BU-746-M
July 1981

Coverage Probabilities for Multivariate Normal Confidence Sets Centered at James-Stein Estimators

George Casella*

* George Casella is Assistant Professor, Biometrics Unit, Cornell University, Ithaca, NY 14853. Research was supported by National Science Foundation Grant No. MCS79-05771.

ABSTRACT

The usual confidence set, based on an observation, X, from a multivariate normal distribution with mean θ and identity covariance matrix is $C_X(\theta) = \{\theta : (\theta - X)'(\theta - X) \leq c^2\}$. We consider confidence sets of the form $C_\delta(\theta) = \{\theta : [\theta - \delta(X)]'[\theta - \delta(X)] \leq c^2\}$, where $\delta(X)$ is either the James-Stein estimator or its positive part version. The exact formulas for the coverage probability of $C_\delta(\theta)$ are derived and evaluated numerically. The numerical evidence suggests that, for all $|\theta|$, the coverage probability of $C_\delta(\theta)$ exceeds that of $C_X(\theta)$.

KEY WORDS: James-Stein estimator; Multivariate normal distribution; Confidence sets.

1. INTRODUCTION

The classic confidence set for the mean, θ, of a p-variate normal distribution with identity covariance matrix, based on the observation X, is given by

$$C_X(\theta) = \{\theta : (\theta - X)'(\theta - X) \le c^2\} \quad . \qquad (1.1)$$

If c^2 is chosen to satisfy $P(\chi^2_p \le c^2) = 1-\alpha$, where χ^2_p denotes a central chi-square random variable with p degrees of freedom, then the sphere $C_X(\theta)$ has probability $1-\alpha$ of covering the true value of θ . $C_X(\theta)$ has the optimal property that, among the class of procedures with coverage probability at least $1-\alpha$, $C_X(\theta)$ minimizes the maximum expected volume.

Stein (1962) wondered if $C_X(\theta)$ was unique in having this property. He conjectured that, for $p \ge 3$, confidence sets of the form

$$C_\delta(\theta) = \{\boldsymbol{\theta} : [\boldsymbol{\theta} - \delta(X)]'[\theta - \delta(X)] \le c^2\} \quad , \qquad (1.2)$$

where $\delta(X) = [1-(p-2)/X'X]X$, should have coverage probability which exceeds $1-\alpha$ for all values of $\boldsymbol{\theta}$. Joshi (1969) proved that for $p=1$ or 2, $C_X(\boldsymbol{\theta})$ is admissible, which is to say, there does not exist a procedure with coverage probability at least $1-\alpha$ and smaller volume than $C_X(\boldsymbol{\theta})$. For $p \ge 3$, Brown (1966) and Joshi (1967) independently established the existence of a dominating procedure. They showed that if $\delta(X) = \{1-[a/(b+X'X)]\}X$, then $C_\delta(\boldsymbol{\theta})$ dominates $C_X(\theta)$ for sufficiently small a and sufficiently large b . This result was not established for any definite values of a and b, but Olshen (1977), for selected values of a and b, calculated the coverage probability of these sets by simulation. The numerical evidence showed that these confidence sets had higher coverage probability than $C_X(\theta)$ for small $|\boldsymbol{\theta}|$.

More recent work on this problem has been done by Faith (1976), Morris (1977), and Berger (1980). Faith derives Bayes confidence sets and shows that, for $p = 3$ or 5, these sets have smaller volume and greater coverage probability than $C_X(\theta)$ for all $|\theta|$ excepting a small interval of middle values. His numerical evidence suggests that, even for the middle values, $C_X(\theta)$ can be dominated. Morris starts with a generalized Bayes estimator of θ and, using the posterior variances, constructs confidence intervals for each coordinate of θ. His simulations of coverage probabilities suggest that these intervals are superior to the usual ones. Berger proceeds by first developing a robust generalized Bayes estimator, $\delta^B(X)$, of θ, then considering a confidence set of the form

$$\{\theta : [\theta - \delta^B(X)]'\Sigma^{-1}(X)[\theta - \delta^B(X)] \le c^2\} \quad ,$$

where $\Sigma(X)$ is the posterior covariance matrix. He shows that this set can have smaller volume than $C_X(\theta)$ for all $|\theta|$, and greater coverage probability for sufficiently large $|\theta|$.

We proceed here in a relatively simple fashion, and consider confidence sets of the form (1.2) where $\delta(X)$ is either

$$\delta(X) = [1 - (a/X'X)]X \quad \text{or} \quad \delta^+(X) = [1 - (a/X'X)]^+X \quad ,$$

where '+' denotes the positive part. We derive formulas for the exact coverage probability of these confidence sets. Since the volumes of these sets are the same as that of $C_X(\theta)$, only coverage probability need be considered. As a result of these formulas it is immediately seen that if $|\theta|^2 \le c^2$, the confidence set centered at $\delta^+(X)$ is superior to $C_X(\theta)$. The integrals are difficult to deal with analytically, but can be evaluated numerically. Tables are constructed which give coverage probabilities for selected values of p and $|\theta|$. In all cases, the numerical evidence suggests that these confidence sets are superior to $C_X(\theta)$.

2. EVALUATION OF COVERAGE PROBABILITIES

In evaluating the coverage probability of the confidence set

$$C_\delta(\theta) = \{\theta : [\theta - \delta(X)]'[\theta - \delta(X)] \le c^2\} \quad , \tag{2.1}$$

it is easier to work with the θ section

$$C_\theta(\delta) = \{X : [\theta - \delta(X)]'[\theta - \delta(X)] \le c^2\} \quad . \tag{2.2}$$

Since $X \in C_\theta(\delta)$ if and only if $\theta \in C_\delta(\theta)$, it follows that

$$P_\theta[C_\delta(\theta)] = P_\theta[C_\theta(\delta)] \quad . \tag{2.3}$$

To evaluate $P_\theta[C_\theta(\delta)]$ we proceed in a manner similar to Faith (1976). First consider the case $|\theta| > 0$. For fixed θ and r, the intersection of $C_\theta(\delta)$ and the shell $S_{\theta,r} = \{X : |X - \theta|^2 = r^2\}$, is the set

$$I_{\theta,r}(\delta) = \{X : X \in C_\theta(\delta), X \in S_{\theta,r}\} \quad . \tag{2.4}$$

Given θ and r, the distribution of X is uniform on the shell $S_{\theta,r}$. Hence, the conditional coverage probability, given θ and r, is the ratio of the surface area of $I_{\theta,r}(\delta)$ to that of $S_{\theta,r}$. If we denote this ratio by $A_{\theta,r}(\delta)$, it then follows that

$$P_\theta[C_\delta(\theta)] = \int_0^\infty A_{\theta,r}(\delta)\,dG(r^2) \quad , \tag{2.5}$$

where $G(\cdot)$ is the cdf of a central χ^2 random variable with p degrees of freedom.

We will work with the estimator $\delta(X) = [1 - (a/X'X)]X$. The equivalent formulas for the positive part version are obtained in a similar manner. For fixed θ and r we can write

$$X = \left(1 + \frac{z}{|\theta|}\right)\theta + y \quad , \tag{2.6}$$

where $y'\theta = 0$. Then for $X \in S_{\theta,r}$, we have

$$|\theta - \delta(X)|^2 = r^2 - 2a + \frac{2a(|\theta|^2 + z|\theta|) + a^2}{|\theta|^2 + 2z|\theta| + r^2} \stackrel{\text{def}}{=} g(z,r) \quad .$$

Notice that, for $r^2 < |\theta|^2 + a$, $g(z,r) \downarrow z$ and, for $r^2 > |\theta|^2 + a$, $g(z,r) \uparrow z$. Let $z_0 = z_0(r)$ satisfy $g(z_0, r) = c^2$. The form of $I_{\theta,r}(\delta)$ depends on the location of z_0 with respect to r . A little algebra will verify that

$$g(r,r) = \frac{[r(r+|\theta|)-a]^2}{(r+|\theta|)^2} \quad .$$

It then follows that $g(r,r) = c^2$ if and only if r is equal to one of the following four roots:

$$r_1 = [\tau_- + (\tau_+^2 + 4a)^{\frac{1}{2}}]/2, \qquad r_2 = [\tau_- - (\tau_+^2 + 4a)^{\frac{1}{2}}]/2 \quad ,$$
$$r_3 = [\tau_+ + (\tau_-^2 + 4a)^{\frac{1}{2}}]/2, \qquad r_4 = [\tau_+ - (\tau_-^2 + 4a)^{\frac{1}{2}}]/2 \quad ,$$

where $\tau_+ = |\theta| + c$ and $\tau_- = |\theta| - c$. The ordering of the roots depends on the relationship among $|\theta|^2$, c^2, and a^2/c^2 . We ignore the case where $|\theta|^2 > c^2$ and $|\theta|^2 < a^2/c^2$ since in most applications it will be vacuous. We obtain the following representation for $I_{\theta,r}(\delta)$:

i) If $|\theta|^2 > c^2$, $|\theta|^2 > a^2/c^2$,

$$I_{\theta,r}(\delta) = \begin{cases} S_{\theta,r} & r^2 \le r_4^2 \\ A_R & r_4^2 < r^2 \le r_2^2 \\ A_L & r_1^2 < r^2 < r_3^2 \\ \phi & \text{otherwise} \end{cases} \quad .$$

ii) If $|\theta|^2 \le c^2$, $|\theta|^2 > a^2/c^2$,

$$I_{\theta,r}(\delta) = \begin{cases} S_{\theta,r} & r^2 \le r_4^2 \quad \text{or} \quad r_1^2 < r^2 < r_2^2 \\ A_R & r_4^2 < r^2 \le r_1^2 \\ A_L & r_2^2 < r^2 \le r_3^2 \\ \phi & \text{otherwise} \end{cases} \quad .$$

iii) If $|\theta|^2 \le c^2$, $|\theta|^2 \le a^2/c^2$,

$$I_{\theta,r}(\delta) = \begin{cases} S_{\theta,r} & r_1^2 < r^2 \le r_2^2 \\ A_R & r_4^2 < r^2 \le r_1^2 \\ A_L & r_2^2 < r^2 \le r_3^2 \\ \phi & \text{otherwise} \end{cases} \quad , \qquad (2.7)$$

where $A_R = \{X : z \ge z_0, z^2 + |y|^2 = r^2\}$ and $A_L = \{X : z \le z_0, z^2 + |y|^2 = r^2\}$.

The last step in computing the coverage probability is finding the ratio of the surface area of A_R and A_L to $S_{\theta,r}$. Notice that A_R and A_L are, respectively, the surface of the shell $S_{\theta,r}$ to the right and left of a plane perpendicular to the z-axis through z_0 . By elementary calculus we find that the surface ratio of A_R to $S_{\theta,r}$, which we denote by $h(z_0, r^2)$, is given by

$$h(z_0, r^2) = \frac{\Gamma(p/2) \int_{z_0}^{r} (r^2-x^2)^{(p-3)/2} dx}{r^{p-2}\Gamma[(p-1)/2]\Gamma(\frac{1}{2})} \quad .$$

This can also be expressed in terms of the incomplete Beta function ratio, or if p is odd, the binomial formula can be used to express $h(z_0, r^2)$ as a finite sum. The surface area ratio of A_L to $S_{\theta,r}$ is given by $1 - h(z_0, r^2)$.

Let $G(\cdot)$ be the cdf of a central chi-square random variable with p degrees of freedom. From (2.7) we obtain the following formulas for the coverage probabilities.

i) If $|\theta|^2 > c^2$, $|\theta|^2 > a^2/c^2$,

$$P_\theta[C_\delta(\theta)] = P(\chi_p^2 \le r_4^2) + \int_{r_4^2}^{r_2^2} h(z_0, x)\,dG(x) + \int_{r_1^2}^{r_3^2} [1 - h(z_0, x)]\,dG(x) \quad .$$

ii) If $|\theta|^2 \le c^2$, $|\theta|^2 > a^2/c^2$,

$$P_\theta[C_\delta(\theta)] = P(\chi_p^2 \le r_4^2) + P(r_1^2 \le \chi_p^2 \le r_3^2) + \int_{r_4^2}^{r_1^2} h(z_0, x)\,dG(x) + \int_{r_2^2}^{r_3^2} [1 - h(z_0, x)]\,dG(x) \quad .$$

iii) If $|\theta|^2 \le c^2$, $|\theta|^2 \le a^2/c^2$,

$$P_\theta[C_\delta(\theta)] = P(r_1^2 \le \chi_p^2 \le r_2^2) + \int_{r_4^2}^{r_1^2} h(z_0, x)\,dG(x) + \int_{r_2^2}^{r_3^2} [1 - h(z_0, x)]\,dG(x) \quad .$$

If $|\theta| = 0$, the coverage probability is given by

$$P_0[C_\delta(0)] = P(w_0 \le \chi_p^2 \le w_1) \quad ,$$

where w_0 and w_1 are the roots of $(w - a)^2 = c^2 w$. It is straightforward to verify that $0 \le w_0 \le c^2 \le w_1$.

As $|\theta| \to \infty$, both r_4^2 and r_2^2 converge to c^2, while r_1^2 and r_3^2 approach infinity. Thus, from the above formulas it follows that

$$\lim_{|\theta| \to \infty} P_\theta[C_\delta(\theta)] = P(\chi_p^2 \le c^2) = 1 - \alpha \quad .$$

It is also true, however, that $r_4^2 \le c^2$, so that the above formulas do not explicitly show that the set $C_\delta(\theta)$ has higher coverage probability than $C_X(\theta)$. For $a = p - 2$, these integrals have been evaluated numerically for selected values of $|\theta|$ and p. In all cases examined the coverage probability of $C_\delta(\theta)$ is at least as high as that of $C_X(\theta)$ and, in some cases, the difference is quite substantial. These coverage probabilities are given in Table 1.

Insert Table 1 here

The positive part James-Stein estimator is defined by

$$\delta^+(X) = [1 - (a/X'X)]^+ X \quad . \tag{2.8}$$

The advantage over the ordinary James-Stein estimator is that the singularity at $X'X = 0$ has been removed, and the coordinates of $\delta^+(X)$ have the same sign as those of X. While $\delta^+(X)$ is not an admissible estimator of θ, it is known to be difficult to improve upon.

The derivation of the coverage probability of $C_{\delta^+}(\theta)$ follows quickly from that of $C_\delta(\theta)$. If we define $g^+(z,r) = |\theta - \delta^+(X)|^2$, it follows from the monotonicity of $g(z,r)$ and the fact that $g(z,r) = |\theta|^2$ at $|X|^2 = a$, that

$$g^+(z,r) = \begin{cases} \min[|\theta|^2, g(z,r)] & r^2 \le |\theta|^2 + a \\ \max[|\theta|^2, g(z,r)] & r^2 > |\theta|^2 + a \end{cases} .$$

The intersection $I_{\theta,r}(\delta^+)$, of the sphere $C_{\delta^+}(\theta)$ and the shell $S_{\theta,r}$ is given by

$$I_{\theta,r}(\delta^+) = \begin{cases} S_{\theta,r} & r^2 \le |\theta|^2 + a & \\ I_{\theta,r}(\delta) & r^2 > |\theta|^2 + a & \Bigr\} \ |\theta|^2 \le c^2 \\ I_{\theta,r}(\delta) & r^2 \le |\theta|^2 + a & \\ \emptyset & r^2 > |\theta|^2 + a & \Bigr\} \ |\theta|^2 > c^2 \end{cases} ,$$

where $I_{\theta,r}(\delta)$ is defined in (2.7). Recall that $G(\cdot)$ is the cdf of a central chi-square random variable with p degrees of freedom. Using (2.7), the coverage probability of $C_{\delta^+}(\theta)$ is

i) If $|\theta|^2 \le c^2$,

$$P_\theta[C_{\delta^+}(\theta)] = P(\chi_p^2 \le r_3^2) + \int_{r_2^2}^{r_3^2} [1 - h(z_0, x)]dG(x) \quad ,$$

ii) If $|\theta|^2 > c^2$,

$$P_\theta[C_{\delta^+}(\theta)] = P(\chi_p^2 \le r_4^2) + \int_{r_4^2}^{r_2^2} h(z_0, x)dG(x) \quad .$$

If $|\theta| = 0$, the coverage probability is given by

$$P_0[C_{\delta^+}(0)] = P(\chi_p^2 \le w_1) \quad ,$$

where w_1 is the largest root of $(w - a)^2 = c^2 w$.

These formulas are slightly simpler than those for the ordinary James-stein estimator, and one can easily see that $C_{\delta^+}(\theta)$ dominates both $C_\delta(\theta)$ and $C_X(\theta)$ if $|\theta|^2 \le c^2$. If $|\theta|^2 > c^2$ then, surprisingly, $C_\delta(\theta)$ dominates $C_{\delta^+}(\theta)$; however, the difference in coverage probabilities is negligible. $C_\delta(\theta)$ dominates because if θ has large positive coordinates and X has small, negative coordinates ($X'X \ll a$), then $\delta(X)$ will be close to θ but $\delta^+(X)$ will be zero. The fact that this type of occurrence happens with small probability is reflected in the negligible gain in the coverage probability.

For $a = p - 2$, $P_\theta[C_{\delta^+}(\theta)]$ was computed for selected values of $|\theta|$ and p . In all cases examined the coverage probability exceeded that of $C_X(\theta)$. The results are presented in Table 2.

Insert Table 2 here

4. CONCLUSIONS

The calculations of coverage probabilities for confidence sets centered at James-Stein type estimators provide strong evidence that there is much to be gained over the usual confidence set. The implementation and interpretation of these confidence sets is straightforward, and even lends itself to coordinatewise interpretations.

REFERENCES

BERGER, JAMES O. (1980), "A Robust Generalized Bayes Estimator and Confidence Region for a Multivariate Normal Mean," Annals of Statistics, 8, 716-761.

BROWN, L. D. (1966), "On the Admissibility of Invariant Estimators of One or More Location Parameters, Annals of Mathematical Statistics, 37, 1087-1136.

FAITH, R. E. (1976), "Minimax Bayes Set and Point Estimators of a Multivariate Normal Mean," Technical Report #66, Ann Arbor, Michigan: University of Michigan.

JOSHI, V. M. (1967), "Inadmissibility of the Usual Confidence Sets for the Mean of a Multivariate Normal Population," Annals of Mathematical Statistics, 38, 1868-1875.

___________ (1969), "Admissibility of the Usual Confidence Sets for the Mean of a Univariate or Bivariate Normal Population," Annals of Mathematical Statistics, 40, 1042-1067.

MORRIS, C. (1977), "Interval Estimation for Empirical Bayes Generalizations of Stein's Estimator," The Rand Paper Series, California: Rand Corporation.

OLSHEN, R. (1977), "Comments on 'A Note on a Reformation of the S-Method of Multiple Comparisons' by H. Scheffé," Journal of the American Statistical Association, 72, 144-146.

STEIN, C. (1962), "Confidence Sets for the Mean of a Multivariate Normal Distribution," Journal of the Royal Statistical Society, Ser. B, 24, 265-296.

Table 1. Coverage Probabilities for the Set $C_\delta(\theta) = \{\theta : [\theta - \delta(X)]'[\theta - \delta(X)] \le c^2\}$, where $\delta(X) = [1 - (p-2)/X'X]X$ and c^2 Satisfies $P(\chi^2 \le c^2) = 1 - \alpha$

$\alpha = .10$

$\lvert\theta\rvert$ \ p	3	5	7	9	11	13	15	25
0	.9455	.9752	.9869	.9925	.9955	.9972	.9982	.9998
2	.9356	.9698	.9846	.9917	.9952	.9971	.9982	.9998
4	.9062	.9343	.9622	.9808	.9903	.9948	.9970	.9998
6	.9026	.9162	.9337	.9510	.9661	.9780	.9866	.9993
8	.9014	.9093	.9202	.9323	.9443	.9556	.9657	.9943
10	.9009	.9060	.9133	.9218	.9307	.9397	.9484	.9819
15	.9004	.9027	.9061	.9102	.9148	.9196	.9247	.9502
20	.9002	.9015	.9035	.9058	.9085	.9114	.9145	.9317
25	.9001	.9010	.9022	.9038	.9055	.9075	.9095	.9214
50	.9000	.9002	.9006	.9010	.9014	.9019	.9024	.9058
100	.9000	.9001	.9001	.9003	.9003	.9005	.9006	.9015
500	.9000	.9000	.9000	.9000	.9000	.9000	.9000	.9001
1000	.9000	.9000	.9000	.9000	.9000	.9000	.9000	.9001

$\alpha = .05$

$\lvert\theta\rvert$ \ p	3	5	7	9	11	13	15	25
0	.9703	.9849	.9916	.9950	.9970	.9981	.9988	.9999
2	.9605	.9846	.9918	.9953	.9972	.9983	.9990	.9999
4	.9540	.9706	.9842	.9925	.9960	.9978	.9987	.9999
6	.9516	.9600	.9701	.9794	.9868	.9921	.9955	.9998
8	.9509	.9558	.9623	.9692	.9757	.9818	.9865	.9984
10	.9505	.9538	.9582	.9631	.9682	.9731	.9778	.9938
15	.9502	.9517	.9538	.9562	.9589	.9618	.9647	.9785
20	.9501	.9510	.9521	.9536	.9552	.9569	.9588	.9685
25	.9500	.9506	.9574	.9523	.9534	.9545	.9558	.9627
50	.9500	.9501	.9503	.9506	.9509	.9512	.9515	.9535
100	.9500	.9500	.9501	.9501	.9502	.9503	.9504	.9509
500	.9500	.9500	.9500	.9500	.9500	.9500	.9500	.9501
1000	.9500	.9500	.9500	.9500	.9500	.9500	.9500	.9500

Table 2. Coverage Probabilities for the Set $C_{\delta^+}(\theta) = \{\theta : [\theta - \delta^+(X)]'[\theta - \delta^+(X)] \le c^2\}$, where $\delta^+(X) = [1 - (p-2)/X'X]^+X$ and c^2 Satisfies $P(\chi^2 \le c^2) = 1 - \alpha$

$\alpha = .10$

$\lvert\theta\rvert$ \ p	3	5	7	9	11	13	15	25
0	.9565	.9879	.9959	.9985	.9994	.9998	.9999	.9999
2	.9458	.9809	.9926	.9972	.9989	.9995	.9998	.9999
4	.9062	.9343	.9622	.9808	.9949	.9977	.9989	.9999
6	.9026	.9162	.9337	.9510	.9661	.9780	.9866	.9993
8	.9014	.9093	.9202	.9323	.9443	.9556	.9567	.9943
10	.9009	.9060	.9133	.9218	.9307	.9397	.9484	.9819
15	.9004	.9027	.9061	.9102	.9147	.9196	.9247	.9502
20	.9002	.9015	.9035	.9059	.9085	.9114	.9145	.9317
25	.9001	.9010	.9022	.9038	.9055	.9075	.9095	.9214
50	.9000	.9002	.9006	.9010	.9014	.9019	.9024	.9058
100	.9000	.9001	.9001	.9002	.9004	.9005	.9006	.9015
500	.9000	.9000	.9000	.9000	.9000	.9000	.9000	.9001
1000	.9000	.9000	.9000	.9000	.9000	.9000	.9000	.9000

$\alpha = .05$

$\lvert\theta\rvert$ \ p	3	5	7	9	11	13	15	25
0	.9788	.9945	.9983	.9994	.9998	.9999	.9999	.9999
2	.9745	.9917	.9970	.9989	.9996	.9998	.9999	.9999
4	.9540	.9706	.9847	.9960	.9982	.9992	.9997	.9999
6	.9516	.9600	.9701	.9794	.9868	.9920	.9955	.9999
8	.9509	.9558	.9623	.9692	.9757	.9816	.9865	.9984
10	.9505	.9538	.9582	.9631	.9682	.9731	.9778	.9938
15	.9502	.9517	.9538	.9562	.9589	.9618	.9647	.9785
20	.9501	.9510	.9521	.9536	.9552	.9569	.9588	.9685
25	.9500	.9506	.9514	.9523	.9534	.9545	.9558	.9627
50	.9500	.9501	.9503	.9506	.9509	.9512	.9515	.9535
100	.9500	.9500	.9501	.9501	.9502	.9503	.9504	.9509
500	.9500	.9500	.9500	.9500	.9500	.9500	.9500	.9501
1000	.9500	.9500	.9500	.9500	.9500	.9500	.9500	.9500