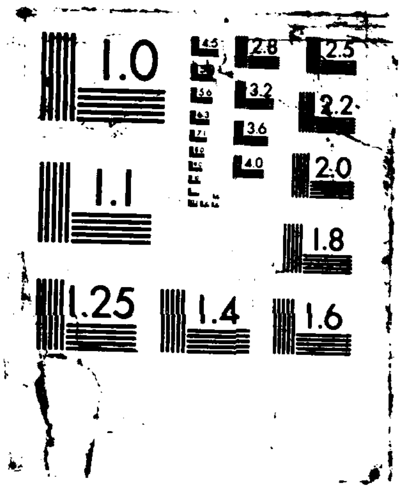
AD-A191 108 IMPROVED ESTIMATORS IN SIMULTANEOUS ESTIMATION OF SCALE PARAMETERS(U) STANFORD UNIV CA DEPT OF STATISTICS D K DEY ET AL. 10 DEC 87 TR-399 N00014-86-K-0156 1/1

UNCLASSIFIED F/G 12/3 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

IMPROVED ESTIMATORS IN SIMULTANEOUS ESTIMATION OF SCALE PARAMETERS

BY

DIPAK K. DEY and ALAN E. GELFAND

TECHNICAL REPORT NO. 399
DECEMBER 10, 1987

Prepared Under Contract
N00014-86-K-0156 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

IMPROVED ESTIMATORS IN SIMULTANEOUS ESTIMATION OF SCALE PARAMETERS

Dipak K. Dey and Alan E. Gelfand

Key Words and Phrases: simultaneous estimation; scale parameters; inadmissibility.

ABSTRACT

A general class of estimators is developed for improving upon best scale invariant estimators of two or more arbitrary scale parameters (or powers thereof) for arbitrary positive distributions with sufficient moments under weighted squared error loss function. The technique is to compute the risk difference in terms of moments of the distribution. Some conditions are obtained under which the maximum improvement is possible, and the form of the estimator can be chosen to achieve this maximum along any specified ray.

The result is then extended to the estimation of a linear transform of the parameter vector. Finally, some examples are given with numerical calculations to obtain the amount of risk improvement.

1. INTRODUCTION

There has been considerable interest in improving upon the standard estimators in multivariate estimation problems. Since the celebrated work of Stein (1955), numerous results have been proved which show that the presence of Stein effect is just part of a very general phenomenon, which has little to do with the exact form of the loss function or the underlying distribution.

The simultaneous estimation of location parameters was investigated by several authors during the past decade culminating in a unified discussion by Shinozaki (1984). The improved estimation of scale parameters from exponential families has also been studied recently. The major results in this direction are obtained by Hudson (1978), Berger (1980), and Ghosh, Hwang and Tsui (1984). The basic technique for obtaining improved estimators of scale parameters in exponential families is to obtain an unbiased estimator of the risk, using integration by parts technique and solving a differential inequality. Recently DasGupta (1986), however, obtained improved estimators of gamma scale parameters without a variational argument by proposing an estimator which is a function of the geometric mean.

In this paper we extend DasGupta (1984, 1986) in several directions. We consider the problem of simultaneous estimation of arbitrary powers of scale parameters from arbitrary independent positive valued distributions under weighted squared error loss. All that is required is that moments of a specified order be finite.

Suppose $X = (X_1,\dots,X_p)$, $p \geq 2$, where X_i has density $f_{\theta_i}(x_i) = \frac{1}{\theta_i} f_i(x_i/\theta_i)$, $x_i > 0$, $\theta_i > 0$, $i = 1,\dots,p$. Also suppose $E_{\theta_i}(X_i^{2s_i}) < \infty$, $i = 1,\dots,p$. Our problem is to estimate $\theta^s = (\theta_1^{s_1},\dots,\theta_p^{s_p})$ under the loss given as

$$L(a,\theta^s) = \sum_{i=1}^{p} w_i \theta_i^{c_i} (a_i - \theta_i^{s_i})^2 . \tag{1.1}$$

It can be easily checked that, under the loss (1.1), the best invariant estimator of θ^s is $\delta^0(X) = (\delta_1^0(X),\dots,\delta_p^0(S))$, where $\delta_i^0(X)$ is given as

$$\delta_i^0(X) = r_{i,s_i,2s_i} X_i^{s_i}, \quad i = 1,\dots,p, \tag{1.2}$$

where $r_{i,s_i,2s_i} = E(X_i^{s_i})/E(X_i^{2s_i})$, the expectation $E(\cdot)$ taken under $\theta_i = 1$. In Section 2, we propose a class of improved estimators which dominate $\delta^0(X)$. Another class of estimators are obtained which incorporates a prior guess about the parameter vector.

In Section 3, we study the component risk behavior of certain "best" estimators in the class obtained in Section 1. By aggregating the component risk, we can obtain an explicit form for the improvement in risk and identify the direction in which maximum improvement is attained.

Section 4 is devoted to the estimation of a linear combination $\ell'\theta^s$ of the parameter vector, where the vector $\ell = (\ell_1, \ldots \ell_p)'$ is $p \times 1$. Several estimators are obtained which dominate aX^s where $a = (a_1, , , . a_p)'$ is $p \times 1$ vector and $X^s = (X_1^{s_1}, \ldots, X_p^{s_p})$. Some examples are given which include estimation of $\sum_i \theta_i$.

Finally, in Section 5 we give examples of simultaneous estimation of ratio of independent normal variances and estimation of ranges of rectangular distributions. Some numerical studies are performed which give the percentage improvements in risks.

2. IMPROVED ESTIMATORS OF θ^s

In this section we will develop estimators of θ^s which will dominate the best invariant estimator (1.2) under the loss given in (1.1) for $p \geq 2$.

Brown (1966) and Brown and Fox (1974) provide perhaps the most general conditions for admissibility of the best invariant scale parameter in the one dimensional case. By transforming the scale problem to the equivalent location problem, it is straightforward to verify that $r_{s,2s}X^s$ is an admissible estimation for θ^s under squared error loss if $EX^{2s}(\ell n X)^2 < \infty$ where the expectation is taken at $\theta = 1$. That is, referring to conditions (a) - (d) on p. 808 of Brown and Fox, since the best invariant estimator is unique (a) and (b) are apparent. The expectation condition immediately implies

(c) and (d) as well by appealing to Lemma 2.3.3, p. 1105, of Brown (1966).

Now we will develop the following notation:

$$\begin{aligned}
&(1)\quad m_{i,\alpha} = E_1 X_i^{\alpha};\\
&(2)\quad r_{i,\alpha_i,\beta_i} = m_{i,\alpha_i}/m_{i,\beta_i};\\
&(3)\quad g_\alpha(X) = \prod_j X_j^{\alpha_j};\\
&(4)\quad E_1 g_\alpha(X) = \prod_j m_{j,\alpha_j} = \rho_\alpha;\\
&(5)\quad E_1 g_\alpha^2(X) = \prod_j m_{j,2\alpha_j} = \rho_{2\alpha}.
\end{aligned} \tag{2.1}$$

With the above notation, it follows that $E_{\theta_i}(X_i^r) = \theta_i^r m_{i,r}$, $i = 1,\ldots,p$. Classes of improved estimators are given in Theorem 2.1, but first we note

<u>Lemma 2.1</u>. For any α_j, $j = 1,\ldots,p$,

$$p\prod_j \theta_j^{\alpha_j/p} / \sum_j \theta_j^{\alpha_j} \le 1.$$

<u>Proof</u>. The arithmetic mean-geometric mean inequality.

<u>Lemma 2.2</u>. Under one of the following conditions,

(i) $c > 0$ and $\tau \ge 0$,

(ii) $c < 0$ and $\tau \le 0$,

(iii) $c = 0$ and arbitrary τ,

the following inequality holds:

$$\prod_j \theta_j^{\tau/p} \sum_j \theta_j^{c} / \sum_j \theta_j^{c+\theta} \le 1.$$

<u>Proof</u>. The proof for (i) and (ii) follows from Hölder's inequality.

<u>Lemma 2.3</u>. $r_{i,t+s,t}$ as defined in (2.1) is an increasing function of s if $t > 0$, a decreasing function of s if $t < 0$.

<u>Proof</u>. It is clear that X_i^s is monotone in X_i and $h_i(x,t) = x_i^t f_i(x) / \int x_i^t f_i(x)\, dx$ is MLR in X_i.

Thus EX_i^{t+s}/EX_i^t is increasing in s if $t > 0$, and decreasing in s if $t < 0$.

We are now ready for

Theorem 2.1. Consider the estimator $\delta(X)$ given componentwise as

$$\delta_i(X) = \delta_i^0(X) - b\,\mathrm{sgn}[(\alpha_i + \beta_i - s_i)s_i]X_i^{p_i} g_\alpha(X), \qquad (2.2)$$

$$i = 1,\dots,p(\geq 2),$$

where $\alpha_i + \beta_i - s_i \neq 0$ for all $i = 1,\dots,p$. Then $\delta(X)$ dominates $\delta^0(X)$ in terms of risk if either (1) or (1') and (2) hold:

(1) $\alpha_i = (c_i + \beta_i + s_i)/p$ and $\beta_i = -c_i/2$, $i = 1,\dots,p$; (2.3a)

(1') All c_i equal, all β_i equal, all s_i equal and $\alpha_i = \alpha = (s-\beta)/p$; (2.3b)

and

(2) $0 < b < 2\rho_\alpha d^{(1)}/\rho_{2\alpha}d^{(2)}$ (2.4)

where $d^{(1)}$ and $d^{(2)}$ are positive and defined below.

Proof. Let $\Delta(\theta) = R(\theta,\delta) - R(\theta,\delta_0)$, the risk difference. It is sufficient to show that $\Delta(\theta) < 0$. Now, using the loss (1.1), it follows that

$$\Delta(\theta) = E_\theta[\sum_i w_i\theta_i^{c_i}\{b_i^2 X_i^{2\beta_i} g_\alpha^2(X) + 2r_{i,s_i,2s_i}X_i^{s_i}\, b_i X_i^{\beta_i}\, g_\alpha(X)$$

$$- 2b_i\theta_i^{s_i}X_i^{\beta_i}g_\alpha(X)\}], \quad b_i = -b\ \mathrm{sgn}[(\alpha_i + \beta_i - s_i)s_i],$$

$$= \sum_i w_i\theta_i^{c_i}\{b_i^2\theta_i^{2\beta_i}g_{2\alpha}(\theta)(\prod_i^p m_{j,2\alpha_j})m_{i,2(\alpha_i+\beta_i)}/m_{i,2\alpha_i}$$

$$+ 2r_{i,s_i,2s_i}\theta_i^{\beta_i+s_i}b_i g_\alpha(\theta)(\prod_j m_{j,\alpha_j})m_{i,\alpha_i+\beta_i+s_i}/m_{i,\alpha_i}$$

$$- 2b_i\theta_i^{\beta_i+s_i}g_\alpha(\theta)(\prod_j m_{j,\alpha_j})m_{i,\alpha_i+\beta_i}/m_{i,\alpha_i}\}$$

$$= \sum_i w_i \theta_i^{c_i} \{ b_i^2 \theta_i^{2\beta_i} g_{2\alpha}(\theta) \rho_{2\alpha} r_{i,2(\alpha_i+\beta_i),2\alpha_i}$$

$$+ 2 r_{i,s_i,2s_i} b_i \theta_i^{\beta_i+s_i} g_\alpha(\theta) \rho_\alpha r_{i,\alpha_i+\beta_i+s_i,\alpha_i}$$

$$- 2 b_i \theta_i^{\beta_i+s_i} g_\alpha(\theta) \rho_\alpha r_{\alpha_i+\beta_i,\alpha_i} \}$$

$$= g_{2\alpha}(\theta)\rho_{2\alpha} \sum_i w_i b_i^2 \theta_i^{2\beta_i+c_i} r_{i,2(\alpha_i+\beta_i),2\alpha_i}$$

$$+ 2 g_\alpha(\theta)\rho_\alpha \sum_i w_i \theta_i^{c_i+\beta_i+s_i} b_i d_i \tag{2.5}$$

where $d_i = r_{i,s_i,2s_i} \cdot r_{i,\alpha_i+\beta_i+s_i,\alpha_i} - r_{i,\alpha_i+\beta_i,\alpha_i}$, $i = 1,\ldots,p$.

Now, using Lemma 2.3, it follows that for $s_i > 0$ if $\alpha_i + \beta_i < s_i$, then $d_i < 0$; and if $\alpha_i + \beta_i > s_i$, then $d_i > 0$. Similarly for $s_i < 0$, if $\alpha_i + \beta_i < 0$, then $d_i < 0$; and if $\alpha_i + \beta_i > s_i$, then $d_i < 0$. Note that $b_i d_i$ is always negative. Now, define

$$d^{(1)} = \min_i w_i |d_i|$$

$$d^{(2)} = \max_i w_i \, r_{i,2(\alpha_i+\beta_i),2\alpha_i}.$$

It follows from (2.5) that

$$\Delta(\theta) < g_{2\alpha}(\theta)\rho_{2\alpha} b^2 d^{(2)} \Sigma \theta_i^{c_i+2\beta_i} - 2 g_\alpha(\theta)\rho_\alpha b d^{(1)} \Sigma \theta_i^{c_i+\beta_i+s_i}$$

$$= g_\alpha(\theta) \Sigma \theta_i^{c_i+\beta_i+s_i} \{ b^2 \rho_{2\alpha} d^{(2)} K(\theta) - 2bd^{(1)} \rho_\alpha \}, \tag{2.6}$$

where

$$K(\theta) = g_\alpha(\theta) \Sigma \theta_i^{c_i+2\beta_i} / \Sigma \theta_i^{c_i+\beta_i+s_i}.$$

If (1) of (2.3a) holds, then from Lemma 2.1 we have $K(\theta) < 1$, i.e., $\Delta(\theta) < 0$ if (2.4) holds. If (1') of (2.3b) holds, then with Lemma 2.2 we take $\tau = s - \beta$ and $\alpha = (s-\beta)/p$. Note that $\alpha - \beta \neq s$, i.e., $(s-\beta)/p + \beta \neq s$ requires $\beta \neq s$ and if $c+2\beta > 0 \Rightarrow \tau \geq 0$, i.e., $\beta < s \Rightarrow -c < 2\beta < 2s$, if $c > -2s$

if $c+2\beta < 0 \Rightarrow \tau \leq 0$, i.e., $\beta > s \Rightarrow 2s < 2\beta < -c$, if $c < -2s$

$c+2\beta = 0 \Rightarrow c \neq -2s$.

Again from (2.6), it follows $K(\theta) < 1$ so that $\Delta(\theta) < 0$ if (2.4) holds. This completes the proof of the theorem.

Remark 2.1. Theorem 2.1 is a more general version of Theorem 2.1 in Das Gupta (1984).

Remark 2.2. From the proof of Theorem 2.1, it follows that if for any i, $\alpha_i + \beta_i = s_i$, i.e., $c_i = -2s_i$, the dominance result does not follow. Thus, under the invariant quadratic loss, our result does not quite give an improved estimator. However, if a minor restriction is taken on the parameter space, say $\min_i \theta_i^{s_i} > \varepsilon$, where ε is a preassigned positive number, then estimators of our form will dominate the best invariant estimator under the invariant quadratic loss.

Now we will consider further generalization of Theorem 2.1. For example, suppose we have a prior guess for θ, say, $\theta_0 = (\theta_{01},\ldots,\theta_{0p})$. How might we incorporate this information in our improved estimator? The following theorem gives the improved estimator in this case.

Theorem 2.2. Consider the estimator $\delta^{\theta_0}(X)$ given componentwise as

$$\delta_i^{\theta_0}(X) = \delta_i^0(X) - b\,\mathrm{sgn}[(\alpha_i+\beta_i-s_i)s_i](X_i/\theta_{0i})^{\beta_i}\, g_\alpha(X/\theta_0),$$

$$i = 1,\ldots,p. \tag{2.7}$$

Then $\delta^{\theta_0}(X)$ dominates $\delta^0(X)$ in terms of risk if either (2.3a) or (2.3b) holds and

$$0 < b < 2g_\alpha(\theta_0)\rho_\alpha d^{*(1)}/\rho_{2\alpha} d^{*(2)}, \tag{2.8}$$

where $d^{*(1)}$ and $d^{*(2)}$ are positive and given below.

Proof. Define $k_i = \theta_{0i}^{-\beta_i}$, $i = 1,\ldots,p$. Now, we can write

$$\delta_i^{\theta_0}(X) = r_{i,s_i,2s_i} X_i^{s_i} + b_i k_i g_\alpha^{-1}(\theta_0) g_\alpha(X) X_i^{\beta_i}$$

where

$$b_i = -b \operatorname{sgn}[(\alpha_i+\beta_i-s_i)s_i], \quad i = 1,\ldots,p.$$

Now, it clearly follows that the risk difference

$$\begin{aligned}\Delta(\theta) = \sum_i w_i \theta_i^{c_i} \{ & b_i^2 g_{-\alpha}^2(\theta_0) k_i^2 E X_i^{2\beta_i}(g_\alpha^2(X)) \\ & + 2b_i r_{i,s_i,2s_i} g_{-\alpha}(\theta_0) k_i E(w_i^{s_i} X_i^{\beta_i} g_\alpha(X)) \\ & - 2b_i \theta_i^{s_i} k_i g_{-\alpha}(\theta_0) E(X_i^{\beta_i} g_\alpha(X))\} \\ = {} & g_{2\alpha}(\theta) g_{-2\alpha}(\theta_0) \rho_{2\alpha} \sum_i w_i \theta_i^{c_i+2\beta_i} b_i^2 k_i r_{i,2(\alpha_i+\beta_i),2\alpha_i} \\ & + 2 g_\alpha(\theta) \rho_\alpha g_{-\alpha}(\theta_0) \sum_i w_i \theta_i^{c_i+\beta_i+s_i} \\ & \times b_i \{ r_{i,s_i,2s_i} r_{i,\alpha_i+\beta_i+s_i,\alpha_i} - r_{i,\alpha_i+\beta_i,\alpha_i} \}. \end{aligned} \tag{2.9}$$

Define

$$d^{*(1)} = \min_i w_i k_i |d_i|$$

and

$$d^{*(2)} = \max_i w_i k_i r_{i,2(\alpha_i+\beta_i),2\alpha_i}.$$

Now following calculations similar to that in proof of Theorem 2.1, it follows that

$$\begin{aligned}\Delta(\theta) < {} & g_{2\alpha}(\theta/\theta_0) \rho_{2\alpha} b^2 d^{*(2)} \sum_i \theta_i^{c_i+2\beta_i} \\ & - 2 g_\alpha(\theta/\theta_0) \rho_\alpha d^{*(1)} \sum_i \theta_i^{c_i+\beta_i+s_i}. \end{aligned} \tag{2.10}$$

As in the proof of Theorem 2.1, it follows that if (2.8) holds, then $\Delta(\theta) < 0$. This completes the proof.

Remark 2.3. Das Gupta (1984) considers a special case of estimators which in our notation would extend (2.2) to

$$\delta_i(X) = \delta_i^0(X) + b\ \mathrm{sgn}[(\alpha_i+\beta_i-s_i)s_i]X_i^{i}g_\alpha(X)\theta_{0i}^{s_i}. \tag{2.11}$$

It is clear that by transformation (2.11) is equivalent to (2.7).

Remark 2.4. Notice that if $\theta_{0i} = 1$, $i = 1,\ldots,p$, then the estimator (2.7) is the same as that of (2.2).

Remark 2.5. It is noteworthy that the range of b in (2.5) depends also on the values of θ_0.

Remark 2.6. The dominance result holds even if X_i's arise from different families of distributions, say, one from exponential and the other from uniform.

3. COMPONENT RISK AND MAXIMUM IMPROVEMENTS

In this section we study the component risk behavior of the estimators given in (2.2) and in (2.7) when $f_i(x) = f(x)$, $i = 1,\ldots,p$. In this case, we take $r_{i,\alpha,\beta} = r_{\alpha,\beta}$ and $m_{i,s} = m_s$.

The following theorem gives the maximum improvement of the estimator of the form (2.2) when b is the midpoint of (2.4) and the direction at which the maximum attains.

Theorem 3.1. Suppose X_i is distributed with pdf $f_{\theta_i}(x_i) = \frac{1}{\theta_i} f(x/\theta_i)$, $i = 1,\ldots,p$. Consider the estimator defined componentwise as

$$\delta_i(X) = r_{s,2s}X_i^s - bX_i^\beta \prod_j X_j^\alpha,\ i = 1,\ldots,p, \tag{3.1}$$

with $b = r^p_{\alpha,2\alpha}d/r_{2(\alpha+\beta),2\alpha}$

where

$$d = r_{s,2s}\cdot r_{\alpha+\beta+s,\alpha} - r_{\alpha+\beta,\alpha}. \tag{3.2}$$

Then the maximum improvement in risk when the loss is (1.1) with $c_i = c$ and $w_i = 1$ is attained when θ_i's are equal, say, to θ and the amount of improvement is

$$d^2 r_{2\alpha,2(\alpha+\beta)} m_\alpha^{2p} p\theta^{c+2s}/m_{2\alpha}. \tag{3.3}$$

<u>Proof</u>. The difference in risk for the ith component is $\Delta_i(\theta)$ where

$$\begin{aligned}\Delta_i(\theta) &= R(\delta_i^0,\theta_i^s) - R(\delta_i,\theta_i^s)\\ &= -\theta_i^c[E\{b^2X_i^{2\beta}\prod_j X_j^{2\alpha} - 2bX_i^\beta \prod_j X_j^\alpha(r_{s,2s}X_i^s - \theta_i^s)\}]\\ &= -\theta_i^c[b^2\theta_i^{2(\alpha+\beta)}(\eta_{-i}^{(\alpha)})^2 m_{2\alpha}^p m_{2(\alpha+\beta)}/m_{2\alpha}\\ &\quad -2b\theta_i^{\beta+\alpha+s}(\eta_{-i}^{(\alpha)})m_\alpha^p\{m_{\alpha+\beta+s}\, r_{s,2s}/m_\alpha - m_{\alpha+\beta}/m_\alpha\}]\\ &\quad (\text{where } \eta_{-i}^{(\alpha)} = \prod_{j\neq i} \theta_j^\alpha)\\ &= -\theta_i^c[b^2\theta_i^{2(\alpha+\beta)}(\eta_{-i}^{(\alpha)})^2 m_{2\alpha}^p\, r_{2(\alpha+\beta),2\alpha}\\ &\quad -2b\theta_i^{\alpha+\beta+s}(\eta_{-i}^{(\alpha)})\, m_\alpha^p d].\end{aligned} \tag{3.4}$$

Substituting b in (3.4), we obtain

$$\begin{aligned}\Delta_i(\theta) &= \theta_i^c[2r_{\alpha,2\alpha}^p(d/r_{2(\alpha+\beta),2\alpha})\, \theta_i^{\alpha+\beta+s}\, dm_\alpha^p\eta_{-i}^{(\alpha)}\\ &\quad -\left(r_{\alpha,2\alpha}^{2p}/r_{2(\alpha+\beta),2\alpha}^2\right)d^2(\eta_{-i}^{(\alpha)})^2\, \theta_i^{2(\alpha+\beta)} m_{2\alpha}^p r_{2(\alpha+\beta),2\alpha}]\\ &= \theta_i^{c+2s}\, d^2\, \frac{r_{\alpha,2\alpha}m_\alpha^p}{r_{2(\alpha+\beta),2\alpha}}\, \{\theta_i^{\beta+\alpha-s}\eta_{-i}^{(\alpha)}(2 - \theta_i^{\beta+\alpha-s}\, \eta_{-i}^{(\alpha)})\}.\end{aligned} \tag{3.5}$$

From (3.5) it follows that the maximum improvement occurs when

$$\theta_i^{\alpha+\beta-s}\, \eta_{-i}^{(\alpha)} = 1.$$

Since this is true for each $i = 1,\ldots,p$, maximum overall improvement occurs when all θ_i are equal, say to θ, and the total improvement by summing (3.5) over i is

$$d^2 r_{2\alpha,2(\alpha+\beta)}\, m_\alpha^{2p}\, p\theta^{c+2s}/m_{2\alpha}^p. \tag{3.6}$$

This completes the proof of the theorem.

We note that summing (3.5) over i yields the total improvement of δ for general θ. The component risk associated with δ_0 is clearly $\theta_i^{c+2s}(1 - m_s^2/m_{2s})$ whence the total risk for δ_0 is

$$\left(1 - \frac{m_s^2}{m_{2s}}\right) \Sigma\theta_i^{c+2s} . \tag{3.7}$$

If we define the percent relative improvement in risk as

$$PI = \frac{R(\theta,\delta_0) - R(\theta,\delta)}{R(\theta,\delta_0)} \times 100 , \tag{3.8}$$

then in the case where all θ_i are equal (3.8) does not depend on θ and becomes

$$\frac{d^2}{1 - \frac{m_s^2}{m_{2s}}} \cdot \frac{m_{2\alpha}}{m_{2(\alpha+\beta)}} \cdot \frac{(m_\alpha)^{2p}}{(m_{2\alpha})^p} . \tag{3.9}$$

We recall from (2.3b) that in (3.1) $\alpha = (s-\beta)/p$. Consider the limiting PI as $p \to \infty$. If $EX^{t(\beta-s)} < \infty$ for some $t > 0$, this insures that $\lim_{p\to\infty} (m_{2\alpha})^{-p}\, m_\alpha^{2p} = 1$. Hence the limit of (3.9) becomes

$$\frac{(m_{\beta+s}\, r_{s,2s} - m\ )^2}{1 - m_s\, r_{s,2s}} . \tag{3.10}$$

In the case when $\beta = 0$, $s = 1$, this reduces to $1 - m_1\, r_{1,2}$ as in Proposition 4.1 of Das Gupta (1984).

<u>Remark 3.1</u>. Theorem 3.1 is a more general version of Das Gupta (1984) Proposition 3.1.

Theorem 3.1 obtains maximum improvement along the ray $(\theta,\theta,\ldots,\theta) \equiv \theta\underset{\sim}{1}$. Suppose we have prior information leading us to seek maximum improvement along the ray defined by a specified vector θ_0. If we use the estimators in (2.7) or (2.11), we will find that maximum improvement occurs again along the ray $\theta\underset{\sim}{1}$. Rather, we extend Theorem 3.1 using estimators slightly more general than those in (2.7). We retain the same setup as in Theorem 3.1 and consider estimators given componentwise as

$$\delta_i^0(X) = r_{s,2s}\, X_i^s - b_i(X_i/\theta_{0i})^\beta\, \Pi(X_j/\theta_{0j})^\alpha,\quad i = 1,\ldots,p. \qquad (3.11)$$

In order that such estimators dominate δ_0, we need the risk difference

$$\Delta(\theta) = g_{2\alpha}(\theta)\, g_{-2\alpha}(\theta_0)\, r_{2(\alpha+\beta),2\alpha}\, m_{2\alpha}^p\, \Sigma\, b_i\, \theta_i^{c+2\beta}$$

$$+\, 2g_\alpha(\theta)\, g_{-\alpha}(\theta_0)\, m_\alpha^p\, d\, \Sigma\, b_i \theta_i^{c+\beta+s} < 0$$

with d as in (3.2). Let $b_i = h(\theta_{0i})b$ where $h(\cdot) > 0$ and let $\underline{h} = \min_i h(\theta_{0i})$, $\bar{h} = \max_i h(\theta_{0i})$. Mimicking the argument of Theorem 2.1, if b is between 0 and $2\underline{h}\, r_{\alpha,2\alpha}^p d/(\bar{h})^2\, r_{2(\alpha+\beta),2\alpha}$, then (3.11) will dominate δ_0.

Turning to component risk, if we take $b = \underline{h}\, r_{\alpha,2\alpha}^p\, d/(\bar{h})^2\, r_{2(\alpha+\beta),2\alpha}$, then the difference in risk for the ith component is, analogous to (3.5),

$$\theta_i^{c+2s} d^2\, \frac{r_{\alpha,2\alpha}\, m_\alpha^p}{r_{2(\alpha+\beta),2\alpha}}\, \{h(\theta_{0i})\frac{\underline{h}}{(\bar{h})^2}\, \theta_i^{\beta+\alpha-s}\eta_{-i}^{(\alpha)}(2-h(\theta_{0i})\frac{\underline{h}}{(\bar{h})^2}\theta_i^{\beta+\alpha-s}\eta_{-i}^{(\alpha)})\}$$

(3.12)

The maximum improvement occurs when $h(\theta_{0i})\theta_i^{\beta-s}$ is constant for all i. If $h(\theta_{0i}) = \theta_{0i}^{s-\beta}$, this occurs along the ray defined by θ_0. We summarize the above as

Theorem 3.2. Suppose X_i distributed with p.d.f. $f_{\theta i}(X_i) = \theta_i^{-1} f(X_i/\theta_i)$. Let θ_0 be an arbitrary positive vector. Estimators of the form (3.11) uniformly improve upon δ_0 for b_i satisfying conditions given below (3.11). Componentwise risk improvement is given by (3.12) and is maximized along the ray defined by θ_0.

4. ESTIMATION OF A LINEAR COMBINATION

In this section we will consider estimation of a linear combination of powers of scale parametric. Suppose the parameter function to be estimated is $\gamma(\theta) = \sum_i \ell_i \theta_i^{s_i}$ (we take $\ell_i \neq 0$ w.l.o.g.)

$$L(\gamma(\theta),a) = (a - \gamma(\theta))^2, \tag{4.1}$$

and

$$\delta_0(X) = \sum_j a_j X_j^{s_j} \tag{4.2}$$

is an arbitrary linear estimator of $\gamma(\theta)$.

The following theorem gives an estimator of $\gamma(\theta)$ which has smaller risk than $\delta_0(X)$ under the loss (4.1).

Theorem 4.1. Consider the estimator

$$\delta(X) = \delta_0(X) + b \prod_j X_j^{s_j/p}. \tag{4.3}$$

Suppose

$$d_j = \ell_j - a_j r_{j,s_j(1+1/p),s_j/p}, \quad j = 1,\ldots,p,$$

and

$$d_{(1)} = \min_j d_j, \quad d_{(p)} = \max_j d_j .$$

Then $\delta(X)$ dominates $\delta_0(X)$ under the loss (4.1) for $p \geq 2$ if one of the following conditions hold:

(1) $d_j > 0$, $\forall j = 1,\ldots,p$, and $0 < b < 2\rho_{s/p}\, p d_{(1)}/\rho_{2s/p}$

(2) $d_j < 0$, $\forall j = 1,\ldots,p$, and $2\rho_{s/p}\, p d_{(p)}/\rho_{2s/p} < b < 0$

If any $d_j = 0$, $j = 1,\ldots,p$, this approach fails to provide an improved estimator.

Proof. Let $\Delta(\gamma(\theta)) = R(\gamma(\theta),\delta(X)) - R(\gamma(\theta),\delta_0(X))$ be the risk difference. Then it clearly follows that

$$\begin{aligned}\Delta(\gamma(\theta)) &= b^2\rho_{2s/p}(\prod_j \theta_j^{2s_j/p}) - 2b\gamma(\theta)\rho_{s/p}(\prod_j \theta_j^{s_j/p}) \\ &\quad + 2b\rho_{s/p}\sum_j a_j\theta_j^{s_j}(\prod_j \theta_j^{s_j/p})r_{j,s_j+s_j/p,s_j/p} \\ &= b^2\rho_{2s/p}\prod_j \theta_j^{2s_j/p} - 2b\rho_{s/p}\prod_j \theta_j^{s_j/p}\sum_j \theta_j^{s_j}d_j\end{aligned}$$

Now the proof follows similarly to Theorem 2.1 using Lemma (2.1).

Remark 4.1. If we write $a_j = \ell_j r_{j,0,s_j}$, we are improving upon the corresponding linear unbiased estimator. Here $d_j = \ell_j(1 - r_{j,0,s_j}\, r_{j,s_j+s_j/p,s_j/p})$.
But by Lemma 2.3, the term in parenthesis is always < 0; whence if all ℓ_j have the same sign, the unbiased estimator is inadmissible. If we write $a_j = \ell_j r_{j,s_j,2s_j}$, we are improving upon the corresponding linear combination of best invariant estimators. Here $d_j = \ell_j(1 - r_{j,s_j,2s_j}\, r_{j,s_j+s_j/p,s_j/p})$. Now by Lemma 2.3 the term in parenthesis is always > 0. Whence, if all ℓ_j have the same sign, this estimator is inadmissible as well.

Remark 4.2. Interestingly, if for example $s_j = 1$ and all ℓ_j are equal, say, $\ell > 0$, i.e., we are estimating $\ell\sum_i\theta$, then with δ_0 as in (4.2), we need an expander if ℓ is sufficiently large; whereas, we need a shrinker if ℓ is sufficiently small. This paradoxical dependence upon an arbitrary ℓ is intuitively reasonable as a response to the relative magnitude of $\gamma(\theta)$.

Remark 4.3. Theorem 4.1 is directly applicable in the estimation of systems reliability when two or more components are connected in series. In those examples, often our problem is to estimate $\sum_i \theta_i$ where θ_i's are scale parameters of the ith life distribution.

Remark 4.4. Theorem 4.1 can be easily extended for the estimation of a linear transform $L\theta^s$ where $L = (\ell_{ij})$, an $r \times p$ matrix.

5. NUMERICAL STUDIES

In this section, we study PI as in (3.8) for our improved estimators (2.2) in three interesting cases.

In the first two, we take $s_i = 1$ and, for simplicity, $c_i = c$, $\beta_i = -c/2$, $i = 1,\ldots,p$. This implies from (2.3a) that $\alpha = p^{-1}(1 + c/2)$ and the improved estimator (2.2) becomes

$$\delta_i(X) = a_i X_i + b X_i^{-c/2} \prod_i X_i^{(1+c/2)/p} \tag{5.1}$$

where $0 < b < 2\rho_\alpha d^{(1)}/\rho_{2\alpha}\, d^{(2)}$. We set b as the midpoint of this interval.

Example 1. F-distributions. Suppose $S = (S_1,\ldots,S_p)$ and $T = (T_1,\ldots,T_p)$ are independent where $S_i \sim \sigma_i^2 \chi^2_{n_{1i}}$ and $T_i \sim \tau_i^2 \chi^2_{n_{2i}}$, $i = 1,\ldots,p$. Our problem is to estimate $\theta_i = \sigma_i^2/\tau_i^2$, $i = 1,\ldots,p$. The best scale invariant estimator of $\theta = (\theta_1,\ldots,\theta_p)$ is $\delta^0(X) = (\delta_1^0(X),\ldots,\delta_p^0(X))$ where

$$\delta_i^0(X) = a_i X_i\ , \quad i = 1,\ldots,p, \tag{5.2}$$

with $X_i = S_i/T_i$ and $a_i = (n_{2i} - 4)(n_{1i} + 2)^{-1}$, $i = 1,\ldots,p$, and individually the X_i are admissible.

For convenience we set $n_{1i} = n_1$, $n_{2i} = n_2$ and

$$m_{i,\alpha} = \left(\frac{n_1}{n_2}\right)^\alpha \Gamma\left(\frac{n_1}{2} + \alpha\right)\Gamma\left(\frac{n_2}{2} - \alpha\right)/\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right) \tag{5.3}$$

provided $-n_1/2 < \alpha < n_2/2$. Thus for a given θ, we can readily evaluate (3.5), (3.7), and ultimately (3.8). In Table I we present PI for various p, c, and (n_1,n_2) combinations. θ was created by selecting each coordinate randomly within the given range. The results for a typical value are presented along with the maximum PI using (3.9).

Example 2. Reciprocal Beta. Suppose Y_i is distributed as $\eta_i \cdot Be(2+\varepsilon,1)$, $\varepsilon > 0$, $i = 1,2,\ldots,p$, and we seek to estimate $\eta^{-1} \equiv (\eta_1^{-1},\ldots,\eta_p^{-1})$, i.e., $s = -1$. For convenience we transform to $X_i = Y_i^{-1}$ and taking $\theta_i = \eta_i^{-1}$ we have at $\theta_i = 1$, $f(X) = (2+\varepsilon) \cdot X^{-(3+\varepsilon)}$. The best invariant estimator of θ_i is $(1+\varepsilon)^{-1}\varepsilon X_i$ and is admissible from the discussion at the beginning of Section 2 since $\varepsilon > 0$. Using (5.1) and noting that $m_\alpha = (\varepsilon+2-\alpha)^{-1}(\varepsilon+2)$, $\alpha < 2$, it is straightforward to calculate (3.10) which becomes, for $c > -2$,

$$\frac{(1+c/2)^2}{(1+c/2+\varepsilon)^2} \cdot \frac{(\varepsilon+2)(\varepsilon+2+c)}{(\varepsilon+2+c/2)^2} . \qquad (5.4)$$

Using (5.4) we see that for p large if either ε or $c \to \infty$, PI $\to 0$ while if ε and c are close to 0 nearly 100% improvement is possible. Table II displays results for small to moderate p using (3.8) with θ as in the previous example.

Example 3. Simultaneous Exponential and Rectangular Distributions. Here we consider X_i arising from two different distributional families. Suppose X_i have p.d.f. as follows:

$$f(X_i|\theta_i) = \begin{cases} \theta_i^{-1} e^{-x_i/\theta_i}, & i = 1,\ldots,p_1, \quad x_i > 0,\ \theta_i > 0 \\ \theta_i^{-1}, & i = p_1+1,\ldots,p, \quad 0 < x_i < \theta_i,\ \theta_i > 0 \end{cases} \qquad (5.5)$$

with $p_1 \geq 1$, $p \geq 2$. In estimating $\theta = (\theta_1,\ldots,\theta_p)$, the best invariant estimator is $\delta^0(X) = (\delta_1^0(X),\ldots,\delta_p^0(X))$ where $\delta_i^0(X) = a_i X_i$, $a_i = 1/2$, $i \leq p_1$, $a_i = 3/2$, $i > p_1$.

TABLE I

F-Distributions
Percent Relative Improvement in Risk

	$n_1 = 5$, $n_2 = 5$	$n_1 = 10$, $n_2 = 10$	$n_1 = 20$, $n_2 = 20$	$n_1 = 5$, $n_2 = 10$	$n_1 = 10$, $n_2 = 20$
<u>c = 0</u>					
<u>p = 2</u>					
(0,12)	25.23	9.75	4.82	10.94	6.20
θ_i equal	25.31	9.78	4.84	10.97	6.22
<u>p = 5</u>					
(0,12)	39.90	18.24	9.23	21.54	12.20
θ_i equal	53.24	24.33	12.32	28.74	16.28
<u>p = 10</u>					
(0,12)	44.24	21.06	10.72	25.43	14.33
θ_i equal	64.21	30.52	15.57	36.91	20.79
<u>p = 30</u>					
(0,12)	49.85	24.28	12.41	29.80	16.70
θ_i equal	72.07	35.11	17.95	43.09	24.14
<u>c = 1</u>					
<u>p = 2</u>					
(0,12)	29.80	17.92	9.93	19.83	12.65
θ_i equal	30.02	18.05	10.01	19.98	12.74
<u>p = 5</u>					
(0, 12)	33.81	24.25	14.38	27.37	18.60
θ_i equal	57.17	41.00	24.31	46.28	31.46
<u>p = 10</u>					
(0,12)	34.02	25.77	15.59	29.09	20.27
θ_i equal	66.02	50.02	30.26	56.45	39.34
<u>p = 30</u>					
(0,12)	35.62	28.05	17.19	31.51	22.41
θ_i equal	71.54	56.34	34.53	63.29	45.01

TABLE I (Continued)

			$c = 2$		
$p = 2$	$n_1 = 5$ $n_2 = 5$	$n_1 = 10$ $n_2 = 10$	$n_1 = 20$ $n_2 = 20$	$n_1 = 5$ $n_2 = 10$	$n_1 = 10$ $n_2 = 20$
(0,12)	17.91	23.14	15.01	24.55	18.96
θ_i equal	18.14	23.44	15.51	24.87	19.20
$p = 5$					
(0,12)	15.78	20.97	15.82	19.31	19.01
θ_i equal	34.87	46.35	34.97	42.68	42.01
$p = 10$					
(0,12)	12.57	19.84	15.72	15.60	18.48
θ_i equal	33.96	53.59	42.45	42.13	49.90
$p = 30$					
(0,12)	10.07	19.64	16.12	12.64	18.59
θ_i equal	29.53	57.98	47.60	37.33	54.90

TABLE II

Reciprocal Beta
Percent Relative Improvement in Risk

			$c = 0$	
$p = 2$	$\varepsilon = .01$	$\varepsilon = .1$	$\varepsilon = 1$	$\varepsilon = 10$
(0,12)	75.95	56.36	10.21	0.22
θ_i equal	76.20	56.54	10.24	0.23
$p = 5$				
(0,12)	68.74	56.00	14.43	0.41
θ_i equal	91.74	74.03	19.25	0.55
$p = 10$				
(0,12)	65.57	54.43	15.28	0.47
θ_i equal	95.16	79.00	22.17	0.68
$p = 30$				
(0,12)	67.18	56.36	16.65	0.54
θ_i equal	97.13	81.49	24.07	0.78

TABLE II (Continued)

	ε = .01	ε = .1	ε = 1	ε = 10
c = 1				
p = 2				
(0,12)	61.16	52.46	16.07	0.48
θ_i equal	61.61	52.84	16.19	0.49
p = 5				
(0,12)	50.91	44.68	16.68	0.68
θ_i equal	86.08	75.55	28.20	1.15
p = 10				
(0,12)	46.82	41.48	16.41	0.73
θ_i equal	90.87	80.50	31.84	1.41
p = 30				
(0,12)	46.58	41.49	17.01	0.78
θ_i equal	93.55	83.33	34.14	1.60
c = 2				
p = 2				
(0,12)	1.91	14.16	18.51	0.81
θ_i equal	1.93	14.34	18.75	0.82
p = 5				
(0,12)	32.81	30.34	15.14	0.86
θ_i equal	72.50	67.05	33.45	1.89
p = 10				
(0,12)	30.13	27.77	;3.92	0.86
θ_i equal	81.37	75.01	37.72	2.31
p = 30				
(0,12)	29.14	26.87	13.68	0.88
θ_i equal	86.03	79.33	40.38	2.61

In order to insure the existence of necessary expectations in (2.2), we take $\beta = 1/2$ so that $\alpha = 1/2p$ whence our estimator has ith coordinate

$$a_i X_i + b X_i^{1/2} \pi X_j^{1/2p} \qquad (5.6)$$

In (5.5) $m_{i,\alpha} = \Gamma(\alpha + 1)$, $i \leq p_1$, while $m_{i,\alpha} = (\alpha + 1)^{-1}$, $i > p_1$. Again taking b to be the midpoint of (2.4), we can calculate $\Delta(\theta)$ exactly using (2.5). Straightforwardly,

$$R(\theta,\delta^0) = \frac{1}{2} \sum_1^{p_1} \theta_i^{c+2} + \frac{1}{4} \sum_{p_1+1}^{p} \theta_i^{c+2}$$

so that PI in (3.8) may be obtained. Some values are presented in Table III with θ as in the previous examples. Generally, small improvement is observed. In fact, if $p_1 \geq 2$ and $p_0 - p_1 \geq 2$, one would do better overall by creating a dominating estimator of the form (2.2) for the first p_1 coordinates, a different dominating estimator of the form (2.2) for the remaining $p - p_1$ coordinates and then combining the two. We do not pursue details and extensions here.

TABLE III

Exponential/Uniform

Percent Relative Improvement in Risk

		$p_1 = 1$, $p = 2$	$p_1 = 2$, $p = 4$	$p_1 = 2$, $p = 10$	$p_1 = 8$, $p = 10$	$p_1 = 5$, $p = 10$	$p_1 = 10$, $p = 20$
(0,12)	c = 0	0.31	1.46	1.79	2.53	1.79	2.71
	c = 1	0.27	1.26	1.76	2.39	1.68	2.65
	c = 2	0.69	1.09	1.75	2.32	1.67	2.55
θ_i equal	$c \neq -1$	0.69	1.84	2.22	3.44	2.96	3.42

ACKNOWLEDGMENT

The authors acknowledge Brad Carlin for performing the computations.

REFERENCES

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.

Brown, L. D. (1966). On the admissibility of invariant estimators of one or more location parameters. Ann. Math. Statist., 37, 1087-1136.

Brown, L. D., and Fox, M. (1974). Admissibility in statistical problems involving a location or scale parameter. Ann. Statist., 4, 807-14.

Das Gupta, A. (1984). Simultaneous estimation of arbitrary scale parameters under arbitrary quadratic loss. Purdue Univ. Tech. Rep. #84-32.

Das Gupta, A. (1986). Simultaneous estimation in the multiparameter gamma distribution under weighted quadratic loss. Ann. Statist., 14, 206-19.

Das Gupta, A., and Sinha, B. K. (1984). Estimation in the multiparameter exponential family: admissibility and inadmissibility results. Tech. Rep. #84-8, Department of Statistics, Purdue.

Ghosh, M., Hwang, J. T., and Tsui, K. W. (1984). Construction of improved estimators in multiparameter estimation for continuous exponential families. J. Mult. Anal., 14, 212-20.

Hudson, H. M. (1978). A natural identity for exponential families with applications in multiparameter estimation. Ann. Statist., 6, 473-84.

Shinozaki, N. (1984). Simultaneous estimation of location parameters under quadratic loss. Ann. Statist., 12, 322-35.

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.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 399	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Improved Estimators In Simultaneous Estimation Of Scale Parameters		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Dipak K. Dey and Alan E. Gelfand		8. CONTRACT OR GRANT NUMBER(s) N00014-86-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics, Stanford University, Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research, Statistics & Probability Program Code 1111		12. REPORT DATE December 10, 1987
		13. NUMBER OF PAGES 24
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Simultaneous estimation; Scale parameters; Inadmissibility.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE FOLLOWING PAGE.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TECHNICAL REPORT NO. 399

20. ABSTRACT

A general class of estimators is developed for improving upon best scale invariant estimators of two or more arbitrary scale parameters (or powers thereof) for arbitrary positive distributions with sufficient moments under weighted squared error loss function. The technique is to compute the risk difference in terms of moments of the distribution. Some conditions are obtained under which the maximum improvement is possible, and the form of the estimator can be chosen to achieve this maximum along any specified ray.

The result is then extended to the estimation of a linear transform of the parameter vector. Finally, some examples are given with numerical calculations to obtain the amount of risk improvement.

END
DATE
FILMED
5-88
DTIC