BOUNDS ON THE PERFORMANCE OF SPHERICALLY SYMMETRIC ESTIMATORS WHICH DOMINATE THE JAMES-STEIN ESTIMATOR

by

CARL SPRUILL
Georgia Institute of Technology
and
Purdue University

Technical Report # 82-19

AMS Subject Classifications: Primary 62C15; Secondary 62F10, 62C07.

Key Words and Phrases: Admissibility, estimation, normal distribution, inequalities, Bessel functions.

1. Introduction. Although the James-Stein positive part estimator ξ_o^+ of a multivariate normal mean is inadmissible it seems widely held that the available improvement is quite small. The same view was expressed by James and Stein (1960) concerning its precursor ξ_o. Now known as the James-Stein estimator ξ_o is defined by

$$\xi_o(X) = (1 - \frac{p-2}{||X||^2})X,$$

where $p \geq 3$ is the dimension of the unknown mean vector μ, $X \sim \mathcal{N}(\mu,I)$, and $||X||^2 = \sum_{j=1}^{p} X_j^2$. The positive part estimator is

$$\xi_o^+(X) = (1 - \frac{p-2}{||X||^2})_+X$$

and dominates ξ_o in the sense that its risk function lies strictly below that of ξ_o for all values of $\mu \in \mathbb{R}^p$.

What is the basis for this opinion as it pertains to ξ_o^+? Surely no part of it is by direct comparison for there appear to be no admissible estimators in the literature known to dominate either ξ_o or ξ_o^+. There is a powerful technique due to Stein (1973) and developed by Hudson (1978), Berger (1980), and Hwang (1980) for constructing estimators which improve upon a given estimator but it has been shown by Moore and Brook (1978) that it fails for both ξ_o and ξ_o^+ to produce a dominant estimator. In their words "this shows that these estimators are so good that very powerful methods will be needed to obtain alternatives which dominate them". In his original paper which proved the inadmissibility of X in $p \geq 3$ dimensions Stein (1956) also contributed to the evidence that ξ_o could not be greatly improved uniformly in $||\mu||$ large. He proved that if $c > (p-2)^2$ then no spherically symmetric estimator ξ can be found which satisfies for some $\lambda_o > 0$

$$R(\xi,\lambda) \leq p - \frac{c}{\lambda} \qquad \text{for all } \lambda \geq \lambda_o \text{, where } \lambda = ||\mu||^2 .$$

Since

$$R(\xi_o,\lambda) = p - (p-2)^2 \, E[1/\chi_p^2(\lambda)] \leq p - \frac{(p-2)^2}{p-2+\lambda}$$

any spherically symmetric estimator, and consequently any estimator, which dominates ξ_o will have a risk function which comes close to that of ξ_o at infinitely many μ values as $||\mu||$ gets large. Efron and Morris (1973) concluded that ξ_o^+ was a good estimator in its own right as one member of a class of good rules which are minimax and have Bayesian properties. The class of rules to which they refer is found in section five of their paper. The members of this class of extended Stein rules have certain desirable properties including a rather favorable comparison, in terms of Bayes risk against certain normal priors, with an admissible minimax estimator due to Strawderman (1971).

We sought more direct evidence of the status of ξ_o^+ in relation to admissible estimators which dominate it. Our results are far from definitive but we count them as further evidence in favor of the general view that the positive part estimator can not be greatly improved in terms of its risk function. We prove that if an estimator ξ dominates ξ_o then the area between the graphs of their risk functions plotted as a function of $\lambda > 0$ does not exceed 2(p-2). This upper bound is simple but crude and numerical computations yield smaller values, of which a selected few are given below. In the case of ξ_o^+ the area between its risk function and that of any dominating estimator does not exceed $2(p-2 + R(\xi_o^+,0) - R(\xi_o,0))$. Again this is a simple but crude bound and more exact bounds are provided for certain values of p. For example, the area between the graphs of the risk function of ξ_o^+ and any estimator which dominates it does not exceed .40 in p = 3 dimensions. In order to place this figure in perspective the area bounded by the graphs of the risk functions of ξ_o and ξ_o^+ is 1.12.

In addition to the global measure of improvement we present the following results concerning local improvement. For any given $\lambda \geq 0$ the maximum possible improvement on the risk of ξ_o using a bounded risk admissible spherically symmetric (with respect to the origin) estimator is no more than 2 units. The improvement by an admissible estimator which dominates ξ_o would presumably be much smaller and is of course the quantity of interest. However the bound above is the best we are currently able to provide. Denoting by

$$g_{p,\lambda}(t) = \sum_{j\geq 0} \left(\frac{\lambda}{2}\right)^j \frac{e^{-\lambda/2}}{j!} \frac{t^{p/2+j-1}e^{-t/2}}{\Gamma(p/2+j)2^{p/2+j}}, \quad t > 0 > \lambda > 0,$$

the density of the non-central chi square random variable with p degrees of freedom and non-centrality parameter λ the maximal possible improvement in risk over ξ_o^+ at a fixed point λ is no more than

$$(R(\xi_o^+,0) - R(\xi_o,0))\frac{g_{p,\lambda}(p-2\sqrt{p-2})}{g_{p,0}(p-2\sqrt{p-2})} + 2 \overset{\Delta}{=} B(\lambda).$$

The global measures also provide some additional information about local behavior for if ξ is spherically symmetric and dominates ξ_o^+ then the Lebesgue measure $m(S(\varepsilon))$ of the set

$$S(\varepsilon) = \{\lambda\colon R(\xi,\lambda) \leq (1-\varepsilon)R(\xi_o^+,\lambda)\}$$

satisfies

$$(1.1) \qquad m(S(\varepsilon)) < [p-2+R(\xi_o^+,0) - R(\xi_o,0)]/\varepsilon.$$

The stronger version of this bound yields for example the fact that in p = 3 dimensions the measure of the set of λ values at which a dominating estimator has a risk of no more than 90% that of ξ_o^+ is no greater that 2.0.

Of course the calculations have all been made based upon a single p-dimensional observation. If n independent observations are available then the maximal improvement at a point over ξ_{on} is bounded by 2/n and that for ξ_{on}^+ is bounded by $n^{-1}B(n\lambda)$. The measure of the set in (1.1) will also be reduced by a factor of n^{-1} and all areas by a factor of n^{-2}. The estimator ξ_{on} is defined by

$$\xi_{on}(\overline{X}) = \left(1 - \frac{p-2}{n||\overline{X}||^2}\right)\overline{X}$$

and ξ_{on}^+ is its positive part.

2. Preliminaries.

Let $X_1,\ldots,X_p$ be independent normal random variables, $X_i \sim \mathcal{N}(\mu_i,1)$. The class of (randomized) estimators of the vector μ is the collection of all regular conditional Borel probability measures $\delta(\cdot|x)$ on $\mathbb{R}^p \times \mathbb{R}^p$. Denoting by G the group of orthogonal transformations on $\mathbb{R}^p$, a measure ν is termed invariant if it satisfies $\nu(gB) \equiv \nu(B)$ for all $g \in G$ and all Borel sets B. An estimator δ is spherically symmetric if for each x $\delta(\cdot|x)$ is invariant. An estimator ξ is spherically symmetric and non-randomized if and only if there is a measurable function h such that

$$\xi(x) = [1-h(||x||^2)]x.$$

The risk function of an arbitrary estimator δ is

$$\tilde{R}(\delta,\mu) = \int\int ||a-\mu||^2\delta(da|x)\ P_\mu(dx).$$

The risk function of a spherically symmetric estimator depends only on $\lambda = ||\mu||^2$ and will often be written $\tilde{R}(\delta,\mu) = R(\delta,\lambda)$.

Lemma 2.1 The class of non-randomized spherically symmetric admissible estimators is a non-empty complete class in the subclass of all spherically symmetric estimators.

Proof: We apply an elegant and powerful result of Balder (1980). The reader is referred to sections two and three of that paper. First replace the action space $\mathbb{R}^p$ with a compact action space $A = \hat{\mathbb{R}}^p$, where $\hat{\mathbb{R}}^p$ is an Alexandroff compactification (Bourbaki (1966)). In order to preserve the notion of invariance we choose $\hat{\mathbb{R}}^p$ to be a unit sphere in $\mathbb{R}^{p+1}$. The homeomorphism from $\mathbb{R}^p$ onto $\hat{\mathbb{R}}^p - \{\infty\}$ is

$$\varphi(x) = \left(\frac{2}{4+||x||^2}\right)(||x||^2_; \ 2x)$$

at $x \in \mathbb{R}^p$. Define a loss function $\hat{\ell}$ on $A \times \mathbb{R}^p$ by $\hat{\ell}(a,\mu) = ||\varphi_a^{-1}-\mu||^2$ if $a \neq \infty$ and $\hat{\ell}(a,\mu) = +\infty$ if $a = \infty$. The function $\hat{\ell}$ is Borel measurable and in Balder's terminology a normal integrand. To each regular conditional probability δ on $\mathbb{R}^p \times \mathbb{R}^p$ is associated a regular conditional probability $\hat{\delta}$ on $A \times \mathbb{R}^p$ by setting $\hat{\delta}(\infty|x) \equiv 0$ and

(2.1) $$\hat{\delta}(\varphi B|x) = (1-\hat{\delta}(\infty|X))\ \delta(B|x).$$

Given the regular conditional probability $\hat{\delta}$ employ (2.1) to define δ for all Borel sets B in $\mathbb{R}^p$ and all x such that $\hat{\delta}(\infty|x) < 1$. If $\hat{\delta}(\infty|x) = 1$ set $\delta(0|x) = 1$. Thus for every δ $\tilde{R}(\delta,\mu) \equiv \hat{R}(\hat{\delta},\mu)$ where

$$\hat{R}(\hat{\delta},\mu) = \int \hat{\ell}\,(a,\mu)\ \hat{\delta}(da|x)\ P_\mu(dx).$$

Conversely given $\hat{\delta}$ there is a δ satisfying $\hat{R}(\hat{\delta},\mu) \equiv \tilde{R}(\delta,\mu)$ if $\hat{R}(\hat{\delta},\mu) < \infty$ for all μ and $\tilde{R}(\delta,\mu) \leq \hat{R}(\hat{\delta},\mu)$ otherwise. For each $g \in G$ define a transformation $\hat{g}_g$ on A by $\hat{g}_g\varphi(x) = \varphi(gx)$ and $\hat{g}_g\,(2,0) \equiv (2,0)$. Let $\hat{G}$ be the collection of all such transformations. The measure δ is invariant under G if and only if $\hat{\delta}$ is invariant under $\hat{G}$.

The topological dual of the separable Banach space C(A) of continuous functions on A is isometrically isomorphic to the set of signed Baire measures on A and M_1^+ to the probability measures. Since the Baire sets and Borel sets coincide in A, given $m \in M_1^+$ there is an associated Borel measure $\hat{\nu}$ on A such that for all $f \in C(A)$ $< m,f > = \int f d\hat{\nu}$. Define the set $M_I^+ \subset M_1^+$ as the image under this isometry of the invariant measures on A. Define $\Gamma(x) \equiv M_I^+$ for all $x \in \mathbb{R}^p$.

Balder's Theorem 1 and Corollary 2 apparently remain valid without the assumption that the values of Γ are extremal in M_1^+. The set M_I^+ is manifestly convex so in order to appeal to these results it suffices to prove that M_I^+ is a closed subset of M_1^+ in the weak* topology on M_1. Since it is metrizable it suffices to check that if m_n converges to m weak* and if $m_n \in M_I^+$ for all n then $m \in M_I^+$, or equivalently in terms of the associated measures $\hat{\nu}_n$, $\hat{\nu}$ that $\hat{\nu}$ is invariant. It is enough to carry this demonstration forward for compact sets. Therefore let K be a compact set in A. Since $\hat{\nu}(K) = \inf\{\hat{\nu}(0): 0 \supset K, 0 \text{ open}\}$ there is a sequence 0_k of open sets such that $\hat{\nu}(0_k) \downarrow \hat{\nu}(K)$. The space A is normal so by Urysohn's lemma there is a sequence of functions $\{f_k\} \subset C(A)$ such that f_k is one on K, zero on 0_k^c, and $0 \leq f_k(a) \leq 1$ for all $a \in A$. Thus for all k

(2.2) $$\hat{\nu}(K) \leq \int f_k d\hat{\nu} \leq \hat{\nu}(0_k).$$

The functions $\tilde{f}_k = f_k \circ \hat{g}^{-1}$ are also continuous and the inequality (2.2) holds when K, f_k, and 0_k are replaced by $\hat{g}K$, $\tilde{f}_k$, and $\hat{g}0_k$. For all k and n

(2.3) $$\left| \int (f_k-\tilde{f}_k)d\hat{\nu} \right| \leq \left| \int (f_k-\tilde{f}_k)(d\hat{\nu}-d\hat{\nu}_n) \right| + \left| \int (f_k-\tilde{f}_k)d\hat{\nu}_n \right| = |\langle m-m_n, f_k-\tilde{f}_k \rangle|$$

where we have used the fact that $\int (f_k-\hat{f}_k)d\hat{\nu}_n$ vanishes owing to the invariance of $\hat{\nu}_n$. Taking the limit in (2.3) first on n and then on k and using (2.2) shows that $\hat{\nu}$ is invariant and therefore that M_I^+ is closed.

Let δ be a given spherically symmetric estimator and $\hat{\delta}$ be its associate. Balder's corollary 2 yields an invariant admissible $\hat{\delta}^*$ such that

$\hat{R}(\hat{\delta}^*,\cdot) \leq \hat{R}(\hat{\delta},\cdot)$. Using the convexity of the loss function one can now easily deduce the existence of a non-randomized spherically symmetric admissible ξ which satisfies $\tilde{R}(\xi,\mu) \leq \tilde{R}(\delta,\mu)$ for all μ. □

Stein (1956) proved that if a spherically symmetric estimator is admissible in the class of spherically symmetric estimators then it is admissible in the full class. Brown (1971) proved that the admissible estimators are generalized Bayes. Since Strawderman and Cohen (1971) proved that a generalized Bayes estimator ξ is spherically symmetric if and only if its generalized prior is spherically symmetric we conclude that if ξ is an admissible spherically symmetric estimator then (see equation (1.2.2) of Brown (1971))

$$\xi(x)-x = \nabla \log f^*(x),$$

where

$$f^*(x) = \int e^{-\frac{1}{2}||x-\mu||^2} d\Pi(\mu)$$

and Π is a spherically symmetric positive Borel measure on $\mathbb{R}^p$ for which $f^*(x) < \infty$ for all $x \in \mathbb{R}^p$.

Define the Borel measure π on $\mathbb{R}^1$ by $\pi(a,b] = \int_{a<||\mu||^2\leq b} d\Pi(\mu)$ and

$$g_{p,\pi}(t) = \int_0^\infty g_{p,\lambda}(t)d\pi(\lambda).$$

For the cases of interest the funtion $g_{p,\pi}(t)$ will be finite for all $t > 0$ and possess derivatives of all orders which may be passed beneath the integral (see Spruill (1979)). With $\frac{d}{dt} g_{p,\pi}(t) = g'_{p,\pi}(t)$ we have the following lemma.

<u>Lemma 2.2</u> If $\xi(x) = [1-h(||x||^2)]x$ is an admissible spherically symmetric estimator then there is a non-negative Borel measure π on $[0,\infty)$ for which

$g_{p,\pi}(t) < \infty$ for all $t > 0$ and

$$h(t) = \frac{p-2}{t} - 2\,\frac{g'_{p,\pi}(t)}{g_{p,\pi}(t)}\,.$$

Proof: It suffices to prove that

$$(2.4) \quad \nabla \log(f^*(x)) = -\left[\frac{p-2}{t} - \frac{g'_{p,\pi}(t)}{g_{p,\pi}(t)}\right]x.$$

Let ν be uniformly distributed on the surface of the sphere of radius r in p dimensions, $S_r = \{x: ||x||=r\}$. One can show (see Gihman and Skorohod (1974) equation IV.2.8 for example) that the moment generating function of ν is

$$\phi_r(w) = E[e^{w\cdot\nu}] = \Gamma(p/2)\sum_{j\geq 0}\left(\frac{||w||^2 r^2}{4}\right)^j \frac{1}{j!\,\Gamma(p/2+j)}$$

where $\Gamma(\cdot)$ is the gamma function. Thus

$$f^*(x) = e^{\frac{-||x||^2}{2}} \int e^{x\cdot\mu} e^{\frac{-||\mu||^2}{2}}\, d\Pi(\mu)$$

$$= e^{\frac{-||x||^2}{2}} \int_0^\infty e^{-\frac{\lambda}{2}} \int_{S_{\sqrt{\lambda}}} e^{x\cdot\mu} dS(\mu)\ \ d\pi(\lambda)$$

$$= e^{\frac{-||x||^2}{2}} \int_0^\infty \phi_{\sqrt{\lambda}}(x) e^{-\lambda/2} d\pi(\lambda).$$

Setting $||x||^2 = t$ we have

$$f^*(x) = \left(\frac{2}{t}\right)^{p/2-1} \Gamma(p/2) g_{p,\pi}(t)$$

and (2.4) follows readily. □

Brown shows that the generalized Bayes spherically symmetric estimator ξ has bounded risk if and only if there is a constant B satisfying for all $t > 0$

$$\left| \frac{p-2}{t} - 2\phi(t) \right| \leq B/\sqrt{t},$$

where $\phi(t) = g'_{p,\pi}(t)/g_{p,\pi}(t)$. Using this fact and other properties of ϕ one can show that equation (6) of section 3 of Stein (1973) can be employed to yield, for any bounded risk spherically symmetric admissible estimator ξ,

$$R(\xi,\lambda) - R(\xi_o,\lambda) = 4E[S\phi^2(S)+2\phi(S)+2S\phi(S)],$$

where S is non-central chi-square with p degrees of freedom and non-centrality parameter λ and that the final term may be integrated by parts to yield

$$(2.5)\quad R(\xi,\lambda) = R(\xi_o,\lambda)+4\int_0^\infty (t\phi^2(t)g_{p,\lambda}(t)-2t\phi(t)g'_{p,\lambda}(t))dt.$$

Finally, writing $g_p(x,t) = \int_0^\infty g_{p,\lambda}(t)d\lambda$, it can be shown that if ξ is an admissible estimator which dominates ξ_o then for all $x > 0$

$$(2.6)\quad \int_0^x [R(\xi_o,\lambda)-R(\xi,\lambda)]d\lambda = -4\int_0^\infty [t\phi^2(t)g_p(x,t)-2t\phi(t)g'_p(x,t)]dt.$$

For details see Spruill (1979).

3. <u>Local bounds.</u> Denote by $\mathcal{A}$ the class of admissible bounded risk spherically symmetric estimators. We shall prove that, given $\lambda \geq 0$

$$\inf_{\xi\in\mathcal{A}} R(\xi,\lambda) \geq R(\xi_o,\lambda)-2.$$

One can prove the following lemma (see Spruill (1980).

<u>Lemma</u> 3.1 Suppose that $\sum_{j\geq 0} a_j t^j$ and $\sum_{j\geq 0} b_j t^j$ converge for all $t > 0$ and that for j sufficiently large $b_j > 0$. If $\lim_{j\to\infty} \frac{a_j}{b_j} = \gamma$ (including $\gamma = \pm\infty$)

then

$$\lim_{t\to\infty} \frac{\sum_{j\geq 0} a_j t^j}{\sum_{j>0} b_j t^j} = \gamma.$$

Theorem 3.1. For all $\lambda \geq 0$,

$$\inf_{\xi \in \mathcal{A}} R(\xi,\lambda) \geq R(\xi_o,\lambda)-2.$$

Proof: Let $\xi \in \mathcal{A}$. Using (2.5) and the fact that $t\phi^2 g_{p,\lambda}(t) - 2t\phi g'_{p,\lambda}(t)$ is minimized by the choice

$$\phi(t) = \frac{g'_{p,\lambda}(t)}{g_{p,\lambda}(t)} \tag{3.1}$$

we have $R(\xi,\lambda) \geq R(\xi_o,\lambda)-4I_{p,\lambda}$

where $I_{p,\lambda} = \int_o^\infty t\left(\frac{g'_{p,\lambda}(t)}{g_{p,\lambda}(t)}\right)^2 g_{p,\lambda}(t)dt.$

Writing $g'_{p,\lambda}(t) = 1/2(g_{p-2,\lambda}(t)-g_{p,\lambda}(t))$ and expanding we obtain

$$I_{p,\lambda} = 1-\left(\frac{p+\lambda}{4}\right)+(1/4)\int_o^\infty t\frac{g_{p-2,\lambda}(t)}{g_{p,\lambda}(t)}g_{p-2,\lambda}(t)dt.$$

Now using $tg_{p-2,\lambda}(t) = (p-2)g_{p,\lambda}(t)+\lambda g_{p+2,\lambda}(t)$

we have

$$I_{p,\lambda} = 1/2+\frac{\lambda}{4}\int_o^\infty \left(\frac{g_{p-2,\lambda}(t)g_{p+2,\lambda}(t)}{g^2_{p,\lambda}(t)}-1\right)g_{p,\lambda}(t)dt.$$

One can write

$$\frac{g_{p-2,\lambda}(t)g_{p+2,\lambda}(t)}{g^2_{p,\lambda}(t)} = \frac{I_{\nu-1}(\sqrt{\lambda t})I_{\nu+1}(\sqrt{\lambda t})}{I^2_\nu(\sqrt{\lambda t})} \overset{\wedge}{=} \varphi(\sqrt{\lambda t})$$

where I_ν is a modified Bessel function of the first kind and $\nu = p/2-1$. Using the development in series of the product of two Bessel functions due to Schläfli found in Watson (1962) ((5) of section 5.41) one has

$$\varphi(x) = \frac{\sum_{j\geq 0} \frac{(x/2)^{2(\nu+j)}\Gamma(2(\nu+j)+1)}{j!\,\Gamma(2\nu+j+1)\,\Gamma(\nu+j)\,\Gamma(\nu+j+2)}}{\sum_{j\geq 0} \frac{(x/2)^{2(\nu+j)}\Gamma(2(\nu+j)+1)}{j!\,\Gamma(2\nu+j+1)\Gamma^2(\nu+j+1)}}$$

The ratio of coefficients is $\frac{\nu+j}{\nu+j+1}$ which is increasing in j. By a result of Lehmann (1959) φ is increasing in x. A consequence of lemma 3.1 is that $\lim_{x\to\infty} \varphi(x) = 1$. Thus $I_{p,\lambda} \leq 1/2$ and the assertion has been verified. □

Actual computations for p = 4,5,6,10, and 20 and λ = .1,.2,.5,.7, 1,2,5,10,15,20,50 and 100 show that $I_{p,\lambda}$ is very close to .5 in all cases. In every case it is greater than .4994.

If $\{h_\lambda\}_{\lambda\geq 0}$ is a family of functions $h_\lambda: (a,b) \to (0,\infty)$ satisfying $\frac{h_{\lambda_1}(x)}{h_{\lambda_0}(x)} \uparrow$ in x for all $0 \leq \lambda_0 < \lambda_1 < \infty$ and if $\psi: (a,b) \to \mathbb{R}$ satisfies $\psi(x) \geq 0$ for $x \in (a,x_o]$ and $\psi(x) \leq 0$ for $x \in [x_o,b)$ then

$$\int_a^b \psi h_\lambda \leq \frac{h_\lambda(x_o)}{h_0(x_o)} \int_a^b \psi h_0 \tag{3.2}$$

for all $\lambda \geq 0$ as the reader can verify.

<u>Lemma 3.2</u> If $\Delta_o = R(\xi_o^+,0)-R(\xi_o,0)$ then $\Delta_o < 0$ and for all $\lambda \geq 0$

$$R(\xi_o^+,\lambda) - R(\xi_o,\lambda) \leq \frac{g_{p,\lambda}(p-2\sqrt{p-1})}{g_{p,o}(p-2\sqrt{p-1})}\,\Delta_o. \tag{3.3}$$

<u>Proof:</u> Use Stein's (1973) formula (6) of section 3 to get

$$R(\xi_o^+,\lambda)-R(\xi_o,\lambda) = \int_0^{p-2} [\frac{(p-2)^2}{t} + t -2p]g_{p,\lambda}(t)dt. \tag{3.4}$$

Using (3.2) on (3.4) with $h_\lambda = g_{p,\lambda}$ and $\psi(x) = \frac{(p-2)^2}{x} + x-2p$ yields (3.3).

Obviously $\Delta_o < 0$. □

It now follows that

$$\inf_{\xi \in \mathcal{A}} R(\xi,\lambda) \geq R(\xi_o,\lambda) - \left(2+ \frac{g_{p,\lambda}(p-2\sqrt{p-1})}{g_{p,o}(p-2\sqrt{p-1})} \Delta_o\right).$$

4. Global Bounds. We first prove that if ξ is an admissible spherically symmetric estimator which dominates ξ_o then $\int_0^\infty [R(\xi_o,\lambda)-R(\xi,\lambda)]d\lambda < 2(p-2)$. Suppose that $z_0(x), z_1(x), \ldots$ are positive, increasing, differentiable functions, $\{b_j\}_{j\geq o} > \{d_j\}_{j\geq o}$ are sequences of positive constants and the sums $\sum_{j\geq o} b_j z_j(x)$, $\sum_{j\geq o} d_j z_j(x)$, $\sum_{j\geq o} b_j z_j'(x)$, $\sum_{j\geq o} d_j z_j'(x)$ all converge for x in an interval I contained in the domain of all the z_i's. One can prove (see Spruill(1979)) that if b_j/d_j is monotone non-decreasing in j and $\frac{d}{dx} \ell n\, z_j(x)$ is monotone non-decreasing in j for each $x \in I$ then the function

$$(4.1) \qquad f(x) = \frac{\sum_{j\geq 0} b_j z_j(x)}{\sum_{j\geq 0} d_j z_j(x)}$$

is monotone non-decreasing in x. Let $g_p(t) = \lim_{x\to\infty} g_p(x,t)$.

Lemma 4.1 If ξ is a spherically symmetric estimator satisfying $R(\xi,\lambda) \leq R(\xi_o,\lambda)$ for all $\lambda \geq 0$ then

$$(4.2) \quad \int_0^\infty [R(\xi_o,\lambda) - R(\xi,\lambda)]d\lambda \leq 4 \int_0^\infty t\left(\frac{g_p'(t)}{g_p(t)}\right)^2 g_p(t)dt.$$

Proof: According to (2.6) and (3.1) for all $x > 0$

$$(4.3) \quad \int_0^x [R(\xi_o,\lambda)-R(\xi,\lambda)]d\lambda \leq 4 \int_0^\infty t\left(\frac{g_p'(x,t)}{g_p(x,t)}\right)^2 g_p(x,t)dt.$$

For t fixed

$$\frac{t g_{p-2}(x,t)}{g_p(x,t)} = \frac{\sum_{j\geq 0} \eta_j(x)\left(\frac{t}{4}\right)^j \frac{1}{j!} \frac{1}{\Gamma(p/2+j-1)}}{\sum_{j\geq 0} \eta_j(x)\left(\frac{t}{4}\right)^j \frac{1}{j!} \frac{1}{\Gamma(p/2+j)}}$$

where $\eta_j(x) = \int_0^x \lambda^j e^{-\lambda/2}\, d\lambda$. Identifying $z_j(x) = \eta_j(x)$, $\frac{b_j}{d_j} = \frac{p}{2} + j-1$ in (4.1)

it follows that $\frac{tg_{p-2}(x,t)}{g_p(x,t)}$ is non-decreasing in x. Thus $\frac{g_p'(x,t)}{g_p(x,t)}$ is non-decreasing in x for t fixed. Let $I_x(t)$ be the indicator function of the set $\{t:\ g_p'(x,t)>0\}$. We have

$$I_x(t)\ \left(\frac{g_p'(x,t)}{g_p(x,t)}\right)^2\ g_p(x,t)$$

non-decreasing in x for t fixed. The monotone convergence theorem now yields the desired result upon taking the limit as $x \to \infty$ on both sides of (4.3). □

Theorem 4.1 If ξ is a spherically symmetric estimator which satisfies $R(\xi,\lambda) \leq R(\xi_o,\lambda)$ for all $\lambda \geq 0$ then $\int_0^\infty [R(\xi_o,\lambda)-R(\xi,\lambda)]d\lambda < 2(p-2)$.

Proof: We have

$$g_p(t) = \left(\frac{t}{2}\right)^{p/2-1} e^{-t/2} \sum_{j\geq 0}\left(\frac{t}{2}\right)^j \frac{1}{\Gamma(p/2+j)}$$

and $g_{p-2}(t) = \left(\frac{t}{2}\right)^{p/2-1} e^{-t/2} \sum_{j\geq -1}\left(\frac{t}{2}\right)^j \frac{1}{\Gamma(p/2+j)}$

so

$$\int_0^\infty t\ \left(\frac{g_{p-2}(t)-g_p(t)}{g_p(t)}\right)^2 dt = \int_0^\infty t\left((t/2)^{p/2-2}\ \frac{e^{-t/2}}{\Gamma(p/2-1)}\right)^2 \frac{1}{g_p(t)}\ dt$$

$$(4.4)\qquad = 4\int_0^\infty \frac{t^{p/2-2}e^{-t/2}}{\Gamma(p/2-1)2^{p/2-1}}\ \left(\Gamma(p/2-1)\sum_{j\geq o}(t/2)^j\ \frac{1}{\Gamma(p/2+j)}\right)^{-1}dt$$

$$< \frac{4\ \Gamma(p/2)}{\Gamma(p/2-1)} = 2(p-2).$$

The assertion now follows directly from (4.2). □

Corollary. If ξ is a spherically symmetric estimator which satisfies

$R(\xi,\lambda) \leq R(\xi_o^+,\lambda)$ for all $\lambda \geq 0$ then

(4.5) $\int_0^\infty [R(\xi_o^+,\lambda)-R(\xi,\lambda)]d\lambda < 2(p-2+\Delta_o)$.

Proof: We also have $R(\xi,\lambda) \leq R(\xi_o,\lambda)$ and expressing the integrand on the left hand side of (4.5) as the sum of $R(\xi_o,\lambda) - R(\xi,\lambda)$ and $R(\xi_o^+,\lambda) - R(\xi_o,\lambda)$ it suffices to prove

(4.6) $\int_0^\infty [R(\xi_o^+,\lambda) - R(\xi_o,\lambda)]d\lambda < 2\Delta_o$.

Using (3.3) and $\int_0^\infty g_{p,\lambda}(t_o)\, d\lambda > 2g_{p,0}\ (t_o)$ where $t_o = p-2\sqrt{p-1}$ verifies (4.6). □

Numerical evaluation of (4.4) provides the entries in the first row of table 1 which are the bounds on $\int_0^\infty [R(\xi_o,\lambda) - R(\xi,\lambda)]d\lambda$ whenever $R(\xi,\lambda) \leq R(\xi_o,\lambda)$ for all $\lambda \geq 0$. The areas $\int_0^\infty [R(\xi_o^+,\lambda)-R(\xi_o,\lambda)]d\lambda$ were numerically evaluated utilizing (3.4) with $g_{p,\lambda}$ replaced by g_p and appear in the second row of the table. The final row is the difference and is an upper bound on $\int_0^\infty [R(\xi_o^+,\lambda)-R(\xi,\lambda)]d\lambda$ for any spherically symmetric estimator ξ which satisfies $R(\xi,\lambda) \leq R(\xi_o^+,\lambda)$ for all $\lambda \geq 0$.

Acknowledgement: I thank the referees whose comments contributed to the improvement of the original version of this paper and James O. Berger for discussions on this topic.

p

3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1.52	2.58	3.41	4.11	4.72	5.28	5.79	6.27	6.71	7.13	7.53	7.92	8.28	8.63	8.97	9.30	9.62	9.93
1.12	1.76	2.26	2.68	3.05	3.38	3.69	3.97	4.23	4.48	4.72	4.95	5.16	5.37	5.57	5.77	5.95	6.13
.40	.82	1.15	1.43	1.67	1.90	2.10	2.30	2.48	2.65	2.81	2.97	3.12	3.26	3.40	3.53	3.67	3.80

Table 1

Bounds on areas between risk functions

References

Balder, E. J. (1980). An extension of the usual model in statistical decision theory with applications to stochastic optimization problems. *J. Multivariate Anal.* **10** 385-397.

Berger, J. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma scale parameters. *Ann. Statist.* **8** 545-571.

Bourbaki, N. (1966). *Elements of Mathematics, General Topology, Part 1*. Addison-Wesley, Reading, Massachussetts.

Brown, L. D. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Statist.* **42** 855-903.

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

Gihman, I. I. and Skorohod, A. V. (1974). *The Theory of Stochastic Processes I.* Springer-Verlag, New York.

Hudson, H. M. (1978). A natural identity for exponential families with applications in multiparameter estimation. *Ann. Statist.* **6** 473-484.

Hwang, J. T. (1980). Semi tail upper bounds on the class of admissible estimators with applications to Poisson and negative binomial families. Technical report, Cornell University.

James, W. and Stein, C. (1960). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 361-379. Univ. of Calif. Press.

Lehmann, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

Moore, T. and Brook, R. J. (1978). Risk estimate optimality of James-Stein estimators. *Ann. Statist.* **6** 917-919.

Spruill, M. C. (1979). Bounds on the performance of admissible spherically symmetric estimators better than the James-Stein estimator. Technical Report, School of Mathematics, Georgia Institute of Technology.

Spruill, M. C. (1980). Estimation of the non-centrality parameter of a chi-squared distribution. Technical Report, School of Mathematics, Georgia Institute of Technology.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Prob.* 1 197-206.

Stein, C. (1973). Estimation of the mean of a multivariate distribution. *Proc. Prague Symp. Asymptotic Statist.* 345-381.

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

Strawderman, W. E. and Cohen, A. (1971). Admissibility of estimators of the mean vector of a multivariate normal distribution with quadratic loss. *Ann. Math. Statist.* **42** 270-296.

Watson, G. N. (1962). *A Treatise on the Theory of Bessel Functions.* Cambridge University Press, Cambridge.