66530
15	15	0.90213	0.90194	0.90018	15	15	0.79525	0.79525	0.79516

Appendix

In this appendix, the conditions in (2.9) on $\phi(r)$ are assumed to be satisfied.

LEMMA A.1. *Let* $S=t^{p-2}\exp\{-a(t-1/t)/2\}$, *where* $t=r_+/r_-$ *and* $r_\pm$ *are roots of equation* (2.5). *Then* $\partial S/\partial a<0$.

PROOF. First observe that $\partial t/\partial a<0$, by direct differentiation and conditions (2.9) (ii) and (iii). Next, if $a\leqq p-2$, then

$$\begin{aligned}\frac{\partial S}{\partial a}&\leqq\frac{1}{2}t^{p-3}\exp\left\{-\frac{1}{2}a\left(t-\frac{1}{t}\right)\right\}\left\{-(t^2-1)\right.\\&\qquad\left.+\frac{a\phi(r_-)}{r_-^2}\left(\frac{\partial ru(r)}{\partial r}\bigg|_{r=r_-}\right)^{-1}(t^2+1-2t)\right\}\\&<\frac{1}{2}t^{p-3}\exp\left\{-\frac{1}{2}a\left(t-\frac{1}{t}\right)\right\}\{-(t^2-1)+(t^2+1)-2t\}\leqq 0\ .\end{aligned}$$

The second inequality is correct because we are interested in the r_- such that $(1-a\phi(r_-)/r_-^2)>0$, and $\phi(r)/r$ is decreasing, which leads to $\partial ru(r)/\partial r>1$.

LEMMA A.2. *Let* r_+ *and* r_- *be the roots of equation* (2.5) *and let*

$$\alpha(r_\pm)=r_\pm^2-2r_\pm|\theta|\cos\beta+|\theta|^2\ . \tag{A.1}$$

Then $\partial\alpha(r_+)/\partial\beta<0$ *and* $\partial\alpha(r_-)/\partial\beta>0$.

PROOF. Letting $\Delta=c^2-|\theta|^2\sin^2\beta$, we have

$$\frac{\partial\alpha(r_\pm)}{\partial\beta}=\frac{2|\theta|\sin\beta}{(1-a(\partial/\partial r)(\phi(r)/r)|_{r=r_\pm})}\left[\frac{a\phi(r_\pm)}{r_\pm}\left(1\mp\frac{|\theta|\cos\beta}{\sqrt{\Delta}}\right)-a\phi'(r_\pm)\right]$$

by differentiating (A.1) and (2.5), and using the chain rule. Thus, $\partial\alpha(r_+)/\partial\beta<0$, since $\Delta<|\theta|^2\cos^2\beta$ when $|\theta|>c$, $\phi'(r_\pm)\geqq0$ and $\phi(r)/r$ is decreasing. Furthermore, $\dfrac{\partial\alpha(r_-)}{\partial\beta}\geqq\dfrac{2a|\theta|\sin\beta}{(1-a(\partial/\partial r)(\phi(r)/r)|_{r=r_-})}\left[\dfrac{\phi(r_-)}{r_-}-\phi'(r_-)\right]$ >0, since $\phi(r)/r$ is decreasing.

LEMMA A.3. *Let* $\alpha(r_\pm)$ *be defined as in Lemma* A.2. *Suppose* $\beta=0$. *Then* (i) $\partial\alpha(r_+)/\partial|\theta|<0$, (ii) $\partial\alpha(r_-)/\partial|\theta|>0$ *if* $a<c^2$.

PROOF. For $\beta=0$, denote $\alpha(r_\pm)=(r_\pm-|\theta|)^2$ by $\alpha_{|\theta|}(r_\pm)$. Differentiating (2.5) and using the chain rule yields

$$\frac{\partial\alpha_{|\theta|}(r_\pm)}{\partial|\theta|}=\frac{[2a(r_\pm-|\theta|)]}{1-a(\partial/\partial r)(\phi(r)/r)|_{r=r_\pm}}\left\{\frac{\partial}{\partial r}\left(\frac{\phi(r)}{r}\right)\bigg|_{r_\pm}\right\}\ . \tag{A.2}$$

By condition (ii) of (2.9), the sign of (A.2) is the opposite of the sign of the expression in the square brackets. When $\beta=0$, $r_\pm u(r_\pm)=|\theta|\pm c$. Since $u(r_+)\leqq1$, $r_+>|\theta|$ and hence $\partial\alpha_{|\theta|}(r_+)/\partial|\theta|<0$. This establishes (i). To establish (ii) observe that $r_-u(r_-)=r_--a\phi(r_-)/r_-=|\theta|-c$, that is r_-^2-

$(|\theta|-c)r_- - a\phi(r_-)=0$ which implies $r_-^2-(|\theta|-c)r_- - a\leqq 0$. Thus, we have $0<r_-<|\theta|-c+\sqrt{a}$. If $a<c^2$, $r_-<|\theta|$ and hence $\partial\alpha_{|\theta|}(r_-)/\partial|\theta|>0$.

Acknowledgement

The authors would like to thank the referee for his careful reading of the original manuscript. His comments helped us to clarify the presentation.

UNIVERSITY OF WISCONSIN-MADISON

REFERENCES

[1] Baranchik, A. J. (1970). A family of minimax estimators of the mean of a multivariate normal distribution, *Ann. Math. Statist.*, **41**, 642-645.

[2] Casella, G. (1981). Coverage probabilities for multivariate normal confidence sets centered at James-Stein estimators, *Technical Report BU-767-M*, Biometric Unit, Cornell University.

[3] Casella, G. and Hwang, J. T. (1983). Empirical Bayes confidence sets for the mean of a multivariate normal distribution, *J. Amer. Statist. Ass.*, **78**, 688-698.

[4] Efron, B. and Morris, C. N. (1973). Stein's estimation rule and its competitors—an empirical Bayes approach, *J. Amer. Statist. Ass.*, **68**, 117-130.

[5] Hwang, J. T. and Casella, G. (1982a). Minimax confidence sets for the mean of a multivariate normal distribution, *Ann. Statist.*, **10**, 868-881.

[6] Hwang, J. T. and Casella, G. (1982b). Improved set estimators for the mean of a multivariate normal distribution, *Technical Report*, Department of Mathematics, Cornell University.

[7] Hwang, J. T. and Chen, J. (1983). How do the improved confidence sets for the mean of multivariate normal distribution perform when the true distribution is non-normal?, *Technical Report*, Department of Mathematics, Cornell University.

[8] James, W. and Stein, C. (1961). Estimation with quadratic loss, *Proc. Fourth Berkeley Symp. Math. Statist. Probl.*, **1**, 361-379, University of California Press.

[9] Stein, C. (1962). Confidence sets for the mean of a multivariate normal distribution, *J. R. Statist. Soc.*, B, **24**, 265-296.