CONSTRUCTING CONDITIONALLY ACCEPTABLE RECENTERED SET ESTIMATORS

GEORGE CASELLA[1]

Cornell University

BU-889-M September, 1985

[1] Research supported by National Science Foundation Grant No.MCS83-02541.

AMS 1980 *subject classifications.* Primary 62C99; secondary 62F10.

Key words and Phrases. Betting procedures, confidence sets, multivariate normal.

Summary

The usual confidence sphere for a multivariate normal mean can be uniformly improved upon, in terms of coverage probability, by recentering it at a positive-part James-Stein estimator. However, these improved sets can have poor conditional performance. Using the theory of relevant betting procedures, which provides an objective means for assessing the conditional performance of a statistical procedure, a criterion for conditional acceptability can be established. A method of constructing such sets is outlined, and applied to some recentered confidence sets. In particular, recentering at the positive-part James-Stein estimator yields a conditionally acceptable confidence set.

CONSTRUCTING CONDITIONALLY ACCEPTABLE RECENTERED SET ESTIMATORS

By GEORGE CASELLA[1]

Cornell University

1. **Introduction.** The usual frequentist theory of statistics is only concerned with long-run (averaged over the sample space) performance. In particular, if $C(X)$ is a set estimator for a parameter θ, where $X \sim F(X|\theta)$, then frequentist theory measures the performance of $C(X)$ according to its confidence coefficient, $1-\alpha$, given by

$$\inf_{\theta} P_\theta[\theta \varepsilon C(X)] = 1-\alpha \quad , \tag{1.1}$$

where $P_\theta[\theta \varepsilon C(X)]$ is the probability that the random set $C(X)$ covers θ. If there exists a subset, S, of the sample space (a recognizable subset in the terminology of Fisher, 1956), that satisfies either

$$\begin{aligned} &\text{i)}\quad P_\theta(\theta \ \varepsilon\ C(X)|X \ \varepsilon\ S) > 1-\alpha+\varepsilon \qquad \text{for all } \theta, \\ &\text{or} \\ &\text{ii)}\quad P_\theta(\theta \ \varepsilon\ C(X)|X \ \varepsilon\ S) < 1-\alpha-\varepsilon \qquad \text{for all } \theta, \end{aligned} \tag{1.2}$$

for some $\varepsilon > 0$, the one should have doubts about assigning confidence $1-\alpha$ to the set $C(X)$. A subset S that satisfies (1.2) is called a relevant subset for $C(X)$, and provides a winning betting strategy against $C(X)$. More precisely, Buehler (1959) argued that one can make the statement "the probability that the set $C(X)$ covers θ is $1-\alpha$" only if one is willing to accept bets that $\theta \ \varepsilon\ C(X)$ at odds $1-\alpha:\alpha$, and accept bets that $\theta \ \varepsilon\ C(X)$ at odds $\alpha:1-\alpha$. If $C(X)$ satisfies 1.2i, for example, and the betting strategy "bet for coverage if $X \ \varepsilon\ S$" is adopted, then the bettor will have positive expected gain for all θ.

In practice, one is usually willing to forgive errors in the direction of 1.2i, i.e., erring on the conservative side. The fact that the stated (nominal) confidence coefficient $1-\alpha$ may be smaller than the actual coverage probability (conditional or unconditional) is forgivable statistically. Errors in the direction of 1.2ii, however, are not forgivable and cast serious doubt on the validity of assigning $1-\alpha$ confidence to $C(X)$.

Since frequentist measure are pre-experimental, and the inferences are unconditional, one cannot expect the data-dependent precision that both Bayesian and likelihood inference has. However, one has a right to expect reasonable conditional performance, which is certainly not the case if 1.2ii is satisfied. Robinson (1976) and Bondar (1977) argue for the adoption of conditional confidence criteria, one reasonable candidate criterion being: use no set $C(X)$ for which 1.2ii is satisfied.

We will be concerned here, in general, with the conditional performance of frequentist confidence procedures, and we will refer to the pair $\langle C(X), 1-\alpha\rangle$ as a confidence set for the parameter θ. (In general, α may be a function of X, $\alpha=\alpha(X)$, but our concern here is measurement of conditional performance using the frequentist confidence coefficient, which is always independent of X.) Specifically, we consider set estimation of the mean of a multivariate normal distribution. If $X \sim N_p(\theta, I)$, a p-variate normal with mean θ and identity covariance matrix, the usual confidence set is given by

(1.3) $$C_0(X) = \{\theta: |\theta-X| \leq c\} \quad ,$$

a sphere of radius C centered at X, where c satisfies $P(\chi_p^2 < c^2) = 1-\alpha$. This set estimator $\langle C_0(X), 1-\alpha\rangle$ is relatively free of conditional defects (Robinson, 1979b); in particular, there are no recognizable subsets for which 1.2ii is satisfied. But $\langle C_0(X), 1-\alpha\rangle$ can be improved upon, in the frequentist sense, by a set estimator $\langle C_\delta(X), 1-\alpha\rangle$, where

$$C_\delta(X) = \{\theta : |\theta - \delta(X)| \le c\} \quad , \tag{1.4}$$

and $\delta(X)$ is a Stein-type estimator. We will refer to sets such as C_δ as recentered confidence sets. Results of Hwang and Casella (1982,1984) show that recentering at a positive-part Stein estimator will uniformly improve coverage probability over $C_0(X)$, while clearly maintaining the same volume.

The major goal of this paper is to find out whether the good conditional properties of $\langle C_0(X), 1-\alpha\rangle$ are retained by $\langle C_\delta(X), 1-\alpha\rangle$. We will mainly be concerned with sets centered at the ordinary and positive-part James-Stein estimator, but in Section 4 we will also consider point estimators based on Strawderman's (1971) prior distributions.

In Section 2 we give some background for the betting scenario of confidence set estimation, and in Section 3 we outline the method of constructing confidence sets with acceptable conditional properties. The construction in Section 3 depends on a technical lemma, Lemma 3.1, whose proof is given in the Appendix. In Section 4 we apply the results of Section 3 to obtain recentered confidence sets with acceptable conditional properties.

2. **Betting Procedures.** Buehler's concept of relevant subsets was extended and formalized by Robinson (1979a) to the concept of relevant betting procedures (i.e., functions). Betting strategies exist which cannot be expressed in terms of subsets, so Robinson's extension was intended to include all possible betting strategies. Thus, a betting procedure, s(X), is defined to be any bounded function of X. Without loss of generality we take this bound to be unity. We can think of $|s(x)|$ as the probability that a bet of one unit is made when X=x is observed, with the sign of s(X) giving the direction of the bet.

Definition 2.1: For the confidence set, $\langle C(X), \beta(X) \rangle$, the betting procedure s(X) is said to be

i. semirelevant if

$$E_\theta\{[I(\theta \in C(X)) - \beta(X)]s(X)\} \geq 0 \quad \text{for all } \theta$$

and is strictly positive for some θ,

ii. relevant if, for some $\varepsilon > 0$,

$$E_\theta\{[I(\theta \in C(X)) - \beta(X)]s(X)\} \geq \varepsilon E_\theta|s(X)| \quad \text{for all } \theta$$

with strict inequality for some θ,

iii. super-relevant if, for some $\varepsilon > 0$

$$E_\theta\{[I(\theta \in C(x)) - \beta(X)]s(X)\} \geq \varepsilon \quad \text{for all } \theta.$$

Notice that if s(X) is the indicator function of some set, then the above definition of a relevant betting procedure reduces to that of Buehler.

In this paper we will only deal with the case $\beta(X) = 1-\alpha$, that is, the same confidence is asserted for every X. The main reason for this restriction is that our concern is with the conditional evaluation of frequentist confidence sets, where the reported confidence level is the infimum of the coverage probabilities. We enquire as to whether this reported confidence will remain valid when evaluated conditionally. It is, of course, possible to perform conditional evaluations of confidence sets whose confidence levels are a nonconstant function of X. However, unconditional (averaged over X) evaluations of such sets are, in general, extremely difficult, which is one reason why such sets have seen little use within frequentist theory.

If X is an observation from a p-variate normal distribution with mean vector θ and covariance matrix I, the usual frequentist $1-\alpha$ confidence set is

$$C_0(X) = \{\theta : |\theta - X| \leq c\} \tag{2.1}$$

where c satisfies $P(\chi_p^2 < c^2) = 1-\alpha$. From the results of Robinson (1979b), it follows that $C_0(X)$ does not allow relevant betting procedures, which is the strongest conditional property that one can expect $C_0(X)$ to have. (Essentially, only proper Bayes posterior confidence sets do not allow the existence of semirelevant betting procedures.)

When evaluating procedures on both frequentist and conditional grounds, there seems to be a trade-off in the sense that procedures with the strongest conditional properties (i.e., proper Bayes procedures) have weaker frequentist properties, and vice versa. Since we are addressing conditional properties from a frequentist standpoint, and want to use procedures that are $1-\alpha$ procedures in a frequentist sense, we cannot hope that these procedures have the strongest conditional properties. What we can hope for, instead, is that our frequentist procedures have a conditional property that is strong enough to eliminate any aberrant behavior.

The type of betting procedure that causes the most concern about the worth of a confidence set is a negatively-biased betting procedure.

Definition 2.2: A betting procedure s(X) is *negatively biased* if $-1 \leq s(X) \leq 0$ for all X, and *positively biased* if $0 \leq s(X) \leq 1$ for all X.

A negatively-biased relevant betting procedure, for example, will always bet against coverage and have a positive expected gain for all θ, giving us the interpretation that, conditionally, the confidence set is not achieving its stated level of confidence. If we can identify a negatively-biased relevant subset (for example, bet against C(X) if $X \varepsilon S$), then the second inequality in (1.2) will obtain, namely

$$P_\theta[\theta \varepsilon C(X) | X \varepsilon S] \leq 1-\alpha-\varepsilon \qquad \text{for all } \theta \quad . \tag{2.2}$$

If a negatively-biased betting procedure exists for a confidence set, we should be concerned about the statistical validity of asserting $1-\alpha$ confidence.

We now turn to the subject of our main concern, the conditional properties of confidence sets recentered at the ordinary and positive-part James-Stein Estimator,

$$\begin{aligned} C^{\delta^{JS}} &= \left\{\theta : |\theta - \delta^{JS}(X)| \leq c\right\}, \qquad & C^{\delta^+} &= \left\{\theta : |\theta - \delta^+(X)| \leq c\right\} \\ \delta^{JS}(X) &= \left(1 - \frac{a}{|X|^2}\right) X \qquad &, \quad \delta^+(X) &= \left(1 - \frac{a}{|X|^2}\right)^+ X \quad . \end{aligned} \tag{2.3}$$

Since both of these sets collapse to the usual confidence set as $|\theta| \to \infty$, it can be shown that

$$\lim_{|\theta|\to\infty} E_\theta\{[I(\theta \varepsilon C^\delta) - (1-\alpha)]s(X)\} = \lim_{|\theta|\to\infty} E_\theta\{[I(\theta \varepsilon C^o) - (1-\alpha)]s(X)\} \quad , \tag{2.4}$$

where, here, C^{δ} stands for either $C^{\delta^{JS}}$ or $C^{\delta^{+}}$. (Indeed, (2.4) holds for a confidence set recentered at any minimax estimator.) From Robinson (1979b), C^{o} has no super-relevant betting procedures, so the right-hand side of (2.4) must be zero. Therefore, it follows that there can be no super-relevant betting procedures for $C^{\delta^{JS}}$ or $C^{\delta^{+}}$ or, in general, any confidence set recentered at a minimax estimator.

As mentioned before, the absence of super-relevant betting procedures is an extremely weak conditional property: the real dividing line between good and bad conditional performance coming at relevant betting procedures. Therefore, we now investigate the question of existence of relevant betting procedures and, in particular, negatively-biased relevant betting procedures. We find that, as one might expect, $C^{\delta^{JS}}$ allows negatively-biased procedures.

The coverage probability of $C^{\delta^{JS}}(X)$ can be calculated by integrating over the region

$$\left\{|X|,\beta: |\theta|^2 - 2|\theta|\left(1 - \frac{a}{|X|^2}\right)|X|\cos\beta + \left(1 - \frac{a}{|X|^2}\right)^2 |X|^2 \le c^2\right\} ,$$

where $\cos\beta = \theta'X/|\theta||X|$. It is easy to see that this region is contained in

(2.5) $$\left\{|X|: \left[|\theta| - \left(1 - \frac{a}{|X|^2}\right)|X|\right]^2 \le c^2\right\} .$$

Consider the intersection of the region in (2.5) with $\{|X|:|X|^2 < h\}$, where $0 < h < a$. If $|\theta| > c$, this intersection is empty. If $|\theta| \le c$, the intersection is given by the region

(2.6) $$\{|X|: r_{-}(|\theta|) < |X| < h^{\frac{1}{2}}\},\ r_{-}(|\theta|) = \tfrac{1}{2}\{|\theta| - c + [(|\theta|-c)^2 + 4a]^{\frac{1}{2}}\} .$$

Since $r_{-}(|\theta|) \ge r_{-}(0) = \frac{1}{2}[-c + (c^2+4a)^{\frac{1}{2}}] > 0$ for $0 \le |\theta|$ c, the intersection is empty for all $|\theta|$ if $h^{\frac{1}{2}} < r_{-}(0)$. Therefore, the betting procedure $s(X) = -I_{[0,r_-(0)]}(|X|^2)$ is a negatively-biased relevant betting procedure for $C^{\delta^{JS}}(X)$; that is, there exists $\varepsilon > 0$ for which

$$(2.7) \qquad P_\theta[\theta \varepsilon C^{\delta^{JS}}(X) \,\big|\, |X|^2 \leq r\ (0)] < 1-\alpha-\varepsilon \qquad \text{for all } \theta \ .$$

The key defect in $C^{\delta^{JS}}$, that is exploited by the betting procedure, is the unboundedness of $\delta^{JS}(X)$ for X near 0. Since $\delta^+(X)$ does not suffer from this, one might hope that C^{δ^+} does not allow negatively biased betting procedures. That this is, in fact, true is the subject of the next section.

3. **Eliminating Negatively-Biased Betting.** A confidence set that admits a negatively-biased relevant betting procedure cannot be regarded as acceptable from a conditional viewpoint. The existence of such a betting procedure can be interpreted as saying that in some situations (i.e., for some subsets of the sample space) we are certain that the stated unconditional confidence level is not being attained. Therefore, we take as a minimal requirement for conditional acceptability that a confidence set should not admit negatively-biased relevant betting. This conditional criteria agrees with that of Bondar (1977) and Robinson (1976).

One way of guaranteeing that a confidence set does not allow negatively-biased betting to verify that it is a Bayes credible region against some (possibly improper) prior. More precisely, if X has density $f(x|\theta)$ and there is a distribution $\pi(\theta)$ resulting in a posterior density $\pi(\theta|x)$ for which the $1-\alpha$ frequentist confidence set $C(X)$ satisfies

$$P_x[\theta \varepsilon C(x)] = \int_{\theta \varepsilon C(x)} \pi(\theta|x)d\theta \geq 1-\alpha \qquad \text{for all } x \tag{3.1}$$

then if $0 \leq s(x) \leq 1$ and $m(x)$ is the marginal distribution of X, interchanging orders of integration gives

$$\int_{\Theta} E_{\theta}\{[I(\theta \varepsilon C(x)) - (1-\alpha)]s(X)\}\pi(\theta)d\theta$$

$$= \int_{\mathcal{X}} \int_{\Theta} [I(\theta \varepsilon C(x)) - (1-\alpha)]\pi(\theta|x)d\theta s(x)m(x)dx \geq 0 \quad . \tag{3.2}$$

It then follows that C has no negatively-biased semirelevant betting procedures, which implies that C has no negatively-biased relevant betting procedures.

The calculation in (3.2) is justified only if the interchange of integrals is justified. This is clearly the case of $\pi(\theta)$ (and hence $m(x)$) is a proper density. If $\pi(\theta)$ is not proper, the interchange may not be justified. Since good frequentist procedures often arise from improper priors, we will be especially concerned with this case, and must therefore pay more attention to the interchange of integrals in (3.2).

From Fubini's Theorem the calculation in (3.2) is justified if

$$\int_{X} |s(x)m(x)|dx < \infty \quad , \tag{3.3}$$

but since an improper $\pi(\theta)$ will lead to an improper $m(x)$, the inequality in (3.3) need not hold. However, for the case of a normal distribution, it is possible to get a relatively simple characterization of all relevant betting procedures against a class of recentered confidence sets. This characterization, given in the following lemma, will be helpful in verifying (3.3).

<u>Lemma 3.1</u>: Let $X \sim N_p(\theta, I)$ and let α and c satisfy $P_\theta(|X-\theta| \le c) = P(\chi_p^2 \le c^2) = 1-\alpha$. Let $\delta(X) = [1-\gamma(|X|)]X$, where $\gamma(|X|)$ satisfies

i) $0 \le \gamma(|X|) \le 1$

ii) $|\gamma(|X|)| \le K_1/|X|^r$ for $|X|^2 > K_0$, where K_0, K_1, and $r > 1$ are positive constants.

If $s(X)$ is a relevant betting procedure against the set estimator $\langle C^\delta(X), 1-\alpha\rangle$, where $C^\delta(X) = \{\theta: |\theta-\delta(X)| \le c\}$, then

$$\int_{X} |s(X)|dX < \infty \quad . \tag{3.4}$$

<u>Proof</u>: Given in the Appendix.

Lemma 3.1 gives us some flexibility in applying the operations in (3.2) to a particular confidence set. If the confidence set in question satisfies the conditions of the lemma, then we only need consider betting procedures with finite Lebesgue integrals, and consideration of improper priors becomes less troublesome. For example, an improper prior that leads to a bounded $m(x)$ will satisfy (3.3) for all relevant betting procedures, allowing the integrals in (3.2) to be interchanged.

Note that although the argument outlined at the beginning of this section led to the conclusion of no negatively-biased semirelevant betting procedures, the characterization in Lemma 3.1 only applies to relevant betting. Thus, the strongest conclusions we can hope to get will apply to relevant, but not semirelevant, betting procedures.

4. Conditionally Acceptable Confidence Sets.

We now apply the results of Section 3 to establish, in particular, that a confidence sphere recentered at the positive-part James-Stein estimator does not allow negatively-biased relevant betting procedures. We will use a slight modification of the priors used by Strawderman (1971):

$$
\begin{aligned}
& X|\theta \sim N(\theta, I) \\
(4.1)\qquad & \theta|\lambda \sim N[0, \lambda^{-1}(1-\lambda)I], \qquad 0 < \lambda < 1 \\
& \lambda \sim (1-a)\lambda^{-a}, \qquad 0 \le a < 1 .
\end{aligned}
$$

The specification in (4.1) gives a proper prior on θ. Our modification is to take a=2, giving θ an improper prior, but resulting in a proper posterior if $p \ge 3$,

$$\pi(\theta|X) = \frac{C(|X|)e^{-\frac{1}{2}|X-\theta|^2}}{|\theta|^{p-2}} \tag{4.2}$$

where

$$C^{-1}(|X|) = (2\pi)^{p/2} E\left(\frac{1}{\chi_p^2(|X|^2)}\right)^{(p/2)-1} ,$$

and $\chi_p^2(|X|^2)$ denotes a noncentral chi-squared random variable with p degrees of freedom and noncentrality parameter $|X|^2$. The marginal distribution of X is

$$m(X) = K\left(\frac{1}{|X|}\right)^{p-2} P\left(\chi_{p-2}^2 < |X|^2\right) , \tag{4.3}$$

which is a bounded function of X for $p \ge 3$. Thus, for betting procedures satisfying (3.4), the interchange of integrals in (3.2) is justified, and we can use this prior structure to identify conditionally acceptable confidence sets. That is, we want to see if there are recentered $1-\alpha$

frequentist confidence sets that are also 1-α Bayes credible sets using (4.2). Such sets would allow no negatively-biased betting.

Since we are using the posterior in (4.2), an obvious place to recenter our confidence set is at the mean of this posterior, which is

$$\delta^S(X) = [1 - h(|X|)]X \tag{4.4}$$
$$h(|X|) = \frac{p-2}{|X|^2} \frac{P(\chi^2_p < |X|^2)}{P(\chi^2_{p-2} < |X|^2)} .$$

Under squared error loss, this estimator is minimax but not admissible. The prior in (4.1) gives rise to admissible estimators only if $p \geq 3$ and $a < 2$ (Berger, 1976).

Since $\delta^S(X)$ can be somewhat difficult to calculate, there is some interest in looking at the performance of recentered confidence sets using other, easier to calculate estimators. In particular, since $h(0) = (p-2)/p$, we can approximate $\delta^S(X)$ reasonably with

$$\delta^{SA}(X) = [1-h_A(|X|)]X \tag{4.5}$$
$$h_A(|X|) = \begin{cases} \frac{p-2}{p} & \text{if } |X|^2 \leq p \\ \frac{p-2}{|X|^2} & \text{if } |X|^2 > p \end{cases} .$$

We will also consider the performance of confidence sets recentered at $\delta^+(X)$, the positive-part estimator.

Let C_{δ^S}, $C_{\delta^{SA}}$, and C_{δ^+} denote the confidence sphere of radius c (where $P(\chi^2_p > c^2) = \alpha$) centered at δ^S, δ^{SA}, and δ^+, respectively. For these confidence sets we want to establish (using C generically)

(4.6)

i) $P_\theta(\theta \varepsilon C) \geq 1-\alpha \;\forall\theta$, i.e., C is a 1-$\alpha$ frequentist confidence set.

ii) $P_X(\theta \varepsilon C) \geq 1-\alpha \;\forall X$, i.e., C is a 1-$\alpha$ Bayes credible set for the posterior in (4.2).

Since the three confidence sets in question satisfy the conditions of Lemma 3.1, and since m(X) is bounded, it will then follow from (4.6) that none of these confidence sets allow negatively-biased relevant betting.

That (4.6) is satisfied for these sets will mainly be verified numerically. Results of Hwang and Casella (1982,1984) show that $C_{\delta^{SA}}$ and C_{δ^+} will satisfy (4.6)i if the constant p-2 is replaced by a slightly smaller value (approximately .8(p-2)). Numerical studies, given in the aforementioned references and also in Casella and Hwang (1983) strongly support the claim that (4.6)i is in fact true for δ^{SA} and δ^+. The case of δ^S is much more complicated, owing to the more complex form of the estimator. Thus, no analytical proof of (4.6)i for δ^S will be attempted here. Figure 1 is a graph of the coverage probabilities of C_{δ^S} for $1-\alpha = .9$ and p = 3,5,7,11,15, calculated using numerical integration. The figure shows that C_{δ^S} maintains $1-\alpha$ confidence in this case. Other cases of α and p were also examined (but are not reported here), and in all cases C_{δ^S} was found to be a $1-\alpha$ confidence set.

The verification that these confidence sets are also $1-\alpha$ Bayes credible sets is also, unfortunately, quite intractable analytically. Again, numerical integration was used to verify these facts. Figures 2-4 give graphs of credible probabilities for C_{δ^S}, $C_{\delta^{SA}}$, and C_{δ^+}, respectively. Again, we only report the case $1-\alpha = .9$ and p = 3,5,7,11,15, although similar results were obtained for other cases. The figures show that these confidence sets are maintaining their $1-\alpha$ credible probability against $\pi(\theta|X)$ of (4.2). (It is interesting to note the dips in Figures 3 and 4, with the more pronounced dip in Figure 4. These dips occur at the join points (points of nondifferentiability) of the estimators δ^{SA} and δ^+).

Two analytical calculations of interest that can be done with $\pi(\theta|X)$ are evaluations at $|X| = 0$ and $|X| = \infty$. It is a straightforward exercise to verify that the distribution of $|\theta|$ at $|X| = 0$ is given by

$$|\theta| \Big| |X| = 0 \sim \chi_2^2 \quad ,$$

independent of p. Since the three estimators under consideration are all 0 at $|X| = 0$, the credible probability of all three confidence sets is $P(\chi_2^2 < c^2)$ at $|X| = 0$. Since c^2 is chosen to satisfy $P(\chi_p^2 < c^2) = 1-\alpha$, it follows that if $p \geq 3$, $P(\chi_2^2 \leq c^2) > 1-\alpha$. As $|X| \to \infty$ all the estimator collapse to the usual confidence set and the credible probabilities all approach $1-\alpha$.

Acknowledgement: The author thanks the referees for valuable comments on an earlier version of this paper.

Appendix: Proof of Lemma 3.1

Suppose s(X) is relevant for $\langle C^\delta(X), 1-\alpha\rangle$. Then there exists $\varepsilon > 0$ such that

$$E_\theta\{[I(\theta \varepsilon C^\delta(X)) - (1-\alpha)]s(X)\} \geq \varepsilon E_\theta|s(X)| \quad \forall\theta \quad . \tag{A.1}$$

Multiply both sides of (A.1) by $\pi_b(\theta)$, a $N\left[0,\left(b^{-1}-1\right)I\right]$ density $(0 < b < 1)$ and integrate over all θ. It follows from (A.1) that

$$\int_\Theta E_\theta\{I[\theta\varepsilon C^\delta(X)] - (1-\alpha)\}s(X)\pi_b(\theta)d\theta > \varepsilon \int_\Theta E_\theta|s(X)|\pi_b(\theta)d\theta \tag{A.2}$$

for $0 < b < 1$. The proof will proceed by showing that for sufficiently small b, the inequality in (A.2) can be violated if

$$\int_{\mathcal{X}} |s(X)|dX = \infty \quad . \tag{A.3}$$

Since $\pi_b(\theta)$ is a proper density and s(X) is bounded, the order of integration in the left-hand side of (A.2) can be reverse, yielding

$$\int_{\mathcal{X}} \left[\int_{\theta\varepsilon C^\delta(X)} \pi_b(\theta|X)d\theta - (1-\alpha)\right] s(X)m_b(X)dX \quad , \tag{A.4}$$

where $\pi_b(\theta|X)$ is the conditional density of θ given X, $N[(1-b)X,(1-b)I]$, and $m_b(x)$ is the marginal density of X, $N(0,b^{-1}I)$.

Since s(X) is not a function of θ, we have

$$\begin{aligned} \int_\Theta E_\theta|s(X)|\pi_b(\theta)d\theta &= \int_{\mathcal{X}} |s(X)|m_b(X)dX \\ &= \frac{b^{p/2}}{(2\pi)^{p/2}} \int_{\mathcal{X}} |s(X)|e^{-\frac{1}{2}b|X|^2}dX \quad . \end{aligned} \tag{A.5}$$

As $b \to 0$, $e^{-\frac{1}{2}|X|^2}$ increases, so from the monotone convergence theorem

$$\text{(A.6)} \qquad \lim_{b\to 0} \int_{\mathcal{X}} |s(X)| e^{-\frac{1}{2}b|X|^2} dX = \int_{\mathcal{X}} |s(X)| dX \quad .$$

Now consider (A.4). The integration over X will be split into three pieces:

$$\text{(A.7)} \qquad \begin{aligned} W_1 &= \{X : |X| < K\} \quad , \\ W_2 &= \{X : K \le |X| \le b^{-q}\} \quad , \\ W_3 &= \{X : |X| > b^{-q}\} \quad , \end{aligned}$$

where K and $q > \frac{1}{2}$ are constants. The exact method of choosing K will be detailed later. (K will depend only on ε.)

Since $s(X)$ is bounded (without loss of generality) by $|s(X)| \le 1$ we have

$$\left| \int_{W_1} \left[\int_{\theta \varepsilon C^{\delta}(X)} \pi_b(\theta|X) d\theta - (1-\alpha) \right] s(X) m_b(X) dX \right|$$

$$\le \int_{W_1} m_b(X) dX = P_b(|X| < K) = P(\chi_p^2 < K^2 b) \quad ,$$

since $b|X|^2 \sim \chi_p^2$. It is straightforward to verify that

$$\lim_{b\to\infty} \frac{P(\chi_p^2 < K^2 b)}{b^{p/2}} < \infty \quad ,$$

and it then follows from (A.3), (A.5), and (A.6) that for sufficiently small b

$$\text{(A.8)} \qquad \int_{W_1} \left[\int_{\theta \varepsilon C^{\delta}(X)} \pi_b(\theta|X) d\theta - (1-\alpha) \right] s(X) m_b(X) dX < (\varepsilon/3) \int_{\mathcal{X}} |s(X)| m_b(X) dX \quad .$$

A similar argument will show that the integral over the region W_3 can be bounded by $P_b(|X| > b^{-q}) = P(\chi_p^2 > b^{1-2q})$. For $q > \frac{1}{2}$, $\lim P(\chi_p^2 > b^{1-2q})/b^{p/2} < \infty$, so the integral over W_3 also satisfies (A.8) for sufficiently small b.

It remains to establish that the integral over W_2 will also satisfy (A.8). Straightforward transformations will establish that

$$\int_{\theta\epsilon C^{\delta}(X)} \pi_b(\theta|X)d\theta = [2\pi(1-b)]^{-\frac{1}{2}} \int_{|\delta(X)|-c}^{|\delta(X)|+c} P\left(\chi_{p-1}^2 \leq \frac{1}{1-b}\{c^2 - [y-|\delta(X)|]^2\}\right)$$

$$\times \exp\{-\tfrac{1}{2}[y - (1-b)|X|]^2/(1-b)\}dy \quad .$$

Make the further transformation $t = [y - (1-b)|X|]/(1-b)^{\frac{1}{2}}$ to obtain

$$\int_{\theta\epsilon C^{\delta}(X)} \pi_b(\theta|X)d\theta$$

(A.9)

$$= (2\pi)^{-\frac{1}{2}}\int_T P\left[\chi_{p-1}^2 \leq \frac{1}{1-b}\left(c^2 - \{(1-b)^{\frac{1}{2}}t - [b - \gamma(|X|)]|X|\}^2\right)\right] e^{-\frac{1}{2}t^2}dt \ ,$$

where $T = \left(t:\{-c+[b-\gamma(|X|)]|X|\}/(1-b)^{\frac{1}{2}} \leq t < \{c+[b-\gamma(|X|)]|X|\}/(1-b)^{\frac{1}{2}}\right)$. (Recall that $|\delta(X)| = |X| - \gamma(|X|)|X|)$. For $K < |X| < b^{-q}$ we have

$$\Big|[b-\gamma(|X|)]|X|\Big| \leq b(X| + \gamma(|X|)|X|$$
$$\leq b^{1-q} + \frac{K_0}{|X|} \leq b^{1-q} + \frac{K_0}{K}$$

where the second inequality follows from assumption ii) on $\gamma(|X|)$. Since

$$(2\pi)^{-\frac{1}{2}}\int_{-c}^{c} P\left(\chi_{p-1}^2 \leq c^2 - t^2\right) e^{-\frac{1}{2}t^2}dt = 1-\alpha \quad ,$$

it follows from (A.9) and (A.10) that we can choose b sufficiently small and K sufficiently large so that

$$\left| \int_{\theta \varepsilon C^{\delta}(X)} \pi_b(\theta|X)d\theta - (1-\alpha) \right| < \varepsilon/3 \quad .$$

Therefore, for the integral over W_2 we have

$$\int_{W_2} \left[\int_{\theta \varepsilon C^{\delta}(X)} \pi_b(\theta|X)d\theta - (1-\alpha) \right] s(X)m_b(X)dX \le (\varepsilon/3)\int_{W_2} s(X)m_b(X)dX$$

$$\le (\varepsilon/3)\int_{\mathcal{X}} |s(X)|m_b(X)dX \quad ,$$

contradicting (A.2) and completing the proof. ||

REFERENCES

BERGER, J. O. (1976). Admissible minimax estimation of a multivariate normal mean with arbitrary quadratic loss. *Ann. Statist.* **4** 223-226.

BONDAR, J. V. (1977). A conditional confidence principle. *Ann. Statist.* **5** 881-891.

BUEHLER, R. J. (1959). Some validity criteria for statistical inferences. *Ann. Math. Statist.* **30** 866-877.

CASELLA, G. and HWANG, J. T. (1983). Empirical Bayes confidence sets for the mean of a multivariate normal distribution. *J. Amer. Statist. Assoc. — Theory and Methods Section* **78** 688-698.

FISHER, R. A. (1956). *Statistical methods and scientific inference*, Oliver and Boyd, Edinburgh.

HWANG, J. T. and CASELLA, G. (1982). Minimax confidence sets for the mean of a multivariate normal distribution. *Ann. Statist.* **10** 868-881.

HWANG, J. T. and CASELLA, G. (1984). Improved set estimators for a multivariate normal mean. *Statistics and Decisions*, Supplemental Issue No. **1** 3-16.

ROBINSON, G. K. (1976). Properties of Student's t and of the Behrens-Fisher solution to the two means problem. *Ann. Statist.* **4** 963-971.

ROBINSON, G. K. (1979a). Conditional properties of statistical procedures. *Ann. Statist.* **7** 742-755.

ROBINSON, G. K. (1979b). Conditional properties of statistical procedures for location and scale parameters. *Ann. Statist.* **7** 756-771.

STRAWDERMAN, W. E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Statist.* **42** 385-388.

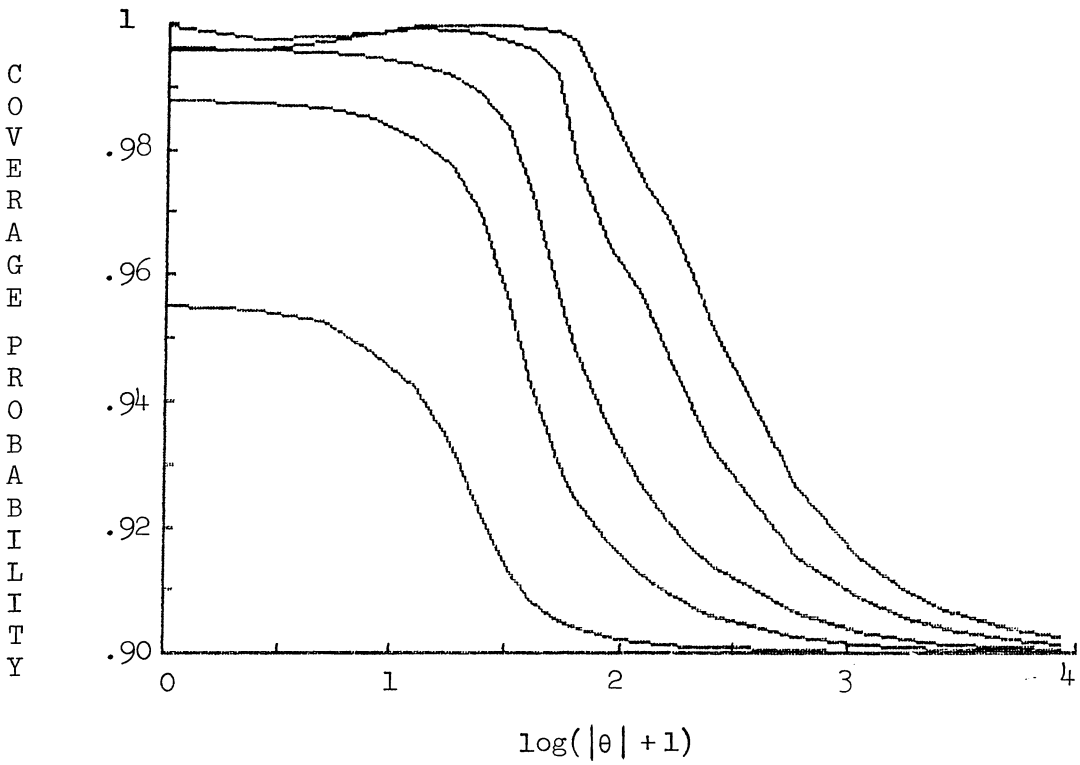

Figure 1: Coverage probabilities for $C_{\delta^{S}}$, $1-\alpha = .9$. The probabilities are increasing in p(dimension), and are shown for $p = 3,5,7,11,15$.

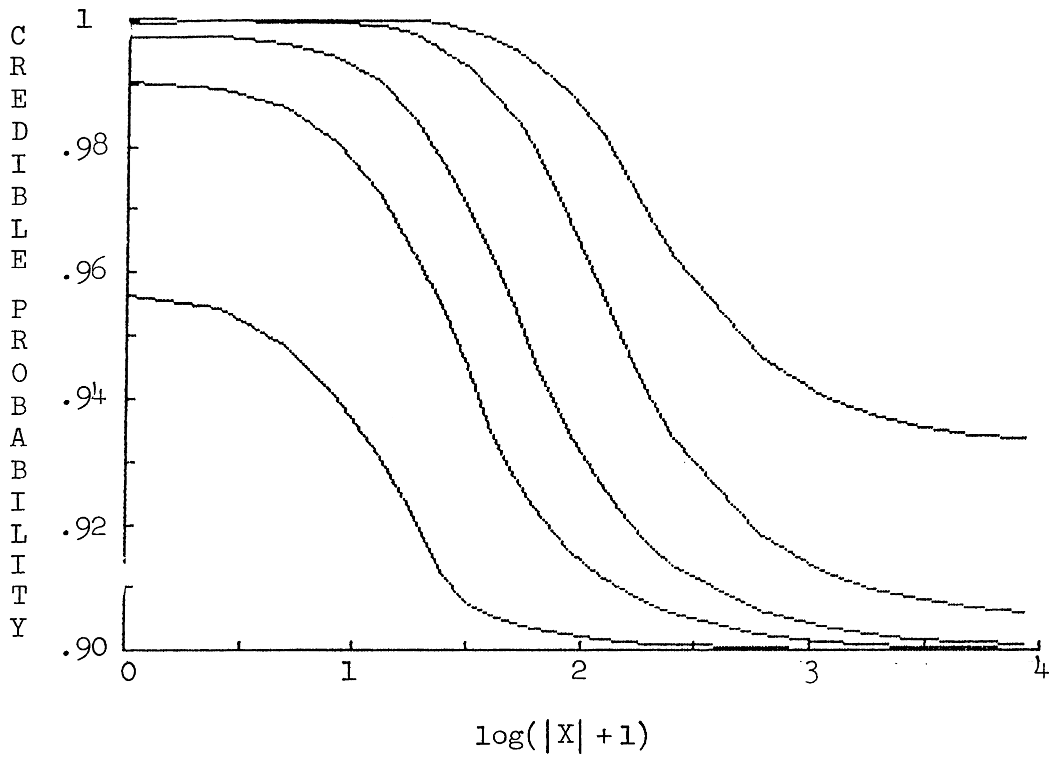

Figure 2. Credible probabilities for C_{δ^S}, $1-\alpha = .9$. The probabilities are increasing in p(dimension), and are shown for p = 3,5,7,11,15.

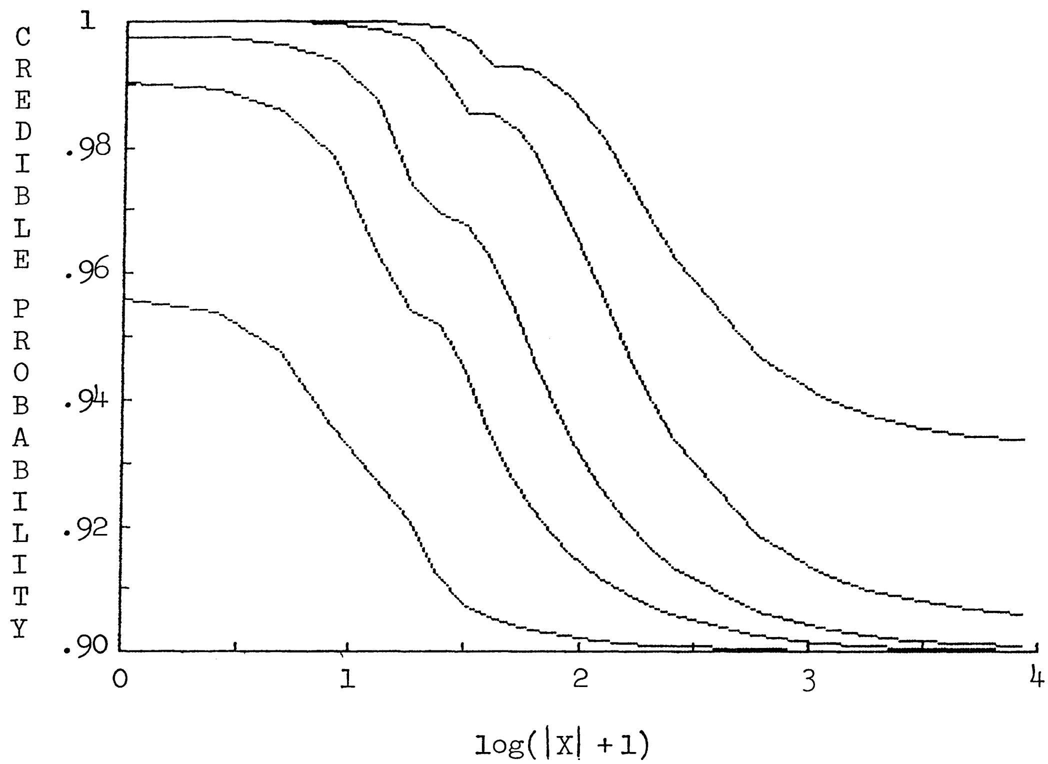

Figure 3. Credible probabilities for $C_{\delta^{SA}}$, $1-\alpha = .9$. The probabilities are increasing in p(dimension), and are shown for p = 3,5,7,11,15.

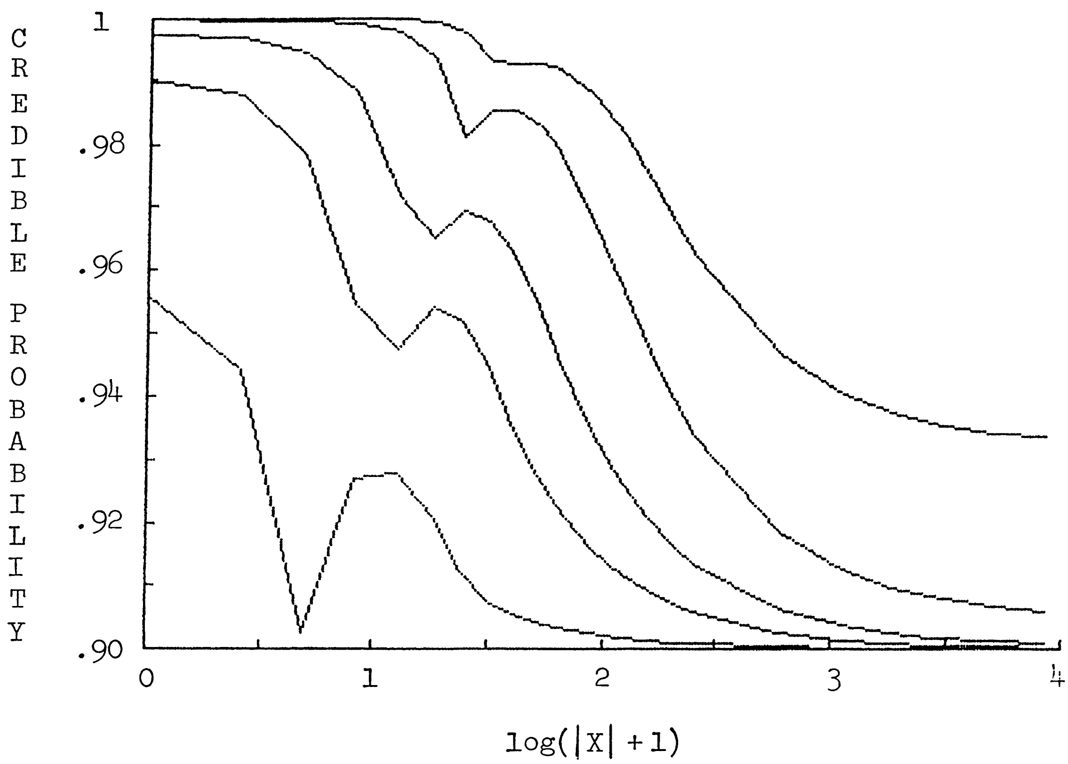

Figure 4. Credible probabilities for C_{δ^+}, $1-\alpha = .9$. The probabilities are increasing in p(dimension), and are shown for p = 3,5,7,11,15.