The Annals of Statistics
1987, Vol. 15, No. 4, 1363–1371

CONDITIONALLY ACCEPTABLE RECENTERED SET ESTIMATORS[1]

BY GEORGE CASELLA

Cornell University

The usual confidence sphere for a multivariate normal mean can be uniformly improved upon, in terms of coverage probability, by recentering it at a Stein-type 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.

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

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

where $P_{\theta}[\theta \in C(X)]$ 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

$$P_{\theta}(\theta \in C(X)|X \in S) > 1 - \alpha + \varepsilon \quad \text{for all } \theta, \tag{1.2i}$$

or

$$P_{\theta}(\theta \in C(X)|X \in S) < 1 - \alpha - \varepsilon \quad \text{for all } \theta, \tag{1.2ii}$$

for some $\varepsilon > 0$, then 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 $C(X)$. If $C(X)$ satisfies (1.2i), for example, and the betting strategy "bet for coverage if $X \in S$" 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 $C(X)$.

Received October 1984; revised March 1987.
[1]Research supported by National Science Foundation grant DMS-85-01973.
AMS 1980 *subject classifications.* Primary 62C99; secondary 62F10.
Key words and phrases. Betting procedures, confidence sets, multivariate normal.

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

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

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

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

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 question of whether these recentered sets maintain good conditional properties has not been settled, however, and is the main concern of this paper.

Buehler's (1959) 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 1.1. For the confidence set $\langle C(X), \beta(X)\rangle$, the betting procedure $s(X)$ is *relevant* if, for some $\varepsilon > 0$,

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

with strict inequality for some θ. If $\varepsilon = 0$, then $s(X)$ is *semirelevant*.

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

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

DEFINITION 1.2. A betting procedure $s(X)$ is *negatively biased* if $-1 \le s(X) \le 0$ for all X and *positively biased* if $0 \le s(X) \le 1$ 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 $X \in S$], then the second inequality in (1.2) will obtain. If a negatively biased betting procedure exists for a confidence set, we should be concerned about the statistical validity of asserting $1 - \alpha$ confidence.

In Section 2, we outline the method of constructing confidence sets with acceptable conditional properties. The construction in Section 2 depends on a technical lemma, Lemma 2.1, whose proof is given in the Appendix. In Section 3, we use the results of Section 2 to exhibit a recentered confidence set with acceptable conditional properties.

2. Eliminating negatively biased betting. The existence of a negatively biased relevant betting procedure can be interpreted as saying that for some subsets of the sample space we are certain that the stated unconditional confidence level is not being attained. Therefore, we take the nonexistence of negatively biased relevant betting to be a minimal requirement for conditional acceptability of a confidence set. This conditional criterion agrees with that of Bondar (1977) and Robinson (1976).

One way of guaranteeing that a confidence set does not allow negatively biased betting is 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

$$P_x[\theta \in C(x)] = \int_{\theta \in C(x)} \pi(\theta|x)\, d\theta \geq 1 - \alpha \quad \text{for all } x, \tag{2.1}$$

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

$$\begin{aligned} &\int_\Theta E_\theta\{[I(\theta \in C(X)) - (1-\alpha)]s(X)\}\pi(\theta)\, d\theta \\ &\quad = \int_X \int_\Theta [I(\theta \in C(x)) - (1-\alpha)]\pi(\theta|x)\, d\theta\, s(x) m(x)\, dx \geq 0. \end{aligned} \tag{2.2}$$

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 (2.2) is justified only if the interchange of integrals is justified. This is clearly the case if $\pi(\theta)$ [and hence $m(x)$] is a proper density, but 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 (2.2).

From Fubini's theorem, the calculation in (2.2) is justified if

$$\int_X |s(x) m(x)|\, dx < \infty, \tag{2.3}$$

but since an improper $\pi(\theta)$ will lead to an improper $m(x)$, the inequality in (2.3) 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 (2.3).

LEMMA 2.1. *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) $\gamma(|X|)|X| \to 0$ *as* $|X| \to \infty$.

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

$$\int_X |s(x)|\, dx < \infty. \tag{2.4}$$

PROOF. Given in the Appendix.

Lemma 2.1 gives us some flexibility in applying the operations in (2.2) 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 (2.3) for all relevant betting procedures, allowing the integrals in (2.2) 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 2.1 only applies to relevant betting. Thus, the strongest conclusions we can hope to get will apply to relevant, but not semirelevant, betting procedures.

3. Conditionally acceptable confidence sets. 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,

$$C_{\delta^{\mathrm{JS}}} = \{\theta\colon |\theta - \delta^{\mathrm{JS}}(X)| \le c\}, \qquad C_{\delta^+} = \{\theta\colon |\theta - \delta^{+}(X)| \le c\},$$
$$\delta^{\mathrm{JS}}(X) = \left(1 - \frac{a}{|X|^2}\right)X, \qquad \delta^{+}(X) = \left(1 - \frac{a}{|X|^2}\right)^{+} X.$$

As might be expected, the confidence set $C_{\delta^{\mathrm{JS}}}$ is relatively easy to dismiss. Since $\delta^{\mathrm{JS}}(X)$ is unbounded near zero, a set of the form $\{X\colon |X| < k\}$ will provide a negatively biased relevant betting procedure against $\langle C_{\delta^{\mathrm{JS}}}, 1-\alpha\rangle$. Since $\delta^+(X)$ does not suffer from this problem, one might hope that $\langle C_{\delta^+}, 1-\alpha\rangle$ does not allow negatively biased betting procedures. The results of Section 2 can be used to establish that $\langle C_{\delta^+}, 1-\alpha\rangle$ does not allow negatively biased relevant betting procedures. We will use a slight modification of the priors used by Strawderman

(1971):

$$\begin{aligned} X|\theta &\sim N(\theta, I), \\ \theta|\lambda &\sim N\left[0, \lambda^{-1}(1-\lambda)I\right], \qquad 0 < \lambda < 1, \\ \lambda &\sim \lambda^{-2}. \end{aligned} \tag{3.1}$$

The specification in (3.1) gives θ an improper prior, but results in a proper posterior if $p \geq 3$. Furthermore, the marginal density function of X is a bounded function of X for $p \geq 3$. Thus, for betting procedures satisfying (2.4), the interchange of integrals in (2.2) 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 $1 - \alpha$ Bayes credible sets using the prior in (3.1). Such sets would allow no negatively biased betting.

Since we are using the posterior in (3.2), an obvious place to recenter our confidence set is at the mean of this posterior. We are mainly concerned, however, with the conditional behavior of $\langle C_{\delta^+}, 1 - \alpha \rangle$, so attention will be confined to this set. The estimator $\delta^+(X)$ is, for large $|X|$, quite close to the posterior mean, so there is some reason to expect C_{δ^+} to perform reasonably against this prior. For the confidence set $\langle C_{\delta^+}, 1 - \alpha \rangle$, we want to establish:

$$P_\theta(\theta \in C_{\delta^+}) \geq 1 - \alpha \qquad \forall\, \theta, \tag{3.2i}$$

i.e., C_{δ^+} is a $1 - \alpha$ frequentist confidence set, and

$$P_x(\theta \in C_{\delta^+}) \geq 1 - \alpha \qquad \forall\, x, \tag{3.2ii}$$

i.e., C_{δ^+} is a $1 - \alpha$ Bayes credible set for the prior in (3.1). Since $\langle C_{\delta^+}, 1 - \alpha \rangle$ satisfies the conditions of Lemma 2.1 and since $m(x)$ is bounded, it will then follow from (3.2) that there can be no negatively biased relevant betting.

That (3.2) is satisfied will mainly be verified numerically. Results of Hwang and Casella (1982, 1984) show C_{δ^+} will satisfy (3.2i) if the constant $p - 2$ is replaced by a slightly smaller value [approximately $0.8(p - 2)$]. Numerical studies, given in the aforementioned references and also in Casella and Hwang (1983) strongly support the claim that (3.2i) is in fact true for C_{δ^+}. The verification that $\langle C_{\delta^+}, 1 - \alpha \rangle$ is also $1 - \alpha$ Bayes credible sets is also, unfortunately, quite intractable analytically. Again, numerical integration was used to verify this fact. Figure 1 gives graphs of credible probabilities for C_{δ^+} for the case $1 - \alpha = 0.9$ and $p = 3, 5, 7, 11, 15$. (Similar results were obtained for other cases.) The figure shows that $\langle C_{\delta^+}, 1 - \alpha \rangle$ is maintaining its $1 - \alpha$ credible probability. [It is interesting to note the dips in Figure 1. These dips occur at the join points (points of nondifferentiability) of the estimator δ^+.]

Two analytical calculations of interest that can be done with the prior of (3.1), $\pi(\theta|X)$, 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

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

independent of p. Since $\delta^+(X) = 0$ at $|X| = 0$, the credible probability of $\langle C_{\delta^+}, 1 - \alpha \rangle$ is $P(\chi_2^2 < c^2)$ at $|X| = 0$. Since c^2 is chosen to satisfy $p(\chi_p^2 < c^2) =$

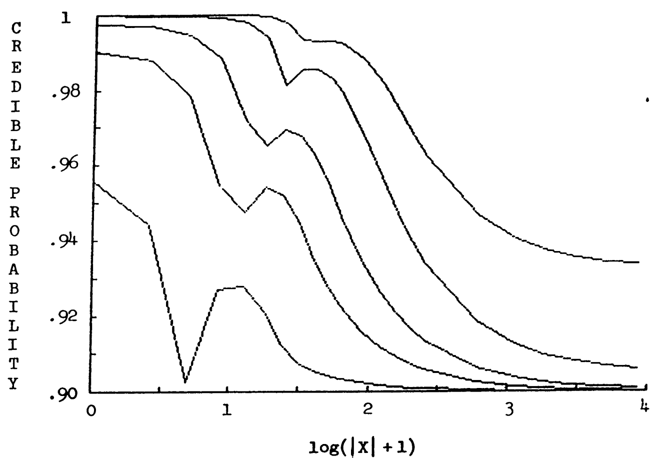

FIG. 1. *Credible probabilities for* $C_{\delta^+}, 1 - \alpha = 0.9$. *The probabilities are increasing in* p *(dimension) and are shown for* $p = 3, 5, 7, 11, 15$.

$1 - \alpha$, it follows that if $p \geq 3$, $P(\chi_2^2 \leq c^2) > 1 - \alpha$. As $|X| \to \infty$, $\langle C_{\delta^+}, 1 - \alpha\rangle$ collapses to the usual confidence set and its credible probabilities all approach $1 - \alpha$.

APPENDIX

PROOF OF LEMMA 2.1. Suppose $s(X)$ is relevant for $\langle C_\delta(X), 1 - \alpha\rangle$. Then there exists $\varepsilon > 0$ such that

$$E_\theta\{[I(\theta \in C_\delta(X)) - (1 - \alpha)]s(X)\} \geq \varepsilon E_\theta|s(X)| \qquad \forall\, \theta. \tag{A.1}$$

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

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

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

$$\int_X |s(x)|\, dx = \infty. \tag{A.3}$$

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 reversed, yielding

$$\int_X \left[\int_{\theta \in C_\delta(X)} \pi_b(\theta|x)\, d\theta - (1 - \alpha)\right] s(x) m_b(x)\, dx, \tag{A.4}$$

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

$$\int E_\theta |s(X)| \pi_b(\theta)\, d\theta = \frac{b^{p/2}}{(2\pi)^{p/2}} \int_X |s(x)| e^{-b/2|x|^2}\, dx. \tag{A.5}$$

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

$$\lim_{b\to 0} \int_X |s(x)| e^{-b/2|x|^2}\, dx = \int_X |s(x)|\, dx. \tag{A.6}$$

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

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

where K and q, $\frac{1}{2} < q < 1$, 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

$$\left| \int_{W_1} \left[\int_{\theta \in C_\delta(x)} \pi_b(\theta|x)\, d\theta - (1-\alpha) \right] s(x) m_b(x)\, dx \right|$$

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

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

$$\lim_{b\to 0} \frac{P(\chi_p^2 < K^2 b)}{b^{p/2}} < \infty,$$

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

$$\begin{aligned} &\int_{W_1} \left[\int_{\theta \in C_\delta(x)} \pi_b(\theta|x)\, d\theta - (1-\alpha) \right] s(x) m_b(x)\, dx \\ &\qquad < (\varepsilon/3) \int_X |s(x)| m_b(x)\, dx. \end{aligned} \tag{A.8}$$

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 $q > \frac{1}{2}$,

$$\lim_{b\to 0} P(\chi_p^2 > b^{1-2q}) / b^{p/2} < \infty,$$

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). Make the transformation $y = |\theta| \cos\beta$ and $Z = |\theta| \sin\beta$, where $\cos\beta = \theta' X / |\theta|\, |X|$.

Note that $Z^2 \sim \chi^2_{p-1}$, independent of y. We have

$$\int_{\theta \in C_\delta(X)} \pi_b(\theta|X)\, d\theta$$

$$= [2\pi(1-b)]^{-1/2} \int_{|\delta(X)|-c}^{|\delta(X)|+c} P\left(\chi^2_{p-1} \le \frac{1}{1-b}\{c^2 - [y - |\delta(X)|]^2\}\right)$$

$$\times \exp\left\{-\frac{1}{2}\frac{[y-(1-b)|X|]^2}{(1-b)}\right\} dy.$$

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

$$\int_{\theta \in C_\delta(X)} \pi_b(\theta|X)\, d\theta$$

$$\text{(A.9)} \quad = (2\pi)^{-1/2} \int_T P\left[\chi^2_{p-1} \le \frac{1}{1-b}\left(c^2 - \left\{(1-b)^{1/2}t - [b-\gamma(|X|)]|X|\right\}^2\right)\right] e^{-t^2/2}\, dt,$$

where $T = \{t\colon \left|(1-b)^{1/2}t - [b-\gamma(|X|)]|X|\,\right| \le c\}$. For $K < |X| < b^{-q}$, we have for some K_0,

$$\begin{aligned} |[b-\gamma(|X|)]|X|\,| &\le b|X| + \gamma(|X|)|X| \\ &\le b^{1-q} + K_0, \end{aligned}$$

where the second inequality follows from assumption (ii) (of Lemma 2.1) on $\gamma(|X|)$. Note also that, since $\lim_{|X|\to\infty}|X|\gamma(|X|) = 0$, K_0 can be made as small as we like by choosing K sufficiently large; moreover, $b^{1-q} \to 0$ as $b \to 0$. Since

$$\text{(A.10)} \qquad (2\pi)^{-1/2} \int_{-c} P\left(\chi^2_{p-1} \le c^2 - t^2\right) e^{-t^2/2}\, dt = 1 - \alpha,$$

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

$$\left| \int_{\theta \in C_\delta(X)} \pi_b(\theta|X)\, d\theta - (1-\alpha) \right| < \varepsilon/3.$$

Therefore, for the integral over W_2, we have

$$\begin{aligned} &\int_{W_2} \left[\int_{\theta \in C_\delta(x)} \pi_b(\theta|x)\, d\theta - (1-\alpha)\right] s(x) m_b(x)\, dx \\ &\quad \le (\varepsilon/3) \int_X |s(x)| m_b(x)\, dx, \end{aligned}$$

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

Acknowledgments. The author thanks Jim Bondar for valuable comments on an earlier version of this paper and the referees for suggestions that led to an improved presentation.

REFERENCES

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** 845–863.

CASELLA, G. and HWANG, J. T. (1983). Empirical Bayes confidence sets for the mean of a multivariate normal distribution. *J. Amer. Statist. Assoc.* **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. *Statist. Decisions* (*Suppl.*) **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.

DEPARTMENT OF PLANT BREEDING
AND BIOMETRY
BIOMETRICS UNIT
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853-7801