A New Inadmissibility Theorem with Applications to Estimation of Survival and Hazard Rates and Means in the Scale Parameter Family

by

Anirban DasGupta*
Purdue University

Dipak K. Dey** and Alan E. Gelfand**
University of Connecticut, Storrs

Technical Report #87-25

Department of Statistics
Purdue University

June 1987

AMS Subject Classification: Primary classification 62C15, 62F10, Secondary classification 62N05.
Key words and phrases: scale parameter, risk, inadmissible, equivariant, survival rate, hazard rate, loss-robustness, Bayes risk.
* Research was supported by the National Science Foundation grant number DMS-8401996.
** Research was supported by the University of Connecticut research foundation.

★

Abstract

The problem of simultaneous estimation of survival or hazard rates and the means in the scale parameter family is considered under the general weighted quadratic loss $\sum_{i=1}^{p} w_i \theta_i^{c_i} (\delta_i - \eta_i)^2$ where $(\eta_1, \ldots, \eta_p)$ are the functionals estimated and $\theta_1, \ldots, \theta_p$ are the scale parameters. The best equivariant estimator is typically proved to be inadmissible for $p \geq 2$. The region of the parameter space where the improved estimators provide the maximum improvement in risk is identified and expressions for the maximum percentage risk - improvements are derived. Fairly extensive numerical studies have been made.

1. Introduction.

Let $X_i, 1 \le i \le p$, have independent scale-parameter distributions $\frac{1}{\theta_i} f_i(\frac{X_i}{\theta_i}), X_i > 0, \theta_i > 0$. The best scale-invariant estimator of $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_p)$ under the loss $L(\underset{\sim}{\theta}, \underset{\sim}{a}) = \sum_{i=1}^{p} w_i \theta_i^{c_i} (\theta_i - a_i)^2$ is $\delta_0(\underset{\sim}{x}) = (\delta_{0,1}(\underset{\sim}{x}), \ldots, \delta_{0,p}(\underset{\sim}{x}))$, where $\delta_{0,i}(\underset{\sim}{x}) = \frac{E\theta_i=1(X_i)}{E\theta_i=1(X_i^2)} \cdot X_i$. It follows easily from Brown (1966) and Brown and Fox (1974) that $\delta_0(\underset{\sim}{x})$ is an admissible estimator of $\underset{\sim}{\theta}$ under the aforementioned loss for $p = 1$. It follows from a general nonparametric theorem in DasGupta (1986) that under the special squared-error loss, $\delta_0(\underset{\sim}{x})$ is inadmissible for $p \ge 2$ and a dominating estimator is $\delta_i(\underset{\sim}{x}) = \delta_0(\underset{\sim}{x}) + c \cdot \prod_j X_j^{\frac{1}{p}} \cdot \underset{\sim}{1}$, where c is a suitable positive number and $\underset{\sim}{1} = (1, \ldots, 1)$. In this paper we obtain a two-fold generalization of the above result. There is a good amount of literature on simultaneous estimation of location/scale parameters. Stein (1959), Farrell (1964), Brown (1968) are some of the earlier works on this topic; more recently, Berger (1980), DasGupta (1984, 1986), Shinozaki (1984) have obtained improved estimators in the context of simultaneous estimation of several location/scale parameters under various losses. Two interesting parametric functions of great importance in clinical trials and reliability studies are the so-called survival and hazard rates, where the survival rate is defined as $S_\theta(t) = P_\theta(X > t)$, and the hazard rate is defined as $\lambda_\theta(t) = \frac{f_\theta(t)}{S_\theta(t)}$, where t is an arbitrary number, and f stands for the survival density. The survival rate $\lambda_\theta(t)$ may be interpreted as the conditional survival density of a patient given that the patient has survived for time t. When the original survival distribution f comes from the scale-parameter family (e.g., if f is Exponential, Gamma, Weibull etc.), then the functions $S_\theta(t)$ and $\lambda_\theta(t)$ very often assume the form $t^\gamma \theta^s$ for suitable values of γ and s. Of special interest is the value $t = x$, where x is the observed value of X in the sample. In this paper, with this motivation, we address the problem of simultaneous estimation of p functionals $x_i^{\gamma_i} \theta_i^{s_i}$, $1 \le i \le p$, where X_i's are independently distributed with scale-parameters θ_i; since $\theta_i^{s_i}$ is again a scale-parameter for $X_i^{s_i}$, we can assume without any loss of generality that $s_i = 1$. The loss is assumed to be $\sum_{i=1}^{p} \omega_i \theta_i^{c_i} (x_i^{\gamma_i} \theta_i - a_i)^2$. In the next section, we derive the best estimator of $x_i^{\gamma_i} \theta_i$ of the form $a_i X_i^{\gamma_i + 1}$; we then show that subject to the existence of certain moments of the X_i's, this best scale-invariant estimator is inadmissible for $p \ge 2$ and we obtain explicit improved estimators. In section 3, the direction along which these new estimators provide the maximum improvement in risk is identified and the amount of maximum percentage risk-improvement is obtained. In section 4, some general remarks about a few related problems are made; section 5 gives numerical figures for the percentage risk-improvements for a variety of scale-parameter distributions and different values of c_i in the loss function.

2. Improved estimators.

We first introduce the following notation:

$$\text{let } m_{i,\alpha} = EX_i^\alpha,$$
$$r_{i,\alpha,\beta} = \frac{EX_i^\alpha}{EX_i^\beta}, \tag{2.1}$$

whenever such expectation are finite, and where $E(\cdot)$ denotes expectation under $\underset{\sim}{\theta} = \underset{\sim}{1}$. We will first state a few Lemmas for subsequent use in the course of derivation of the improved estimators.

Lemma 1. Under the loss $L(\underset{\sim}{\theta}, \underset{\sim}{a}) = \sum_{i=1}^{p} \omega_i \theta_i^{c_i} (x_i^{\gamma_i}\theta_i - a_i)^2$, the best estimator of $x_i^{\gamma_i}\theta_i$ of the form $a_i X_i^{\gamma_i+1}$ is given by $\delta_{0,i}(X_i) = r_{i,2\gamma_i+1,2\gamma_i+2} X_i$.

Proof: Easy.

Lemma 2. For any $q_j, j = 1, \ldots, p$, $\prod_{j=1}^{p} \theta_j^{\frac{q_j}{p}} \leq \frac{1}{p} \sum_{j=1}^{p} \theta_j^{q_j}$.

Proof: Follows from the arithmetic mean-geometric mean inequality.

Lemma 3. $r_{i,t+s,t}$ is an increasing function of t for $s > 0$, and a decreasing function of t for $s < 0$.

Proof: We will only consider the case $s > 0$; the proof in the other case is similar. Note that the density $h_i(x,t) = \frac{x^t f_i(x)}{\int x^t f_i(x)dx}$ is MLR in x; consequently, by a well known result, for $s > 0, E(X^s/X \sim h_i(x,t)) = r_{i,t+s,t}$ is increasing in t.

We will now prove the inadmissibility of the best scale-invariant estimators of $x_i^{\gamma_i}\theta_i$; the proof and the algebra are most transparent for the squared error loss case, which we treat in the following theorem.

Theorem 1. Consider the problem of estimating $(x_1^{\gamma_1}\theta_1, \ldots, x_p^{\gamma_p}\theta_p)$ under the sum of squared-error losses. Let $\underset{\sim}{\delta}_0(\underset{\sim}{X})$ denote the best scale-invariant estimator and

$$\text{let } \underset{\sim}{\delta}(\underset{\sim}{X}) = \underset{\sim}{\delta}_0(\underset{\sim}{X}) + \underset{\sim}{c} \cdot \left(\prod_{j=1}^{p} X_j^{\frac{\gamma_j+1}{p}} \right), \tag{2.2}$$

where $\underset{\sim}{c} = (c_1, \ldots, c_p)'$, and $c_i = c\,sgn(\gamma_i + 1)$ for a suitable $c > 0$ (see (2.5)). For $p \geq 2, R(\underset{\sim}{\theta}, \underset{\sim}{\delta}) < R(\underset{\sim}{\theta}, \underset{\sim}{\delta}_0)$ for every $\underset{\sim}{\theta}$, where $R(\cdot,\cdot)$ denotes the risk function, and hence $\underset{\sim}{\delta}_0(\underset{\sim}{X})$ is inadmissible.

Proof: By familiar calculations,

$$\Delta(\underset{\sim}{\theta}) = R(\underset{\sim}{\theta}, \underset{\sim}{\delta}) - R(\underset{\sim}{\theta}, \underset{\sim}{\delta}_0)$$
$$= pc^2 \left(\prod_j m_j, \frac{2}{p}(\gamma_j + 1)\right) \cdot \left(\prod_j \theta_j^{\frac{2}{p}(\gamma_j+1)}\right)$$
$$+2c \sum_i r_{i,2\gamma_i+1,2\gamma_i+2} \cdot r_{i,(\gamma_i+1)(1+\frac{1}{p}),\frac{\gamma_i+1}{p}} sgn(\gamma_i + 1)\theta_i^{\gamma_i+1}$$
$$\cdot \left(\prod_j m_j, \frac{\gamma_j+1}{p}\right) \cdot \left(\prod_j \theta_j^{\frac{\gamma_j+1}{p}}\right)$$
$$-2c \sum_i r_{i,\gamma_i(1+\frac{1}{p})+\frac{1}{p},\frac{\gamma_i+1}{p}} sgn(\gamma_i + 1)\theta_i^{\gamma_i+1} \cdot \left(\prod_j m_j, \frac{\gamma_j+1}{p}\right) \cdot \left(\prod_j \theta_j^{\frac{\gamma_j+1}{p}}\right). \tag{2.3}$$

$$\text{Let } \varepsilon = \min_i \left[\left\{m_i, \gamma_i(1+\frac{1}{p}) + \frac{1}{p} - m_i, (\gamma_i+1)(1+\frac{1}{p}) r_{i,2\gamma_i+1,2\gamma_i+2}\right\} \times \frac{sgn(\gamma_i+1)}{m_i, \frac{\gamma_i+1}{p}}\right]. \tag{2.4}$$

Note that it follows from Lemma 3 that $\varepsilon > 0$. It now follows from (2.2), (2.3), and Lemma 2, that for $c > 0$,

$$\Delta(\underset{\sim}{\theta}) < \left\{c^2 \cdot \left(\prod_j m_j, \frac{2}{p}(\gamma_j+1)\right) - 2c\varepsilon \left(\prod_j m_j, \frac{\gamma_j+1}{p}\right)\right\} \left(\sum_i \theta_i^{\gamma_i+1} \cdot \prod_j \theta_j^{\frac{\gamma_j+1}{p}}\right) < 0$$
$$\text{for } 0 < c < 2\varepsilon \left(\prod_j m_j, \frac{\gamma_j+1}{p}\right) \Big/ \left(\prod_j m_j, \frac{2}{p}(\gamma_j+1)\right). \tag{2.5}$$

Remarks, generalizations, and discussion of the assumptions

1. For $\gamma_i \equiv 0$, the problem is one of simultaneous estimation of $\theta_1, \ldots, \theta_p$. Theorem 1 implies that in the general scale-parameter family uniform mean squared error improvement can be obtained by shifting by multiples of the geometric mean $t = \prod_j X_j^{\frac{1}{p}}$ whenever $p > 1$. This was first obtained in DasGupta (1986). Shifting by the geometric mean in a scale-parameter problem is much like shifting by the arithmetic mean of the coordinates in a location problem because it is well known that if x_i's have a scale parameter family of distributions, then log x_i's have a location parameter family of distributions. It is also interesting that Theorem 1 holds even if the coordinate distributions do not come from the same parametric family, because the functional form of the density of X_i was allowed to depend on i.

2. Several generalizations of Theorem 1 are fairly easily obtained. First, the vector $\underset{\sim}{1}$ in the improved estimator can be generalized to an arbitrary positive vector $\underset{\sim}{\alpha}$. This will enable us to

give unequal shifts in different coordinates. This scope of choice in the vector $\underset{\sim}{\alpha}$ also leads to the natural question of selecting an $\underset{\sim}{\alpha}$ that best suits the available prior knowledge on $\underset{\sim}{\theta}$ so that good risk-improvements would be obtained in that part of the parameter space where $\underset{\sim}{\theta}$ is likely to lie.

Next, the statement of Theorem 1 holds for more general losses of the form $(\underset{\sim}{a}-\underset{\sim}{\theta})'Q(\underset{\sim}{\theta})\ (\underset{\sim}{a}-\underset{\sim}{\theta})$ where $Q(\theta)$ is any positive definite matrix such that $\inf_{\theta} \lambda_{min}(Q) > 0$ and $\sup_{\theta} \lambda_{max}(Q) < \infty$ where $\lambda_{min}, \lambda_{max}$ are the minimum and the maximum eigenvalues of Q.

3. The estimators in Theorem 1 are <u>somewhat</u> loss-robust, but not immensely so (for a beautiful treatment of the problem of constructing loss-robust estimators in the normal (and some others) distribution, see Hwang (1983); however, robustness with respect to such a large class of losses may often be unachievable and is perhaps also a conservative formulation of the problem). Consider first the case of $Q = I$ and $\gamma_i \equiv 0$; the allowed range of c in (2.5) is

$$0 < c < 2c_0, \text{ where } c_0 = \varepsilon \cdot \left(\prod_j r_j, \frac{1}{p}, \frac{2}{p} \right).$$

Taking $c = c_0$, one will have uniformly smaller frequentist risk for the estimate $\delta_0(\underset{\sim}{X}) + c_0 t \underset{\sim}{1}$ under squared-error loss. It is natural to ask if one can still achieve uniform domination with this <u>same</u> estimator when the loss is $(\underset{\sim}{a}-\underset{\sim}{\theta})'Q(\underset{\sim}{a}-\underset{\sim}{\theta})$ where Q is not necessarily the identity matrix. We have been able to prove that uniform domination can be obtained by using the <u>same</u> estimator described above for all Q such that

$$\lambda_{min} \geq \frac{trQ}{2p} \tag{2.6}$$

Comparable loss-robustness was achieved in Berger (1976) for estimating a multi-normal mean. It was proved in Berger (1976) that if $\underset{\sim}{X} \sim N(\underset{\sim}{\theta}, I)$, then for estimating $\underset{\sim}{\theta}$ under loss $L_Q(\underset{\sim}{\theta}, \underset{\sim}{a}) = (\underset{\sim}{a}-\underset{\sim}{\theta})'$ $Q(a-\underset{\sim}{\theta})$, the usual James-Stein estimate continues to be minimax for $p \geq 3$, if

$$\lambda_{max}(Q) \leq \frac{2trQ}{p+2} \tag{2.7}$$

Both (2.6) and (2.7) essentially mean that the eigenvalues of Q should not be very scattered. Thus a moderate amount of loss-robustness can be achieved with our estimators.

4. As mentioned in section 1, the survival and hazard rates are of the form $x^{\gamma}\theta^{s}$ for many standard scale-parameter distributions. The following table gives expressions for these functionals in some standard cases.

Table 1

Distribution	Density	$S_\theta(x)$	$\lambda_\theta(x)$
Pareto	$\frac{\alpha\theta^\alpha}{x^{\alpha+1}}(x>\theta)$	$\theta^\alpha x^{-\alpha}$	x^{-1}
Uniform	$\frac{1}{\theta}(0<x<\theta)$	$1-x\theta^{-1}$	$\frac{1}{\theta-x}(*)$
Exponential	$\theta e^{-\theta x}(x>0)$	$e^{\frac{-x}{\theta}}(*)$	θ^{-1}
Weibull	$\frac{\alpha}{\beta}x^{\alpha-1}e^{-x^2/\beta}(x>0)$	$e^{-\frac{x^\alpha}{\theta}}(*)$	$x^{\alpha-1}\theta^{-1}$

Note that Theorem 1 will not apply to the cases marked (*) in the table above, although it will apply to a large number of other cases.

We now state a generalization of Theorem 1 for the more general loss $L(\underset{\sim}{\theta}, \underset{\sim}{a}) = \sum_{i=1}^{p} \omega_i \theta_i^{c_i}(a_i - x_i^{\gamma_i}\theta_i)^2$. The proof of this result is similar to that of Theorem 1 and is omitted.

Theorem 2. Consider the problem of estimating $(x_1^{\gamma_1}\theta_1, \ldots, x_p^{\gamma_p}\theta_p)$ under the loss $L(\underset{\sim}{\theta}, \underset{\sim}{a}) = \sum_{i=1}^{p} \omega_i \theta_i^{c_i}(a_i - x_i^{\gamma_i}\theta_i)^2$.

$$\text{Let } \varepsilon = \min_i \Big[\Big\{m_{i,\gamma_i - \frac{c_i}{2} + \frac{\gamma_i + 1 + \frac{c_i}{2}}{p}} - m_{i,\gamma_{i+1} - \frac{c_i}{2} + \frac{\gamma_i + 1 + \frac{c_i}{2}}{p}} \cdot r_{i,2\gamma_i+1,2\gamma_i+2}\Big\} \cdot \frac{sgn(r_i + 1 + \frac{c_i}{2})}{m_{i,\frac{\gamma_i+1+\frac{c_i}{2}}{p}}}\Big].$$

Let $\underset{\sim}{\delta}_0(\underset{\sim}{X})$ be the best scale-invariant estimator and let $\underset{\sim}{\delta}(\underset{\sim}{X}) = (\delta_1(\underset{\sim}{X}), \ldots, \delta_p(\underset{\sim}{X}))$, with

$$\delta_i(\underset{\sim}{X}) = \delta_{0,i}(\underset{\sim}{X}) + b\,sgn(\gamma_i + 1 + \frac{c_i}{2}) X_i^{\frac{-c_i}{2}} \left(\prod_{j=1}^{p} X_j^{\frac{\gamma_j + 1 + \frac{c_j}{2}}{p}}\right), \tag{2.8}$$

$$\text{where } 0 < b < \frac{2p\varepsilon\left(\min_i \omega_i\right)\left(\prod_j m_{j,\frac{\gamma_{j+1}+\frac{c_j}{2}}{p}}\right)}{\left(\sum_i \omega_i r_{i,\frac{2(\gamma_i+1+\frac{c_i}{2})}{p} - c_i,\frac{2(\gamma_i+1+\frac{c_i}{2})}{p}}\right)\left(\prod_j m_{j,\frac{2(\gamma_j+1+\frac{c_j}{2})}{p}}\right)}. \tag{2.9}$$

For $p \geq 2$, $R(\underset{\sim}{\theta}, \underset{\sim}{\delta}) < R(\underset{\sim}{\theta}, \underset{\sim}{\delta}_0)$ for every $\underset{\sim}{\theta}$ and hence $\underset{\sim}{\delta}_0(\underset{\sim}{X})$ is inadmissible.

3. Component Risks and Maximum Risk Improvements

In this section we study the component risk behavior of the estimators given in (2.2) and in (2.8) when $f_i(x) = f(x), i = 1, \ldots, p$. In this case, we take $r_{i,\alpha,\beta} = r_{\alpha,\beta}$ and $m_i = m$.

The following theorem gives the maximum improvement of the estimator of the form (2.8) when $\gamma_i \equiv 0$ and b is the midpoint of the interval in (2.9); we also find the direction at which this maximum attains.

Theorem 3. Suppose X_i is distributed with $pdf\, f_{\theta_i}(x_i) = \frac{1}{\theta_i} f(x/\theta_i), i = 1, \ldots, p$. Assume that $c_i \equiv c$ and $\omega_i \equiv 1$ in the loss $L(\underset{\sim}{\theta}, \underset{\sim}{a})$; consider the estimator defined in (2.8) with $\gamma_i \equiv 0$ and $b = \varepsilon r^p_{\alpha,2\alpha}/r_{2(\alpha+\beta),2\alpha}$, where $\alpha = \frac{1+\frac{c}{2}}{p}, \beta = -\frac{c}{2}$, and ε is as defined in Theorem 2. Then the maximum improvement in risk is attained when the θ_i's are equal, say to θ, and the amount of improvement is

$$\varepsilon^2 r_{2\alpha,2(\alpha+\beta)} m^{2p}_\alpha p \theta^{c+2}/m_{2\alpha}. \tag{3.1}$$

Proof: The difference in risk for the ith component is $\Delta_i(\theta)$ where

$$\begin{aligned}
\Delta_i(\theta) &= R(\delta_{0,i}, \theta_i) - R(\delta_i, \theta_i) \\
&= -\theta_i^c [E\{b^2 X_i^{2\beta} \prod_j X_j^{2\alpha} - 2b X_i^\beta \prod_j X_j^\alpha (r_{1,2} X_i - \theta_i)\}] \\
&= -\theta_i^c [b^2 \theta_i^{2(\alpha+\beta)} (\eta^{(\alpha)}_{-1})^2 m^p_{2\alpha} m_{2(\alpha+\beta)}/m_{2\alpha} \\
&\quad - 2b\theta_i^{\beta+\alpha+1} (\eta^{(\alpha)}_{-1}) m^p_\alpha \{m_{\alpha+\beta+1} r_{1,2}/m_\alpha - m_{\alpha+\beta}/m_\alpha\}] \\
&\quad (\text{where } \eta^{(\alpha)}_{-1} = \prod_{j \neq i} \theta_j^\alpha) \\
&= -\theta_i^c [b^2 \theta_i^{2(\alpha+\beta)} (\eta^{(\alpha)}_{-i})^2 m^p_{2\alpha} r_{2(\alpha+\beta),2\alpha} \\
&\quad - 2b\theta_i^{\alpha+\beta+1} (\eta^{(\alpha)}_{-i}) m^p_\alpha d].
\end{aligned} \tag{3.2}$$

Substituting b in (3.2), we obtain

$$\begin{aligned}
\Delta_i(\theta) &= \theta_i^c [2 r^p_{\alpha,2\alpha} (\varepsilon / r_{2(\alpha+\beta),2\alpha}) \theta_i^{\alpha+\beta+1} \varepsilon m^p_\alpha \eta^{(\alpha)}_{-i} \\
&\quad - (r^{2p}_{\alpha,2\alpha}/r^2_{2(\alpha+\beta),2\alpha}) \varepsilon^2 (\eta^{(\alpha)}_{-i})^2 \theta_i^{2(\alpha+\beta)} m^p_{2\alpha} r_{2(\alpha+\beta),2\alpha}] \\
&= \theta_i^{c+2} \varepsilon^2 \frac{r_{\alpha,2\alpha} m^p_\alpha}{r_{2(\alpha+\beta),2\alpha}} \{\theta_i^{\beta+\alpha-1} \eta^{(\alpha)}_{-i} (2 - \theta_i^{\beta+\alpha-1} \eta^{(\alpha)}_{-i})\}.
\end{aligned} \tag{3.3}$$

From (3.3) it follows that the maximum improvement occurs when

$$\theta_i^{\alpha+\beta-1} \eta^{(\alpha)}_{-i} = \frac{2(c+\alpha+\beta+1)}{c+2(\alpha+\beta)}.$$

Since this is true for each $i = 1, \ldots, p$, the maximum overall improvement occurs when all θ_i's are equal, say to θ, and the total improvement by summing (3.3) over i is

$$\varepsilon^2 r_{2\alpha,2(\alpha+\beta)} m^{2p}_\alpha p \theta^{c+2}/m^p_{2\alpha}. \tag{3.4}$$

This completes the proof of the theorem.

We note that summing (3.3) over i yields the total improvement of δ for general θ. The component risk associated with δ_0 is clearly $\theta_i^{c+2}(1 - m_1^2/m_2)$ whence the total risk for δ_0 is

$$(1 - \frac{m_1^2}{m_2})\Sigma\theta_i^{c+2}. \tag{3.5}$$

If we define the percent relative improvement in risk as

$$PI = \frac{R(\theta, \delta_0) - R(\theta, \delta)}{R(\theta, \delta_0)} \times 100, \tag{3.6}$$

then in the case where all θ_i are equal, (3.6) is independent of θ and equals

$$\frac{\varepsilon^2}{1 - \frac{m_1^2}{m_2}} \cdot \frac{m_{2\alpha}}{m_{2(\alpha+\beta)}} \cdot \frac{(m_\alpha)^{2p}}{(m_{2\alpha})^p}. \tag{3.7}$$

Consider now the limiting value of PI as $p \to \infty$. If $EX^{t(\beta-1)} < \infty$ for some $t > 0$, then $\lim_{p\to\infty} (m_{2\alpha})^{-p} m_\alpha^{2p} = 1$. Since $\alpha = (1-\beta)/p$, this implies that the limit of (3.7) that as $p \to \infty$ is equal to

$$\frac{(m_{\beta+1}r_{1,2} - m)^2}{1 - m_1 r_{1,2}}. \tag{3.8}$$

If the density f is Gamma with a shape parameter α, then (3.8) equals 1 and if f is Pareto with a shape parameter α, then (3.8) equals $\frac{1}{(\alpha-1)^2}$; these indicate that encouraging risk-improvements are possible, approaching 100% as $p \to \infty$, in the Gamma case and also in the Pareto case for α close to 2.

4. Remarks, discussions, and generalizations.

As pointed out in section 2, under ordinary squared-error loss, the best scale-invariant estimator $\underset{\sim}{\delta}_0(\underset{\sim}{X})$ of $(x_1^{\gamma_1}\theta_1, \ldots, x_p^{\gamma_p}\theta_p)$ can be uniformly dominated by estimators of the form $\underset{\sim}{\delta}_0(\underset{\sim}{X}) + c \cdot \left(\prod_{j=1}^{p} X_j^{\frac{\gamma_j+1}{p}}\right) \cdot \underset{\sim}{\alpha}$, where $\underset{\sim}{\alpha}$ is an arbitrary positive vector. It is natural to ask if $\underset{\sim}{\alpha}$ can be chosen to maximize the risk-improvement in some desirable parts of the parameter space or perhaps to minimize the Bayes risk against a certain prior. It is relatively easy to show that if $\theta_1, \ldots, \theta_p$ are iid, then the Bayes risk is minimized by choosing $\underset{\sim}{\alpha}$ to be proportional to $\underset{\sim}{1} = (1, \ldots, 1)'$; when $\theta_1, \ldots, \theta_p$ have an arbitrary joint distribution π, the minimizing $\underset{\sim}{\alpha}$ can be found by using a sequential algorithm. For both of these results, see DasGupta (1984). For related works on minimization of Bayes risks among minimax estimators, see Berger (1982), DasGupta and Berger (1986), DasGupta and Rubin (1986), DasGupta and Bose (1987), Marazzi (1985), Kempthorne (1987) etc. Essentially similar uniform domination results hold for simultaneous estimation of linear combinations of the scale parameters. For details, see Dey and Gelfand (1986). The main result of this paper implies the surprising fact that the estimates for the gamma scale-parameters proposed in DasGupta (1986) are robust and work equally well for arbitrary scale-parameters. In a recent paper, Klonecki and Zontek (1987) prove that in the gamma case, linear estimators $L\underset{\sim}{X}$ of $\underset{\sim}{\theta}$ which cannot be uniformly dominated by estimators of the form (2.2) are admissible under squared-error loss; in fact, this is a characterization of all linear admissible estimates in the gamma case. In the light of this result and our present article, it certainly seems plausible that a similar striking result may hold in the entire scale-parameter family. Finally, note that although the results proved in this article consider only the case of a single observation from each coordinate distribution, the multiple observation case, theoretically, is included in our article if one restricts attention to $Y_1, \ldots, Y_p$, where Y_i is the Pitman estimate of θ_i based on a sample of size n_i from f_i.

5. Numerical Studies

In this section, we study PI as in (3.6) for our improved estimators (2.8) in two cases. We take $\gamma_i \equiv 0$ and $c_i \equiv c$ in both cases. As noted before, in (3.7), $\beta = -c/2$ and $\alpha = \frac{1+\frac{c}{2}}{p}$. The improved estimator considered is

$$\delta_i(X) = \delta_0, i(X) + bX_i^{-c/2} \prod_i X_i^{(1+c/2)/p}, \tag{5.1}$$

where b is the midpoint of the interval (2.9).

Example 1. F-distributions. Suppose $S = (S_1, \ldots, S_p)$ and $T = (T_1, , , ., T_p)$ are independent where $S_i \sim \sigma_i^2 \chi_{n_{1i}}$ 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 $\delta^0(X)$ where

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

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

For convenience we set $n_{1i} = n_1, n_{2i} = n_2$; note that

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

provided - $n_1/2 < \alpha < n_2/2$. Thus for a given θ, we can readily evaluate (3.3), (3.5), and ultimately (3.6). 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 (3.7).

Example 2. Reciprocal Beta. Suppose Y_i is distributed as $\eta_i \cdot Be(2+\varepsilon, 1), \varepsilon > 0, i = 1, 2, \dots, p$, and we seek to estimate $\eta^{-1} \equiv (\eta_1^{-1}, \dots, \eta_p^{-1})$. For convenience we transform to $X_i = Y_i^{-1}$; 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 $(1+\varepsilon)^{-1}\varepsilon X_i$. Using (5.1) and noting that $m_\alpha = (\varepsilon+2-\alpha)^{-1}(\varepsilon+2), \alpha < 2$, it is straightforward to calculate (3.8) which becomes, for $c > -2$,

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

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 (3.6) with θ as in the previous example.

Table I

F-Distributions

Percent Relative Improvement in Risk

$c = 0$

	$n_1 = 5$, $n_2 = 5$	$n_1 = 10$, $n_2 = 10$	$n_1 = 20$, $n_2 = 20$	$n_1 = 5$, $n_2 = 10$	$n_1 = 10$, $n_2 = 20$
$p = 2$					
(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
$p = 5$					
(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
$p = 10$					
(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
$p = 30$					
(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

$c = 1$

	$n_1 = 5$, $n_2 = 5$	$n_1 = 10$, $n_2 = 10$	$n_1 = 20$, $n_2 = 20$	$n_1 = 5$, $n_2 = 10$	$n_1 = 10$, $n_2 = 20$
$p = 2$					
(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
$p = 5$					
(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
$p = 10$					
(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
$p = 30$					
(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

$c = 2$

	$n_1 = 5$, $n_2 = 5$	$n_1 = 10$, $n_2 = 10$	$n_1 = 20$, $n_2 = 20$	$n_1 = 5$, $n_2 = 10$	$n_1 = 10$, $n_2 = 20$
$p = 2$					
(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
			$c = 1$	
$p = 2$	$\varepsilon = .01$	$\varepsilon = .1$	$\varepsilon = 1$	$\varepsilon = 10$
(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	13.92	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

Acknowledgement

The authors acknowledge Brad Carlin for performing the computations.

References

Berger, J. (1976). Minimax estimation of a multivariate normal mean under arbitrary quadratic loss. *Jour. Mult. Anal.* **6**, 256–264.

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.

Berger, J. (1982). Estimation in continuous exponential families: Bayesian estimation subject to risk restrictions and inadmissibility results. *Stat. Decision Theory and Related Topics* **III**, S. S. Gupta and J. Berger (Eds.), Academic Press, New York.

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

Brown, L. D. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Ann. Math. Statist.* **39**, 29–48.

Brown, L. D. and Fox, M. (1974). Admissibility in statistical problems involving a location or scale parameter. *Ann. Statist.* **4**, 807–814.

DasGupta, A. (1984). Admissibility on the gamma distribution: Two examples. *Sankhya (Series A)*, **Part 3**, 395–407.

DasGupta, A. (1984). Simultaneous estimation of arbitrary scale-parameters under arbitrary quadratic loss. *Technical report* **#84-32**, Department of Statistics, Purdue University.

DasGupta, A. (1986). Simultaneous estimation in the multiparameter gamma distribution under weighted quadratic losses. *Ann. Statist.* **14**, 206–219.

DasGupta, A. and Berger, J. (1986). Estimation of multiple gamma scale parameters: Bayes estimation subject to uniform domination. *Comm. in Stat., Special Volume on Stein-type Multivariate Estimation* **15**, 2065–2086.

DasGupta, A. and Rubin, H. (1986). Bayesian estimation subject to minimaxity of a multivariate normal mean in the case of a common unknown variance. To appear in the Proceedings of the Fourth Purdue Symposium on Stat. Decision Theory and Related Topics.

DasGupta, A. and Bose, A. (1987). Γ-minimax and restricted-risk Bayes estimation of multiple Poisson means under ε-contaminations of the subjective prior. *Technical report* #87-7, Department of Statistics, Purdue University. Submitted for publication.

Dey, D. K. and Gelfand, A. E. (1986). Improved estimators in simultaneous estimation of scale parameters. Unpublished manuscript.

Farrell, R. (1964). Estimators of a location parameter in the absolutely continuous case. *Ann. Math. Statist.* **35**, 949–998.

Hwang, J. T. (1983). Universal domination and stochastic domination: estimation simultaneously under a broad class of loss functions. *Ann. Statist.* **13**, 295–314.

Kempthorne, P. J. (1987). Numerical specification of discrete least favorable prior distributions. *J. Sci. Stat. Comput.* **8**, No. 2.

Klonecki, W. and Zontek, S. (1987). Inadmissibility results for linear simultaneous estimation in the multiparameter gamma distribution. Pre-print.

Marazzi, A. (1985). On constrained minimization of the Bayes risk for the linear model. *Stat. and Decisions* **3**, Nos. 3/4, 277–296.

Shinozaki, N. (1984). Simultaneous estimation of location parameters under quadratic loss. *Ann. Statist.* **12**, 322–335.

Stein, C. (1959). The admissibility of Pitman's estimator of a single location parameter. *Ann. Math. Statist.* **30**, 970–979.