ESTIMATION OF MULTIPLE GAMMA SCALE-PARAMETERS: BAYES ESTIMATION SUBJECT TO UNIFORM DOMINATION

by

Anirban DasGupta and James Berger

Technical Report #85-13

Department of Statistics
Purdue University

Revised April 1986

ESTIMATION OF MULTIPLE GAMMA SCALE-PARAMETERS: BAYES ESTIMATION SUBJECT TO UNIFORM DOMINATION

Anirban DasGupta James Berger

Department of Statistics
Purdue University
West Lafayette, Indiana

ABSTRACT

Simultaneous estimation of p gamma scale-parameters is considered under squared-error loss. The problem of minimizing, subject to uniform risk domination, the Bayes risk (or more generally the posterior expected loss) against certain conjugate or mixtures of conjugate priors is considered. Rather surprisingly, it is shown that the minimization can be done conditionally, thus avoiding variational arguments. Relative savings loss (and a posterior version thereof) are found, and it is found that in the most favorable situations, Bayesian robustness can be achieved without sacrificing substantial subjective Bayesian gains.

1. Introduction

In multiparameter estimation, one often encounters the decision-theoretic problem of estimating a vector $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$

of p parameters under a certain loss, say, squared-error, i.e.,

$$L(\underset{\sim}{\theta},\underset{\sim}{\delta}) = \sum_{i=1}^{p} (\delta_i-\theta_i)^2. \tag{1.1}$$

The development in multiparameter estimation has been of two broad types: frequentist, and Bayesian. The emphasis in the frequentist research has been on identifying the admissible procedures; perhaps the problem which has received the most attention is that of mathematically demonstrating the inadmissibility of certain standard classical estimators in large dimensions (popularly known as the Stein-effect) and finding explicit better estimators. To a pure Bayesian, the important problem is to find the rule that minimizes the posterior expected loss with respect to a certain prior π. Both approaches lead to problems, naturally of different kinds. Since the celebrated work of Stein (1956), numerous results have been proved which show that the presence of Stein effect is just part of a very general phenomenon, having little to do with the exact form of the loss function or the underlying distribution. The class of improved estimators, in many problems, is startlingly large; often times a disappointing feature of these improved estimators is the very nominal risk-improvement which makes the researcher feel that the gains were not worth the effort. Also there exists the very serious problem of actually selecting an alternative estimator for practical use.

The subjective Bayesian problem is easily stated, but has its own problems. All said and done, the prior π determined in a subjective fashion can be no more than an approximation to the "true prior" and sensitivity of the subjective Bayes rule to misspecification of π can be of concern. The problem of finding reasonably insensitive Bayes rules is thus of major importance. Very briefly stated, this is the robust Bayesian problem. See Berger (1980, 1981, 1985) for further discussion. It is in this context that the following restricted risk Bayes problem assumes importance. Many times in multiparameter estimation, there exists a natural estimate $\delta_0(\underset{\sim}{x})$ of the parameter vector $\underset{\sim}{\theta}$. Letting $R(\underset{\sim}{\theta},\delta)$ denote the frequentist risk of any estimate δ, often one can find a class of estimators $\mathcal{D}_\varepsilon$ such that $R(\theta,\delta) \leq R(\theta,\delta_0) + \varepsilon$ for all δ in $\mathcal{D}_\varepsilon$. In particular, every rule in $\mathcal{D}_0$ uniformly dominates the standard estimate δ_0. Let $\pi(\underset{\sim}{\theta})$ denote a subjectively determined prior, and δ_π the estimator which minimizes the integrates Bayes risk $r(\pi,\delta) = \int R(\theta,\delta)d\pi(\underset{\sim}{\theta})$ (or more generally the posterior expected loss). The restricted risk Bayes problem is to find an estimator $\delta_{\pi,\varepsilon}$ from $\mathcal{D}_\varepsilon$ such that $r(\pi,\delta_{\pi,\varepsilon}) - r(\pi,\delta) \leq 0$ for every δ in $\mathcal{D}_\varepsilon$. The resulting estimator $\delta_{\pi,\varepsilon}$ will be especially attractive if it is only marginally worse than the unrestricted Bayes rule δ_π in terms of the average Bayes risk; then we will have an estimator which is never much worse than δ_0 in a frequentist sense (and thus has a degree of built-in robustness), and yet is only marginally worse in terms of the Bayes risk.

The restricted risk Bayes problem, as stated above, is extremely difficult in most cases. In Berger (1982), a modified version of this problem was considered for the symmetric normal case. In many multiparameter problems, one can find an unbiased estimator $D\delta(\underset{\sim}{x})$ of the risk-difference $R(\underset{\sim}{\theta},\delta) - R(\underset{\sim}{\theta},\delta_0)$. If an estimator δ is such that $D\delta(\underset{\sim}{x}) \leq \varepsilon$ for every $\underset{\sim}{x}$, then clearly δ belongs to $\mathcal{D}_\varepsilon$. The modified problem addressed in Berger (1982) was to find an estimator (we will call it $\delta_{\pi,\varepsilon}$ again) with the smallest Bayes risk with respect to π among all estimators δ with $D\delta(\underset{\sim}{x}) \leq \varepsilon$. This slightly different problem was treated in Berger (1982) when $\underset{\sim}{X} \sim N_p(\underset{\sim}{\theta}, \sigma^2 I)$ and $\underset{\sim}{\theta} \sim N_p(\underset{\sim}{\mu}, \tau^2 I)$ and it was found that the solution $\delta_{\pi,\varepsilon}$ is a combination of the James-Stein estimator and the unrestricted linear Bayes estimator; specifically,

$$\delta_{\pi,\varepsilon}(\underset{\sim}{X}) = \underset{\sim}{X} - \frac{2(p-2)\sigma^2}{|\underset{\sim}{X}-\underset{\sim}{\mu}|^2} \cdot (\underset{\sim}{X}-\underset{\sim}{\mu}), \text{ if } |\underset{\sim}{X}-\underset{\sim}{\mu}|^2 > 2(p-2)(\sigma^2+\tau^2)$$
$$= \underset{\sim}{X} - \frac{\sigma^2}{\sigma^2+\tau^2} \cdot (\underset{\sim}{X}-\underset{\sim}{\mu}), \text{ if } |\underset{\sim}{X}-\underset{\sim}{\mu}|^2 \le 2(p-2)(\sigma^2+\tau^2). \tag{1.2}$$

An intuitive way to understand this estimate is that it coincides with the optimal Bayes estimate when data support prior information and otherwise it is a boundary point of the class of improved estimators (note the multiplier 2(p-2) instead of the customary p-2). The proof involves complicated variational arguments; more evidence of this was later found in Chen (1983). The objective of this paper is to show that the hard variational argument can often be avoided by minimizing the posterior loss, pointwise, and then showing that when these restricted Bayes actions for different fixed $\underset{\sim}{X}$ are combined to give a procedure, one does not go outside of the class of estimators one started with. The conditional minimization also has the added appeal of good posterior interpretation. We will mostly deal with estimation of multiple gamma parameters but will first consider the normal example to illustrate the basic idea.

Example. Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{\theta},\sigma^2 I)$, $\underset{\sim}{\theta} \sim N_p(\underset{\sim}{\mu},\tau^2 I)$, $p \ge 3$, σ^2, $\tau^2 > 0$ known. The problem is to find an estimator $\delta^c(\underset{\sim}{X})$ with the smallest Bayes risk in the class of spherically symmetric estimators $\delta(\underset{\sim}{X})$ of $\underset{\sim}{\theta}$ of the form

$$\delta(\underset{\sim}{X}) = \underset{\sim}{X} - \rho\left(\frac{||\underset{\sim}{X}-\underset{\sim}{\mu}||^2}{\sigma^2}\right) \cdot (\underset{\sim}{X}-\underset{\sim}{\mu}) \tag{1.3}$$

satisfying $\overline{D}\,\phi(r) \le 0$, where $\overline{D}$ is as in (4.9) in Berger (1982), $r = \dfrac{||\underset{\sim}{X}-\underset{\sim}{\mu}||^2}{\sigma^2}$, and $\rho(r) = -\dfrac{\phi'(r)}{\phi(r)}$. [$D\phi(r)$ is just the usual un-unbiased estimator of $R(\underset{\sim}{\theta},\delta) - R(\underset{\sim}{\theta},X)$]. It is assumed the loss is ordinary squared-error. The unrestricted Bayes estimator is

$\delta_\pi(\underset{\sim}{X}) = \underset{\sim}{X} - \frac{\sigma^2}{\sigma^2+\tau^2}(\underset{\sim}{X}-\underset{\sim}{\mu})$; by familiar calculations, $\delta^c(\underset{\sim}{X})$ minimizes the Bayes risk among all such δ if

$$E_m||\delta^c(\underset{\sim}{X}) - \delta_\pi(\underset{\sim}{X})||^2 \leq E_m||\delta(\underset{\sim}{X}) - \delta_\pi(\underset{\sim}{X})||^2,$$

where E_m denotes expectation with respect to the marginal distribution of $\underset{\sim}{X}$.

By Lemma 3 in Berger (1982), for $||\underset{\sim}{X}-\underset{\sim}{\mu}||^2 \geq 2(p-2)(\sigma^2+\tau^2)$, and for any δ satisfying (1.3),

$$||\delta(\underset{\sim}{X}) - \delta_\pi(\underset{\sim}{X})||^2 = \left(\frac{\sigma^2}{\sigma^2+\tau^2} - \rho\left(\frac{|\underset{\sim}{X}-\underset{\sim}{\mu}|^2}{\sigma^2}\right)\right)^2 \cdot ||\underset{\sim}{X}-\underset{\sim}{\mu}||^2$$

$$\geq \left(\frac{2}{\sigma^2+\tau^2} - \frac{2(p-2)\sigma^2}{||\underset{\sim}{X}-\underset{\sim}{\mu}||^2}\right) \cdot ||\underset{\sim}{X}-\underset{\sim}{\mu}||^2$$

$$= ||\delta_\pi(\underset{\sim}{X}) - \delta^c(\underset{\sim}{X})||^2,$$

where δ^c is as in Berger (1982).
For $||\underset{\sim}{X}-\underset{\sim}{\mu}||^2 < 2(p-2)(\sigma^2+\tau^2)$, $||\delta_\pi(\underset{\sim}{X}) - \delta^c(\underset{\sim}{X})||^2 = 0$. Consequently, δ^c is the required estimator and moreover, it does minimize the posterior expected loss pointwise. Note that it can be checked directly that δ^c satisfies $\overline{D}\,\phi^c(r) \leq 0$. For a generalization to the case when σ^2 is unknown, see DasGupta and Rubin (1986).

2. ESTIMATION OF MULTIPLE GAMMA PARAMETERS

Let $X_1, X_2, \ldots X_p$ be independently distributed and suppose X_i has density

$$f(x_i|\theta_i) = \frac{e^{-\theta_i x_i} x_i^{\alpha-1}\theta_i^\alpha}{\Gamma(\alpha)}, \quad x_i > 0;\ \alpha > 0 \text{ known.} \tag{2.1}$$

Suppose it is desired to estimate $(\theta_1^{-1}, \theta_2^{-1}, \ldots, \theta_p^{-1})$ under squared-error loss $L(\underset{\sim}{\theta}, \underset{\sim}{\delta}) = \sum_{i=1}^{p}(\delta_i - \theta_i^{-1})^2$. The usual estimate is

$\delta_0(\underset{\sim}{x}) = \frac{\underset{\sim}{X}}{\alpha+1}$. The following result was proved in DasGupta (1986).

Theorem 1. Let $\mathcal{D}_0 = \left\{ \delta : \delta_i(\underset{\sim}{x}) = \frac{x_i}{\alpha+1} + \frac{\alpha r(t) + \frac{tr'(t)}{p}}{\alpha+1} \right\}$,
where $t = (\prod_{i=1}^{p} x_i)^{1/p}$,

i) $0 < r(t) \leq c_0 t$, with $c_0 = \frac{2(p-1)}{p(\alpha + \frac{1}{p})^2}$,

ii) $r(t)$ is differentiable (lack of derivative at a finite number of points causes no problem),

iii) $r(t)$ is non-decreasing,

iv) $\frac{r(t)}{t}$ is non-increasing.

Then for every estimator $\delta(\underset{\sim}{X}) \in \mathcal{D}_0$, $R(\underset{\sim}{\theta},\delta) < R(\underset{\sim}{\theta},\delta_0)$ for all $\underset{\sim}{\theta}$, if $p \geq 2$.

Thus every rule in $\mathcal{D}_0$ uniformly dominates δ_0 and hence cannot have worse Bayes risk than δ_0 no matter how badly the prior π is specified. The problem we will consider is to find a δ^c in $\mathcal{D}_0$ which has the smallest Bayes risk in $\mathcal{D}_0$ with respect to a prior π. The calculations are most easily carried out with a conjugate prior. In the next two theorems, we will assume that the θ_i are iid with density

$$\pi(\theta_i) = \frac{e^{-r\theta_i}\theta_i^{\beta-1}r^{\beta}}{\Gamma(\beta)} \quad . \tag{2.2}$$

The use of a conjugate prior is not of concern from a Bayesian robustness standpoint, because the resulting estimator has a built-in robustness via dominance over δ_0.
As stated earlier, the general finding of this paper is that one can actually minimize the posterior expected loss pointwise in $\mathcal{D}_0$; the ability to minimize the posterior expected loss, according to us, is very appealing (for further reference, see

L. Brown's discussion on Berger (1983)).

THEOREM 2. Let α and β in (2.1), (2.2) respectively be such that

i) $\beta < 2$,

ii) $p(2\beta + \alpha\beta - 2) < 2\alpha + \beta$.

Then $\delta_i^c(X) = \frac{X_i}{\alpha+1} + \frac{(\alpha p+1)c_0 t}{p(\alpha+1)}$, and corresponds to $r^c(t) \equiv c_0 t$.

Proof: The unrestricted Bayes estimate is $\delta_{\pi,i}(\underset{\sim}{x}) = \frac{x_i+r}{\alpha+\beta-1}$. It is enough to show that

$$\sum_{i=1}^p (\delta_i^c(\underset{\sim}{x}) - \delta_{\pi,i}(\underset{\sim}{x}))^2 \le \sum_{i=1}^p (\delta_i(\underset{\sim}{x}) - \delta_{\pi,i}(\underset{\sim}{x}))^2 \quad \forall \underset{\sim}{x}, \forall \delta \in \mathcal{D}_0 .$$

Towards this, note that if $\delta_i(\underset{\sim}{x}) = \frac{x_i}{\alpha+1} + \frac{\alpha r(t) + \frac{tr'(t)}{p}}{\alpha+1}$, then

$$\sum_{i=1}^p (\delta_i(\underset{\sim}{x}) - \delta_{\pi,i}(\underset{\sim}{x}))^2$$

$$= \sum_{i=1}^p \left(\frac{x_i}{\alpha+1} + \frac{\alpha r(t) + \frac{tr'(t)}{p}}{\alpha+1} - \frac{x_i}{\alpha+\beta-1} - \frac{r}{\alpha+\beta-1}\right)^2$$

$$= \sum_{i=1}^p \left(\frac{(\beta-2)x_i}{(\alpha+1)(\alpha+\beta-1)} + \frac{\alpha r(t) + \frac{tr'(t)}{p}}{\alpha+1} - \frac{r}{\alpha+\beta-1} \right)^2$$

$$= \frac{p}{(\alpha+1)^2} \left[\left(\alpha r(t) + \frac{tr'(t)}{p}\right) - \frac{(2-\beta)t}{\alpha+\beta-1}\right]^2$$
$$+ \frac{2(\beta-2)\left(\alpha r(t) + \frac{tr'(t)}{p}\right)\left(\sum_{i=1}^p x_i - pt\right)}{(\alpha+1)^2(\alpha+\beta-1)} - \frac{2pr\left(\alpha r(t) + \frac{tr'(t)}{p}\right)}{(\alpha+1)(\alpha+\beta-1)}$$
$$- \frac{p(2-\beta)^2 t^2}{(\alpha+1)^2(\alpha+\beta-1)^2} . \tag{2.3}$$

Conditions iii) and iv) in Theorem 1 imply that $0 \le r'(t) \le \frac{r(t)}{t}$. Since $\sum_{i=1}^p x_i \ge pt$ and $\beta < 2$, the second term in (2.3) is minimized by $r(t) = c_0(t)$; so is the third term since $r > 0$; and finally, the first term is also minimized by $r(t) = c_0 t$, since,

$$\alpha r(t) + \frac{tr'(t)}{p} \leq (\alpha + \frac{1}{p})c_0 t = \frac{2(p-1)t}{\alpha p+1} \leq \frac{(2-\beta)t}{\alpha+\beta-1}$$

under condition ii) of the present theorem. Hence proved.

Remark.

1. The unrestricted Bayes rule is $\frac{x_i+r}{\alpha+\beta-1}$. For small β, this is substantially larger than $\frac{x_i}{\alpha+1}$. Theorem 2 thus says that we should give the estimate $\frac{x_i}{\alpha+1}$ the maximum possible increment to come closest to the optimal Bayes rule.

2. The potential use of Theorem 2 for large p is not profound. For $p = 2$ and $\alpha = 1$, condition ii) is satisfied if $\beta < 1.2$. Since $\alpha + \beta > 2$ is required for $E(\theta_i^{-2} | \underset{\sim}{X} = \underset{\sim}{x})$ to be finite, the effective range is $1 < \beta < 1.2$ in this case.

Whereas for $\beta < 2$, the restricted Bayes rule often is just a boundary point of $\mathcal{D}_0$, the restricted Bayes rule for $\beta = 2$, as we shall see now, is a combination of the unrestricted Bayes rule δ_π and a limit point of $\mathcal{D}_0$. This was also found in the situations considered in Berger (1982) and Chen (1983). Note that the case $\beta = 2$ has some special importance in that with such a prior for the θ_i's, some of the improved estimates of the gamma means obtained in DasGupta (1986) had an empirical Bayesian justification.

THEOREM 3. Let $\beta = 2$ in (2.2). Then

$$\begin{aligned} \delta_i^c(\underset{\sim}{x}) &= \frac{x_i}{\alpha+1} + \frac{(\alpha + \frac{1}{p})c_0 t}{\alpha+1}, && \text{if } t \leq t_0 = \frac{pr}{(\alpha p+1)c_0} \\ &= \frac{x_i}{\alpha+1} + \frac{r}{\alpha+1}, && \text{if } t > t_0 = \frac{pr}{(\alpha p+1)c_0}; \end{aligned} \tag{2.4}$$

the estimate corresponds to

$$r^c(t) = c_0 t \qquad \text{if } t \le t_0$$
$$= \frac{r}{\alpha} - \frac{r}{\alpha(\alpha p+1)} \left(\frac{t_0}{t}\right)^{\alpha p} \qquad \text{if } t > t_0. \tag{2.5}$$

Proof: The unrestricted Bayes rule now is $\delta_{\pi,i}(\underset{\sim}{X}) = \frac{X_i+r}{\alpha+1}$. Hence, for any $\delta \varepsilon \mathcal{D}_0$,

$$\sum_{i=1}^{p} (\delta_i(\underset{\sim}{X}) - \delta_{\pi,i}(\underset{\sim}{X}))^2 = \frac{1}{p(\alpha+1)^2} (\alpha pr(t) + tr'(t) - pr)^2. \tag{2.6}$$

Note that for $t \le t_0$,

$\alpha pr(t) + tr'(t) \le (\alpha p+1)c_0 t \le pr$

$\Rightarrow (\alpha pr(t) + tr'(t) - pr)^2 \ge (\alpha pr^c(t) + tr'^c(t) - pr)^2.$

Also, for $t > t_0$, $\alpha pr^c(t) + tr'^c(t) - pr \equiv 0$. Hence, for all $\delta \varepsilon \mathcal{D}_0$, $\forall \underset{\sim}{X}$, $\sum_{i=1}^{p} (\delta_i(\underset{\sim}{X}) - \delta_{\pi,i}(\underset{\sim}{X}))^2 \ge \sum_{i=1}^{p} (\delta_i^c(\underset{\sim}{X}) - \delta_{\pi,i}(x))^2$. It now remains to show that $\delta^c \varepsilon \mathcal{D}_0$ and this follows by a straightforward verification of the conditions of Theorem 1.

<u>Remark</u>

1. Again the optimal estimator $\delta^c(\underset{\sim}{X})$ actually minimizes the posterior expected loss over $\mathcal{D}_0$.
2. The optimal estimator δ^c was obtained by setting $r^c(t)$ equal to $c_0 t$ on a certain part and equal to a <u>suitable</u> solution of $\alpha pr(t) + tr'(t) - pr = 0$ on the remaining part. Essentially this idea works in the normal example set out in the introduction. We will now see how this same idea works in our gamma case for certain non-conjugate priors. For $\beta = 2$, each θ_i has density $\pi(\theta_i) = r^2\theta_i e^{-r\theta_i}$. We will now consider more general priors $\pi(\theta_i) = \theta_i g(\theta_i)$.

<u>Lemma 1.</u> The Bayes estimate of θ_i^{-1} has the following representation:

$$\delta_{\pi,i}(x) = \frac{x_i+\gamma(x_i)}{\alpha+1} ,$$

$$\text{where} \quad \gamma(x_i) = \frac{-\int e^{-\theta_i x_i}\theta_i^{\alpha+1} g'(\theta_i)d\theta_i}{\int e^{-\theta_i x_i}\theta_i^{\alpha+1} g(\theta_i)d\theta_i} , \tag{2.7}$$

if $\lim\limits_{\theta_i \to 0,\infty} e^{-\theta_i x_i}\theta_i^{\alpha+1} g(\theta_i) = 0$ for all x_i.

Proof: Integration by parts.

Lemma 2. Let $\mu(t) = E[\gamma(x_i)/t]$, $1 \le i \le p$. The problem of minimizing $\Delta(\delta) = r(\pi,\delta) - r(\pi,\delta_\pi) \overset{\text{def.}}{=} \int\{R(\underset{\sim}{\theta},\delta) - R(\theta,\delta_\pi)\}d\pi(\theta)$ in $\mathcal{D}_0$ is equivalent to minimizing $E[\alpha pr(t) + tr'(t) - p\mu(t)]^2$.

Proof: Familiar calculations.

Lemma 3. If $g(\theta)$ is such that $\delta = \inf\limits_\theta \frac{-g'(\theta)}{g(\theta)} > 0$ and $\frac{\mu(t)}{t}$ is decreasing in t, then there exists a unique t_0 such that for any $r(\cdot)$ as in Theorem 1, $\alpha pr(t) + tr'(t) \le p\mu(t)$ for $t \le t_0$ (note t_0 may be $+\infty$).

Proof: Since $\delta > 0$, $\mu(t) \ge \delta$ for every t. Consequently, $\lim\limits_{t\to 0} \frac{\mu(t)}{t} = \infty$. Again, for any $r(\cdot)$ as in Theorem 1, $\alpha pr(t) + tr'(t) \le (\alpha p+1)c_0 t$. Now, let $t_0 = \sup\{t: \frac{\mu(t)}{t} \ge (\alpha + \frac{1}{p})c_0\}$.

Lemma 4. If g is log-convex and $\frac{\theta g'(\theta)}{g(\theta)}$ is decreasing, then $\mu(t)$ is increasing.

Proof: Let $Z = \sum\limits_{i=1}^{p} \log X_i = p \log t$. Clearly, it is enough to show that $E[\gamma(X_1)/Z = z]$ is increasing in z. It follows from the definition of $\gamma(X_1)$ that

$$\gamma(X_1) = E[\frac{-g'(\theta)}{g(\theta)}] ,$$

where θ has density proportional to $e^{-\theta X}\theta^{\alpha+1}\, g(\theta)$. By a standard

Montone Likelihood Ratio (MLR) argument it follows that $\gamma(X_1)$ is increasing in X_1 if g is log-convex (i.e., if $\frac{g'}{g}$ is increasing). Denoting $Y_i = \log X_i$, it would suffice to show that under the hypotheses of the Lemma, $E[h(Y_1)/Z = z]$ is increasing in z, where $h(Y_1) = \gamma(e^{Y_1})$. Note that h is increasing since γ itself is increasing.

Since the θ_i's are iid, it is clear that the X_i's as well as the Y_i's are marginally iid. Let f(x) denote the marginal density of (each) X_i and g(y) the marginal density of (each) Y_i.

$$\therefore f(x) = \text{constant}.x^{\alpha-1} \int e^{-\theta x}\, \theta^{\alpha+1}\, g(\theta)d\theta. \tag{2.8}$$

Differentiating (2.8) and then integrating by parts, it follows that

$$\frac{xf'(x)}{f(x)} = (\alpha-1) - (\alpha+2) - \frac{\int e^{-\theta x}\theta^{\alpha+2}g'(\theta)d\theta}{\int e^{-\theta x}\theta^{\alpha+1}g(\theta)d\theta}$$

$$= -3 - \frac{\int e^{-\theta x}\theta^{\alpha+2}g'(\theta)d\theta}{\int e^{-\theta x}\theta^{\alpha+1}g(\theta)d\theta}\ ; \tag{2.9}$$

since $\frac{\theta g'(\theta)}{g(\theta)}$ is decreasing, (2.9) implies that $\frac{xf'(x)}{f(x)}$ is decreasing in x, which in turn implies that $\frac{g'(y)}{g(y)}$ is decreasing in y. Using the convolution formula for the density of the sum of independent random variables, it is easy to show that the conditional density of Y_1 given $Z = \sum_{i=1}^{p} Y_i = z$ is MLR in Y_1, if the marginal density of each Y_i is log-concave (i.e., $\frac{g'(y)}{g(y)}$ decreasing). Now recall that $\mu(t) = E[h(Y_1)/Z = z]$ where h is increasing and the lemma follcws.

<u>Remark</u> In specific examples, it may not be any more difficult to directly show that $\frac{xf'(x)}{f(x)}$ is decreasing in x. This turns out to be the case in the example given later in this section.

<u>Lemma 5.</u> Let $\psi(t) = \int_0^t \mu(s)s^{\alpha p-1}ds$. If g is decreasing and

log-convex, $\frac{\theta g'(\theta)}{g(\theta)}$ is decreasing in θ, and $\frac{\mu(t)}{t}$ is decreasing in t, then $r^c(t)$ defined as

$$\begin{aligned} r^c(t) &= c_0 t &&, \text{ if } t \le t_0 \\ &= pt^{-\alpha p}(\psi(t)-\psi(t_0)) + \frac{p\mu(t_0)}{\alpha p+1}\left(\frac{t_0}{t}\right)^{\alpha p}, &&\text{ if } t > t_0 \end{aligned} \tag{2.10}$$

satisfies the conditions of Theorem 1.

Proof: First note that t_0 (as in Lemma 3) is such that $(\alpha p+1)c_0 t_0 = p\mu(t_0)$ and hence $r^c(t)$ is continuous. Clearly $r^c(t)$ satisfies all the conditions of Theorem 1 for $t \le t_0$. We shall now verify the conditions for $t > t_0$. By straightforward calculations

$$r'(t) \ge 0 \quad \forall t > t_0$$

$$\Leftrightarrow -\alpha p(\psi(t)-\psi(t_0)) + (t^{\alpha p}\mu(t) - t_0^{\alpha p}\mu(t_0)) + \frac{t_0^{\alpha p}\mu(t_0)}{\alpha p+1} > 0 \;\forall t > t_0. \tag{2.11}$$

Since g is decreasing, $\mu(t_0) > 0$, and hence it suffices to show that

$$f(t) = t^{\alpha p}\mu(t) - \alpha p\psi(t)$$

is increasing (for $t \ge t_0$). Since $f'(t) = t^{\alpha p}\mu'(t) \ge 0$ by Lemma 4, it follows $r^c(t)$ is increasing. In order to show that $\frac{r^c(t)}{t}$ is decreasing, we will show that $tr'^c(t) \le r^c(t)\; \forall t > t_0$. By direct calculations

$$\begin{aligned} tr'^c(t) - r^c(t) &= -p[(\alpha p+1)t^{-\alpha p}(\psi(t)-\psi(t_0)) \\ &\qquad + \mu(t_0)\left(\frac{t_0}{t}\right)^{\alpha p} - \mu(t)]. \\ &= -pt^{-\alpha p}[(\alpha p+1)\int_{t_0}^{t} \psi'(s)ds - \int_{t_0}^{t} \left(\frac{d}{ds}\{\mu(s)s^{\alpha p}\}\right)ds]. \end{aligned}$$

$$= -pt^{-\alpha p}\int_{t_0}^{t} \{(\alpha p+1)\psi'(s) - \mu'(s)s^{\alpha p} - \alpha p s^{\alpha p-1}\mu(s)\}ds$$

$$= -pt^{-\alpha p}\int_{t_0}^{t} (\psi'(s) - \mu'(s)s^{\alpha p})ds$$

$$= -pt^{-\alpha p}\int_{t_0}^{t} s^{p\alpha-1}\{\mu(s)-s\mu'(s)\}ds$$

$$\leq 0 \text{ since } s\mu'(s) \leq \mu(s). \tag{2.12}$$

(2.12) also shows $r^c(t) \leq c_0 t \,\forall t > t_0$. This verifies all the conditions of Theorem 1.

Theorem 4. If g satisfies the regularity conditions of Lemma 3 and Lemma 5, then δ^c defined as

$$\delta_i^c(\underset{\sim}{x}) = x_i + \frac{\alpha r^c(t) + \frac{tr'^c(t)}{p}}{\alpha+1}$$

minimizes $\int(R(\underset{\sim}{\theta},\delta) - R(\underset{\sim}{\theta},\delta_\pi))d\pi(\underset{\sim}{\theta})$ in $\mathcal{D}_0$.

Proof: Lemma 5 proves $\delta^c \in \mathcal{D}_0$. Lemmas 2, 3, and the fact that $\alpha p r^c(t) + tr'^c(t) \equiv p\mu(t)$ for $t > t_0$ now prove Theorem 4.

Remark.

1. Theorem 4 shows that if the θ_i's have prior density $\theta_i g(\theta_i)$ and g satisfies all the regularity conditions of Theorem 4, then the restricted risk Bayes rule continues to be either a limit point of $\mathcal{D}_0$ or the unrestricted Bayes rule. Theorem 3 follows as a special case since there $g(\theta) = r^2 e^{-r\theta}$ which is decreasing and log-convex and $\mu(t) \equiv r \Rightarrow \frac{\mu(t)}{t}$ is decreasing.
2. One class of priors for which Theorem 4 is likely to work out is a mixture of conjugate gamma priors with $g(\theta) = \sum_{i=1}^{k} \varepsilon_i r_i^2 e^{-r_i\theta}$, where $\sum_{i=1}^{k} \varepsilon_i = 1$. We will shortly give an example.

An example of a $g(\theta)$ which satisfies all the regularity conditions of Lemmas 3 and 5 and yet is not a mixture of exponentials is $g(\theta) = e^{-\theta}\theta^{-\frac{1}{2}}$; note that g cannot be a mixture of exponentials because it is not completely monotonic

3. Of obvious value would be easily verifiable sufficient conditions on g for $\frac{\mu(t)}{t}$ to be decreasing. An anonymous referee has pointed out the following sufficient condition for $\frac{\mu(t)}{t}$ to be decreasing:

Let $\psi(x_1,t)$ denote the conditional density of X_1 given t. By definition, $\mu(t) = \int\gamma(x_1)\psi(x_1,t)dx_1$. In order to show that $\frac{\mu(t)}{t}$ is decreasing in t, it suffices to show that $\mu(t)$ is concave. Now $\mu''(t) = \int\gamma(x_1)p(x_1,t)dx_1$, where $p(x_1,t) = \frac{\partial^2}{\partial t^2}\psi(x_1,t)$. Clearly, $\int p(x_1,t)dx_1 = 0$ for every t since $\int\psi(x_1,t)dx_1 = 1$ for every t. Writing $p(x_1,t) = p^+(x_1,t) - p^-(x_1,t)$ (for fixed t), it would be enough to show that $\int\gamma(x_1)p^+(x_1,t)dx_1 \leq \int\gamma(x_1,t)p^-(x_1,t)dx_1$. If now, for each $t, p(x_1,t)$, as a function of x_1, starts with positive values (at $x_1 = 0$) and changes sign only once from positive to negative, then it is clear that $p^+(x_1,t)$ is stochastically smaller than $p^-(x_1,t)$ and hence $E_{p^+}[\gamma(X_1)] \leq E_{p^-}[\gamma(X_1)]$ since $\gamma(X_1)$ is increasing in X_1 for log-convex g. Unfortunately, except for some special priors, $\lim_{x_1\to 0} p(x_1,t)$ is actually negative so that this sufficient condition often does not hold.

<u>Example</u>. Let θ_i be iid with common density $\sum_{i=1}^{k} \varepsilon_i r_i^2 \theta e^{-r_i\theta}$, where $0 < \varepsilon_i < 1$, $\sum_{i=1}^{k} \varepsilon_i = 1$, $r_i > 0$ known. Thus in the notation of Lemma 1, $g(\theta) = \sum_{i=1}^{k} \varepsilon_i r_i^2 e^{-r_i\theta}$. By direct algebra,

$$\frac{d^2}{d\theta^2} \log g(\theta) = \frac{\Sigma\varepsilon_i r_i^4 e^{-r_i\theta} \cdot \Sigma\varepsilon_i r_i^2 e^{-r_i\theta} - (\Sigma\varepsilon_i r_i^3 e^{-r_i\theta})^2}{(\Sigma\varepsilon_i r_i^2 e^{-r_i\theta})^2} \geq 0$$

$$\Rightarrow g \text{ is log-convex.} \qquad (2.13)$$

Also, clearly g is decreasing; moreover,

$$\frac{-g'(\theta)}{g(\theta)} = \frac{\Sigma\varepsilon_i r_i^3 e^{-r_i\theta}}{\Sigma\varepsilon_i r_i^2 e^{-r_i\theta}} > \min_i r_i > 0, \forall\theta .$$

Hence, for Theorem 4 to work out, it only remains to verify that $\frac{\mu(t)}{t}$ and $\frac{xf'(x)}{f(x)}$ are both decreasing where f is the marginal density of (each) x_i. We will analytically attempt below the case $k = p = 2$ and comment on the more general situations later. Assume without loss $r_1 > r_2$.

By direct integration, $f(x) = \text{constant}\cdot x^{\alpha-1}\sum\frac{\varepsilon_i r_i^2}{(x+r_i)^{\alpha+2}}$. A laborious routine computation shows that $\frac{xf'(x)}{f(x)}$ is decreasing if $r_1 \leq \frac{2\alpha+6}{2\alpha+5} r_2$. In order to show that $\frac{\mu(t)}{t}$ is decreasing in t, first note that the joint marginal density of x_1 and t is given as

$$f(x_1,t) = \text{constant}\cdot\frac{t^{2\alpha-1}}{x_1}\cdot\int e^{-\theta_1 x_1}\theta_1^{\alpha+1}g(\theta_1)d\theta_1$$

$$\cdot\int e^{-\theta_2 x_2}\theta_2^{\alpha+1}g(\theta_2)d\theta_2. \tag{2.14}$$

Using (2.7), (2.14), and the definition of $\mu(t)$, it follows that

$$\mu(t) = -\frac{\int(\int e^{-\theta x_1}\theta^{\alpha+1}g'(\theta)d\theta)\cdot\frac{1}{x_1}\cdot(\int e^{-\theta\frac{t^2}{x_1}}\theta^{\alpha+1}g(\theta)d\theta)dx_1}{\int(\int e^{-\theta x_1}\theta^{\alpha+1}g(\theta)d\theta)\cdot\frac{1}{x_1}\cdot(\int e^{-\theta\frac{t^2}{x_1}}\theta^{\alpha+1}g(\theta)d\theta)dx_1}. \tag{2.15}$$

Now using the actual form of $g(\theta)$, by direct computation,

$$\mu(t) = \frac{\varepsilon_1 r_1^3(\varepsilon_1 r_1^2 f_{11}+\varepsilon_2 r_2^2 f_{12}) + \varepsilon_2 r_2^3(\varepsilon_1 r_1^2 f_{21}+\varepsilon_2 r_2^2 f_{22})}{\varepsilon_1 r_1^2(\varepsilon_1 r_1^2 f_{11}+\varepsilon_2 r_2^2 f_{12}) + \varepsilon_2 r_2^2(\varepsilon_1 r_1^2 f_{21}+\varepsilon_2 r_2^2 f_{22})}$$

$$= \frac{r_1 g_1 + r_2 g_2}{g_1 + g_2} , \tag{2.16}$$

where $f_{ij}(t) = \int \frac{1}{(x_1+r_i)^{\alpha+2}(t^2+r_j x_1)^{\alpha+2}} \, dx_1$,

$$g_1(t) = \varepsilon_1 r_1^2(\varepsilon_1 r_1^2 f_{11}(t) + \varepsilon_2 r_2^2 f_{12}(t)),$$

$$g_2(t) = \varepsilon_2 r_2^2(\varepsilon_1 r_1^2 f_{21}(t) + \varepsilon_2 r_2^2 f_{22}(t)). \tag{2.17}$$

Now note that $|f'_{ij}(t)| = \int \frac{2(\alpha+2)t}{(x_1+r_i)^{\alpha+2}(t^2+r_j x_1)^{\alpha+3}} \, dx_1$

$$\leq \frac{2(\alpha+2)}{t} \cdot f_{ij}(t) \quad \forall t, \forall i,j. \tag{2.18}$$

Consequently, $\left|\frac{g_i'(t)}{g_i(t)}\right| \leq \frac{2(\alpha+2)}{t}$, $\forall t, i = 1,2$. As usual, to show that $\frac{\mu(t)}{t}$ is decreasing, we will show that $t\mu'(t) \leq \mu(t)$, $\forall t$.

First note that $\mu'(t) = \frac{(r_1-r_2)(g_1' g_2 - g_2' g_1)}{(g_1+g_2)^2}$

$$\Rightarrow |t\mu'(t)| = \frac{(r_1-r_2)\left|\frac{g_1'}{g_1} - \frac{g_2'}{g_2}\right| \cdot t}{\frac{(g_1+g_2)^2}{g_1 g_2}}$$

$$\leq (r_1-r_2)\cdot(\alpha+2). \tag{2.19}$$

On the other hand $\mu(t) = \frac{r_1 g_1 + r_2 g_2}{g_1+g_2} > r_2$, $\forall t$. $\quad (2.20)$

Combining (2.19) and (2.20) it follows $\frac{\mu(t)}{t}$ is decreasing if

$r_1 < \frac{\alpha+3}{\alpha+2} \cdot r_2$. Since $\frac{\alpha+3}{\alpha+2} > \frac{2\alpha+6}{2\alpha+5}$, all conditions in Theorem 4 will hold if $r_1 < \frac{2\alpha+6}{2\alpha+5} r_2$.

Remarks.

1. Crude bounds have been used in (2.18) through (2.20) in showing that $\frac{\mu(t)}{t}$ is decreasing in t. That's part of the reason a wider spectrum of values for r_1, r_2 could not be handled. By explicitly evaluating the analogs of the function $f_{ij}(t)$ in the special case $\alpha = 1$, we have been able to show that for small as well as large t, $\frac{\mu(t)}{t}$ is decreasing no matter what mixture is used as a prior. For any given mixture prior, the monotone nature of $\frac{\mu(t)}{t}$ in the middle zone can always be verified on the computer with any degree of confidence.

2. The same phenomenon holds when the prior is of the form $\theta^{\beta-1}\tau(\theta)$, where τ is a mixture like in the example above and $\beta < 2$. For $\beta > 2$, various assumptions made on $g(\theta)$ in Lemmas 1 through 5 (decreasing, log-convex etc.) fail and we have no results in that case. In fact, the simple conjugate gamma prior with $\beta > 2$ causes great problems and we are convinced the form of the restricted risk Bayes rules is basically different in that case.

3. BAYESIAN PERFORMANCE OF THE RESTRICTED RISK BAYES RULES

Finally, in this section, we will get back to the restricted risk Bayes rule obtained for a conjugate gamma prior with $\beta = 2$ in Theorem 3. The form of the optimal estimator in this case is simple enough so that Bayes risk calculations are done fairly easily. All estimators in the class $\mathcal{D}_0$ uniformly dominate the best equivariant estimator to start with; consequently robustness of our restricted risk Bayes rule is of no real concern. It is the sacrifice in potential Bayesian improvement that is the issue of importance. It is conventional to judge the necessary amount of sacrifice by looking at the relative savings loss (RSL) of Efron and Morris (1971) defined as

$$RSL(\pi,\delta) = \frac{r(\pi,\delta) - r(\pi,\delta_\pi)}{r(\pi,\delta_0) - r(\pi,\delta_\pi)}, \tag{3.1}$$

where for any δ, $r(\pi,\delta) - r(\pi,\delta_\pi)$ is defined as $\int(R(\underset{\sim}{\theta},\delta) - R(\underset{\sim}{\theta},\delta_\pi))d\pi(\theta)$ and δ_π is defined as the rule that minimizes the posterior expected loss. It is also meaningful to replace in (3.1) the integrated Bayes risk by the corresponding posterior expected loss. The corresponding savings loss will be called the posterior relative savings loss (PRSL). We will study both. Before deriving the expressions for RSL/PRSL, we briefly remind the reader that low values of these quantities will imply that robustness can be attained without making any serious dent on the subjective Bayesian gains available. The following theorem is straightforward.

Theorem 5. If θ_i's are iid with a prior π as in Theorem 3, then

$$\begin{aligned} PRSL(\pi,\delta^c) &= 0 && \text{if } t \geq \frac{r(\alpha p+1)}{2(p-1)} \\ &= \frac{\left(r - \frac{2(p-1)t}{\alpha p+1}\right)^2}{r^2} && \text{if } t < \frac{r(\alpha p+1)}{2(p-1)} . \end{aligned}$$

Remark. The RSL, being a constant, remains necessarily bounded away from 1. The PRSL, on the other hand, is a random variable and can get arbitrarily close to 1 for "unlikely" data. However, the attractive feature of the PRSL is that it remains zero for a fairly long time and at the same stroke the degree of its departure from zero also gives an idea about how strongly the data at hand support the subjective prior specification. In fact, to many Bayesians, the PRSL will be the appropriate quantity to look at to judge performance of the estimator. In order to get an idea of the magnitude of the PRSL, we have tabulated below values of the PRSL when t is set equal to its marginal expectation, i.e.,

$$t = E(t) = \left(\frac{\Gamma(\alpha+\frac{1}{p})}{\Gamma\alpha}\right)^p \Big/ \left(\frac{\Gamma(2-\frac{1}{p})}{\Gamma 2}\right)^p r. \tag{3.2}$$

Note that if t is close to its marginal mean, then prior opinion is supported and consequently, from a subjective Bayesian viewpoint, δ_π is an eminently reasonable estimate to use. Thus while calculating the PRSL at $t = E(t)$, δ^c is being compared with δ_π on the latter's homeground.

	$\alpha = 1$				$\alpha = 2$		
p:	5	10	20	p:	5	10	20
PRSL:	.15226	.10695	.08746		.02970	.00734	.00184

The values indicate that 10% of the available Bayesian gains must be sacrificed in return for dominance over δ_0 in 10 dimensions when $\alpha = 1$. The numbers, however, are especially encouraging for $\alpha = 2$. In this case even for $p = 5$, full robustness is guaranteed by sacrificing a nominal 3% of the subjective Bayesian gains. We will shortly see that the restricted risk Bayes rules also give best values of RSL for $\alpha = 2$; why $\alpha = 2$ turns out to be the most favorable situation is not clear to us, except that we are tempted to think it may have something to do with equality of α and β. We now proceed to the RSL calculations; the following theorem is also straightforward.

<u>Theorem 6.</u> $\mathrm{RSL}(\pi,\delta^c) = \frac{1}{p^2r^2} E\left\{[(\alpha p+1)c_0t-pr]^2 I_{t\le \frac{pr}{(\alpha p+1)c_0}}\right\}.$

<u>Theorem 7</u>. Let f(x) denote the common marginal density of the x_i's. Let $\log a = E(\log x)$, where $E(\cdot)$ denotes expectation under f. Then

$$\lim_{p\to\infty} \mathrm{RSL}(\pi,\delta^c) \le \left(\frac{2a}{\alpha} - 1\right)^2.$$

Proof. Assume without loss $r = 1$. First note

$f(x_1) = \frac{1}{B(\alpha,2)} \cdot \frac{x^{\alpha-1}}{(1+x)^{\alpha+2}}$. Hence

$$|E(\log x)| = k\left|\int_0^\infty \frac{x^{\alpha-1}\log x}{(1+x)^{\alpha+2}}\right| < \infty \quad \forall\alpha > 0.$$

Since x_i's are iid, the strong law of large numbers implies that

$\frac{1}{p}\sum_{i=1}^{p} \log x_i = \log t \overset{a.s.}{\to} \log a$

$\Rightarrow t \overset{a.s.}{\to} a$, as $p \to \infty$.

From Theorem 6, using the definition of c_0,

$$\lim_{p\to\infty} RSL(\pi,\delta^c) \leq \lim_{p\to\infty} E[\frac{2}{\alpha}\cdot t-1]^2 \tag{3.3}$$

Now note that $E(t^3) = (\frac{\Gamma(\alpha+\frac{3}{p})}{\Gamma\alpha})^{1/p} (\frac{\Gamma(2-\frac{3}{p})}{\Gamma 2})^{1/p}$ is uniformly bounded in p. Consequently, $\{t^2 = t^2(p)\}$ is uniformly integrable. This together with the almost sure convergence to 'a" gives the result.

We will now derive actual expressions for 'a' for integral α's. These are essential to understand the nature of the RSL for large p.

Theorem 8.

i) $\log a = -1$ for $\alpha = 1$

ii) $\log a = 0$ for $\alpha = 2$

iii) $\log a = \frac{1}{\alpha-1} + \frac{1}{\alpha-2} + \ldots + \frac{1}{2}$, for $\alpha \geq 3$.

Proof:

i) See Gradshteyn and Ryzhik (1965).

ii) Straightforward.

iii) $I(\alpha) = \int_0^\infty \frac{x^{\alpha-1}\log x}{(1+x)^{\alpha+2}}\,dx$

$$= \frac{1}{\alpha+1}\int_0^\infty \frac{(\alpha-1)x^{\alpha-2}\log x + x^{\alpha-2}}{(1+x)^{\alpha+1}}\,dx$$

(by integrating by parts)

$$= \frac{\alpha-1}{\alpha+1} I(\alpha-1) + \frac{1}{\alpha+1} B(\alpha-1,2). \tag{3.4}$$

Using induction on α,

$$I(\alpha) = \frac{1}{\alpha+1} B(\alpha-1,2) + \frac{\alpha-1}{(\alpha+1)\alpha} B(\alpha-2,2) + \ldots + \frac{(\alpha-1)(\alpha-2)\ldots 3}{(\alpha+1)\alpha(\alpha-1)\ldots 4} B(2,2)$$
$$= \frac{1}{\alpha(\alpha+1)} \left[\frac{1}{\alpha-1} + \frac{1}{\alpha-2} + \ldots + \frac{1}{2}\right] \tag{3.5}$$

Since $\log a = \frac{1}{B(\alpha,2)} I(\alpha)$, the result follows. Using Theorems 7 and 8, the following upper bounds on the limiting RSL's are easily found.

Limiting RSL (as $p \to \infty$)

α	lim RSL(π,δ^C)
1	.0698
2	0
3	.0098
4	.0226
5	.0330
6	.0412

At first glance it seems undesirable that except for $\alpha = 2$, the limiting RSL's are not exactly zero. But the actual limits are fairly small and the values get stabilized for reasonably large p, so that a non-zero limiting RSL is not going to be serious in applications. The table indicates an increasing trend in the limits as α increases. One may suspect that for large α the RSL may get bad. The following final theorem shows it cannot.

Theorem 9.

$$\lim_{\alpha\to\infty} \lim_{p\to\infty} RSL(\pi,\delta^C) \le .0964.$$

Proof: Since $\log a = \log(\alpha-1) - 1 + \gamma + O(\frac{1}{\alpha})$, where $\gamma(\approx .57722)$ is the Euler's constant, Theorem 7 gives that

$$\lim_{\alpha\to\infty} \lim_{p\to\infty} RSL(\pi,\delta^C) \le (2e^{\gamma-1}-1)^2 \approx .096376.$$

Hence proved.

ACKNOWLEDGMENT

We are deeply indebted to Herman Rubin for his constant advice and suggestions throughout this work. Sincere thanks are also due to William Studden and Malay Ghosh for helpful comments, and to the National Science Foundation, grant DMS-8401996.

REFERENCES

Berger, James (1980). A robust generalized Bayes estimator and confidence region of a multivariate normal mean. Ann. Statist. 8, 716-761.

Berger, James (1981). Bayesian robustness and the Stein-effect. JASA 77, 358-368.

Berger, James (1982). Estimation in continuous exponential families; Bayesian estimation subject to risk restrictions and inadmissibility results. Statistical Decision Theory and Related Topics III (Vol. 2). Academic Press, New York.

Berger, James (1983). The robust Bayesian viewpoint. Robustness in Bayesian Statistics, J. Kadane (ed.). North-Holland, Amsterdam.

Berger, James (1985). Statistical Decision Theory and Bayesian Analysis. Springer-Verlag, New York.

Chen, S. Y. (1983). Restricted risk Bayes estimation for the means of the multivariate normal distribution. Technical Report #33, Department of Statistics, Purdue University.

DasGupta, Anirban (1986). Simultaneous estimation in the multiparameter gamma distribution under weighted quadratic losses. (to appear in Ann. Statist.)

DasGupta, Anirban and Rubin, Herman (1986). Bayes estimation of the mean of a multivariate normal distribution subject to full minimaxity in the case of a common unknown variance: a case for Bayesian robustness. Technical Report #86-10, Department of Statistics, Purdue University.

Efron, B. and Morris, C. (1971). Limiting the risk of Bayes and empirical Bayes estimators - Part I: The Bayes case. JASA 66, 807-815.

Gradshteyn, I. S. and Ryzhik, I. M. (1965). Tables of integrals, series, and products. Academic Press, New York.