Γ-Minimax Estimation of Vector Parameters In Restricted Parameter Spaces

by

Anirban DasGupta* and Mohan Delampady**
Purdue University
University of British Columbia
Vancouver

Technical Report #90-42

Department of Statistics
Purdue University

August 1991

*Research supported by NSF Grant DMS: 89-230-71.
**Research supported by NSERC (Canada) Grant A9250.

Γ-Minimax Estimation of Vector Parameters
In Restricted Parameter Spaces

Abstract

For arbitrary distributions P_θ in a k-dimensional space indexed by a p-dimensional parameter θ, the Γ-minimax estimator is derived when it is known that θ lies in a sufficiently small bounded convex set in $\mathbb{R}^p$. The family Γ consists of priors of the form $(1-\varepsilon)G_0 + \varepsilon Q$, where $0 \leq \varepsilon \leq \varepsilon_0$ for some specified ε_0, G_0 is fixed and Q belongs to a suitable family $\mathcal{L}$. Three choices of $\mathcal{L}$ are discussed: the family of all distributions, the family of spherically symmetric distributions and the family of spherically symmetric unimodal distributions. The theory is applied to four specific examples, including estimation of a single normal mean and estimation of a multivariate normal mean. Stein's identity is used here in a novel way to check subharmonicity of risk functions.

Key words: Risk, Bayes risk, minimax, subharmonic, spherically symmetric, unimodal.

1. Introduction

Let $\underset{\sim}{X}$ be a random vector with probability distribution $P_{\underset{\sim}{\theta}}$ where $\underset{\sim}{\theta}$ is a p dimensional unknown parameter. In this article, we consider the problem of estimating $\underset{\sim}{\theta}$ under squared error loss when $\underset{\sim}{\theta}$ is known to lie in a bounded convex set C in $\mathbb{R}^p$. The information that $\underset{\sim}{\theta}$ lies in a bounded set is often available in practice. In fact, it may even be argued that in virtually all real world problems, the parameters can be assumed to belong to appropriate bounded sets and the standard theoretical assumption that they can take any value in $\mathbb{R}^p$ (or such unbounded sets) is really an assumption of convenience. The literature on estimation of a bounded parameter includes Ghosh (1964), Casella and Strawderman (1981), Bickel (1981), DasGupta (1985), Berry (1989), Eichenauer (1989a,b), Mintz and Zeytinoglu (1984), Zeytinoglu and Mintz (1988), Donoho, Liu and MacGibbon (1990) etc. Γ-minimax estimation of the parameter $\underset{\sim}{\theta}$ will be considered. In the Γ-minimax approach, we assume that the prior $G(\underset{\sim}{\theta})$ for $\underset{\sim}{\theta}$ belongs to a suitable family Γ. We then minimize, over decision rules δ, the quantity $r_\Gamma(\delta) = \sup_{G\in\Gamma} r(G,\delta)$ where $r(G,\delta)$ denotes the Bayes risk of δ with respect to the prior $G(\underset{\sim}{\theta})$. If Γ is the family of all priors, then Γ-minimaxity coincides with usual minimaxity. However, if Γ contains only reasonable priors which are considered plausible in the context of a given problem, then a Γ-minimax estimator indeed provides a reasonable middle ground between a classical minimax estimator which may be too conservative and a subjective Bayes estimator which may lack robustness against misspecification of prior opinion. For a lucid discussion of the Γ-minimax approach and its relation to restricted risk Bayesian estimation, see Berger (1985); also see DasGupta and Bose (1988).

We consider the family of priors

$$\Gamma = \{G\colon\ G = (1-\varepsilon)G_0 + \varepsilon Q\}, \tag{1.1}$$

where $0 \le \varepsilon \le \varepsilon_0$ for some prespecified ε_0, G_0 is fixed and the contamination Q belongs to an appropriate family $\mathcal{L}$. Three different choices of $\mathcal{L}$ are considered:

$\mathcal{L}_1 =$ family of all distributions

$\mathcal{L}_2 =$ family of spherically symmetric priors

$\mathcal{L}_3 =$ family of spherically symmetric unimodal priors.

Notice $\mathcal{L}_3 \subseteq \mathcal{L}_2 \subseteq \mathcal{L}_1$; it can thus be argued that $\mathcal{L}_2$ or $\mathcal{L}_3$ may be used when prior information is more precise and $\mathcal{L}_1$ may be used when it is less precise. For $\mathcal{L} = \mathcal{L}_i$, $1 \le i \le 3$, the initial prior G_0 is also assumed to belong to $\mathcal{L}_i$. This is clearly reasonable. A very attractive property of the family (1.1) is that, in general, by allowing a mass of ε to be redistributed in any of a variety of ways, flexibility in the tail of the prior G is achieved. But notice that if $\underset{\sim}{\theta}$ is known to lie in a small bounded set, the problem of specifying the tail of the prior is no longer of concern. Thus, in our context, the contamination simply plays the role of allowing a possibly small mass ε to be distributed according to some other convenient distribution. For results on Γ-minimax estimation of a one dimensional normal mean when $\mathcal{L} = \mathcal{L}_1$ or when $\varepsilon = 1$ and $\mathcal{L} = \mathcal{L}_3$ (i.e., when Γ is simply the class of all symmetric unimodal priors), see Eichenauer (1989a,b). The principal advantages of our results are that ε is permitted to vary, that we do not need practically any assumption on the likelihood function and that our results are for parameters of arbitrary dimension. The flexibility in the choice of ε is attractive and indeed has long been considered rather desirable.

In section 2, we first give a simple result on subharmonicity of risk functions if the parameter space is a sufficiently small bounded set in $\mathbb{R}^p$. We give a few technical lemmas about subharmonic functions, followed by an explicit derivation of the Γ-minimax rules.

In section 3, the theory is applied to four specific examples: estimation of a single normal mean, estimation of a binomial success probability and estimation of a multivariate normal mean. In each case, in order to apply the theory of section 2, one needs to check that the risk function of the Γ-minimax rule is subharmonic (convex in one parameter cases); in the normal problems, a novel application of Stein's identity (see Stein (1973, 1981)) is made for this purpose. Section 4 contains some concluding remarks.

2. Derivation of the Γ-minimax rule

A principal result needed for derivation of the Γ-minimax rule is that the risk function of the suggested rule be subharmonic. Towards this end, we assume that $\underset{\sim}{\theta}$ lies in a bounded convex set $mS + \underset{\sim}{\theta}_0$ where S is a fixed bounded convex set containing 0 in its interior, and $\underset{\sim}{\theta}_0$ is a fixed point in $\mathbb{R}^p$; here

$$mS + \underset{\sim}{\theta}_0 = \{m\underset{\sim}{\theta} + \underset{\sim}{\theta}_0 \colon \underset{\sim}{\theta} \in S\}. \tag{2.1}$$

Notice that $mS + \underset{\sim}{\theta}_0 \downarrow \{\theta_0\}$ as $m \downarrow 0$. Thus for small $m > 0$, $mS + \underset{\sim}{\theta}_0$ is a small bounded set containing $\underset{\sim}{\theta}_0$. Let $G(\underset{\sim}{\theta})$ be a fixed prior distribution on $mS + \underset{\sim}{\theta}_0$ and let $\delta_m(\underset{\sim}{X})$ denote the Bayes rule with respect to G for estimating $\underset{\sim}{\theta}$ under sum of squared error loss. Let $R(\underset{\sim}{\theta}, \delta_m)$ denote the risk function of $\delta_m(X)$ and define

$$g(\underset{\sim}{\theta}, m) = \Delta R(\underset{\sim}{\theta}, \delta_m)$$

where $\Delta(\cdot)$ denotes Laplacian; we assume $g(\underset{\sim}{\theta}, m)$ exists for all $\underset{\sim}{\theta}$ in $mS + \underset{\sim}{\theta}_0$. The following theorem is implicit in DasGupta (1985); indeed a number of useful generalizations are also given in that article.

Theorem 2.1. Assume $g(\underset{\sim}{\theta}, m)$ is jointly continuous in $\underset{\sim}{\theta}$ and m at $(\underset{\sim}{0}, 0)$. Then there exists $m_0 > 0$ such that $R(\underset{\sim}{\theta}, \delta_m)$ is subharmonic for $\underset{\sim}{\theta}$ in $mS + \underset{\sim}{\theta}_0$ for all $m \leq m_0$.

Discussion: Note that the value of m_0 depends on the choice of the prior. It is possible to write down a set of sufficient conditions under which $\Delta R(\underset{\sim}{\theta}, \delta_m)$ will exist and be jointly continuous in $\underset{\sim}{\theta}$ and m at $(\underset{\sim}{0}, 0)$. However we do not attempt to do it here since that is not the main focus of this article. What is important is to realize that as long as the parameter space is sufficiently small, the risk function of δ_m will usually be subharmonic; in any specific problem, one can directly verify how small the parameter space needs to be for subharmonicity to hold. Section 3 contains a number of such illustrations.

We now give two technical lemmas later needed for the derivation of the Γ-minimax rule.

Lemma 2.2. Let $f(\underset{\sim}{\theta})$ be a continuous subharmonic function on a bounded sphere $S(\underset{\sim}{\theta}, R)$ in $\mathbb{R}^p$, where for any $r > 0$, $S(\underset{\sim}{0}, r)$ denotes the sphere of radius r with center at $\underset{\sim}{0}$. Let $r_1 < r_2 \leq R$. If $m(r)$ denotes $\int f(\underset{\sim}{\theta}) dSr(\underset{\sim}{\theta})$ where $Sr(\cdot)$ is the uniform measure on the surface of $S(\underset{\sim}{0}, r)$, then $m(r_1) \leq m(r_2)$.

Proof: See pp. 363, Rudin (1974).

Lemma 2.3. Let $f(\underset{\sim}{\theta})$ be as in Lemma 2.2. If $m^*(r)$ denotes $\int f(\underset{\sim}{\theta}) d\mu_r(\underset{\sim}{\theta})$ where $\mu_r(\cdot)$ is the uniform measure on the whole sphere $S(\underset{\sim}{0}, r)$, then $m^*(r_1) \leq m^*(r_2)$.

Proof: First assume f is also a spherical function, i.e., f depends on $\underset{\sim}{\theta}$ only through $\underset{\sim}{\theta}'\underset{\sim}{\theta}$. In this case,

$$m^*(r) = \text{ constant } \cdot \frac{\int_0^r f(s)s^{p-1}}{r^p}. \tag{2.2}$$

Since Lemma 2.1 implies $f(s)$ is nondecreasing in s, calculus now shows that $m^*(r)$ is nondecreasing in r. If f is not a spherical function, define

$$h(\underset{\sim}{\theta}'\underset{\sim}{\theta}) = \int f(P\underset{\sim}{\theta})d\Gamma(P), \tag{2.3}$$

where $\Gamma(\cdot)$ is the invariant probability measure on the group of $p \times p$ orthogonal matrices. It follows from an easy application of Fubini's theorem that

$$\int h(\underset{\sim}{\theta}'\underset{\sim}{\theta})d\mu_r(\underset{\sim}{\theta}) = \int f(\underset{\sim}{\theta})d\mu_r(\underset{\sim}{\theta}). \tag{2.4}$$

Since h is subharmonic if f is, the result now follows from the earlier case.

We now present the main results. We make no assumptions on the distribution of the observation X and no reference to it will be made.

Theorem 2.4. Assume $\underset{\sim}{\theta}$ belongs to $\Theta = mS + \underset{\sim}{\theta}_0$ for some $m > 0$, where S is a fixed bounded convex set containing $\underset{\sim}{0}$ and $\underset{\sim}{\theta}_0$ is a fixed vector in $\mathbb{R}^p$. Suppose $\underset{\sim}{\theta}$ has a prior G of the form (1.1) where $\mathcal{L} = \mathcal{L}_1$. Let Q^* denote the least favorable prior with respect to the risk $\tilde{R}(\underset{\sim}{\theta}, \delta) = (1 - \varepsilon_0)r(G_0, \delta) + \varepsilon_0 R(\underset{\sim}{\theta}, \delta)$ when $\underset{\sim}{\theta}$ is restricted to lie on the boundary of $mS + \underset{\sim}{\theta}_0$ and let $\delta^*(X)$ denote the Bayes rule with respect to $Q^*(\underset{\sim}{\theta})$ under this new risk. Then $\delta^*(\underset{\sim}{X})$ minimizes $\sup_{G \in \Gamma} r(G, \delta)$, provided $R(\underset{\sim}{\theta}, \delta^*)$ is subharmonic.

Discussion: First notice that a least favorable prior Q^* exists if $\underset{\sim}{\theta}$ is constrained to lie on the boundary of the parameter space. Also, note that δ^* is precisely the Bayes rule under the original risk $R(\theta, \delta)$ with respect to the prior $(1 - \varepsilon_0)G_0 + \varepsilon_0 Q^*$. The hypothesis that $R(\underset{\sim}{\theta}, \delta^*)$ is subharmonic will usually be satisfied if m is sufficiently small. Evidence for this was given in Theorem 2.1. Again, we emphasize that in any given problem, one can directly verify the subharmonicity. Finally, we would like to caution the reader that even though Theorem 2.4 gives a general method for finding the Γ-minimax rule when $\mathcal{L} = \mathcal{L}_1$ and m is small, identifying Q^* in a high dimensional problem need not be easy. Finding

Q^* in one dimensional situations is a relatively simple task. We will say more about this later. Finally also note the rather interesting fact that the Γ-minimax estimator $\delta^*(\underset{\sim}{X})$ is really just the Γ-minimax estimator for the case $\varepsilon = \varepsilon_0$.

Proof of Theorem 2.4: We will prove that for any rule δ,

$$\sup_{G\in\Gamma} r(G,\delta) \geq \sup_{G\in\Gamma} r(G,\delta^*).$$

However,

$$\begin{aligned}
\sup_{G\in\Gamma} r(G,\delta) &= \sup_{Q\in\mathcal{L}_1} \sup_{0\leq\varepsilon\leq\varepsilon_0} [(1-\varepsilon)r(G_0,\delta)+\varepsilon r(Q,\delta)] \\
&= \sup_{Q\in\mathcal{L}_1} [\max\{r(G_0,\delta),(1-\varepsilon_0)r(G_0,\delta)+\varepsilon_0 r(Q,\delta)\}] \\
&= \max\left\{r(G_0,\delta),\ \sup_{Q\in\mathcal{L}_1}[(1-\varepsilon_0)r(G_0,\delta)+\varepsilon_0 r(Q,\delta)]\right\} \\
&= \sup_{Q\in\mathcal{L}_1}[(1-\varepsilon_0)r(G_0,\delta)+\varepsilon_0 r(Q,\delta)] \\
&= \sup_{\underset{\sim}{\theta}\in\Theta} \tilde{R}(\underset{\sim}{\theta},\delta) \\
&\geq \sup_{\underset{\sim}{\theta}\in\partial\Theta} \tilde{R}(\underset{\sim}{\theta},\delta) \\
&\geq \sup_{\underset{\sim}{\theta}\in\partial\Theta} \tilde{R}(\underset{\sim}{\theta},\delta^*) \\
&= \sup_{\underset{\sim}{\theta}\in\Theta} \tilde{R}(\underset{\sim}{\theta},\delta^*) \\
&= \sup_{G\in\Gamma} r(G,\delta^*);
\end{aligned}$$

here the first five equalities and the first subsequent inequality are trivial; the next inequality is due to the fact that δ^* is (by construction) minimax with respect to the risk $\tilde{R}$ when $\underset{\sim}{\theta}$ lies on the boundary $\partial\Theta$; the following inequality is a consequence of the subharmonicity of $R(\theta,\delta^*)$ (or equivalently, the subharmonicity of $\tilde{R}(\theta,\delta^*)$) and the Maximum modulus theorem, and the final equality follows from repeating the first five equalities in the proof using δ^* for δ.

Remark: In one dimensional problems, if θ lies in the interval $a \leq \theta \leq b$, the required Q^* is simply a two point distribution Q_{λ_0} with mass, say λ_0 and $1-\lambda_0$, at a and b respectively. Suppose we denote the Bayes rule with respect to $(1-\varepsilon_0)G_0+\varepsilon_0 Q_\lambda$ under the risk R by

δ_λ. One can find the required λ_0 by computing $\tilde{R}(\theta, \delta_\lambda)$ and solving for λ that satisfies $\max\{\tilde{R}(a, \delta_\lambda), \tilde{R}(b, \delta_\lambda)\} = \lambda \tilde{R}(a, \delta_\lambda) + (1-\lambda)\tilde{R}(b, \delta_\lambda)$. Of course, one must also check that $R(\theta, \delta_{\lambda_0})$ is convex on $[a, b]$ or that its maximum is attained at the support of Q_{λ_0}.

Theorem 2.5. Assume $\underset{\sim}{\theta}$ belongs to the sphere $S(\underset{\sim}{0}, m)$. Let $\underset{\sim}{\theta}$ have a prior G of the form (1.1) where $\mathcal{L} = \mathcal{L}_2$. Let Q_S denote the uniform distribution on the surface of $S(\underset{\sim}{0}, m)$ and let $\delta_S(\underset{\sim}{X})$ denote the Bayes rule (under R) with respect to the prior $G_S(\underset{\sim}{\theta}) = (1 - \varepsilon_0)G_0(\underset{\sim}{\theta}) + \varepsilon_0 Q_S(\underset{\sim}{\theta})$. Then $\delta_S(\underset{\sim}{X})$ minimizes $\sup\limits_{G \in \Gamma} r(G, \delta)$ provided $R(\underset{\sim}{\theta}, \delta_S)$ is subharmonic.

Proof: By a well known result, it is sufficient to prove that $\sup_{G \in \Gamma} r(G, \delta_S) = r(G_S, \delta_S)$. Towards this end, first notice that $\underset{\sim}{\theta}$ has a spherically symmetric prior Q if and only if it has the representation

$$\underset{\sim}{\theta} = r\underset{\sim}{u},$$

where $\underset{\sim}{u}$ is distributed uniformly on the boundary of $S(\underset{\sim}{0}, 1)$, and $0 \le r \le m$ is distributed independently of $\underset{\sim}{u}$ (see Schoenberg (1938)). Denoting the distribution of r by F, we then have,

$$\begin{aligned}
&\sup_Q r(Q, \delta_S) \\
&= \sup_Q E_Q(R(\theta, \delta_S)) \\
&= \sup_F \int\int R(\theta, \delta_S) dS_r(\theta) dF(r) \\
&= \sup_{0 \le r \le m} \int R(\theta, \delta_S) dS_r(\theta) \\
&= r(Q_S, \delta_S), \text{ by Lemma 2.2.}
\end{aligned}$$

Hence,

$$\begin{aligned}
\sup_{G \in \Gamma} r(G, \delta_S) &= \sup_{0 \le \varepsilon \le \varepsilon_0} \sup_Q [(1-\varepsilon) r(G_0, \delta_S) + \varepsilon r(Q, \delta_S)] \\
&= \sup_{0 \le \varepsilon \le \varepsilon_0} [(1-\varepsilon) r(G_0, \delta_S) + \varepsilon r(Q_S, \delta_S)] \\
&= \max\{r(G_0, \delta_S), (1-\varepsilon_0) r(G_0, \delta_S) + \varepsilon_0 r(Q_S, \delta_S)\} \\
&= (1-\varepsilon_0) r(G_0, \delta_S) + \varepsilon_0 r(Q_S, \delta_S)
\end{aligned}$$

(since G_0 is assumed to be spherically symmetric)

$$= r(G_S, \delta_S),$$

as was required. This proves the Theorem.

Theorem 2.6. Assume $\underset{\sim}{\theta}$ belongs to the sphere $S(\underset{\sim}{0}, m)$. Let $\underset{\sim}{\theta}$ have a prior G of the form (1.1) where $\mathcal{L} = \mathcal{L}_3$. Let Q_u denote the uniform distribution on the entire sphere $S(\underset{\sim}{0}, m)$ and let $\delta_u(\underset{\sim}{X})$ denote the Bayes rule (under R) with respect to the prior $G_u(\underset{\sim}{\theta}) = (1 - \varepsilon_0)G_0(\underset{\sim}{\theta}) + \varepsilon_0 Q_u(\underset{\sim}{\theta})$. Then $\delta_u(\underset{\sim}{X})$ minimizes $\sup_{G \in \Gamma} r(G, \delta)$ provided $R(\underset{\sim}{\theta}, \delta_u)$ is subharmonic.

Proof: The proof follows exactly in the lines of the proof of Theorem 2.5 by writing $\underset{\sim}{\theta} = rv$, where $\underset{\sim}{v}$ is distributed uniformly on the whole sphere $S(\underset{\sim}{0}, 1)$ and $0 \le r \le m$ and $\underset{\sim}{v}$ are independent.

Remark: Notice that for $\mathcal{L} = \mathcal{L}_2$ and $\mathcal{L}_3$ as well, the Γ-minimax estimator is simply the Γ- minimax estimator corresponding to $\varepsilon = \varepsilon_0$.

3. Examples and Applications

In this section, we will give four examples illustrating the theory of section 2. Interest lies in finding how small the parameter space needs to be in order that the risk subharmonicity holds. Since the results of Section 2 imply that consideration of only $\varepsilon = \varepsilon_0$ suffices, we will assume ε is fixed in our examples.

Example 1. Let $X \sim$ Bin (n, θ) and suppose it is known $|\theta - \frac{1}{2}| \le m$. Assume θ has a prior $G = (1 - \varepsilon_0)G_0 + \varepsilon_0 Q$, where Q is symmetric and unimodal about $\frac{1}{2}$ and G_0 is uniform on $[\frac{1}{2} - m, \frac{1}{2} + m]$. This choice of G_0 should be taken just as an artifact; but also observe that for small m, this choice of G_0 may indeed be a reasonable initial prior to use. Theorem 2.6 implies that the Bayes rule with respect to G_0 itself is Γ-minimax provided its risk function is convex for $|\theta - \frac{1}{2}| \le m$. Observe that the exact value of ε_0 is rendered unimportant by choosing G_0 to be uniform on $\frac{1}{2} \pm m$. We will denote the Bayes rule with respect to G_0 by $\delta_{G_0}(X)$, and its value for $X = j$ by d_j. Direct computation gives

$$d_j = \frac{\sum_{i=0}^{n-j} \binom{n-j}{i}(-1)^i \frac{1}{i+j+2}[(.5+m)^{i+j+2} - (.5-m)^{i+j+2}]}{\sum_{i=0}^{n-j} \binom{n-j}{i}(-1)^i \frac{1}{i+j+1}[(.5+m)^{i+j+1} - (.5-m)^{i+j+1}]}. \tag{3.1}$$

Computation of $R(\theta, \delta_{G_0})$ is simplified by the fact that it is simply a polynomial of degree

n; indeed

$$R(\theta,\delta_{G_0}) = \sum_{j=0}^{n}(d_j-\theta)^2\binom{n}{j}\theta^j(1-\theta)^{n-j}. \tag{3.2}$$

(3.2) was analytically differentiated twice and checked for nonnegativity, using MAPLE. The following table gives the values of m up to which $R(\theta,\delta_{G_0})$ is convex for $|\theta-\frac{1}{2}|\le m$, for $n=2,5,10,15,20$.

n	2	5	10	15	20
m	.5	.315	.228	.2	.1895

The numbers are indeed quite satisfactory; for small n, such as $n = 5$, if θ lies in a symmetric subinterval of [.185, .815], then the uniform prior Bayes rule is Γ-minimax. Even for $n=20$, the uniform prior rule is Γ-minimax if θ lies in a symmetric subinterval of [.3105, .6895], a reasonably large interval. If m exceeds the values given above, $R(\theta,\delta_{G_0})$ ceases to be convex, but in fact for somewhat larger values for m, $\delta_{G_0}(X)$ remains Γ-minimax.

Example 2. Let $X \sim N(\theta,1)$ and suppose $|\theta|\le m$. Assume again θ has a prior of the form $G=(1-\varepsilon_0)G_0+\varepsilon_0 Q$, where Q is symmetric and unimodal about 0 and G_0 is uniform on [−m,m]. Again, other choices of G_0 such as a normal density proportional to $e^{-\frac{1}{2}\theta^2}I_{|\theta|\le m}$ can also be handled. Theorem 2.6 again implies that the uniform prior Bayes rule $\delta_{G_0}(X)$ is Γ-minimax if its risk is convex on $[-m,m]$. Easy integration by parts yields

$$\delta_{G_0}(x) = x + \delta(x),$$

where

$$\delta(x) = \frac{\phi(m+x)-\phi(m-x)}{\Phi(m+x)+\Phi(m-x)-1}, \tag{3.3}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the normal pdf and cdf respectively.

It follows by an application of Stein's Identity (Stein (1981)) that

$$R(\theta,\delta_{G_0}) = 1 + E_\theta[u(x)], \tag{3.4}$$

where

$$u(x) = \delta^2(x) + 2\delta'(x). \tag{3.5}$$

Using Stein's Identity twice more, one obtains

$$\frac{d^2}{d\theta^2}R(\theta, \delta_{G_0}) = E_\theta[u''(X)]. \tag{3.6}$$

(3.6) enables one to check convexity of $R(\theta, \delta_{G_0})$ by doing a single dimensional numerical integration instead of second order numerical differentiation; notice the latter is considerably harder and often less reliable.

$u''(x)$ was again obtained using MAPLE and the expression for $\delta(x)$ in (3.3). Numerical results show that (3.6) is nonnegative for $|\theta| \leq m$ if $m \leq 2.065$. Thus, if θ lies in a symmetric subinterval of $[-2.065, 2.065]$, then the uniform prior Bayes rule has a convex risk and is Γ-minimax.

Again, as in Example 1, actually this rule remains Γ-minimax for a somewhat larger value of m. The reason for this is that if $\theta \sim Q$ where Q is symmetric and unimodal, then θ has the representation $\theta = uZ$, when $u \sim u[-1,1]$, $0 \leq Z \leq m$, and u and Z are independent. Denoting the cdf of Z by F, we then have,

$$\begin{aligned}&\sup_Q r(Q, \delta_{G_0})\\ &= \sup_F \frac{1}{2}\int_0^m \int_{-1}^1 R(uz)du\ dF(z),\end{aligned} \tag{3.7}$$

where $R(\theta)$ denotes the risk of $\delta_{G_0}(X)$ at θ. But (3.7) equals $\sup_{0\leq z\leq m} \frac{1}{z}\int_0^z R(v)dv$, since $R(\theta)$ is symmetric in θ. Thus, as long as $\frac{1}{z}\int_0^z R(v)dv$ is maximized at $z = m$, $r(Q, \delta_{G_0})$ is maximized at $Q = G_0$ and $\delta_{G_0}(X)$ remains Γ-minimax. This was checked by finding $R(v)$ by using (3.4) and then numerically integrating $R(v)$. Our results indicate that for $m \leq 2.54$, $\frac{1}{z}\int_0^z R(v)dv$ is maximized at $z = m$.

If X has known standard deviation σ, the interval should be multiplied by σ.

Example 3. In each of the examples above, the exact value of ε_0 becomes immaterial for the determination of m. We now give an example where the value of ε_0 does matter.

Let again $X \sim N(\theta, 1)$ with $|\theta| \leq m$. However, we now let θ have priors of the form $G = (1-\varepsilon_0)G_0 + \varepsilon_0 Q$, where Q is only assumed symmetric about 0. We assume G_0 is uniform on [−m,m]. Again, if Q^* is the prior assigning mass $\frac{1}{2}$ at each of $\pm m$, then the Bayes rule with respect to $(1-\varepsilon_0)G_0 + \varepsilon_0 Q^*$ is Γ-minimax provided its risk function is convex on $[-m, m]$. The value of m upto which this convexity holds now depends on the value of ε_0 and is given below for some selected values of ε_0:

$\underset{\sim}{\varepsilon_0}$	0	.005	.01	.05	.1
$\underset{\sim}{m}$	2.065	2.064	2.04	1.95	1.85

This seems to suggest that as ε_0 increases, the region of convexity decreases.

Example 4. We will now give a multidimensional example. Let $X \sim N_p(\underset{\sim}{\theta}, I)$ and suppose $||\underset{\sim}{\theta}|| \leq m$. Assume $\underset{\sim}{\theta}$ has a prior of the form $G = (1-\varepsilon_0)G_0 + \varepsilon_0 Q$, where G_0 is the uniform distribution on the sphere $S(\underset{\sim}{0}, m)$ and Q is spherically symmetric unimodal. It follows from Theorem 2.6 that the Bayes rule with respect to G_0 is Γ-minimax if its risk function is subharmonic on the sphere $S(\underset{\sim}{0}, m)$.

Denote $||\underset{\sim}{X}||^2$ by v; familiar calculations, using integration by parts, give

$$\delta_{G_0}(\underset{\sim}{x}) = \underset{\sim}{x} + \frac{\nabla_x(P(||\underset{\sim}{\theta}|| \leq m | \underset{\sim}{\theta} \sim N(\underset{\sim}{x}, I))}{P(||\underset{\sim}{\theta}|| \leq m | \underset{\sim}{\theta} \sim N(\underset{\sim}{x}, I))}, \tag{3.8}$$

where $\nabla_x(\cdot)$ denotes gradient with respect to x.

Using the facts that $||\underset{\sim}{\theta}||^2$ is a noncentral χ^2 with p degrees of freedom and noncentrality parameter v if $\underset{\sim}{\theta} \sim N(\underset{\sim}{x}, I)$ and that a noncentral χ^2 distribution is a Poisson (v) mixture of central χ^2 distributions, (3.8) simplifies to

$$\delta_{G_0}(\underset{\sim}{x}) = (1 + h(v))\underset{\sim}{x}, \tag{3.9}$$

where

$$h(v) = \frac{2N'(v)}{N(v)}, \tag{3.10}$$

and

$$N(v) = \sum_{i=0}^{\infty} \frac{e^{-v} v^i}{i!} b(i), \tag{3.11}$$

with

$$b(i) = \frac{1}{\Gamma(\frac{p+2i}{2})} \int_0^{\frac{m}{2}} e^{-x} x^{\frac{p+2i}{2}-1} dx. \tag{3.12}$$

By an application of Stein's Identity (Stein (1981)), one gets

$$R(\underset{\sim}{\theta}, \delta_{G_0}) = E_{\underset{\sim}{\theta}}[u(v)],$$

where

$$u(v) = p + 2ph(v) + vh^2(v) + 4vh'(v). \tag{3.13}$$

In order to check if $R(\underset{\sim}{\theta}, \delta_{G_0})$ is subharmonic, we need to check the Laplacian of $R(\underset{\sim}{\theta}, \delta_{G_0})$ for nonnegativity. However, it follows from Stein's Identity again that

$$\Delta R(\underset{\sim}{\theta}, \delta_{G_0}) = E_{\underset{\sim}{\theta}}[\Delta u(v)], \tag{3.14}$$

i.e., $\Delta(\cdot)$ and $E_{\underset{\sim}{\theta}}(\cdot)$ can be interchanged. By direct computation,

$$\Delta u(v) = 4vu''(v) + 2pu'(v). \tag{3.15}$$

Again, direct computation gives

$$u'(v) = h^2(v) + h'(v)[4 + 2p + 2vh(v)] + 4vh''(v),$$

and

$$u''(v) = h'(v)[4h(v) + 2vh'(v)] + h''(v)[8 + 2p + 2vh(v)] + 4vh^{(3)}(v).$$

h', h'' and $h^{(3)}$ are in turn given by

$$h'(v) = \frac{2N''(v)}{N(v)} - \frac{1}{2}h^2(v)$$
$$h''(v) = \frac{2N^{(3)}(v)}{N(v)} - \frac{3}{2}h(v)h'(v) - \frac{1}{4}h^3(v)$$

and

$$h^{(3)}(v) = \frac{2N^{(4)}(V)}{N(v)} - 2h''(v)h(v) - \frac{3}{2}h'(v)[h^2(v) + h'(v)] - \frac{1}{8}h^4(v). \tag{3.16}$$

The idea is to compute $N(v)$ and its derivatives and then sequentially compute h, h', h'' and $h^{(3)}$, which are then used to compute u' and u'', required for (3.15).

Finally, notice that computing $N'(v)$ is facilitated by the fact that

$$N'(v) = \sum_{i=0}^{\infty} \frac{e^{-v}v^i}{i!}[b(i+1) - b(i)],$$

and simple integration by parts yields the simplification

$$b(i+1) - b(i) = -\frac{e^{-\frac{m}{2}}(\frac{m}{2})^{\frac{p}{2}+i}}{\Gamma(\frac{p}{2}+i+1)}. \tag{3.17}$$

Similar simplifications occur for N'' etc. Finally, (3.14) was computed numerically by a single numerical integration $\int \Delta u(v) f(v)\, dv$ using MAPLE, where $f(v)$ was obtained using the formula

$$f(v) = \frac{\exp\left(-(v + ||\theta||^2)/2\right)}{\Gamma(\frac{p-1}{2})2^{(p+1)/2}\sqrt{2\pi}} \int_0^v \left(\exp(||\theta||\sqrt{r}) + \exp(-||\theta||\sqrt{r})\right) r^{-1/2}(v-r)^{(p-3)/2}\, dr. \tag{3.18}$$

The value of the radius m upto which risk subharmonicity holds depends on the dimension p and is given below for $p \leq 6$. Notice the apparent monotonicity in p.

<u>p</u>	1	2	3	4	5	6
<u>m</u>	2.065	3.09	3.68	4.10	4.445	4.74

4. Concluding Remarks

The emphasis in this article is on applying the theory to some commonly occurring situations. We find it encouraging that our methods apply to multiparameter situations even if we go beyond the computationally convenient frameworks such as conjugate priors. The calculations are, of course, necessarily harder because of the restriction on the parameter.

Other choices of $\mathcal{L}$ may also be considered; for techniques that apply to the case when $\mathcal{L}$ is the class of star-unimodal or symmetric and star-unimodal distributions, see DasGupta and Delampady (1990). Another practically important question is how does the least favorable prior change when the parameter space is not sufficiently small so that subharmonicity arguments are no longer valid. See Vidakovich and DasGupta (1991) for results in this case.

Acknowledgment

The authors are grateful to Bill Studden, Burgess Davis, Richard Lockhart and Richard Routledge for a variety of instructive discussions.

References

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

Berry, J. Calvin (1989). Bayes minimax estimation of a binomial success probability. *Commun. Statist. -Theory Meth.* **18**, 4607–4616.

Bickel, P. J. (1981). Minimax estimation of the mean of a normal distribution when the parameter space is restricted. *Ann. Statist.* **9**, 1301–1309.

Casella, G. and Strawderman, W. E. (1981). Estimating a bounded normal mean. *Ann. Statist.* **9**, 870–878.

DasGupta, A. (1985). Bayes minimax estimation in multiparameter families when the parameter space is restricted to a bounded convex set. *Sankhyā (A)*, 326–332.

DasGupta, A. and Bose, A. (1988). Γ-minimax and restricted risk Bayes estimation of multiple Poisson means under ε-contaminations of the subjective prior. *Stat. and Decisions* **6**, 311–341.

DasGupta, A. and Delampady, M. (1990). Bayesian hypothesis testing with symmetric and unimodal priors. Technical report.

Donoho, D., Liu, R. and MacGibbon, B. (1990). Minimax risk over hyperrectangles. *Ann. Statist.* **18**, 1416–1437.

Eichenauer, J.(1989a). A gamma minimax result for the class of symmetric and unimodal priors. Preprint.

Eichenauer, J. (1989b). Restricted risk Bayes estimation when the parameter interval is bounded. Preprint.

Ghosh, M. N. (1964). Uniform approximation of minimax point estimates. *Ann. Math. Statist.* **35**, 1031–1047.

Rudin, W. (1974). *Real and complex analysis.* McGraw Hill, New York.

Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *Ann. Math.* **39**, 811–841.

Stein, C. (1973). Estimation of the mean of a multivariate distribution. Proc. Prague Symp. on Asymp. Stat., 345–381.

Stein, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9**, 1135–1151.

Vidakovich, B. and DasGupta, A. (1991). Linear versus nonlinear rules: estimation of a multivariate normal mean (in preparation).

Zeytinoglu M. and Mintz, M. (1984). Optimal fixed size confidence procedures for a restricted parameter space. *Ann. Statist.* **12**, 945–957.

Zeytinoglu, M. and Mintz, M. (1988). Robust fixed size confidence procedures for a restricted parameter space. *Ann. Statist.* **16**, 1241–1253.