The Annals of Statistics
1978, Vol. 6, No. 4, 917–919

RISK ESTIMATE OPTIMALITY OF JAMES–STEIN ESTIMATORS

BY TERRY MOORE AND RICHARD J. BROOK

Massey University

This note extends a result of Efron and Morris on domination of the maximum likelihood estimator for the mean of a multivariate normal distribution. We show that this result and our extension follow from a certain differential inequality.

In a certain class of estimators having a unique unbiased estimator for the quadratic risk we find necessary and sufficient conditions for risk estimate dominance of a particular set of estimators. We show that, in the sense of risk estimates, these conditions imply that there are no estimators in this class which dominate the James–Stein or truncated James–Stein estimators.

We wish to estimate the unknown mean, ξ, of a $k \geqq 3$ dimensional random normal vector, X, with variance matrix $\sigma^2 I$. We suppose that we have observed X and S where S/σ^2 has a chi-squared distribution with n degrees of freedom. Our results also include the case of known variance in a similar way to Efron and Morris (1976).

Let $\lambda = \|\xi\|^2/(2\sigma^2)$, let the loss function be $L(\hat{\xi}, \xi) = \|\hat{\xi} - \xi\|^2/\sigma^2$ and let $\tau = \tau(F)$ be absolutely continuous with derivative τ'. We consider estimators of the form

(1) $\hat{\xi}_\tau = (1 - (\tau(F) + 1)/F)X$ where $F = [\|X\|^2/(k-2)]/[S/(n+2)]$, and where the risk $R_\tau(\lambda)$ is finite and the expectation of each term in (2) exists.

Efron and Morris prove the following extension of a theorem of Stein.

THEOREM 1. *There is a unique unbiased estimator based on F for $R_\tau(\lambda)$ given by*

$$\hat{R}_\tau(F) = k - [(k-2)/F](1 - \tau^2(F)) - 4\tau'(F)(1 + c + c\tau(F)) \tag{2}$$

where $c = (k-2)/(n+2)$.

Note that when $c \neq 0$ the expectation of each term in (2) does not exist if

$$\tau^{-1}(F) = O(F^{-(n+2)/4}) \quad \text{as} \quad F \to \infty \qquad \text{or} \tag{3}$$
$$\tau^{-1}(F) = O(F^{(k-2)/4}) \quad \text{as} \quad F \to 0\,.$$

Suppose the estimator $\hat{\xi}_f$ is in the class considered, then we can compare $\hat{\xi}_\tau$ with $\hat{\xi}_f$ by comparing their risk estimates. If these risk estimates are such that $\hat{R}_\tau < \hat{R}_f$, i.e., $\hat{R}_\tau(F) \leqq \hat{R}_f(F)$ for all F and it is false that $\hat{R}_\tau(F) = \hat{R}_f(F)$ almost

Received April 1977; revised June 1977.

AMS 1970 *subject classifications.* Primary 62F10; Secondary 62C99.

Key words and phrases. James–Stein estimator, truncated James–Stein estimator, mean of a multivariate normal distribution, unbiased estimator of the risk, risk estimate dominance.

everywhere, then $\hat{\xi}_\tau$ is said to dominate $\hat{\xi}_f$ in terms of risk estimates. This risk estimate dominance property is clearly weaker than dominance, since, although $\hat{R}_\tau < \hat{R}_f \Rightarrow R_\tau < R_f$, it is not true that $R_\tau < R_f \Rightarrow \hat{R}_\tau < \hat{R}_f$. However, as Efron and Morris (1976) show, this property is strong enough to imply that certain estimators do in fact dominate the maximum likelihood estimator, in cases for which this fact was not previously known. Our purpose is to show that James–Stein estimators are optimal in the class considered in terms of risk estimates. This shows that these estimators are so good that very powerful methods will be needed to obtain alternatives which dominate them. We remark that James–Stein estimators have negative estimates of the risk for small F.

Now $\hat{R}_\tau \leqq \hat{R}_f$ if and only if

$$(1 + b\tau(F))\tau'(F) - a\tau^2(F)/F \geqq (1 + bf(F))f'(F) - af^2(F)/F \tag{4}$$

where

$$a = (k-2)/(4(1+c)), \qquad b = c/(1+c).$$

In the case of known variance, $b \neq 0$, this differential inequality has not been solved in general, while for $b = 0$ it has. The former case is Abel's differential inequality, the latter is Riccati's. Abel's differential equation is to be found in Kamke (1948) and Riccati's appears in Ince (1939). The case most easily solved is that for which $f = t =$ constant, or more generally

$$(1 + bf(F))f'(F) - af^2(F)/F = \alpha a/F \tag{5}$$

where (3) implies that $\alpha \geqq 0$.

In this case, solution of the inequality (4) gives

THEOREM 2. *A necessary and sufficient condition that $\hat{R}_\tau \leqq \hat{R}_t$ is that*

$$\begin{aligned} \psi(F) &= F^{-a}|\tau^2(F) - t^2|^{b/2}|(\tau(F) - |t|)/(\tau(F) + |t|)|^{1/(2|t|)} && t \neq 0 \\ &= F^{-a}\tau^b(F)\exp(-\tau^{-1}(F)) && t = 0 \end{aligned} \tag{6}$$

is a nonincreasing function of F when $\tau^2(F) < t^2$ and nondecreasing when

$$\tau^2(F) > t^2.$$

This theorem together with the next generalise Theorem 2 of Efron and Morris (1976).

THEOREM 3. *Necessary conditions that $\hat{R}_\tau \leqq \hat{R}_t$ are that*

$$-|t| \leqq \tau \leqq |t| \tag{7}$$

$$\text{if } \exists F_1 \text{ such that } \tau(F_1) = -|t| \text{ then } \tau(F) = -|t| \quad \forall F < F_1 \tag{8}$$

$$\text{if } \exists F_2 \text{ such that } \tau(F_2) = |t| \text{ then } \tau(F) = |t| \quad \forall F > F_2 \tag{9}$$

(condition (8) *is not necessary if $|t| \geqq 1 + (1/c)$, $c \neq 0$).*

The proof of Theorem 3 is similar to that of Efron and Morris ((1976), Theorem 2), but uses (3) when $c \neq 0$ which avoids quoting a result, true only for $|t| = 1$, in Efron and Morris (1973).

Since $\hat{\xi}_0$ is the James–Stein estimator, condition (7) implies

THEOREM 4. *No estimator in the class considered dominates the James–Stein estimator in terms of risk estimate.*

Our final result is that this also applies to truncated James–Stein estimators. Let

$$\begin{aligned} f(F) &= t & F &\geqq 1 + t \\ &= F - 1 & F &\leqq 1 + t \qquad \text{where} \quad 0 \leqq t \leqq 1 \,. \end{aligned}$$

Denote $\hat{\xi}_f$, $\hat{R}_f$ and R_f respectively by $\hat{\xi}_t{}^+$, $\hat{R}_t{}^+$ and $R_t{}^+$.

THEOREM 5. *If $\hat{R}_\tau \leqq \hat{R}_t{}^+$ then $\tau = f$.*

The proof of this uses the argument of Theorem 3 to show that if $\exists\, F_0$ such that $\tau(F_0) < f(F_0)$ then $\tau(0) < -1$. Thus by Theorem 3, $\hat{R}_\tau \nleqq \hat{R}_{-1}$ (where $\hat{\xi}_{-1}$ is the maximum likelihood estimator). However, by Theorem 3, $\hat{R}_t{}^+ \leqq \hat{R}_{-1}$ and this is a contradiction.

REFERENCES

[1] EFRON, BRADLEY and MORRIS, CARL (1973). Stein's estimation rule and its competitors—an empirical Bayes approach. *J. Amer. Statist. Assoc.* **68** 117–130.

[2] EFRON, BRADLEY and MORRIS, CARL (1976). Families of minimax estimators for the mean of a multivariate normal distribution. *Ann. Statist.* **4** 11–21.

[3] INCE, E. L. (1939, reprinted 1963). *Integration of Ordinary Differential Equations.* Oliver and Boyd, Edinburgh.

[4] KAMKE, E. (1948). Differentialgleichungen: Losungsmethoden und Losungen. Chelsea, New York.

DEPARTMENT OF MATHEMATICS
MASSEY UNIVERSITY
PALMERSTON NORTH
NEW ZEALAND