Combining Independent Normal Mean Estimation Problems with Unknown Variances

by

James Berger and M. E. Bock
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #404

January 1975

Summary

Let $X = (X_1,\ldots X_p)^t$ be a p-variate normal random vector with unknown mean $\theta = (\theta_1,\ldots\theta_p)^t$ and unknown positive definite diagonal covariance matrix A. Assume that estimates V_i of the variances A_i are available, and that V_i/A_i is $\chi^2_{n_i}$. Assume also that all X_i and V_i are independent. It is desired to estimate θ under the quadratic loss

$$[\sum_{i=1}^{p} q_i(\delta_i-\theta_i)^2]/[\sum_{i=1}^{p} q_iA_i], \text{ where } q_i > 0,\ i=1,\ldots p.$$

Defining $W_i = V_i/(n_i-2)$, $W = (W_1,\ldots W_p)^t$, and $||\ X\ ||_W^2 = \sum_{j=1}^{p} [X_j^2/(q_jW_j^2)]$, it is shown that under certain conditions on r(X,W), the estimator given componentwise by

$$\delta_i(X,W) = (1-r(X,W)/[||\ X\ ||_W^2\ q_iW_i])X_i$$

is a minimax estimator of θ. (The conditions on r require $p \geq 3$.) Good practical versions of this estimator are developed and examples of their use are given.

1. Introduction.

Let $X=(X_1,\ldots X_p)^t$ be a p-variate normal random vector with unknown mean $\theta=(\theta_1,\ldots\theta_p)^t$ and positive definite covariance matrix Σ. Consider the problem of estimating θ, when the loss incurred in estimating θ by $\delta=(\delta_1,\ldots\delta_p)^t$ is the quadratic loss

$$L(\delta,\theta,\Sigma) = (\delta-\theta)^t Q(\delta-\theta)/\mathrm{tr}(Q\Sigma).$$

Here Q is a $p \times p$ positive definite matrix and "tr" denotes the trace. Note that $\mathrm{tr}(Q\Sigma)$ is just a normalizing constant.

The above problem has been of considerable interest since Stein (1955) demonstrated that if $Q = \Sigma = I$ (the $p \times p$ identity matrix) and if $p \geq 3$, then the usual estimator $\delta_0(X)=X$ is inadmissible for estimating θ. Indeed he found minimax estimators which significantly improved upon the risk of δ_0. The generalization of these results to arbitrary Q and Σ was of obvious interest. For the case of known Σ, wide classes of minimax estimators have now been developed. (See Bhattacharya (1966), Hudson (1974), Berger (1974a), Bock (1975), and Berger (1975).) For unknown Σ, however, the results that have been obtained are very incomplete. For the special case $Q = \Sigma^{-1}$, James and Stein (1960) did obtain good minimax estimators better than δ_0. Bhattacharya (1966) obtained results for the situation $\Sigma = \sigma^2 B$, where B is a known $p \times p$ matrix and σ^2 is unknown. The subsequent literature considering unknown Σ has dealt with one or the other of the above two special situations.

American Mathematical Society 1970 subject classification: Primary 62C99; Secondary 62F10, 62H99.

Key words and phrases: Minimax, independent, unknown variances, quadratic loss, normal mean.

In this paper, a first step is made in dealing with arbitrary Q and unknown $\ddagger$. Results are obtained under the assumptions that the X_i are independent and that Q is diagonal with diagonal elements $q_i > 0$, $i=1,\ldots p$. Since the X_i are independent, it is clear that $\ddagger = A$, where A is an unknown $p \times p$ diagonal matrix with diagonal elements $A_i > 0$. It will be assumed that estimates V_i for A_i are available, where V_i/A_i has a chi-square distribution with n_i degrees of freedom. It will also be assumed that $n_i \geq 3$, that all V_i are independent of V_j for $i \neq j$, and that the V_i are independent of the X_j.

Throughout the paper, E[ ] will stand for the expectation of the argument. Subscripts on E (usually θ or A) will denote parameter values under which the expectation is taken. Superscripts on E will be used to clarify the random variable with respect to which the expectation is being taken. When obvious, no subscripts or superscripts will be given.

For notational convenience, let W be the $p \times p$ diagonal matrix with diagonal elements $W_i = V_i/(n_i-2)$. Define

$$|| \, X \, ||_W^2 = X^t W^{-1} Q^{-1} W^{-1} X = \sum_{i=1}^{p} [X_i^2/(q_i W_i^2)].$$

Let $|x|$ denote the usual Euclidean norm of x. Finally, let $\chi^2_{n_i}$, $i=1,\ldots p$, denote independent chi-square random variables with n_i degrees of freedom, and define

$$T = \min_{1 \leq i \leq p} [\chi^2_{n_i}/n_i], \text{ and } \tau = \tau(n_1,\ldots n_p) = E[T^{-1}].$$

In Section 2, it is shown that under certain conditions, estimators of the form

(1.1) $$\delta(X,W) = (I - r(X,W) \, || \, X \, ||_W^{-2} \, Q^{-1} W^{-1}) X$$

are minimax and have risks smaller than 1 (the risk of δ_0). Thus, in combining p independent normal mean estimation problems with unknown variances it is often possible to improve upon the risk of the usual estimator.

In Section 3, a simple practically useful version of the above estimator is developed. Section 4 deals with applications of the theory.

2. A Class of Minimax Estimators.

<u>Theorem 1</u>. Assume δ is of the form (1.1), where

(i) $0 \leq r(X,W) \leq 2(p-2\tau)$,

(ii) $r(X,W)$ is nondecreasing in $|X_i|$ for $i=1,\ldots p$,

(iii) $r(X,W)$ is nonincreasing in W_i for $i=1,\ldots p$,

(iv) $r(X,W)||\ X\ ||_W^{-2}$ is nondecreasing in W_i for $i=1,\ldots p$.

Then δ is a minimax estimator of θ.

<u>Proof</u>: Throughout the proof it will be assumed that all first order partial derivatives of r exist. The generalization to r merely nondecreasing or nonincreasing in the various coordinates can be done analagously by treating all integrals as Riemann integrals.

The risk of δ, denoted $R(\delta,\theta,A)$, is given by

$$R(\delta,\theta,A) = E_{\theta,A}\ L(\delta,\theta,A) = E_{\theta,A}\ [(\delta\ -\theta)^t Q(\delta-\theta)/tr(QA)].$$

Writing $[\delta-\theta]$ as $[(X-\theta)-r||\ X\ ||_W^{-2}Q^{-1}W^{-1}X]$, and expanding the above quadratic expression gives

$$R(\delta,\theta,A)=E_{\theta,A}[(X-\theta)^tQ(X-\theta)/tr(QA)]-E_{\theta,A}[2r||X||_W^{-2}(X-\theta)^tW^{-1}X/tr(QA)]$$
$$+\ E_{\theta,A}[r^2||X||_W^{-4}X^tW^{-1}Q^{-1}QQ^{-1}W^{-1}X/tr(QA)]$$

$$= 1\ -\ E_{\theta,A}[2r||X||_W^{-2}\{\sum_{i=1}^{p} X_i(X_i-\theta_i)/W_i\}/tr(QA)]$$
$$+\ E_{\theta,A}[r^2||X||_W^{-2}/tr(QA)].$$

To show that δ is minimax, it is clearly only necessary to verify that

$$(2.1)\qquad E_{\theta,A}[2r||X||_W^{-2}\{\sum_{i=1}^{p} X_i(X_i-\theta_i)/W_i\}]-E_{\theta,A}[r^2||X||_W^{-2}] \geq 0.$$

A simple integration by parts with respect to X_i gives

$$E_{\theta,A}[\{r||X||_W^{-2}X_i\}\{(X_i-\theta_i)/A_i\}] = E_{\theta,A}[\frac{\partial}{\partial X_i}(r||X||_W^{-2}X_i)]$$

$$= E_{\theta,A}[\frac{r}{||X||_W^2} - \frac{2rX_i^2}{||X||_W^4\, q_iW_i^2} + \frac{X_i}{||X||_W^2}\{\frac{\partial}{\partial X_i} r(X,W)\}].$$

Using the above equality in the first term of (2.1), and noting that

$[X_i \frac{\partial}{\partial X_i} r(X,W)] \geq 0$ by assumption (ii), it is clear that δ will be proven

minimax if it can be shown that

$$(2.2)\qquad E_{\theta,A}[(\sum_{i=1}^{p} \frac{2rA_i}{||X||_W^2\, W_i}) - \frac{4r}{||X||_W^4}(\sum_{i=1}^{p} \frac{X_i^2A_i}{q_iW_i^3}) - \frac{r^2}{||X||_W^2}] \geq 0.$$

At this point, the following equality is needed:

$$(2.3)\qquad E_{\theta,A}[\frac{rA_i}{||X||_W^2W_i}] = E_{\theta,A}[\frac{r}{||X||_W^2} - \frac{4A_irX_i^2}{(n_i-2)||X||_W^4q_iW_i^3} - \frac{2A_i}{(n_i-2)||X||_W^2}\{\frac{\partial}{\partial W_i}r(X,W)\}].$$

Proof of (2.3): Let U be χ_n^2, g: $R^1 \to R^1$ be an absolutely continuous function, and g' denote the derivative of g (where it exists). Efron and Morris (1974) noted that an integration by parts will prove

$$(2.4)\qquad E[Ug(U)] = nE[g(U)] + 2E[Ug'(U)],$$

providing all integrals exist and are finite. In each of the integrals of (2.4) make the change of variables $Z = cU/(n-2)$ $(c > 0)$, and define $h(Z) = g([n-2]Z/c)$. Noting that $g'(U) = ch'(Z)/(n-2)$, (2.4) becomes

$$E[(n-2)Zh(Z)/c] = nE[h(Z)] + 2E[Zh'(Z)].$$

Since $W_i(n_i-2)/A_i = V_i/A_i$ is $\chi^2_{n_i}$, it follows that

$$(2.5) \qquad E_{\theta,A}[(n_i-2)W_i h(W_i)/A_i] = E_{\theta,A}[n_i h(W_i)] + 2E_{\theta,A}[W_i h'(W_i)].$$

Choose $h(W_i) = r(X,W)/(||X||^2_W W_i)$, which under the assumptions on r is absolutely continuous unless $X_i = 0$ (which of course has measure 0). Noting that

$$h'(W_i) = -\frac{r}{||X||^2_W W_i^2} + \frac{2rX_i^2}{||X||^4_W q_i W_i^4} + \frac{1}{||X||^2_W W_i}\{\frac{\partial}{\partial W_i} r(X,W)\},$$

the expression (2.5) reduces to (2.3).

Inserting the expression given by (2.3) for $E_{\theta,A}[rA_i/(||X||^2_W W_i)]$ into (2.2), and collecting terms, gives as a sufficient condition for minimaxity

$$(2.6) \quad E_{\theta,A}[\frac{2pr-r^2}{||X||^2_W} - \frac{4r}{||X||^4_W}\{\sum_{i=1}^{p}[\frac{2A_i}{(n_i-2)} + A_i]\frac{X_i^2}{q_i W_i^3}\} - \frac{4}{||X||^2_W}\{\sum_{i=1}^{p}\frac{A_i}{(n_i-2)}(\frac{\partial}{\partial W_i}r)\}] \geq 0.$$

Notice that $\{\frac{\partial}{\partial W_i} r(X,W)\} \leq 0$ by assumption (iii). Hence (2.6) will be satisfied if

$$(2.7) \quad E_{\theta,A}[\frac{r}{||X||^2_W}\{2p-r-\frac{4}{||X||^2_W}(\sum_{i=1}^{p}\frac{n_i A_i X_i^2}{[n_i-2]q_i W_i^3})\}] \geq 0,$$

and hence if

$$(2.8) \quad E_{\theta,A}[\frac{r}{||X||^2_W}\{2p-r-\frac{4}{||X||^2_W}(\max_{1\leq i\leq p}\frac{n_i A_i}{[n_i-2]W_i})(\sum_{i=1}^{p}\frac{X_i^2}{q_i W_i^2})\}] \geq 0.$$

Since $\sum_{i=1}^{p} X_i^2/(q_i W_i^2) = ||X||_W^2$, it can be concluded that δ is minimax if

(2.9) $E_{\theta,A}[r||X||_W^{-2}\{2p-r-4(\max_{1 \le i \le p} \frac{n_i A_i}{[n_i-2]W_i})\}] \ge 0.$

For notational convenience, define

$$g(W) = 4 \max_{1 \le i \le p} \{n_i A_i/[(n_i-2)W_i]\}.$$

Note that $g(W)$ is nonincreasing in W_1, and by assumption (iii), $r(X,W)$ is nonincreasing in W_1. Hence $\{2p-r(X,W)-g(W)\}$ is nondecreasing in W_1. Assumption (iv) states that $\{r(X,W)||X||_W^{-2}\}$ is also nondecreasing in W_1. Hence

$$E_{A_1}^{W_1}[r||X||_W^{-2}\{2p-r-g(W)\}] \ge (E_{A_1}^{W_1}[r||X||_W^{-2}])(E_{A_1}^{W_1}[2p-r-g(W)]).$$

Since the W_i are independent, it is again clear from assumptions (iii) and (iv) that $E_{A_1}^{W_1}[r||X||_W^{-2}]$ and $E_{A_1}^{W_1}[2p-r-g(W)]$ are nondecreasing in W_2. Hence

$$E_{A_2}^{W_2}\{(E_{A_1}^{W_1}[r||X||_W^{-2}])(E_{A_1}^{W_1}[2p-r-g(W)])\}$$

$$\ge (E_{A_1,A_2}^{W_1,W_2}[r||X||_W^{-2}])(E_{A_1,A_2}^{W_1,W_2}[2p-r-g(W)]).$$

Continuing in the obvious manner verifies that

(2.10) $E_{\theta,A}^X\{E_A^W[r||X||_W^{-2}(2p-r-g(W))]\}$

$$\ge E_{\theta,A}^X\{(E_A^W[r||X||_W^{-2}])(E_A^W[2p-r-g(W)])\}$$

$$= E_{\theta,A}^X\{(E_A^W[r||X||_W^{-2}])(2p-E_A^W[r]-4\tau)\}.$$

(The last step follows since $n_i A_i/[(n_i-2)W_i] = [V_i/(A_i n_i)]^{-1} = [\chi_{n_i}^2/n_i]^{-1}$, and hence $E_A^W[g(W)] = 4E[T^{-1}] = 4\tau$.) Assumption (i) ensures that $r||X||_W^{-2} \ge 0$ and that $(2p-E_A^W[r]-4\tau) \ge 0$. From (2.10), it is thus clear that (2.9) is satisfied, and hence that δ is minimax.||

Obviously, unless $(p-2\tau) > 0$, assumption (i) and hence Theorem 1 is vacuous. Unfortunately, verifying this inequality is calculationally fairly difficult. Section 4 deals with the calculation of τ. The following theorem shows that if $p \geq 3$ and the n_i are large enough, then indeed $(p-2\tau) > 0$.

Theorem 2. Assume $p \geq 3$. There exists an N such that if $n_i \geq N$, i=1,...p, then $(p-2\tau) > 0$.

Proof: Since $p \geq 3$, it clearly suffices to show that $\limsup_{N\to\infty} \tau(n_1,\ldots n_p) \leq 1$. From the definition of T, it is clear that $T^{-q} \leq \sum_{i=1}^{p} (\chi^2_{n_i}/n_i)^{-q}$. For $q > 1$, Jensen's inequality thus gives

$$(2.11) \quad \tau^q = (E[T^{-1}])^q \leq E[T^{-q}] \leq \sum_{i=1}^{p} E[(\chi^2_{n_i}/n_i)^{-q}].$$

An easy calculation shows that

$$E[(\chi^2_{n_i}/n_i)^{-q}] = (n_i/2)^q \Gamma(\frac{n_i}{2} - q)/\Gamma(\frac{n_i}{2}).$$

Together with (2.11) this gives

$$(2.12) \qquad \tau \leq \{\sum_{i=1}^{p} [(n_i/2)^q \Gamma(\frac{n_i}{2} - q)/\Gamma(\frac{n_i}{2})]\}^{1/q}.$$

For fixed $\epsilon > 0$, q can be chosen large enough so that $(2p)^{1/q} < 1 + \epsilon$. For fixed q, it is straightforward to verify by Stirling's approximation that

$$(n_i/2)^q \Gamma(\frac{n_i}{2} - q)/\Gamma(\frac{n_i}{2}) \to 1 \text{ as } n_i \to \infty.$$

Combining these two observations with (2.12) gives the desired result. ||

At this point it should be mentioned that condition (i) of Theorem 1 is undoubtedly stronger than necessary. An examination of the proof of Theorem 1, specifically the passage from (2.7) to (2.8), leads one to think that τ could be replaced by something much closer to one. Indeed, one would guess that τ could be replaced by

$$\max_{1 \leq i \leq p} E(n_i/\chi^2_{n_i}) = \max_{1 \leq i \leq p} [n_i/(n_i-2)].$$

Unfortunately, we were unable to verify any such better condition. The proof of Theorem 2 does indicate, in any case, that if the n_i are large (relative to p), then little is lost by the rougher bound.

3. The Choice of r.

When it comes to suggesting an estimator to use in practice, the choice determined by $r(X,W) \equiv c$, $0 \leq c \leq 2(p-2\tau)$, is attractive because of its simplicity. (It is trivial to verify that the conditions of Theorem 1 are satisfied for this choice of r.) The trouble with constant r, and many other simple choices, is that the resulting estimator has a singularity at zero. Such estimators can usually be considerably improved upon. Attractive methods of eliminating this singularity are discussed in Berger and Bock (1975). The simplest for use in this paper is the use of "positive part" versions of the estimators. If δ is given componentwise by $\delta_i(X) = \phi_i(X)X_i$, the positive part version, $\bar{\delta}(X)$, is defined by $\bar{\delta}_i(X) = \phi_i^+(X)X_i$, where + stands for the usual positive part.

In Berger and Bock (1975) it is shown that if the X_i are independent with densities $f_i(x_i-\theta_i)$ which are symmetric and unimodal, if the loss is a diagonal quadratic loss and if the estimator, δ, is diagonal (i.e. $\delta_i(X) = \phi_i(X)X_i$), then the positive part version, $\bar{\delta}$, has smaller risk than δ. These conditions are clearly satisfied for the situation of this paper. Hence the estimators δ, given in (1.1), can often be considerably improved upon by using $\bar{\delta}$, defined componentwise by

$$\bar{\delta}_i(X,W) = [1 - r(X,W)/(||X||_W^2\, q_iW_i)]^+X_i.$$

For practical purposes, it has already been indicated that the choice $r(X,W) \equiv c$, $0 \leq c \leq 2(p-2\tau)$, is desirable because of its simplicity. The positive part version of such a choice seems to provide a simple, yet quite good minimax estimator. Though such an estimator is probably not admissible, it is unlikely, as in the corresponding situation with known variances, that admissible choices of r would be significantly better. The desired minimax estimators are thus given componentwise by

$$(3.1)\quad \delta_i^c(X,W) = [(1-c/(||X||_W^2\, q_i W_i)]^+ X_i, \text{ for } 0 \leq c \leq 2(p-2\tau).$$

In actual applications, it is clear that the above estimators should be centered at what is considered, apriori, to be the "most likely" parameter value θ_0. This will result in a minimax estimator where the major improvement in risk (over the usual estimator $\delta_0(X)=X$) is at θ_0. The centered version of δ^c is simply

$$(3.2)\quad \delta_{\theta_0}^c(X,W) = \delta^c(X-\theta_0,W) + \theta_0.$$

Finally, the question of choosing c arises. It is reasonable to choose c to minimize the risk at the "most likely" point θ_0. It is easy to see that $c = 2(p-2\tau)$ is the desired minimizing value. This choice of c also gives rise to the most attractive looking risk function, in the sense that the region of considerable improvement seems to be largest. In conclusion, the recommended minimax estimator for practical use (among the class covered by Theorem 1) is $\delta_{\theta_0}^{2(p-2\tau)}$.

4. Applications.

To apply Theorem 1, it is clearly necessary to determine

$$\tau = E[T^{-1}] = E[(\min_{1 \leq i \leq p} \chi_{n_i}^2/n_i)^{-1}].$$

It can be shown that if the n_i are even, then

$$(4.1)\quad \tau = \sum_{k=1}^{p} \sum \{ [\prod_{i \neq k} \frac{m_i^{(m_i - j(i))}}{(m_i - j(i))!}] \frac{(m-J(k)-2)!\, m_k^{m_k}}{(m_k - 1)!\, m^{(m-J(k)-1)}} \},$$

where $m_i = n_i/2$, $m = \sum_{i=1}^{p} m_i$, $J(k) = \sum_{i \neq k} j(i)$, and the inner summation is over all combinations $\{j(1), j(2), \ldots j(k-1), j(k+1), \ldots j(p)\}$ where the $j(\ell)$ are integers between 1 and m_ℓ inclusive. The calculation is relatively straightforward though tedious. It is based upon observing that F_i, the c.d.f. of $\chi^2_{n_i}/n_i$, can be expressed in terms of a cumulative sum of a Poisson distribution.

Tables of τ are clearly desirable. The difficulty in constructing such tables is, however, obvious - the parameters p, $n_1, n_2, \ldots n_p$ are all involved. As a compromise between completeness and conciseness, a table (table 1) is included only for $n_1 = n_2 = \ldots = n_p = n$. Recall that $(p-2\tau)$ must be positive in order for Theorem 1 to guarantee the existence of minimax estimators better than δ_0. For example, if p=5, then n must be at least 8 for the theory to apply.

If the n_i are not all equal, and it is not desired to use the exact formula (4.1), the following two bounds on τ might prove useful. The first bound follows from the observation that

$$T^{-1} = \max_{1 \leq i \leq p} [n_i/\chi^2_{n_i}] \leq \sum_{\text{all } j} \max_{i \in \Omega_j} [n_i/\chi^2_{n_i}],$$

where the Ω_j are disjoint sets of integers whose union is $\{1, 2, \ldots p\}$. Clearly

$$(4.2)\quad \tau \leq \sum_{\text{all } j} \tau_j, \text{ where } \tau_j = E\,[\max_{i \in \Omega_j} (n_i/\chi^2_{n_i})].$$

The second possible bound on τ was conjectured by the referee. It is that

$$(4.3)\quad \tau(n_1, \ldots n_p) \leq \tau(n^*, \ldots n^*), \text{ where } n^* = \min_{1 \leq i \leq p} n_i.$$

Unfortunately, we were unable to prove the above bound. Since intuitive considerations and numerical calculations indicate it is valid, however, it seemed desirable to include it.

As an example of the use of (4.2) and (4.3), assume that p=10 with $n_i=25$ for $1 \le i \le 8$, $n_9=30$, and $n_{10}=5$. From (4.2), with $\Omega_1 = \{1,2,\ldots 9\}$ and $\Omega_2 = \{10\}$, it follows $\tau \le \tau_1 + \tau_2$. Clearly $\tau_2 = E[5/\chi_5^2] = 5/(5-2) = 1.67$. By (4.3), $\tau_1 \le \tau(25,\ldots 25) = 1.67$ (from the table with p=9 and n=25). Hence $\tau \le 3.34$.

The obvious example of the above theory which should be mentioned is that of combining p completely independent normal mean estimation problems. Thus suppose that for j=1,...p, we observe a random sample $Y_1^j, Y_2^j, \ldots Y_{m_j}^j$ of size m_j from a normal distribution with unknown mean θ_j and unknown variance σ_j^2. Assume that the observations are all independent. This problem can be put into the setting of Section 1 by letting

$$X_j = \overline{Y}^j = \frac{1}{m_j}\sum_{i=1}^{m_j} Y_i^j,\ V_j = S_j^2 = \frac{1}{m_j}\sum_{i=1}^{m_j} (Y_i^j - \overline{Y}^j)^2,\ A_j = \sigma_j^2/m_j, \text{ and } n_j = m_j-1.$$

(Note that it is necessary to have $m_j \ge 4$ to ensure that $n_j \ge 3$.) As an example of the improvement in risk (over the usual estimator δ_0) that can be obtained, the risk of the estimator $\delta_0^{2(p-2\tau)}$ was calculated for p=4, $m_1=m_2=m_3=m_4=17$, and $\sigma_j^2=q_j=1 (j=1,\ldots 4)$. Figure 1 depicts this risk (as a function of $|\theta|$). The constant line c=1 is the risk of the usual estimator $\overline{Y} = (\overline{Y}^1, \overline{Y}^2, \ldots \overline{Y}^p)$. Clearly, significant savings are achievable.

Acknowledgements: The authors are grateful to the referee for greatly simplifying the proof of Theorem 2 and for suggesting the given bounds on τ.

References

[1] Alam, Khursheed (1973). A family of admissible minimax estimators of the mean of a multivariate normal distribution. Ann. Statist. 1, 517-525.

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

[3] Berger, J. (1974a). Admissible minimax estimation of a multivariate normal mean with arbitrary quadratic loss. Technical Report 376, Purdue University. (To appear in the Ann. Statist.)

[4] Berger, J. (1974b). Tail minimaxity in location vector problems and its applications. Technical Report 377, Purdue University. (To appear in the Ann. Statist.)

[5] Berger, J. (1975). Minimax estimation of a multivariate normal mean under arbitrary quadratic loss. Technical Report 413, Purdue University. (Submitted to the J. Multivariate Anal.)

[6] Berger, J. and Bock, M. E. (1975). Eliminating singularities of Stein-type estimators of location vectors. Technical Report 414, Purdue University. (Submitted to the J. Roy. Statist. Soc. Ser. B.)

[7] Bhattacharya, P. K. (1966). Estimating the mean of a multivariate normal population with general quadratic loss function. Ann. Math. Statist. 37, 1819-1824.

[8] Bock, M. E. (1975). Minimax estimators of the mean of a multivariate normal distribution. Ann. Statist. 3, 209-218.

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

[10] Efron, B. and Morris, C. (1974). Families of minimax estimators of the mean of a multivariate normal distribution. Rand Paper Series. The Rand Corporation, Santa Monica, Calif.

[11] Hudson, M. (1974). Empirical Bayes estimation. Technical Report #58, Stanford University.

[12] James, W. and Stein, C. (1960). Estimation with quadratic loss. Proc. Fourth Berkeley Symposium Math. Stat. Prob. 1, 361-379. University of California Press.

Table 1

Values of τ

n	DIMENSION-p												
	3	4	5	6	7	8	9	10	11	12	13	14	15
4	3.66	4.22	4.79	5.26	5.15	6.15	6.63	7.06	7.50	7.99	8.34	8.70	8.99
5	2.88	3.29	3.65	3.97	4.30	4.57	4.85	5.09	5.31	5.53	5.74	5.92	6.09
6	2.45	2.78	3.01	3.27	3.51	3.69	3.89	4.05	4.22	4.35	4.48	4.61	4.74
7	2.21	2.46	2.66	2.85	3.03	3.17	3.31	3.44	3.57	3.67	3.78	3.89	4.01
8	2.04	2.26	2.42	2.59	2.74	2.85	2.97	3.08	3.18	3.26	3.34	3.43	3.52
9	1.94	2.14	2.24	2.39	2.51	2.61	2.71	2.79	2.89	2.96	3.03	3.10	3.17
10	1.84	2.01	2.11	2.24	2.35	2.43	2.52	2.58	2.66	2.73	2.78	2.85	2.90
11	1.76	1.89	2.02	2.13	2.23	2.30	2.38	2.44	2.51	2.57	2.62	2.69	2.73
12	1.69	1.83	1.93	2.04	2.12	2.19	2.26	2.32	2.38	2.43	2.47	2.53	2.58
13	1.64	1.77	1.86	1.96	2.04	2.10	2.17	2.22	2.28	2.32	2.37	2.41	2.46
14	1.60	1.72	1.81	1.90	1.97	2.03	2.09	2.14	2.19	2.23	2.27	2.31	2.35
15	1.57	1.68	1.77	1.85	1.91	1.97	2.03	2.07	2.12	2.15	2.19	2.23	2.27
16	1.54	1.64	1.72	1.80	1.86	1.91	1.96	2.01	2.05	2.08	2.12	2.16	2.19
17	1.51	1.61	1.69	1.76	1.82	1.87	1.91	1.96	2.00	2.03	2.06	2.09	2.13
18	1.49	1.59	1.66	1.73	1.78	1.83	1.87	1.91	1.95	1.98	2.01	2.04	2.07
19	1.47	1.56	1.63	1.69	1.75	1.79	1.83	1.87	1.91	1.93	1.96	1.99	2.02
20	1.45	1.53	1.60	1.66	1.71	1.76	1.80	1.84	1.87	1.90	1.93	1.95	1.98
21	1.43	1.51	1.58	1.64	1.69	1.73	1.77	1.80	1.84	1.86	1.89	1.92	1.94
22	1.41	1.50	1.56	1.62	1.66	1.70	1.74	1.77	1.81	1.83	1.86	1.88	1.91
23	1.40	1.48	1.55	1.60	1.64	1.68	1.72	1.75	1.78	1.80	1.83	1.85	1.87
24	1.39	1.45	1.53	1.58	1.62	1.65	1.69	1.72	1.75	1.77	1.80	1.82	1.84
25	1.37	1.45	1.51	1.56	1.60	1.64	1.67	1.70	1.73	1.75	1.77	1.80	1.82
26	1.36	1.44	1.49	1.54	1.58	1.62	1.65	1.68	1.71	1.73	1.75	1.77	1.79
27	1.35	1.42	1.48	1.53	1.57	1.63	1.66	1.66	1.68	1.71	1.73	1.75	1.77
28	1.35	1.41	1.47	1.52	1.55	1.58	1.61	1.64	1.66	1.68	1.71	1.73	1.74
29	1.34	1.41	1.46	1.51	1.54	1.57	1.60	1.63	1.65	1.67	1.69	1.71	1.73
30	1.33	1.40	1.45	1.49	1.53	1.56	1.59	1.61	1.63	1.65	1.67	1.69	1.71
35	1.30	1.36	1.40	1.44	1.48	1.50	1.53	1.55	1.57	1.59	1.60	1.62	1.63
40	1.27	1.32	1.37	1.40	1.43	1.45	1.48	1.50	1.52	1.53	1.54	1.56	1.57
45	1.25	1.30	1.34	1.37	1.40	1.42	1.44	1.46	1.48	1.49	1.50	1.51	1.53
50	1.23	1.28	1.31	1.35	1.37	1.39	1.41	1.43	1.44	1.45	1.47	1.48	1.49
55	1.22	1.26	1.29	1.33	1.35	1.37	1.38	1.40	1.41	1.43	1.44	1.45	1.46
60	1.21	1.25	1.28	1.31	1.33	1.35	1.36	1.38	1.39	1.40	1.41	1.42	1.43
65	1.20	1.24	1.27	1.29	1.31	1.33	1.34	1.36	1.37	1.38	1.39	1.40	1.41
70	1.19	1.23	1.25	1.28	1.30	1.31	1.33	1.34	1.35	1.36	1.37	1.38	1.39
75	1.18	1.22	1.24	1.27	1.28	1.30	1.31	1.33	1.34	1.34	1.35	1.36	1.37

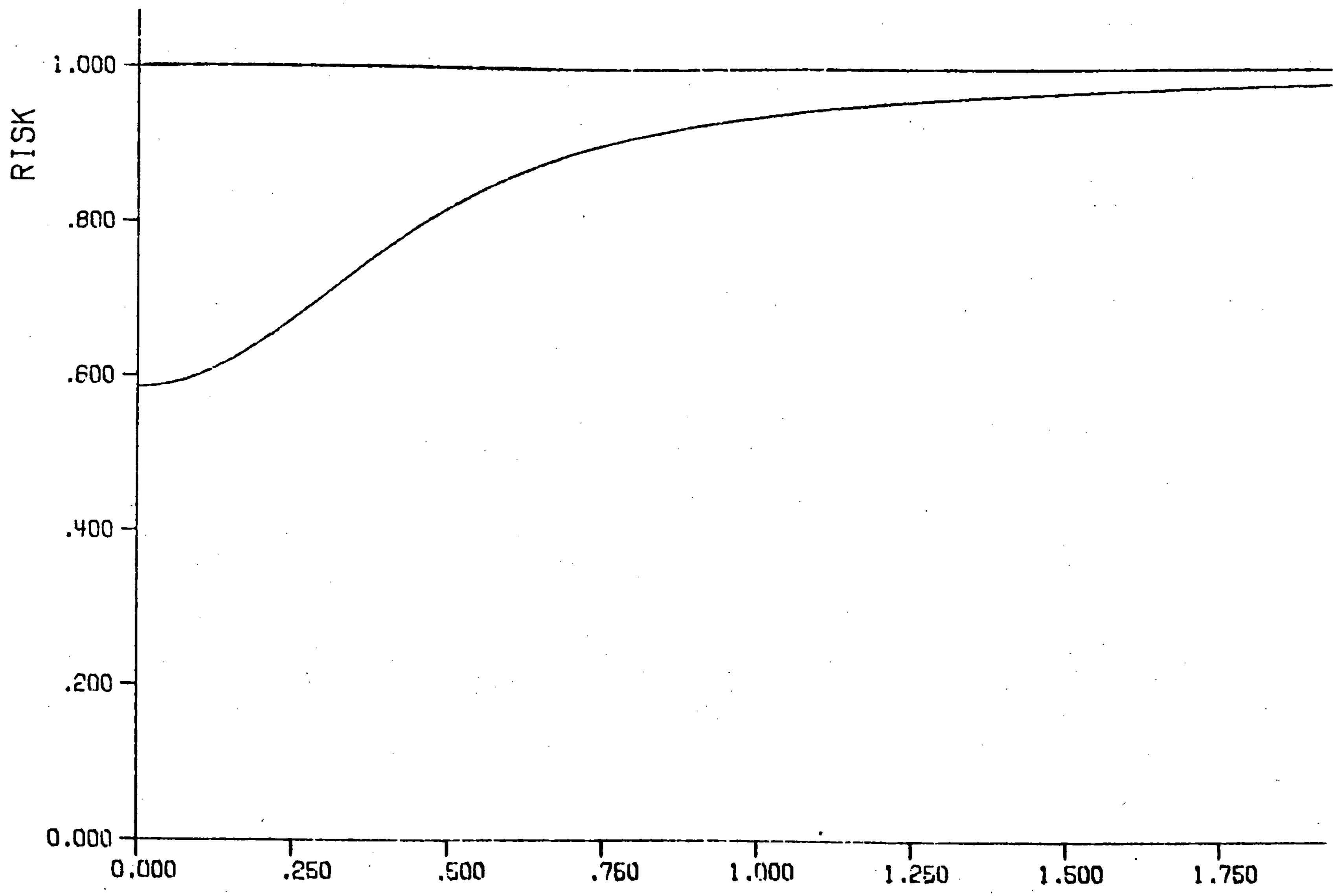

Figure 1. Risk of $\delta_0^{2(p-2\tau)}$.

[13] Lin, Pi-Erh, and Tsai, Hui-Liang (1973). Generalized Bayes minimax estimators of the multivariate normal mean with unknown covariance matrix. Ann. Statist. 1, 142-145.

[14] Stein, C. (1955). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Proc. Third Berkeley Symp. Math. Statist. Prob. 1, 197-206. University of California Press.

[15] Strawderman, W. E. (1973). Proper Bayes minimax estimators of the multivariate normal mean vector for the case of common unknown variances. Ann. Statist. 1, 1189-1194.