# Γ-MINIMAX AND RESTRICTED-RISK BAYES ESTIMATION OF MULTIPLE POISSON MEANS UNDER $\varepsilon$-CONTAMINATIONS OF THE SUBJECTIVE PRIOR

by

Anirban DasGupta and Arup Bose
Purdue University

Technical Report#87-7

Department of Statistics
Purdue University

March 1987

*

# Γ-MINIMAX AND RESTRICTED-RISK BAYES ESTIMATION OF MULTIPLE POISSON MEANS UNDER ε-CONTAMINATIONS OF THE SUBJECTIVE PRIOR

by

Anirban DasGupta and Arup Bose
Purdue University

**Abstract**

For simultaneous estimation of $p$ Poisson means under the normalized squared-error loss $\sum_{i=1}^{p} \frac{(\theta_i - a_i)^2}{\theta_i}$, it is shown that the Γ-minimax rule under an ε-contamination class of priors coincides with the restricted-risk Bayes rule under a subjectively elicited prior under very general conditions. The restricted-risk Bayes rules are explicitly derived for a wide class of priors, including conjugate priors and their mixtures, and point priors. It is shown that these restricted-risk Bayes rules frequently minimize the Posterior Bayes risk itself and they enjoy desirable frequentist as well as Bayesian properties.

The class of priors considered allows subjective specification of such intuitive features as the mean and the standard deviation (or any two percentiles) of the unknown parameters and subsequent use of the rule which is restricted-risk Bayes with respect to a prior which is consistent with these subjective inputs. A characterization result on the form of the prior given the form of the Bayes estimates is also derived.

## 1. Introduction

Let $X_1, X_2, \ldots, X_p$ be $p$ independently distributed Poisson random variables with means $\theta_1, \theta_2, \ldots, \theta_p$ respectively, where $p \geq 2$; consider the problem of estimating $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)'$ under the normalized quadratic loss

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

Clevenson and Zidek (1975) proved that $\underset{\sim}{X}$, the UMVUE of $\underset{\sim}{\theta}$, is inadmissible for $p \geq 2$, and an improved estimator is

$$\delta_{CZ}(\underset{\sim}{X}) = (1 - \frac{p-1}{Z+p-1}) \cdot \underset{\sim}{X}, \tag{1.2}$$

where $Z = \sum_{i=1}^{p} X_i$. Since $\underset{\sim}{X}$ is minimax for all $p \geq 2$, so is $\delta_{CZ}(\underset{\sim}{X})$. Larger classes of minimax estimators have since been obtained; for example, it was proved in Tsui and Press (1982) that every estimator $\delta(\underset{\sim}{X})$ belonging to the class

$$\begin{aligned} \mathcal{D} = \{\delta : \ & \delta(\underset{\sim}{X}) = (1 - \frac{\phi(Z)}{Z+p-1}) \cdot \underset{\sim}{X}, \text{ where} \\ & 0 \leq \phi(Z) \leq 2(p-1), \text{ and } \phi(Z) \text{ is monotone non-decreasing}\}, \end{aligned} \tag{1.3}$$

is minimax. See also Hwang (1982) and Ghosh and Parsian (1981). In view of the fact that one can explicitly write down such a large class of minimax estimators, one needs to build up a well-formulated and systematic theory for selecting an estimator for actual use. The goal of this paper is to address this issue using two well known frequentist Bayes criteria, namely, the $\Gamma$-minimax and the restricted-risk Bayes criteria; we formally prove a mathematical equivalence between the two approaches in our context and give a closed form analytical solution under appropriate conditions. Note that the class $\mathcal{D}$ is certainly not the class of all minimax estimators in the Poisson problem. However, a <u>closed form</u> analytical solution to the restricted-risk Bayes problem is often not possible for larger classes of minimax estimators (see Theorem 2 in Berger (1982)). Restriction to $\mathcal{D}$ seems justifiable on this ground.

In the restricted-risk Bayes approach, one elicits a prior distribution $\pi_0(\underset{\sim}{\theta})$ and minimizes over $\delta \varepsilon \mathcal{D}$ the Bayes risk $r(\pi_0, \delta)$, which is the average of the risk function of $\delta(\underset{\sim}{X})$ with respect to the prior $\pi_0(\underset{\sim}{\theta})$. More generally, one may consider the problem of minimizing the posterior expected loss with respect to $\pi_0$ for estimators $\delta \varepsilon \mathcal{D}$. The restricted-risk Bayes approach has been considered by several authors; see Berger (1982), Bickel (1980), Efron and Morris (1971), DasGupta and Berger (1986), DasGupta and Rubin (1986), Hodges and Lehmann (1952), Marazzi (1985), etc. L. Brown writes in the discussion of Berger (1983): "In brief, there are many possible minimax rules. The only sensible way to choose among them seems to be to construct some crude prior distribution and then minimize the Posterior risk among the minimax rules ...." The resultant estimator will be attractive if it performs reasonably well with respect to the subjectively elicited prior $\pi_0(\underset{\sim}{\theta})$, because in this case the estimator has good risk properties by virtue of its minimaxity and yet the possible subjective gains in the overall Bayes risk are not substantially sacrificed to earn the minimax status. That this is indeed the case has been observed by several authors in continuous cases, notably the Normal and the Gamma cases; see Berger (1982), DasGupta and Berger (1986), and DasGupta and Rubin (1986).

The Γ-minimax approach stems from a desire for an added degree of conservativeness; unlike in the restricted-risk Bayes case, we now allow for the prior to belong to a family of probability distributions Γ. We then minimize over $\delta \varepsilon \mathcal{D}$, the quantity $r_\Gamma(\delta) = \sup_{\pi \varepsilon \Gamma} r(\pi, \delta)$. If Γ is very large so that $\sup_{\pi \varepsilon \Gamma} r(\pi, \delta) = \sup_\theta R(\theta, \delta)$ for every $\delta$, then the Γ-minimax problem would not have a unique solution, because each $\delta \varepsilon \mathcal{D}$ is minimax and hence $\sup_\theta R(\theta, \delta) = p$ for every $\delta \varepsilon \mathcal{D}$. The problem of selecting one estimator out of $\mathcal{D}$ then remains unsolved and we are back to square one. The idea, thus, is to allow a reasonably large class Γ and yet make sure that there is a unique solution to the Γ-minimax problem. We show in the subsequent sections that this can be done.

Again, several authors have considered the Γ-minimax approach; most results, thus far, on the Γ-minimax approach have been obtained when Γ consists of priors of a specific functional form (like, for example, a sub-class of the conjugate priors) or priors with a few specified moments. Most of these classes of priors are unappealing because the functional form or the moments of a prior are very difficult to specify. On the other hand, the Γ-minimax problem typically becomes untractable unless Γ is one of the classes described above. In this paper, we show that it is possible to work with a much more realistic class of priors, and yet give a unique closed form solution to the Γ-minimax problem. We consider the family of priors

$$\Gamma = \{\pi(\underset{\sim}{\theta}) : \pi(\underset{\sim}{\theta}) = (1-\varepsilon)\pi_0(\theta) + \varepsilon q(\underset{\sim}{\theta}), q\varepsilon\mathcal{L}\},$$

where $\pi_0(\underset{\sim}{\theta})$ is a fixed (perhaps subjectively elicited) prior, $0 < \varepsilon < 1$ is a fixed number, and $q$ belongs to a suitable class $\mathcal{L}$ of probability distributions on the parameter space (H). Γ is known as the $\varepsilon$-contamination family of priors and has recently received considerable attention from many authors; see Berger (1986), Berger and Berliner (1986), Berger and Sivaganesan (1986), Huber (1973, 1981) among others. The motivation of the class Γ is that $|P_\pi(\underset{\sim}{\theta}\varepsilon A) - P_{\pi_0}(\underset{\sim}{\theta}\varepsilon A)| \leq \varepsilon$ for every $\pi\varepsilon\Gamma$, and every (measurable) subset A of (H). The $\varepsilon$-contamination class thus allows a maximum error of amount $\varepsilon$ in subjectively specifying the probability of a set. Various choices of $\mathcal{L}$ have been proposed in one dimension: for example, the class of all possible distributions, the class of all unimodal distributions, or the class of all symmetric unimodal distributions, etc. For a general treatment of the Γ-minimax problem in various cases, see Berger (1979), Gupta and Hsiao (1981), Jackson et al. (1970), Kudo (1967), and Robbins (1951, 1964).

In Section 2, we show that the Γ-minimax problem under very general choices for $\mathcal{L}$ is formally equivalent to the restricted-risk Bayes problem under the subjective prior $\pi_0$. We then derive in sections 2 and 3 the explicit form of the restricted-risk Bayes rule under suitable conditions on $\pi_0$. In particular, conjugate, conjugate mixtures, and point priors have been considered. In the process, an interesting result on characterization of priors from the form of the Bayes estimates emerges. In contrast with earlier works on restricted-risk Bayesian estimation (e.g., Berger (1982), Chen (1983), Marazzi (1985)), we can actually minimize the posterior expected loss for a large class of priors. In Section 4, we investigate the extent to which the restricted-risk Bayes rule fails of being unconstrained Bayes. In particular, upper bounds on the RSL's (see Efron and Morris (1971)) of the restricted Bayes rule are derived for quite general priors $\pi_0$ and exact expressions for the

RSL's in a conjugate example are worked out. The results show that for $p \geq 3$, the $\Gamma$-minimax and the restricted-risk Bayes rules perform virtually as well as the unrestricted Bayes rule and even for $p = 2$, the performance is very respectable.

## 2. Derivation of the Γ-minimax and the restricted-risk Bayes rules

We first show the formal equivalence of the $\Gamma$-minimax problem for the $\varepsilon$-contamination class of priors and the restricted-risk Bayes problem for the initial prior $\pi_0$, under very general choices of $\mathcal{L}$.

Lemma 1. Let $\delta(\underset{\sim}{X})$ be as in (1.3). Then $\lim_{\Sigma\theta_i \to \infty} R(\underset{\sim}{\theta}, \underset{\sim}{\delta}) = p$.

**Proof.** It suffices to show that $|E\sum_i \frac{(\delta_i(x)-\theta_i)^2}{\theta_i} - E\sum_i \frac{(x_i-\theta_i)^2}{\theta_i}| \to 0$ as $\Sigma\theta_i \to \infty$.

Observe that
$$|E\sum_i \frac{(\delta_i(X)-\theta_i)^2}{\theta_i} - E\sum_i \frac{(X_i-\theta_i)^2}{\theta_i}|$$
$$= |E\sum_i \frac{(\delta_i(X)-X_i)^2}{\theta_i} + 2E\sum_i \frac{(\delta_i(X)-X_i)(X_i-\theta_i)}{\theta_i}|$$
$$= |E\frac{\phi^2(Z)}{(Z+p-1)^2}\cdot\sum_i \frac{X_i^2}{\theta_i} - 2E\frac{\phi(Z)}{Z+p-1}\cdot\sum_i \frac{X_i(X_i-\theta_i)}{\theta_i}|$$
$$\leq 4(p-1)^2\cdot E\frac{1}{(Z+p-1)^2}\cdot\sum_i \frac{X_i^2}{\theta_i} + 4(p-1)E\frac{1}{Z+p-1}\cdot \sum_i \frac{X_i(X_i-\theta_i)}{\theta_i} \tag{2.1}$$

The result now follows from (2.1) on using the facts that conditional on $Z = z, x_i$ is Bin $(z, \frac{\theta_i}{\Sigma\theta_j})$, and that if $Z$ has a Poisson distribution with mean $\lambda$, then $E\frac{1}{Z+C} \leq e^{-\lambda} + \frac{1}{\lambda}$ for $c \geq 1$. We will now prove that the $\Gamma$-minimax and the restricted-risk Bayes approaches are equivalent if the class $\mathcal{L}$ contains all uniform distributions on compact sets as (weak) limit points.

Theorem 1. For $K, M \geq 0$, let $Q_{K,M}(\underset{\sim}{\theta})$ denote the uniform distribution on the set $C = \{\underset{\sim}{\theta} : K \leq \Sigma\theta_i \leq K+M\}$. Suppose for every $K, M$, there exist $Q_n \varepsilon \mathcal{L}$ such that

$Q_n \xrightarrow{\omega} Q_{K,M}$. Then

$$\begin{aligned} &\inf_{\delta \varepsilon \mathcal{D}} \sup_{\pi \varepsilon \Gamma} r(\pi, \delta) = \sup_{\pi \varepsilon \Gamma} r(\pi, \delta_o) \\ iff \quad &\inf_{\delta \varepsilon \mathcal{D}} r(\pi_0, \delta) = r(\pi_0, \delta_0). \end{aligned}$$

**Proof:** First note that for every $\delta \varepsilon \mathcal{D}$,

$$\begin{aligned} \sup_{\pi \varepsilon \Gamma} r(\pi, \delta) &= (1-\varepsilon) r(\pi_0, \delta) + \varepsilon \sup_{q \varepsilon \mathcal{L}} r(q, \delta) \\ &\leq (1-\varepsilon) r(\pi_0, \delta) + \varepsilon \sup_{\underset{\sim}{\theta}} R(\underset{\sim}{\theta}, \delta) \\ &= (1-\varepsilon) r(\pi_0, \delta) + \varepsilon p. \end{aligned} \tag{2.2}$$

Next, note that for any $\delta \varepsilon \mathcal{D}$, and any $\gamma > 0$, Lemma 1 implies that $\exists K > 0$ such that $R(\underset{\sim}{\theta}, \delta) \geq p - \gamma$ if $\sum_{i=1}^{p} \theta_i \geq K$. Let $M > 0$ and let $Q_n \varepsilon \mathcal{L}$ converge weakly to $Q_{K,M}$.

$$\begin{aligned} \text{Hence }, \sup_{\pi \varepsilon \Gamma} r(\pi, \delta) &\geq (1-\varepsilon) r(\pi_0, \delta) + \varepsilon \int R(\underset{\sim}{\theta}, \delta) dQ_n(\underset{\sim}{\theta}) \\ &\geq (1-\varepsilon) r(\pi_0, \delta) + \varepsilon (p - \gamma) \int_{K \leq \Sigma \theta_i \leq K+M} dQ_n(\underset{\sim}{\theta}) \end{aligned} \tag{2.3}$$

Since $Q_n \xrightarrow{\omega} Q_{K,M}$, and $Q_{K,M}(\partial c) = 0$, it follows that $Q_n(C) \to Q_{K,M}(c) = 1$; it then follows from (2.3) that $\sup_{\pi \varepsilon \Gamma} r(\pi, \delta) \geq (1-\varepsilon) r(\pi_0, \delta) + \varepsilon (p - \gamma)$. Since $\gamma > 0$ is arbitrary,

$$\sup_{\pi \varepsilon \Gamma} r(\pi, \delta) \geq (1-\varepsilon) r(\pi_0, \delta) + \varepsilon p. \tag{2.4}$$

The assertion of the Theorem now follows from (2.2) and (2.4).

Remark. Actually, Theorem 1 remains valid if for large $K$, one can find a sequence of distributions from $\mathcal{L}$ converging weakly to some measure $Q$ whose support is contained in $\{\underset{\sim}{\theta} : \Sigma \theta_i \geq K\}$. The uniform distributions on compact sets $\{\underset{\sim}{\theta} = K \leq \Sigma \theta_i \leq K + M\}$ are just convenient distributions of such type. For most choices of $\mathcal{L}$, one will be able to find a sequence $Q_n \varepsilon \mathcal{L}$ having such a measure $Q$ as its weak limit, and the uniform measures described above will work in most cases. Theorem 1 thus asserts that the $\Gamma$-minimax and the restricted-risk Bayes approaches are equivalent for very general choices of $\mathcal{L}$, so long as one restricts attention to estimators in the class $\mathcal{D}$. Note that an arbitrary minimax estimator may not necessarily satisfy the limiting risk property stated in Lemma 1.

In view of Theorem 1, in order to find a $\Gamma$-minimax rule with respect to the $\varepsilon$-contamination class of priors, it is enough to find the restricted Bayes rule with respect to $\pi_o$ in $\mathcal{D}$. The following analysis dwells on deriving the form and the properties of the restricted-risk Bayes rule under various types of prior $\pi_o(\underset{\sim}{\theta})$. The analysis is easier if $\pi_o(\underset{\sim}{\theta})$ is such that the resulting unrestricted Bayes rule is of the form

$$\delta_{\pi_o}(\underset{\sim}{x}) = (1 - f(Z)) \cdot \underset{\sim}{X}, \text{ where } Z = \sum_{i=1}^{p} X_i \tag{2.5}$$

We now prove a result which states that subject to smoothness and regularity conditions, the Bayes rule is of the form (2.5) if and only if the prior density is a function of $\sum_{i=1}^{p} \theta_i$ alone.

**Theorem 2.** Let $x_i \overset{\text{indep.}}{\sim} Poi(\theta_i), 1 \le i \le p$. Consider the problem of estimating $\underset{\sim}{\theta}$ under the loss

$$L(\underset{\sim}{\theta}, \underset{\sim}{a}) = \sum_{i=1}^{p} \frac{(\theta_i - a_i)^2}{\theta_i}.$$

Suppose $\pi_o(\underset{\sim}{\theta})$ is a prior density on $(0,\infty)^p$ such that $\frac{\partial}{\partial \theta_i}\pi_o(\underset{\sim}{\theta})$ exists and is continuous on $(0,\infty)^p$ for every $i$. Also assume

(a) For $1 \le i \le p$, $\lim_{\theta_i \to 0,\infty} \prod_{j=1}^{p} \theta_j^{x_j} e^{-\sum_{j=1}^{p} \theta_j} \pi_0(\underset{\sim}{\theta}) = 0$ whenever $x_i > 0$

(b) $\int \prod_j \theta_j^{x_j}\, e^{-\Sigma\theta_j} \pi_0(\underset{\sim}{\theta})\, d\underset{\sim}{\theta}$ and $\int \prod_j \theta_j^{x_j} e^{-\Sigma\theta_j}\, \frac{\partial}{\partial\theta_i}\pi_0\, (\underset{\sim}{\theta}) d\underset{\sim}{\theta}$ are finite for each $i$ and all $x_1, \ldots, x_p \varepsilon Z_+ = \{0, 1, 2, \ldots\}$.

(c) $\int \theta_i^{-1} \prod_{j \ne i} \theta_j^{x_j} e^{-\Sigma\theta_j} \pi_0(\underset{\sim}{\theta}) d\underset{\sim}{\theta} = \infty$ for all $x_1, x_2, \ldots, x_p \varepsilon Z_+$.

(d) For some $t_1, t_2, \ldots, t_p > o$

$\int \cdots \int e^{\Sigma t_j \theta_j} e^{-\Sigma\theta_j} |\frac{\partial}{\partial\theta_i}\pi_0(\underset{\sim}{\theta})| d\underset{\sim}{\theta} < \infty$ for $i = 1, 2, \ldots, p$.

Then $\delta_{\pi_0}(\underset{\sim}{X})$ is of the form (2.5) iff $\pi_0$ is a function of $\Sigma\theta_j$ alone.

**Proof:** We first prove the sufficient part. Since the loss function is $\sum_{i=1}^{p} \frac{(\theta_i - a_i)^2}{\theta_i}$, for $X_i > 0$,

$$E\delta_{\pi_0,i}(\underset{\sim}{X}) = (E(\theta_i^{-1}/\underset{\sim}{X}))^{-1}$$

$$= \frac{\int \prod_j \theta_j^{x_j} e^{-\Sigma\theta_j} \pi_0(\Sigma\theta_j) d\underset{\sim}{\theta}}{\int \theta_i^{x_i - 1} \prod_{j \neq i} \theta_j^{x_j} e^{-\Sigma\theta_j} \pi_0(\Sigma\theta_j) d\underset{\sim}{\theta}}. \tag{2.6}$$

The denominator of (2.6), on integration by parts, and by virtue of assumption (a), equals

$$\frac{1}{x_i} \int \prod_j \theta_j^{x_j} \left\{ e^{-\Sigma\theta_j} \pi_o(\Sigma\theta_j) - e^{\Sigma\theta_j} \pi_o'(\Sigma\theta_j) \right\} d\underset{\sim}{\theta} \tag{2.7}$$

One now has from (2.6),

$$\delta_{\pi_o}, i(\underset{\sim}{x}) = x_i \left[ 1 - \frac{\int \prod_j \theta_j^{x_j} e^{-} \Sigma\theta_j \pi_o'(\Sigma\theta_j) d\underset{\sim}{\theta}}{\int \prod_j \theta_j^{x_j} e^{\,-\Sigma\theta_j} \pi_o(\Sigma\theta_j) d\underset{\sim}{\theta}} \right]^{-1} \tag{2.8}$$

Transforming $(\theta_1, \theta_2, \ldots, \theta_p)$ to $(u_1, u_2, \ldots, u_p)$ where $u_i = \sum_{j=1}^{i} \theta_j$, one has, for any function $h(\Sigma\theta_j)$,

$$\begin{aligned}
&\int \prod_j \theta_j^{X_j} e^{-\Sigma\theta_j} h(\Sigma\theta_j) d\underset{\sim}{\theta} \\
&= \int_0^\infty e^{-u_p} h(u_p) du_p \ldots \int_0^{u_3} (u_3 - u_2)^{x_3} du_2 \int_0^{u_2} u_1^{x_1} (u_2 - u_1)^{x_2} du_1 \\
&= B(x_1 + 1, x_2 + 1) B(x_1 + x_2 + 2, x_3 + 1) \ldots B(x_1 + \ldots + x_{p-1} + p - 1, x_p + 1) \\
&\times \int_0^\infty e^{-u_p} u_p^{x_1 + x_2 + \ldots + x_p + p - 1} h(u_p) du_p
\end{aligned} \tag{2.9}$$

One now has from (2.8) and (2.9),

$$\delta_{\pi_{0,i}}(\underset{\sim}{X}) = \left[ 1 - \frac{\int_0^\infty e^{-u} u^{z+p-1} \pi_o'(u) du}{\int_0^\infty e^{-u} u^{z+p-1} \pi_0(u) du} \right]^{-1} X_i \tag{2.10}$$

This is clearly of the form (2.5). In order to prove the necessary part, we will have to show that $\pi_0(\underset{\sim}{\theta})$ is a function of $\Sigma\theta_j$ alone if $\delta_{\pi_0}(\underset{\sim}{X})$ is of the form (2.5). Note that by the Mean-value theorem, for any $\underset{\sim}{\theta}, \underset{\sim}{\eta}$,

$$\pi_0(\underset{\sim}{\theta}) - \pi_0(\underset{\sim}{\eta}) = (\underset{\sim}{\theta} - \underset{\sim}{\eta})' \cdot \nabla\pi_0(\underset{\sim}{\theta}^*), \tag{2.11}$$

where $\underset{\sim}{\theta}^*$ lies on the line segment joining $\underset{\sim}{\theta}$ and $\underset{\sim}{\eta}$. If now $\frac{\partial}{\partial\theta_i}\prod_0(\underset{\sim}{\theta})$ is independent of $i$ (i.e., for each $i$, the *ith* partial derivative is the same function), then $\nabla\pi_0(\underset{\sim}{\theta}^*)$ is proportional to the vector $\underset{\sim}{1} = (1, 1, \ldots, 1)'$ and (2.11) implies that $\pi_0(\underset{\sim}{\theta}) = \pi_0(\underset{\sim}{\eta})$ if $\Sigma\theta_j = \Sigma\eta_j$. It will thus suffice to show that $\frac{\partial}{\partial\theta_i}\pi_0(\underset{\sim}{\theta})$ is independent of $i$ if $\delta_{\pi_0}(\underset{\sim}{X})$ is of the form (2.5). We will prove this for the case $p = 2$. The proof for a general $p$ is exactly similar. Towards this end, first note that the argument leading to (2.7) and (2.8) give that

$$\delta_{\pi_o,i}(\underset{\sim}{x}) = x_i\left[1 - \frac{\int\prod_j \theta_j^{x_j} e^{-\Sigma\theta_j}\frac{\partial}{\partial\theta_i}\pi_o(\underset{\sim}{\theta})d\underset{\sim}{\theta}}{\int\pi_j\theta_j^{x_j}e^{-\Sigma\theta_j}\pi_0(\underset{\sim}{\theta})d\underset{\sim}{\theta}}\right]^{-1} \tag{2.12}$$

Observe that (2.12) holds even when $x_i = 0$, if one assumes that $\int\prod_j\theta_j^{x_j}\ e^{-\Sigma\theta_j}\frac{\partial}{\partial\theta_i}\ \pi_0(\underset{\sim}{\theta})d\underset{\sim}{\theta}$ and $\int\prod_j\theta_j^{x_j}\ e^{-\Sigma\theta_j}\pi_0(\underset{\sim}{\theta})d\underset{\sim}{\theta}$ are finite for all $x_1, \ldots, x_p$ and $\int\theta_i^{-1}\prod_{j\neq i}\theta_j^{x_j}e^{-\Sigma\theta_j}\pi_0(\underset{\sim}{\theta})d\underset{\sim}{\theta} = \infty$ for each $i$. In this case, it can be directly checked that the Bayes action is $\delta_{\pi_0,i}(\underset{\sim}{X}) = 0$ and the right side of (2.12) is also 0. This, together with the fact that $\delta_{\pi_o}(\underset{\sim}{x})$ is of the form (2.5), implies that $\int\prod_j\theta_j^{x_j}e^{-\Sigma\theta_j}\frac{\partial}{\partial\theta_i}\pi_0(\underset{\sim}{\theta})d\underset{\sim}{\theta}$ is independent of $i$. Hence,

$$\int\prod_j\theta_j^{x_j}e^{-\Sigma\theta_j}\left(\frac{\partial}{\partial\theta_1}\pi_0(\underset{\sim}{\theta}) - \frac{\partial}{\partial\theta_2}\pi_0(\underset{\sim}{\theta})\right)d\underset{\sim}{\theta} = 0 \tag{2.13}$$

$\forall\ x_1, x_2 \varepsilon Z_+$. Let $\frac{\partial}{\partial\theta_1}\pi_0(\underset{\sim}{\theta}) - \frac{\partial}{\partial\theta_2}\pi_0(\underset{\sim}{\theta}) = h(\theta_1, \theta_2)$. Then, one has from (2.13),

$$\begin{aligned}
&\int\theta_2^{x_2}e^{-\theta_2}d\theta_2\int e^{-\theta_1}\theta_1^{x_1}h(\theta_1,\theta_2)d\theta_1 = 0\forall x_1, x_2\varepsilon Z_+.\\
&\Rightarrow\int\theta_2^{x_2}e^{-\theta_2}p_{x_1}(\theta_2)d\theta_2 = 0\ \forall x_2\varepsilon Z_+,\\
&\text{where } p_{x_1}(\theta_2) = \int e^{-\theta_1}\theta_1^{x_1}h(\theta_1,\theta_2)d\theta_1.
\end{aligned} \tag{2.14}$$

By virtue of assumption (d), (2.14) implies that $\forall x_1\varepsilon Z_+, p_{x_1}(\theta_2) = 0$ for almost all $\theta_2$. The null set outside of which $p_{x_1}(\theta_2) = 0$ can be assumed to be the same for all values of $x_1$ because of the countability of $Z_+$. Now note that assumption (d) also implies that for almost all $\theta_2, \int e^{t_1\theta_1}e^{-\theta_1}|h(\theta_1,\theta_2)|d\theta_1 < \infty$; in view of this and the fact that $p_{x_1}(\theta_2) = 0$ for almost all $\theta_2$, one has $h(\theta_1,\theta_2) = 0$ a.e. $(\theta_1)$ for almost all $\theta_2$. Fubini's theorem now implies that the product measure of the set $\{(\theta_1,\theta_2) : h(\theta_1,\theta_2) \neq 0\}$ must be zero. The continuity of $h(\theta_1, \theta_2)$ now gives $\frac{\partial}{\partial\theta_1}\pi_0(\underset{\sim}{\theta}) \equiv \frac{\partial}{\partial\theta_2}\pi_0(\underset{\sim}{\theta})$ and hence that $\pi_0(\underset{\sim}{\theta})$ must be a function of $\Sigma\theta_j$ alone.

In view of Theorem 2, we will for the present consider only those priors $\pi_0$ which are functions of $\Sigma\theta_j$ alone. We next derive a sufficient condition for an estimator in $\mathcal{D}$ to be Bayes with respect to $\pi_0(\underset{\sim}{\theta}) = \pi_0(\Sigma\theta_j)$.

**Lemma 2.** Assume $\pi_0$ is a function of $\Sigma\theta_j$ alone and conditions (a), (b), and (c) of Theorem 2 hold. Let the unrestricted Bayes rule $\delta_{\pi_0}(\underset{\sim}{X})$ be of the form (2.5). Then, in order to find the restricted Bayes rule within $\mathcal{D}$ it is sufficient to minimize $\frac{z\cdot(f(z)-\frac{\phi(z)}{z+p-1})^2}{1-f(z)}$ for every fixed $z$; here $\phi(z)$ is as defined in (1.3).

**Proof:** First note that minimizing $r(\pi_0, \delta)$ is equivalent to minimizing $r(\pi_0, \delta) - r(\pi_0, \delta_{\pi_0})$. Now

$$\begin{aligned} & r(\pi_o, \delta) - r(\pi_o, \delta_{\pi_0}) \\ &= E_{\pi_0} E_{\underset{\sim}{x}|\underset{\sim}{\theta}} \left[ \sum_i \frac{(\theta_i - \delta_i(\underset{\sim}{x}))^2}{\theta_i} - \sum_i \frac{(\theta_i - \delta_{\pi_o,i}(\underset{\sim}{x}))^2}{\theta_i} \right] \\ &= E_{\pi_0} E_{\underset{\sim}{x}|\underset{\sim}{\theta}} \left[ \sum_i \frac{(\delta_{\pi_o,i}(\underset{\sim}{x}) - \delta_i(\underset{\sim}{x}))^2}{\theta_i} + 2\sum_i \frac{(\delta_{\pi_o,i}(\underset{\sim}{x}) - \delta_i(\underset{\sim}{x}))(\theta_i - \delta_{\pi_o,i}(\underset{\sim}{x}))}{\theta_i} \right] \end{aligned} \tag{2.15}$$

Since $\delta_i(\underset{\sim}{x}) = \delta_{\pi_o,i}(\underset{\sim}{x}) = 0$ if $x_i = 0$, (2.15) gives, on interchanging the order of the expectations,

$$\begin{aligned} & r(\pi_0, \delta) - r(\pi_0, \delta_{\pi_0}) \\ &= E_{m(\underset{\sim}{x})} \sum_{i:x_i>0} \frac{(\delta_{\pi_0,i}(\underset{\sim}{x}) - \delta_i(\underset{\sim}{x}))^2}{\delta_{\pi_0,i}(\underset{\sim}{x})}, \end{aligned} \tag{2.16}$$

where $m(\underset{\sim}{x})$ stands for the marginal distribution of $\underset{\sim}{X}$. Since it is enough to minimize the inside sum for each fixed $\underset{\sim}{x}$ in order to minimize the expectation, the lemma now follows on using $\delta(\underset{\sim}{X}) = (1 - \frac{\phi(Z)}{Z+p-1}) \cdot \underset{\sim}{X}$ and $\delta_{\pi_o}(\underset{\sim}{X}) = (1 - f(Z)) \cdot \underset{\sim}{X}$, and observing that $\sum_{i:X_i>0} X_i = \sum_{i=1}^{p} X_i = Z$ for every $\underset{\sim}{X}$.

Interestingly, it turns out that in a large number of cases, it is actually possible to minimize $\frac{z[f(z)-\frac{\phi(z)}{z+p-1}]^2}{1-f(z)}$ pointwise for every $z$; thus the restricted-risk Bayes rules also have the attractive property that they actually minimize the Posterior Bayes risk for every fixed $\underset{\sim}{x}$. We are now in a position to derive explicitly the form of the Γ-minimax rule (or equivalently, the restricted-risk Bayes rule) for various different priors $\pi_o(\Sigma\theta_j)$. An important special prior which is a function of $\Sigma\theta_j$ alone is the prior $\pi_o(\Sigma\theta_j) = e^{-\lambda\Sigma\theta_j}$; this corresponds to the case when $\theta_i$ are iid exponential with mean $\frac{1}{\lambda}$ (note that the prior $\pi_o$ here is unnormalized). We first derive the restricted-risk Bayes rule with respect to this prior. This derivation will motivate the form of the restricted Bayes rule for a more general $\pi_0$ and also illustrate the key ideas involved in the proof.

**Theorem 3.** Let $\theta_i's$ have a joint prior $\pi_o(\underset{\sim}{\theta}) = \pi_o(\Sigma\theta_j) = e^{-\lambda\Sigma\theta_j}, \lambda > 0$. Then the estimator $\delta^*_{\pi_o}(\underset{\sim}{X})$ defined as

$$\delta^*_{\pi_o}(\underset{\sim}{X}) = \frac{\underset{\sim}{X}}{\lambda+1} \text{ if } Z \leq \frac{\lambda+2}{\lambda}(p-1)$$

$$= (1 - \frac{2(p-1)}{Z+p-1}) \cdot \underset{\sim}{X} \text{ if } Z > \frac{\lambda+2}{\lambda}(p-1) \quad (2.17)$$

minimizes the Posterior Bayes risk with respect to $\pi_o$ in the class of estimators $\mathcal{D}$.

**Proof:** The proof involves two steps:

(i) Proving that $\delta^*_{\pi_o} \varepsilon \mathcal{D}$,

(ii) Proving that $\delta^*_{\pi_o}$ minimizes $\frac{z[f(z)-\frac{\phi(z)}{z+p-1}]^2}{1-f(z)}$, where $\delta_{\pi_o}(\underset{\sim}{x}) = (1 - f(z)) \cdot \underset{\sim}{x} = \frac{x}{\lambda+1}$ is the unrestricted Bayes rule and $\phi(z)$ is as in (1.3). First note that $\delta^*_{\pi_o}(\underset{\sim}{X}) = (1 - \frac{\phi^*(Z)}{Z+p-1}) \cdot \underset{\sim}{X}$, where

$$\phi^*(z) = \frac{\lambda}{\lambda+1}(z+p-1) \text{ if } z \leq \frac{\lambda+2}{\lambda}(p-1)$$

$$= 2(p-1) \text{ if } z > \frac{\lambda+2}{\lambda}(p-1) \quad (2.18)$$

Hence it is clear that $\delta^*_{\pi_o} \varepsilon \mathcal{D}$. To prove (ii), note that $\delta^*_{\pi_o}$ coincides with $\delta_{\pi_o}$ for $z \leq \frac{\lambda+2}{\lambda}(p-1)$; hence,

$$\frac{z[f(z) - \frac{\phi^*(z)}{z+p-1}]^2}{1-f(z)} = 0 \leq \frac{z[f(z) - \frac{\phi(z)}{z+p-1}]^2}{1-f(z)}, \tag{2.19}$$

for $z \leq \frac{\lambda+2}{\lambda}(p-1)$, and for every $\phi$.

To prove (ii) for $z > \frac{\lambda+2}{\lambda}(p-1)$, note that for each $\phi$ as in (1.3),

$$\frac{\phi(z)}{z+p-1} \leq \frac{2(p-1)}{z+p-1} \leq \frac{\lambda}{\lambda+1} \equiv f(z), \text{ for } z > \frac{\lambda+2}{\lambda}(p-1);$$

this implies that

$$\left(\frac{\lambda}{\lambda+1} - \frac{2(p-1)}{z+p-1}\right)^2 \leq \left(\frac{\lambda}{\lambda+1} - \frac{\phi(z)}{z+p-1}\right)^2, \tag{2.20}$$

$\forall z > \frac{\lambda+2}{\lambda}(p-1)$. This in turn implies (ii). The Theorem now follows.

The targeted result for more general priors $\pi_o(\Sigma\theta_j)$ will be to show that the restricted Bayes rule in $\mathcal{D}$ is of the form

$$\delta^*_{\pi_o}(\underset{\sim}{x}) = (1 - f(z)) \cdot \underset{\sim}{x} \text{ if } (z+p-1)f(z) \leq 2(p-1)$$

$$= (1 - \frac{2(p-1)}{z+p-1}) \cdot \underset{\sim}{x} \text{ if } (z+p-1)f(z) > 2(p-1). \tag{2.21}$$

In general, proving (ii) is easier than proving (i); in fact, $\delta^*_{\pi_o}$ again coincides with $\delta_{\pi_o}$ for $(z+p-1)f(z) \leq 2(p-1)$ and for $(z+p-1)f(z) > 2(p-1)$, one still has the inequality

$$\frac{\phi(z)}{z+p-1} \leq \frac{2(p-1)}{z+p-1} \leq f(z), \tag{2.22}$$

for each $\phi$ as in (1.3). Thus (ii) is always proved using the same argument as in Theorem 3. To prove (i), note that $\delta^*_{\pi_o}$ has the representation $\delta^*_{\pi_o}(\underset{\sim}{X}) = (1 - \frac{\phi^*(Z)}{Z+p-1}) \cdot \underset{\sim}{X}$, where

$$\phi^*(z) = (z+p-1)f(z) \text{ if } (z+p-1)f(z) \leq 2(p-1)$$

$$= 2(p-1) \text{ if } (z+p-1)f(z) > 2(p-1). \tag{2.23}$$

It is then clear that in order to prove that $\delta^*_{\pi_0} \epsilon \mathcal{D}$, it will be necessary and sufficient to show that $(z+p-1)f(z)$ is monotone increasing in z. This fact is not true for arbitrary $\pi_0(\underset{\sim}{\theta}) = \pi_0(\Sigma\theta_j)$. The bulk of the subsequent analysis deals with finding sufficient conditions on $\pi_0$ for this property to hold.

**Lemma 3** Let $\prod_0(\Sigma\theta_j)$ be log concave and decreasing as a function of $\Sigma\theta_j$. Then $(z+p-1)f(z)$ is monotone increasing in $z$.

**Proof:** First note that (2.9) implies that

$$f(z) = \frac{-I_1(z)}{I_2(z) - I_1(z)}, \tag{2.24}$$

where

$$I_1(z) = \int e^{-u} u^{z+p-1} \pi_o'(u) du,$$
$$\text{and } I_2(z) = \int e^{-u} u^{z+p-1} \pi_0(u) du \tag{2.25}$$

Hence,

$$f(z) = \frac{-I_1(z)/I_2(z)}{1 - I_1(z)/I_2(z)}$$

$$= \frac{E[-\pi_0'(u)/\pi_0(u)]}{1 + E[-\pi_0'(u)/\pi_0(u)]}. \tag{2.26}$$

where $u$ has the density $e^{-u}u^{z+p-1}\pi_0(u)$. Since this is MLR in $z$, it will follow that $\frac{E[-\pi_0'(u)]}{\pi_0(u)}$ is increasing in $z$ and hence $f(z)$ is increasing in $z$ if $\frac{-\pi_0'(u)}{\pi_0(u)}$ is increasing in $u$, i.e., if $\pi_0$ is log-concave. Also note that $f(z) \geq 0$ if $\pi_0$ is decreasing. Since $f(z) \geq 0$ and increasing, it follows that $(z+p-1)f(z)$ must be increasing too. This establishes the Lemma. We have the following Theorem.

**Theorem 4.** Let $\pi_0(\Sigma\theta_j)$ be log-concave and decreasing. Then $\delta^*_{\pi_0}(\underset{\sim}{x})$ defined in (2.21) minimizes the Posterior Bayes risk with respect to $\pi_0$ in $\mathcal{D}$.

**Example:** A simple example of a log-concave and decreasing prior is $\pi_0(u) = e^{-cu^\alpha}, c > 0, \alpha \geq 1$. Note that $\int e^{-c(\Sigma\theta_j)^\alpha} d\underset{\sim}{\theta} < \infty$ for each $c > 0, \alpha > 0$. In particular, $\alpha = 2$ gives the normal type prior.

Next we give sufficient conditions on suitable log-convex priors $\pi_0$ so that $(z+p-1)f(z)$ is increasing. Establishing such sufficient conditions for a general log-convex $\pi_0$ seems extremely difficult. An interesting subclass of log-convex priors is the family of conjugate mixtures

$$\pi_0(\Sigma\theta_j) = \int_0^\infty e^{-\lambda\Sigma\theta_j} p(\lambda) d\lambda \tag{2.27}$$

The following Lemma gives a sufficient condition on the mixing distribution $p(\lambda)$ for $(z+p-1)f(z)$ to be increasing.

**Lemma 4.** Let $\pi_0(\Sigma\theta_j) = \int e^{-\lambda\Sigma\theta_j} p(\lambda)d\lambda$. Let $g(x)$ denote the density of $x = \frac{\lambda}{1+\lambda}$. Then $(z+p-1)f(z)$ is increasing in $z$ if $\frac{xg'(x)}{g(x)}$ is decreasing in $x$.

**Proof:** Note that for each $u, \pi_0'(u) = -\int \lambda e^{-\lambda u} p(\lambda)d\lambda$. Hence, one has from (2.10),

$$1 - f(z) = \left(1 + \frac{\int\int e^{-(1+\lambda)u} u^{z+p-1} \lambda p(\lambda) du d\lambda}{\int\int e^{-(1+\lambda)u} u^{z+p-1} p(\lambda) du d\lambda}\right)^{-1}$$

$$= \left(1 + \frac{\int \frac{\lambda}{(1+\lambda)^{z+p}} p(\lambda)d\lambda}{\int \frac{1}{(1+\lambda)^{z+p}} p(\lambda)d\lambda}\right)^{-1}$$

$$\Rightarrow f(z) = \frac{\int \frac{\lambda}{1+\lambda} \cdot \frac{1}{(1+\lambda)^{z+p-1}} p(\lambda)d\lambda}{\int \frac{1}{(1+\lambda)^{z+p-1}} p(\lambda)d\lambda} \tag{2.28}$$

Changing the variable of integration to $x = \frac{\lambda}{1+\lambda}$, one has,

$$f(z) = \frac{\int_0^1 x(1-x)^{z+p-1} g(x)dx}{\int_0^1 (1-x)^{z+p-1} g(x)dx} \tag{2.29}$$

Thus, in order to show that $(z+p-1)f(z)$ is increasing in $z$, we will need to show that for $z_2 > z_1 > 0$,

$$z_2 \cdot \frac{\int_0^1 x(1-x)^{z_2} g(x)dx}{\int_0^1 (1-x)^{z_2} g(x)dx} \geq \frac{z_1 \int_0^1 y(1-y)^{z_1} g(y)dy}{\int_0^1 (1-y)^{z_1} g(y)dy}$$

$$\Leftrightarrow \int_0^1 \int_0^1 (z_2 x - z_1 y)(1-x)^{z_2}(1-y)^{z_1} g(x)g(y)dxdy \geq 0$$

$$\Leftrightarrow \int_0^{z_2} \int_0^{z_1} (x-y)(1-\frac{x}{z_2})^{z_2}(1-\frac{y}{z_1})^{z_1} g(\frac{x}{z_2})g(\frac{y}{z_1})dxdy \geq 0 \qquad (2.30)$$

Observe that for $x > y$, the integrand in (2.30) is positive; hence, it will suffice to show that

$$\int_0^{z_1} \int_0^{z_1} h(x,y)(1-\frac{x}{z_2})^{z_2}(1-\frac{y}{z_1})^{z_1} g(\frac{x}{z_2})g(\frac{y}{z_1})dxdy \geq 0, \qquad (2.31)$$

where $h(x,y) = x - y$.

Since $h(x,y) = -h(y,x)$, and $h(x,y) \geq 0$ for $x \geq y$, it will suffice to show that for $x > y$, the averaging measure gives bigger mass at $(x,y)$ than at $(y,x)$, i.e., for $x > y$,

$$\frac{g(\frac{x}{z_2})(1-\frac{x}{z_2})^{z_2}}{g(\frac{y}{z_2})(1-\frac{y}{z_2})^{z_2}} \geq \frac{g(\frac{x}{z_1})(1-\frac{x}{z_1})^{z_1}}{g(\frac{y}{z_1})(1-\frac{y}{z_1})^{z_1}} \qquad (2.32)$$

Inequality (2.32) will follow if we can show that for $x > y$,

$$\gamma(a) = \frac{(1-ax)^{\frac{1}{a}} g(ax)}{(1-ay)^{\frac{1}{a}} g(ay)} \qquad (2.33)$$

is decreasing in $a$.

Now, $log\ \gamma(a) = log\ g(ax) - log\ g(ay) + \frac{1}{a}\ log\ (1-ax) - \frac{1}{a}\ log\ (1-ay)$

$$\Rightarrow (log\ \gamma(a))' = \frac{xg\prime(ax)}{g(ax)} - \frac{yg\prime(ay)}{g(ay)} - \frac{log(1-ax)}{a^2}$$

$$- \frac{x}{a(1-ax)} + \frac{log(1-ay)}{a^2} + \frac{y}{a(1-ay)}$$

$$= \frac{1}{a}\left\{\frac{axg\prime(ax)}{g(ax)} - \frac{ayg\prime(ay)}{g(ay)}\right\} + \frac{1}{a^2}\left\{\frac{ay}{1-ay} + log\ (1-ay) - \frac{ax}{1-ax} - log\ (1-ax)\right\}.$$

$$= T_1 + T_2 \text{ (say).} \qquad (2.34)$$

We first prove that $T_2 \leq 0$. Observe that to do this, it will be enough to prove $\frac{z}{1-z} + log(1-z)$ is increasing for $0 < z < 1$. This follows immediately on differentiation.

Since $T_2 \leq 0$, clearly, for $\gamma(a)$ to be decreasing it will suffice if

$$\frac{ax\ g\prime(ax)}{g(ax)} \leq \frac{ay\ g\prime(ay)}{ay} \text{ whenever } x > y$$

$$\Leftrightarrow \frac{xg\prime(x)}{g(x)} \text{ is decreasing in } x.$$

This proves the Lemma.

**Theorem 5.** Let $\pi_0(\underset{\sim}{\theta}) = \int e^{-\lambda\Sigma\theta_j} p(\lambda)d\lambda$ and suppose $p(\lambda)$ satisfies the condition of Lemma 4. Then $\delta^*_{\pi_0}$ defined in (2.21) minimizes the Posterior Bayes risk in $\underset{\sim}{D}$.

We now give examples of mixing distributions $p(\lambda)$ which satisfy this condition.

**Example 1.** Let $p(\lambda) = \lambda^{m-1}(1+\lambda)^{-(m+n)}, m > 0, n \geq 1$.

$$\text{Then } g(x) = x^{m-1}(1-x)^{n-1}, 0 < x < 1$$

$\therefore \frac{xg\prime(x)}{g(x)} = (m-1) - \frac{(n-1)}{1-x}$, which is decreasing in $x$ if $n \geq 1$.

Note that these are precisely the mixing distributions which result in Proper Bayes minimax estimators in the Poisson problem for $p \geq 3$. See Ghosh and Parsian (1981).

**Example 2.** Let $p(\lambda) = e^{-c\lambda}\lambda^m, m > p-1,\ c \geq m+2$.

$$\begin{aligned}\text{First note that } \int \pi_0(\Sigma\theta_j)d\underset{\sim}{\theta} &= \int e^{-\lambda(c+\Sigma\theta_j)}\lambda^m d\lambda d\underset{\sim}{\theta} \\ &= \int \frac{\Gamma(m+1)}{(c+\Sigma\theta_j)^{m+1}}\, d\underset{\sim}{\theta} \\ &= \text{constant} \cdot \int_0^\infty \frac{u^{p-1}}{(c+u)^{m+1}} du \\ &< \infty \text{ if } m > p-1.\end{aligned}$$

Next, note that the density of $x = \frac{\lambda}{1+\lambda}$ is

$$g(x) = e^{-\frac{cx}{1-x}} \cdot x^m (1-x)^{-m-2}$$

$$\Rightarrow \frac{xg\prime(x)}{g(x)} = -\frac{cx}{(1-x)^2} + m + \frac{(m+2)x}{1-x}$$

$$= -\frac{cx}{(1-x)^2} + \frac{(m+2)}{1-x} - 2 \qquad (2.35)$$

Differentiating once more, $\frac{xg\prime(x)}{g(x)}$ is decreasing

$$\text{iff} - \frac{c}{(1-x)^2}\,\frac{-2cx}{(1-x)^3} + \frac{(m+2)}{(1-x)^2} \le 0 \text{ for } 0 < x < 1$$

$$\text{iff}\{c - (m+2)\}\ (1-x) \ge -2cx \text{ for } 0 < x < 1,$$

which holds if $c \ge m+2$.

Finally, we now give examples of a few prior which are not covered by either Theorem 4 or Theorem 5.

**Example 3.** Let $\pi_0(u) = e^{-u}u^n, 0 < n \le p-1$. Note that $\pi_0(u)$ is log-concave, although not decreasing. The proof of Lemma 3 shows that for $(z+p-1)f(z)$ to be increasing, we merely need to show that $f(z) \ge 0$ in this case, since $f(z)$ is increasing by virtue of the log-concavity of $\pi_0$. Also, (2.24) implies that $f(z) \ge 0$ provided

$$I_1(z) = \int e^{-u}u^{z+p-1}\pi_0'(u)du \le 0 \qquad (2.36)$$

By direct computation,

$$I_1(z) = \frac{n \cdot \Gamma(z+n+p-1)}{2^{z+n+p-1}} - \frac{\Gamma(z+n+p)}{2^{z+n+p}};$$

hence, (2.36) holds iff $n \le \frac{z+n+p-1}{2}\ \forall z \ge 0$, i.e. iff $n \le p-1$.

**Example 4.** Let $\pi_0(u) = e^{-u}u^{-n}, 0 < n < p$. In this case, $\pi_0(u)$ is decreasing so that $f(z) \ge 0$; however, $\pi_0(u)$ is log-convex, but cannot be written in the form $\int e^{-\lambda u}p(\lambda)d\lambda$ for any mixing distribution $p(\lambda)$ (see Berger (1975)). It remains to show that $(z+p-1)f(z)$ is increasing in $z$.

By direct computation,

$$I_1(z) = -\left[\frac{\Gamma(z+p-n)}{2^{z+p-n}} + n \cdot \frac{\Gamma(z+p-n-1)}{2^{z+p-n-1}}\right],$$

$$\text{and } I_2(z) = \frac{\Gamma(z+p-n)}{2^{z+p-n.}} \tag{2.37}$$

Hence, (2.24) gives,

$$f(z) \frac{= \frac{\Gamma(z+p-n)}{2^{z+p-n}} + \frac{n\cdot\Gamma(z+p-n-1)}{2^{z+p-n-1}}}{\frac{2\cdot\Gamma(z+p-n)}{2^{z+p-n}} + \frac{n\cdot\Gamma(z+p-n-1)}{2^{z+p-n-1}}}$$

$$= \frac{\frac{z+p-n-1}{2} + n}{z+p-1}, \tag{2.38}$$

which immediately implies that $(z+p-1)f(z)$ is increasing in $z$. Finally note that $\int \pi_0(\underset{\sim}{\theta})d\underset{\sim}{\theta} = \int_0^\infty e^{-u}u^{p-n-1}du < \infty$ if $p-n>0$, i.e., if $n<p$.

## 3. General conjugate and Point priors

In this section, we address the problem of minimizing the Bayes risk in $\mathcal{D}$ when the prior $\pi_0$ not a function of $\Sigma\theta_j$ alone. Important examples of such priors are point priors and the general conjugate priors with density $e^{-\lambda\Sigma\theta_j}\pi\theta_j^{\alpha-1}$, where $\lambda, \alpha > 0$. If $\pi_0$ is a point mass at $\underset{\sim}{\theta_o}$, the problem is one of minimizing $R(\underset{\sim}{\theta_o}, \delta)$ for $\delta\varepsilon\mathcal{D}$. This problem arises naturally and becomes interesting if one has strong prior belief that $\underset{\sim}{\theta}$ is near $\underset{\sim}{\theta_o}$, but at the same time wants full protection against misspecification of prior information. The analysis is easier if we assume $\underset{\sim}{\theta_o} = \theta_o \cdot \underset{\sim}{1}$, where $\theta_o > 0$, and $\underset{\sim}{1} = (1,1,\ldots,1)'$. In an empirical Bayes scenario when the $\theta_i'$s are thought to be similar or exchangeable, this might be an interesting problem. The general conjugate priors with parameters $\lambda$ and $\alpha$ allow one to subjectively specify such intuitive features as the mean and the standard deviation of the $\theta_i$'s; one can then use a conjugate prior consistent with these inputs and use the resulting restricted-risk Bayes rule. It is for this reason that the special exponential priors (and their mixtures) treated in section 2 are not adequate for our purpose because they do not permit subjective specification of a mean and a standard deviation.

**Theorem 6.** Let $\pi_0(\underset{\sim}{\theta})$ be the point mass at $\underset{\sim}{\theta_0} = \theta_0 \cdot \underset{\sim}{1}$. Then the following estimator $\delta^*_{\pi_0}$ minimizes the Bayes risk with respect to $\pi_0$ in $\underset{\sim}{D}$:

$$\delta^*_{\pi_0}(\underset{\sim}{x}) = (\frac{p\theta_0}{z+p-1}) \cdot \underset{\sim}{x} \text{ if } z \leq p-1+p\theta_0$$

$$= (1 - \frac{2(p-1)}{z+p-1}) \cdot \underset{\sim}{x} \text{ if } z > p-1+p\theta_0. \tag{3.1}$$

**Proof:** First note that the unrestricted Bayes rule $\delta_\pi(\underset{\sim}{X}) = \underset{\sim}{\theta_o}$ a.s. $(P_{\underset{\sim}{\theta}})$ for every $\underset{\sim}{\theta}$. Mimicking the argument of Lemma 2, one has that for minimizing the Bayes risk $r(\pi_0, \delta)$, it suffices to minimize

$$E_{m(\underset{\sim}{x})} \sum_{i=1}^{p} \frac{(\delta_i(\underset{\sim}{x}) - \theta_o)^2}{\theta_0}$$

$$= E_{m(Z)} E_{\underset{\sim}{x}|Z=z} \sum_{i=1}^{p} \left[ \frac{(1 - \frac{\phi(z)}{z+p-1})x_i - \theta_o)^2}{\theta_0} \right], \tag{3.2}$$

where $m(\underset{\sim}{X})$ and $m(Z)$ denote the marginal distributions of $\underset{\sim}{X}$ and Z respectively. Clearly, to minimize (3.2), it is enough to minimize the inside conditional expectation for every $z$. The conditional expectation is a constant independent of $\phi(z)$ if $z = 0$. We will therefore consider the minimization problem only for $z > 0$.

Now note that, marginally, $X_i \overset{iid}{\sim} Po_i(\theta_0)$; hence;

$$\mu_1(z) = E(X_i|Z=z) = \frac{z}{p},$$

$$\text{and } \mu_2(z) = E(X_i^2|Z=z) = z \cdot \frac{1}{p}(1 - \frac{1}{p}) + (\frac{z}{p})^2$$

$$= \frac{z(p-1) + z^2}{p^2} \tag{3.3}$$

Also, minimizing the conditional expectation in (3.2) is equivalent to minimizing

$$(1-\frac{\phi(z)}{z+p-1})^2\frac{\mu_2(z)}{\theta_0}-2\,(1-\frac{\phi(z)}{z+p-1})\cdot\mu_1(z)$$

$$=\frac{\mu_2(z)}{\theta_0}(1-\frac{\phi(z)}{z+p-1}-\theta_0\,\frac{\mu_1(z)}{\mu_2(z)})^2-\frac{\theta_0\mu_1^2(z)}{\mu_2(z)}. \tag{3.4}$$

It thus suffices to minimize (over $\phi$) $(z+p-1-\frac{\theta_0\mu_1(z)(z+p-1)}{\mu_2(z)}-\phi(z))^2$. The expressions for $\mu_1(z)$ and $\mu_2(z)$ in (3.3) give that $\frac{\theta_0\mu_1(z)(z+p-1)}{\mu_2(z)}=p\theta_0$; one thus has to minimize (for $z>0$) $(z+p-1-p\theta_0-\phi(z))^2$.

Now note that the estimator in (3.1) is of the form

$$\delta^*_{\pi_0}(\underset{\sim}{X})=(1-\frac{\phi^*(z)}{z+p-1})\cdot\underset{\sim}{X},\ \text{where}$$
$$\phi^*(z)=z+p-1-p\theta_0\ \text{if}\ z\le p-1+p\theta_0$$
$$=2(p-1)\ \text{if}\ z>p-1+p\theta_0 \tag{3.5}$$

Clearly, $\phi^*$ satisfies the properties described in (1.3); also, the argument provided in Theorem 3 gives that $\phi^*$ indeed minimizes $(z+p-1-p\theta_0-\phi(z))^2$. This proves the Theorem.

**Remark.** One major difference between Theorem 6 and the preceding Theorems is that the restricted-risk Bayes rule **never** coincides with the unrestricted Bayes rule $\delta_{\pi_0}(\underset{\sim}{X})$ when $\pi_0$ is a point mass. The reason for this is apparent; when $\pi_0$ is a point mass, $\delta_{\pi_0}(\underset{\sim}{X})$ is **not** an estimator of the form (2.5). We now derive the restricted-risk Bayes rules under general conjugate priors $\pi_0$ with shape parameter $\alpha$ and scale parameter $\lambda$. An important difference between the exponential priors and the general conjugate priors is that the unrestricted Bayes rule $\delta_{\pi_0}(\underset{\sim}{X})$ for a general conjugate prior is $\delta_{\pi_0,i}(\underset{\sim}{x})=\frac{x_i+\alpha-1}{\lambda+1}$, and is thus not of the form (2.5).

Theorem 7. Let $\pi_0(\underset{\sim}{\theta})=e^{-\lambda\sum_j\theta_j}\prod_j\theta_j^{\alpha-1}$, $\lambda,\alpha>0$. Then the following estimator $\delta^*_{\pi_0}$

minimizes the Bayes risk with respect to $\pi_0$ in $\mathcal{D}$ :

$$\delta^*_{\pi_0}(\underset{\sim}{X}) = \underset{\sim}{X} \quad \text{if} \quad \frac{\lambda(Z+p-1)+p(1-\alpha)}{\lambda+1} < 0$$

$$= \left(1 - \frac{p(1-\alpha)}{Z+p-1}\right) \cdot \frac{\underset{\sim}{X}}{\lambda+1} \quad \text{if} \quad 0 \le \frac{\lambda(Z+p-1)+p(1-\alpha)}{\lambda+1} \le 2(p-1)$$

$$= \left(1 - \frac{2(p-1)}{Z+p-1}\right) \cdot \underset{\sim}{X} \quad \text{if} \quad \frac{\lambda(Z+p-1)+p(1-\alpha)}{\lambda+1} > 2(p-1).$$

**Proof:** As in (3.2), in order to minimize the Bayes risk, it suffices to minimize for every $z$

$$E_{\underset{\sim}{X}|Z=z} \sum_{i=1}^{p} \frac{[(1-\frac{\phi(z)}{z+p-1}) \cdot X_i - \frac{X_i+\alpha-1}{\lambda+1}]^2(\lambda+1)}{X_i+\alpha-1}. \tag{3.6}$$

If we let $\mu_1(z) = E(X_i|Z=z)$ and $\mu_2(z) = E(\frac{X_i^2}{X_i+\alpha-1}|Z=z)$, then an argument similar to the one leading to (3.4) shows that minimizing (3.6) is equivalent to minimizing

$$\left(z+p-1-\frac{(z+p-1)\mu_1(z)}{(\lambda+1)\mu_2(z)} - \phi(z)\right)^2$$

$$= \left(z+p-1-\frac{z(z+p-1)}{p(\lambda+1)\mu_2(z)} - \phi(z)\right)^2, \tag{3.7}$$

since $\mu_1(z) = \frac{z}{p}$.

We need to explicitly calculate $\mu_2(z)$ in order to minimize (3.7) for every $z$.

Towards this end, first note that, marginally, $X_i$'s are iid Negative Binomial with

$$P(X_i = x) = \binom{\alpha+x-1}{\alpha-1} \frac{\lambda^\alpha}{(\lambda+1)^{\alpha+x}}, x = 0, 1, 2, \dots \tag{3.8}$$

$$\therefore P(X_i = x|Z=z)$$

$$= \frac{P(X_i = x, \sum_{j \ne i} X_j = z - x)}{P(\sum_j x_j = z)}$$

$$= \frac{\binom{\alpha+x-1}{x}\binom{(p-1)\alpha+z-x-1}{z-x}}{\binom{p\alpha+z-1}{p\alpha-1}}$$

$$= \frac{\binom{\alpha+x-1}{x}\binom{b-\alpha+z-x}{z-x}}{\binom{b+z}{b}}, x = 0,1,2,\ldots,z, \tag{3.9}$$

where $b = p\alpha - 1$. Next, note that

$$\mu_2(z) = E\left(\frac{X_i^2}{X_i+\alpha-1}|Z=z\right)$$
$$= E\left(\frac{X_i(X_i-1)}{X_i+\alpha-1}|Z=z\right) + E\left(\frac{X_i}{X_i+\alpha-1}|Z=z\right) \tag{3.10}$$

Now,

$$E\left(\frac{X_i(X_i-1)}{X_i+\alpha-1}|Z=z\right)$$

$$= \frac{1}{\binom{b+z}{b}} \sum_{x=0}^{z} \frac{x(x-1)}{x+\alpha-1}\binom{\alpha+x-1}{x}\binom{b-\alpha+z-x}{z-x}$$

$$= \frac{\alpha}{\binom{b+z}{b}} \sum_{x=0}^{z^*} \frac{(\alpha^*+x-1)!}{(\alpha^*-1)!x!}\frac{(b^*-\alpha^*+z^*-x)!}{(b^*-\alpha^*)!(z^*-x)!}$$

(where $\alpha^* = \alpha+1, b^* = b+1, z^* = z-2$)

$$= \frac{\alpha}{\binom{b+z}{b}} \cdot \binom{b^*+z^*}{b^*} = \frac{\alpha z(z-1)}{(b+z)(b+1)}. \tag{3.11}$$

Also, by a similar argument,

$$E\left(\frac{X_i}{X_i+\alpha-1}|Z=z\right) = \frac{z}{b+z}. \tag{3.12}$$

Using (3.11) and (3.12), one now has from (3.10) that

$$\mu_2(z) = \frac{\alpha z(z-1) + zp\alpha}{p\alpha(z+p\alpha-1)} = \frac{z(z+p-1)}{p(z+p\alpha-1)}. \tag{3.13}$$

Hence, (3.7) now implies that the restricted-risk Bayes rule will be obtained on minimizing $(z+p-1-\frac{z+p\alpha-1}{\lambda+1}-\phi(z))^2$; since this is equivalent to minimizing $(\frac{\lambda(z+p-1)+p(1-\alpha)}{\lambda+1}-\phi(z))^2$, the theorem now follows on using the familiar argument used previously in Theorems 3 and 6.

## 4. RSL's and Bayesian Performances

In this section, we investigate the amount of possible subjective Bayesian gains one has to sacrifice by using the restricted Bayes rule $\delta^*_{\pi_0}$ instead of the unrestricted Bayes rule $\delta_{\pi_0}$. Traditionally, this is done by calculating the Efron-Morris RSL's (see Efron and Morris (1971)), defined as

$$\text{RSL } (\delta^*_{\pi_0}, \pi_0) = \frac{r(\pi_0, \delta^*_{\pi_0}) - r(\pi_0, \delta_{\pi_0})}{r(\pi_{0,X}) - r(\pi_0, \delta_{\pi_0})} \tag{4.1}$$

Since low values of the RSL would indicate marginal sacrifices in subjective Bayesian gains, we derive upper bounds on the RSL's for a general prior $\pi_0(\Sigma\theta_j)$; we also exactly calculate the RSL's in the conjugate case, i.e., when $\pi_0(\Sigma\theta_j) = e^{-\lambda\Sigma\theta_j}, \lambda > 0$. First we need a few Lemmas.

**Lemma 5.** Let $\pi_0(\Sigma\theta_j)$ be such that $\delta_{\pi_{o,i}}(\underset{\sim}{X}) = 0$ if $X_i = 0$; let also $\delta_{\pi_o}(\underset{\sim}{X}) = (1-f(Z))\cdot \underset{\sim}{X}$; then for any estimator $\delta\varepsilon\mathcal{D}$,

$$r(\pi_0, \delta) - r(\pi_0, \delta_{\pi_0}) = E_{m(Z)} \frac{Z\cdot[f(Z) - \frac{\phi(Z)}{z+p-1}]^2}{1-f(Z)}$$

**Proof:** Apply the argument leading to (2.16) in Lemma 2.

**Lemma 6.** Let $K(p) = \int \pi_0(\Sigma\theta_j)d\underset{\sim}{\theta}$. Then the marginal probability mass function of $\underset{\sim}{X}$ is given as

$$m(\underset{\sim}{x}) = \frac{\int_0^\infty e^{-u}u^{z+p-1}\pi_0(u)du}{(z+p-1)!K(p)}$$

**Proof:** By definition,

$$m(\underset{\sim}{x}) = \frac{\int e^{-\Sigma\theta_j}\prod_j \theta_j^{x_j}\pi_0(\Sigma\theta_j)d\underset{\sim}{\theta}}{\prod_j x_j! K(p)} \tag{4.2}$$

Transforming from $(\theta_1, \theta_2, \ldots, \theta_p)$ to $(\mu_1, \mu_2, \ldots, \mu_p)$, where $u_j = \sum_{i=1}^{j}\theta_i$, and using the steps leading to (2.9) in Theorem 2, one has,

$$m(\underset{\sim}{X}) = B(x_1+1, x_2+1)B(x_1+x_2+2, x_3+1)\ldots B(x_1+\ldots+x_{p-1}+p-1, x_p+1)$$

$$\times \frac{\int e^{-u}u^{z+p-1}\pi_0(u)du}{\prod_j x_j! K(p)} \tag{4.3}$$

The result now follows from (4.3).

**Lemma 7.** The marginal probability mass function of $Z = \sum_{i=1}^{p} X_i$ is given as

$$P(z) = P(Z = z) = \frac{\int_0^\infty e^{-u}u^{z+p-1}\pi_0(u)du}{z!(p-1)!K(p)}.$$

**Proof:** Since $m(\underset{\sim}{x})$ depends on only $z$, clearly, $P(z) = m(\underset{\sim}{x})\cdot\{\#(x_1, x_2, \ldots, x_p) : \sum_{i=1}^{p} x_i = z\} = m(\underset{\sim}{x}) \cdot N_p(z)$ (say). For $p = 2, N_p(z) = z + 1$. We claim that in general $N_p(z) = \binom{z+p-1}{p-1}$. Suppose the result is true for $p$; we will prove it for $p+1$. Fix $x_1 = j$, where $0 \le j \le z$. We have to now find out the number of ways in which $x_2, \ldots, x_p$ can add up to $z - j$; but, by the induction hypothesis, this is $\binom{z-j+p-1}{p-1}$. Hence,

$$N_{p+1}(z) = \sum_{j=0}^{z}\binom{z-j+p-1}{p-1} = \binom{z+p}{p}$$

(see Feller (1957), page 61). The Lemma now follows immediately.

**Lemma 8.** $r(\pi_0, X) - r(\pi_0, \delta_{\pi_0}) = \sum_{z=0}^{\infty} \frac{zf^2(z)P(z)}{1-f(z)}$.

**Proof.** Follows from Lemma 5 and Lemma 7 on setting $\phi(z) \equiv 0$.

**Lemma 9.** Assume $\pi_0(\Sigma\theta_j)$ is such that $(z+p-1)f(z)$ is monotone increasing in $z$. Let $z_0+1 = inf\{z \geq 0 : (z+p-1)f(z) > 2(p-1)\}$. Then,

$$r(\pi_0, \delta^*_{\pi_0}) - r(\pi_0, \delta_{\pi_0}) = \sum_{z>z_0} z \cdot \frac{[f(z) - \frac{2(p-1)}{z+p-1}]^2 P(z)}{1-f(z).}$$

**Proof:** Apply the definition of $\delta^*_{\pi_0}$, Lemma 5, and Lemma 7.

**Lemma 10.**

$$r(\pi_0, X) - r(\pi_0, \delta_{\pi_0}) = \sum_{z=1}^{\infty}(z+p-1)P(z-1)[\frac{zP(z)}{(z+p-1)P(z-1)} - 1]^2$$

**Proof:** From (2.24), one has $f(z) = \frac{-I_1(z)}{I_2(z)-I_1(z)}$, where $I_1, I_2$ are defined in (2.25). Since $P(z) = \frac{I_2(z)}{(p-1)!z!K(p)}$, Lemma 8 gives

$$r(\pi_0, X) - r(\pi_0, \delta_{\pi_0}) = \frac{1}{(p-1)!K(p)} \sum_{z=1}^{\infty} \frac{I_1^2(z)}{(z-1)!(I_2(z) - I_1(z))}. \tag{4.4}$$

$$\begin{aligned} \text{Now }, I_1(z) &= \int e^{-u}u^{z+p-1}\pi_0'(u)du \\ &= \int e^{-u}u^{z+p-1}\pi_0(u)du - (z+p-1)\int e^{-u}u^{z+p-2}\pi_0(u)du, \end{aligned} \tag{4.5}$$

if for every $z$, $e^{-u}u^{z+p}\pi_0(u) \to 0$ as $u \to 0, \infty$ (integrate by parts).

$$\begin{aligned} \text{Hence, } I_1(z) &= I_2(z) - (z+p-1)I_2(z-1) \\ &= (p-1)!K(p)(z-1)![zP(z) - (z+p-1)P(z-1)], \\ \text{on using } P(z) &= \frac{I_2(z)}{(p-1)!K(p)z!} \end{aligned} \tag{4.6}$$

$$\text{Also, from (4.6), } I_2(z) - I_1(z) = (p-1)!K(p)(z-1)!(z+p-1)P(z-1). \tag{4.7}$$

Using (4.6) and (4.7) in (4.4), one has the Lemma.

**Lemma 11.** $r(\pi_0, \delta^*_{\pi_0}) - r(\pi_0, \delta_{\pi_0})$

$$= \sum_{z>z_0} (z+p-1)P(z-1)\left[\frac{zP(z)}{(z+p-1)P(z-1)} - 1\right]^2 - 4(p-1)^2$$

$$\sum_{z>z_0} \frac{P(z)-P(z-1)}{z+p-1} - 4(p-1)P(z_0)$$

**Proof:** From Lemma 9,

$$r(\pi_0, \delta^*_{\pi_0}) - r(\pi_0, \delta_{\pi_0}) = \sum_{z>z_0} \frac{zf^2(z)}{1-f(z)}P(z) - 4(p-1)\sum_{z>z_0} \frac{zf(z)}{(z+p-1)(1-f(z))}P(z)$$

$$+ 4(p-1)^2 \sum_{z>z_0} \frac{z}{(z+p-1)^2(1-f(z))}P(z)$$

$$= A + B + C \text{ (say)}. \tag{4.8}$$

Using exactly the same arguments as in Lemma 10,

$$A = \sum_{z>z_0} (z+p-1)P(z-1)[\frac{zP(z)}{(z+p-1)P(z-1)} - 1]^2 \tag{4.9}$$

Also, $B = 4(p-1)\sum_{z>z_0} \frac{z \cdot I_1(z)}{(z+p-1)(p-1)!K(p)z!}$ (using (2.24) and (4.6)).

$$= 4(p-1)\sum_{z>z_0} P(z-1)[\frac{zP(z)}{(z+p-1)P(z-1)} - 1] \text{ (using (4.6) again)}$$

$$= 4(p-1)\sum_{z>z_0} [\frac{zP(z)}{(z+p-1)} - P(z-1)]$$

$$= 4(p-1)\sum_{z>z_0} [P(z) - P(z-1)] - 4(p-1)^2 \sum_{z>z_0} \frac{P(z)}{(z+p-1)}$$

$$= -4(p-1)P(z_0) - 4(p-1)^2 \sum_{z>z_0} \frac{P(z)}{z+p-1} \tag{4.10}$$

Finally, $C = 4(p-1)^2 \sum_{z>z_0} \frac{z}{(z+p-1)^2} \cdot \frac{I_2(z)-I_1(z)}{I_2(z)} \cdot \frac{I_2(z)}{(p-1)!z!K(p)}$

(using (2.24) and (4.6))

$$= 4(p-1)^2 \sum_{z>z_0} \frac{P(z-1)}{z+p-1} \text{ (using (4.7))} \tag{4.11}$$

Combining (4.9), (4.10), and (4.11), one has the Lemma.

We now need to index the marginal *pmf* of $Z$ with $p$. Henceforth, for each $p$, $P(z,p)$ will denote

$$P(z,p) = P(Z=z) = \frac{\int e^{-u} u^{z+p-1} \pi_0(u)\,du}{(p-1)!z!K(p)}.$$

With this notation, we have the following Lemma.

**Lemma 12.** For every $p$ and $z$, $(z-1)P(z-1,p) = p \cdot \frac{K(p+1)}{K(p)} P(z-2,p+1)$.

**Proof:** Follows from definition of $P(z-1,p)$ and noting that $\int e^{-u} u^{z+p-2} \pi_0(u)\,du = p!(z-2)!K(p+1)P(z-2,p+1)$.

This Lemma enables us to write down a convenient upper bound for $r(\pi_0, \delta^*_{\pi_0}) - r(\pi_0, \delta_{\pi_0})$.

**Lemma 13.**

$$r(\pi_0, \delta^*_{\pi_0}) - r(\pi_0, \delta_{\pi_0}) \le p \left[ \sum_{z>z_0} \{P(z-1,p) + \frac{K(p+1)}{K(p)} P(z-2,p+1)\} \right.$$
$$\left. \times \left( \frac{zP(z,p)}{(z+p-1)P(z-1,p)} - 1 \right)^2 \right]$$
$$+ \frac{4(p-1)^2 P(z_0,p)}{z_0+p-1} - 4(p-1)P(z_0,p).$$

**Proof:** In the expression for $r(\pi_0, \delta^*_{\pi_0}) - r(\pi_0, \delta_{\pi_0})$ in Lemma 11, apply the result of Lemma 12 in the first term by writing $(z+p-1)P(z-1,p)$ as $p \cdot P(z-1,p) + (z-1)P(z-1,p)$.

Next, on the second term use the bound

$$4(p-1)^2 \sum_{z>z_0} \frac{P(z-1,p) - P(z,p)}{z+p-1}$$

$$= 4(p-1)^2 \left[ \sum_{z>z_0-1} \frac{P(z,p)}{z+p} - \sum_{z>z_0} \frac{P(z,p)}{z+p-1} \right]$$

$$\leq 4(p-1)^2 \left[ \sum_{z>z_0-1} \frac{P(z,p)}{z+p-1} - \sum_{z>z_0} \frac{P(z,p)}{z+p-1} \right]$$

$$= 4(p-1)^2 \frac{P(z_0,p)}{z_0+p-1}.$$

The Lemma now follows.

We now specialize to the special conjugate case when $\pi_0(\underset{\sim}{\theta}) = e^{-\lambda \Sigma \theta_j}$. Note that in this case, $K(p) = \int \pi_0(\underset{\sim}{\theta}) d\underset{\sim}{\theta} = \frac{1}{\lambda^p}$. First we give a formula for the marginal pmf of $Z$.

**Lemma 14.** Let $\pi_0(\underset{\sim}{\theta}) = e^{-\lambda \Sigma \theta_j}, \lambda > 0$.

$$\text{Then } P(z,p) = \binom{z+p-1}{p-1} \cdot \left(\frac{\lambda}{\lambda+1}\right)^p \left(\frac{1}{\lambda+1}\right)^z, z = 0,1,2,\ldots.$$

**Proof:** Apply Lemma 7.

**Lemma 15.** If $\pi_0(\underset{\sim}{\theta}) = e^{-\lambda \Sigma \theta_j}$, then

$$\frac{zP(z,p)}{(z+p-1)P(z-1,p)} - 1 \equiv -\frac{\lambda}{\lambda+1}.$$

**Proof:** Apply Lemma 14.

**Lemma 16.** Let $\pi_0(\underset{\sim}{\theta}) = e^{-\lambda \Sigma \theta_j}$. Then $r(\pi_0, X) - r(\pi_0, \delta_{\pi_0}) = \frac{p\lambda}{\lambda+1}$.

**Proof:** From Lemma 10 and Lemma 15,

$$r(\pi_0, X) - r(\pi_0, \delta_{\pi_0}) = (\frac{\lambda}{\lambda+1})^2 \sum_{z=1}^{\infty} \{pP(z-1,p) + (z-1)P(z-1,p)\}$$

$$= p\ (\frac{\lambda}{\lambda+1})^2 [\sum_{z=0}^{\infty} P(z,p) + \frac{K(p+1)}{K(p)} \sum_{z=0}^{\infty} P(z,p+1)]$$

(using Lemma 12)

$$= \frac{p\lambda}{\lambda+1} (\text{since } \frac{K(p+1)}{K(p)} = \frac{1}{\lambda} \text{ and } \sum_{z=0}^{\infty} P(z,p) = 1 \text{ for any } p).$$

**Theorem 8.** Let $\pi_0(\underset{\sim}{\theta}) = e^{-\lambda\Sigma\theta_j}, \lambda > 0$. Then

$$\begin{aligned}\text{RSL}(\delta^*_{\pi_0}, \pi_0) \leq (p\varepsilon)^{-1}\{p\big[P[NB(p,\varepsilon) \geq z_0] + \frac{1}{\lambda}P[NB(p+1,\varepsilon) \geq z_0 - 1]\big]\varepsilon^2 \\ + \frac{4(p-1)^2}{z_0+p-1}P[NB(p,\varepsilon) = z_0] - 4(p-1)P[NB(p,\varepsilon) = z_0]\},\end{aligned}$$

where $\varepsilon = \frac{\lambda}{\lambda+1}$, and $NB(p,\varepsilon)$ denotes a Negative Binomial random variable $Z$ with pmf $P(Z = z) = \binom{z+p-1}{p-1} \varepsilon^p (1-\varepsilon)^z, z = 0,1,2,\ldots.$

<u>Proof:</u> Apply Lemma 13 and Lemma 16.

The following Table gives values of the RSL in the case $\pi_0(\underset{\sim}{\theta}) = e^{-\lambda\Sigma\theta_j}$ for various values of $\lambda$. In computing these RSL's the exact formula of Lemma 11 for $r(\pi_0, \delta^*_{\pi_0}) - r(\pi_0, \delta_{\pi_0})$ rather than the bound of Lemma 13 was used. Theorem 7 above gives an upper bound on the RSL's (and not the exact RSL's as such) because Lemma 13 is used in the proof of Theorem 7. We have found that the convenient upper bound of Theorem 7 agrees very well with the exact values of the RSL's for most of the cases tabulated below.

Table 1: Table for $\text{RSL}(\delta^*_{\pi_0})$ when $\underset{\sim}{\theta} \sim e^{-\lambda \Sigma \theta_j}$

| $\lambda$ \ $p$ | 2 | 3 | 4 | 5 | 6 | 8 | 10 |
|---|---|---|---|---|---|---|---|
| 0 | .1353 | .0430 | .0170 | .0090 | .0040 | .0011 | .0004 |
| 0.25 | .119681 | .035904 | .013982 | .006223 | .003007 | .000817 | .000250 |
| 0.50 | .103314 | .028328 | .010256 | .004276 | .001944 | .000470 | .000129 |
| 1.00 | .082297 | .019120 | .005992 | .002187 | .000876 | .000167 | .000036 |
| 2.00 | .060441 | .011022 | .002766 | .000817 | .000267 | .000034 | .000005 |
| 3.00 | .048814 | .007541 | .001594 | .000400 | .000112 | .000010 | .000001 |
| 4.00 | .042273 | .005395 | .001020 | .000220 | .000055 | .000004 | .000000 |
| 5.00 | .037563 | .004114 | .000703 | .000137 | .000029 | .000002 | .000000 |

The values are indeed encouraging. They show that for $p \geq 3$, one loses at most 4.3% of the possible subjective gain in return for Γ-minimaxity of $\delta^*_{\pi_0}$ under a wide class of priors; moreover, $\delta^*_{\pi_0}$ is also actually minimax. The subjective Bayesian thus has no real reason to worry because he does practically as good using $\delta^*_{\pi_0}$ as he had done using $\delta_{\pi_0}$, and yet he is assured of a good amount of protection against misspecification of the subjective prior. Especially encouraging is the fact that from 3 dimension itself, the RSL's start to get nominal. In contrast, in the normal problems one needs 5 or 6 dimensions before the RSL's get very close to zero (see Berger (1982) and DasGupta and Rubin (1986)). Table 1 shows that the RSL's tend to increase as $\lambda$ gets closer to zero. At first sight, it seems possible that the RSL's may be quite big as $\lambda$ goes to zero. The following Theorem shows that this fear is unfounded.

**Theorem 9.** Let $\pi_0(\underset{\sim}{\theta}) = e^{-\lambda \Sigma \theta_j}, \lambda > 0$. Then $\lim_{\lambda \to 0} \text{RSL}(\delta^*_{\pi_0}, \pi_0)$ exists for every $p$ and is given as

$$\lim_{\lambda \to 0} \text{RSL}(\delta^*_{\pi_0}, \pi_0) = P[X \leq p-1] - P[X = p],$$

where $X$ has a Poisson distribution with mean $2(p-1)$.

**Remark:** The values corresponding to $\lambda = 0$ in Table 1 are actually these limiting RSL's.

**Proof of Theorem 9:** We will use the upper bound of Theorem 8 on RSL $(\delta^*_{\pi_0}, \pi_0)$, although one can show that the same limit is obtained by using the exact expression obtained from Lemma 11. First note that the Negative Binomial and Binomial probabilities are related by the identity $P[NB(p,\varepsilon) \geq K] = P[Bin(K+p-1,\varepsilon) \leq p-1]$

$$\Rightarrow \sum_{z=K}^{\infty} \binom{z+p-1}{p-1} \varepsilon^p (1-\varepsilon)^z = \sum_{r=0}^{p-1} \binom{K+p-1}{r} \varepsilon^r (1-\varepsilon)^{K+p-1-r}. \qquad (4.12)$$

The first term in the upper bound of Theorem 8 on RSL equals $R_1 = \varepsilon\big[P[NB(p,\varepsilon) \geq z_0] + \frac{1-\varepsilon}{\varepsilon} \cdot P[NB(p+1,\varepsilon) \geq z_0 - 1]\big]$. Note that in the conjugate case, $z_0 = \frac{\lambda+2}{\lambda}(p-1)$ (see (2.17)). Now using (4.12),

$$R_1 = \varepsilon \cdot \sum_{r=0}^{p-1} \binom{z_0+p-1}{r} \varepsilon^r (1-\varepsilon)^{z_0+p-1-r} + (1-\varepsilon) \sum_{r=0}^{p} \binom{z_0+p-1}{r} \varepsilon^r (1-\varepsilon)^{z_0+p-1-r} \qquad (4.13)$$

Using Stirling's approximation on the factorials ($m! \approx \sqrt{2\pi} e^{-m} m^{m+\frac{1}{2}}$), it follows after some algebra that

$$\sum_{r=0}^{p} \binom{z_0+p-1}{r} \varepsilon^r (1-\varepsilon)^{z_0+p-1-r} = 0\left[e^{-2(p-1)} \cdot \sum_{r=0}^{p} \frac{(2(p-1))^r}{r!}\right],$$

$$\text{and} \sum_{r=0}^{p-1} \binom{z_0+p-1}{r} \varepsilon^r (1-\varepsilon)^{z_0+p-1-r} = 0\left[e^{-2(p-1)} \cdot \sum_{r=0}^{p-1} \frac{(2(p-1))^r}{r!}\right] \qquad (4.14)$$

Here, $a_n = 0(b_n)$ means $\frac{a_n}{b_n} \to 1$ as $n \to \infty$. Using (4.14), one has from (4.13) and the fact that $\varepsilon \to 0$ as $\lambda \to 0$,

$$R_1 = 0[P(X \leq p)], \qquad (4.15)$$

where $X$ is as in the statement of the Theorem.

The second term in the upper bound of Theorem 8 equals

$$R_2 = \frac{4(p-1)^2}{p\varepsilon(z_0+p-1)} \cdot \binom{z_0+p-1}{p-1} \varepsilon^p (1-\varepsilon)^{z_0},$$

which on applying Stirling's approximation, is found to be

$$0\left[\frac{4(p-1)^2}{p} \cdot \frac{(2(p-1))^{p-2} \cdot e^{-2(p-1}}{(p-1)!} \cdot \varepsilon\right], \qquad (4.16)$$

which converges to zero as $\lambda \to 0$.

Finally, the third term in the upper bound equals

$$R_3 = \frac{-4(p-1)}{p\varepsilon} \binom{z_0+p-1}{p-1} \varepsilon^p (1-\varepsilon)^{z_0}$$

$$= 0\Big[\frac{-2\cdot(2(p-1))^p}{p!} e^{-2(p-1)}\Big]$$

$$= -2P[X=p]. \tag{4.17}$$

Combining (4.15), (4.16), and (4.17), one has the Theorem.

**Acknowledgement.** We are deeply indebted to James Berger for suggesting a significant improvement on an earlier version of Theorem 1. Thanks are due to Dr. Tai-Fang Lu for her help in the computing work. The first author is also grateful to Burgess Davis, Steven Lalley, and T. N. Sriram for helpful discussions.

## References

Berger, James (1975). Minimax estimation of location vectors for a wide class of densities. *Ann. Statist.* **3**, 1313–1328.

Berger, James (1982). Estimation in continuous exponential families: Bayesian estimation subject to risk restrictions and inadmissibility results. *Statistical Decision Theory and Related Topics III* **1**, 109–141, Academic Press, New York.

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

Berger, James (1986). *Statistical Decision Theory and Bayesian Analysis*, Second Edition, Springer-Verlag, New York.

Berger, J., and Berliner, L. M. (1986). Robust Bayes and empirical Bayes analysis with $\varepsilon$-contaminated priors. *Ann. Statist.* **14**, 461–486.

Berger, J., and Sivaganesan, (1986). Ranges of Posterior measures for priors with unimodal contaminations. *Technical Report* **#86-41**, Department of Statistics, Purdue University.

Berger, R. (1979). Gamma minimax robustness of Bayes rules. *Comm. in Statist.* **8**, 543–560.

Bickel, P. J. (1980). Minimax estimation of the mean of a normal distribution subject to doing well at a point. *Technical Report*, Department of Statistics, University of California, Berkeley.

Clevenson, M, and Zidek, J. (1975). Simultaneous estimation of the mean of independent Poisson laws. *Jour. Amer. Statist. Assoc.* **70**, 698–705.

DasGupta, A. and Berger J. (1986). Estimation of multiple Gamma scale-parameters: Bayes estimation subject to uniform domination. *Comm. in Statist.*, special issue on Stein-type multivariate estimation, **15**, 2065–2086.

DasGupta, A. and Rubin, H. (1986). Bayesian estimation subject to minimaxity of the mean of a multivariate normal distribution in the case of a common unknown variance: a case for Bayesian robustness. *Technical Report* **#86-10**, Department of Statistics, Purdue University. To appear in the Proceedings of the Fourth Purdue Symposium on Statistical Decision Theory and Related Topics.

Efron, B. and Morris, C. (1971). Limiting the risk of Bayes and Empirical Bayes estimators - Part I: The Bayes case. *Jour. Amer. Statist. Assoc.* **66**, 807–815.

Feller, William (1957). An Introduction to Probability Theory and Its Applications, Volume 1, Second edition, John Wiley and Sons, Inc., New York.

Gupta, S. S., and Hsiao, P. (1981). On Γ-minimax, minimax, and Bayes procedures for selecting populations close to a control. *Sankhya (Ser. B)* **43**, 291–318.

Ghosh, M., and Parsian, A. (1981). Bayes minimax estimation of multiple Poisson parameters. *Jour. Mult. Anal.* **11**, No. 2, 280–288.

Hodges, J. L., and Lehmann, E. L. (1952). The use of previous experience in reaching statistical decisions. *Ann. Math. Statist.* **23**, 392–407.

Huber, P. J. (1972). Robust statistics: a review. *Ann. Math. Statist.* **43**, 1041–1067.

Huber, P. J. (1981). Robust Statistics, John Wiley and Sons, Inc., New York.

Hwang, J. T. (1982). Improving upon standard estimators in discrete exponential families with applications to Poisson and negative Binomial cases. *Ann. Statist.* **10**, 857–867.

Jackson, D. A., Donovan, T. M., Zimmer, W. J., and Deely, J. J. (1970). Γ-minimax estimators in the exponential family. *Biometrika* **57**, 439-443.

Kudo, H. (1967). On partial prior information and the property of parametric sufficiency. *Proc. Fifth Berkeley Symp. Statist. Probab.* **1**, University of California Press, Berkeley.

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

Robbins, H. (1951). Asymptotically subminimax solutions of compound statistical decision problems. *Proc. Second Berkeley Symp. Math. Statist. Probab.* **1**, University of California Press, Berkeley.

Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. *Ann. Math. Statist.* **35**, 1–20.

Tsui, K. W., and Press, S. J. (1982). Simultaneous estimation of several Poisson parameters under k-normalized squared-error loss. *Ann. Statist.* **10**, 93–100.