Bayesian analysis of a multivariate normal mean with flat tailed priors

by

*Anirban DasGupta
Purdue University

Jayanta K. Ghosh
Indian Statistical Institute
and Purdue University

and

*Mei Mei Zen
Purdue University

Technical Report # 90–22

Department of Statistics
Purdue University

May 1990
Revised July 1990

*Research supported by NSF grant DMS: 8923071.

△

BAYESIAN ANALYSIS OF A MULTIVARIATE NORMAL MEAN WITH FLAT TAILED PRIORS

by

Anirban DasGupta
Purdue University

Jayanta K. Ghosh
Indian Statistical Institute
and Purdue University

and

Mei Mei Zen
Purdue University

ABSTRACT

Bayesian analysis of a multivariate normal mean is considered under nonconjugate flat-tailed priors. It is proved that if $\underset{\sim}{X} \sim N_p(\underset{\sim}{\theta}, \sigma^2 I)$ and $\underset{\sim}{\theta}$ has a prior density $\pi(||\underset{\sim}{\theta} - \underset{\sim}{\mu}||^2/\tau^2)$ where $\underset{\sim}{\mu}, \sigma^2$ and τ^2 are known, then the posterior is starunimodal for all $\underset{\sim}{X}$ if and only if it is logconcave for all $\underset{\sim}{X}$. A necessary and sufficient condition for the posterior to be starunimodal for all $\underset{\sim}{X}$ (or logconcave for all $\underset{\sim}{X}$) is obtained when $\pi(\cdot)$ is a general scale mixture of normal priors. In particular, when $\pi(\cdot)$ is a t prior with m degrees of freedom, it is shown that the posterior is starunimodal for all $\underset{\sim}{X}$ if and only if $\frac{m\tau^2}{\sigma^2} > \frac{m+p}{8}$. For $\frac{m\tau^2}{\sigma^2} \leq \frac{m+p}{8}$, a complete characterization of the set of $\underset{\sim}{X}$ for which the posterior is starunimodal is obtained. This set is shown to be the exterior of a spherical band $a \leq ||\underset{\sim}{X} - \underset{\sim}{\mu}|| \leq b$. The starunimodality property, when it holds, is then used to obtain lower bounds on the posterior probability of subsets of $\mathbb{R}^p$ which are star-shaped about the posterior mode. These lower bounds are then optimized in an appropriate way to construct credible sets under t priors with a guaranteed posterior probability of $1 - \alpha$, for fixed $0 < \alpha < 1$. Evidence is given to show that these credible sets are quite efficient in terms of their size.

Key words: posterior, starunimodal, mode, logconcave, star-shaped sets, Minkowski functionals, posterior probability.

1. Introduction. In this article, we consider Bayesian analysis of a multivariate normal mean $\underset{\sim}{\theta}$ under nonconjugate priors. Special emphasis rests on the multivariate t prior, although many results are proved for more general spherically symmetric priors. The desire to travel outside of the domain of conjugate priors stems from the need or desirability for priors with tails thicker than that of the likelihood function. Notice flat tailed priors such as the t prior often automatically provide a good amount of robustness. See Berger (1985) for general discussions on this. Also see Dawid (1973), Dickey (1976), Meeden and Isaacson (1977), Hill (1980), etc. Although point estimation of $\underset{\sim}{\theta}$ will be peripherally considered, the main goal of this article is to understand the behavior of the posterior distribution, in particular its shape (is the posterior of $\underset{\sim}{\theta}$ unimodal in a suitable sense etc.) and using these shape properties to explicitly construct Bayesian credible sets with a guaranteed posterior probability of $1-\alpha$ (where $0 < \alpha < 1$ is any fixed number). This problem needs attention because as soon as we go outside of conjugate priors, construction of HPD regions in high dimensions becomes an extremely difficult task. The usual numerical methods for constructing high density regions by simulating from the posterior and taking the convex hull of high density points does not work because for the priors we consider (such as a t prior), HPD regions are *not* known to be convex. We provide ample evidence later on that our methods not only explicitly produce credible sets with guaranteed posterior probability, these sets are also typically quite efficient in the sense that the size of the set cannot be much reduced without violating the minimum posterior probability requirement. A principal tool in obtaining these Bayesian credible sets is to establish unimodality of the posterior distribution of $\underset{\sim}{\theta}$ in a suitable sense and then use techniques from the theory of unimodal probability distributions to obtain lower bounds on the probabilities of sets containing the mode of the posterior. Such lower bounds then immediately provide Bayesian credible sets centered at the posterior mode with a guaranteed posterior probability. Unlike in one dimension, there is no single concept of unimodality in higher dimensions. For a lucid treatment of unimodality in high dimensions, we refer the reader to Dharmadhikari and Joag-dev (1988). We will merely mention here that it is possible to index the degree of unimodality by a positive number α in the sense that if a distribution in the p-dimensional space is α-unimodal and if $\alpha < \beta$, then it is also β-unimodal. Thus decreasing the index

of unimodality leads to stronger forms of unimodality. Typically, the best that one can hope for is α-unimodality with $\alpha = p$ (the dimension of the problem). p-unimodality is also known as starunimodality. For (absolutely) continuous distributions with density f, starunimodal distributions are also characterized by the intuitively appealing property that the density f decreases as we move away from the mode along any ray in the p-dimensional space. The concept of strong unimodality (logconcave densities) makes sense in p-dimensional problems also. It is known that distributions which are logconcave are also starunimodal. Thus starunimodality is weaker than logconcavity. We will establish conditions under which the posterior of $\underset{\sim}{\theta}$ given $\underset{\sim}{X} = \underset{\sim}{x}$ is starunimodal. We prove the surprising result that for very general spherically symmetric priors for the mean $\underset{\sim}{\theta}$, the posterior of $\underset{\sim}{\theta}$ given $\underset{\sim}{X} = \underset{\sim}{x}$ is starunimodal for every $\underset{\sim}{x}$ if and only if it is logconcave for every $\underset{\sim}{x}$. Notice there may (and usually do) exist specific $\underset{\sim}{x}$ for which the posterior is starunimodal but not logconcave, but if the posterior is starunimodal for every $\underset{\sim}{x}$, it must be logconcave for every $\underset{\sim}{x}$ too. As an example, suppose $\underset{\sim}{X} \sim N(\underset{\sim}{\theta}, \sigma^2 I)$ and $\underset{\sim}{\theta}$ has a central t distribution with m degrees of freedom, location parameter $\underset{\sim}{\mu}$ and scale parameter τ^2. Assume $\underset{\sim}{\mu}, \sigma^2, \tau^2$ are known. Then it is proved the posterior of $\underset{\sim}{\theta}$ is starunimodal for each $\underset{\sim}{x}$ if and only if $\frac{m\tau^2}{\sigma^2} > \frac{m+p}{8}$. If $\frac{m\tau^2}{\sigma^2} \leq \frac{m+p}{8}$, the posterior is not starunimodal for each $\underset{\sim}{x}$. We have a characterization of the set S_m of $\underset{\sim}{x}$ for which the posterior is not starunimodal. It is proved that S_m is a spherical band on the p-dimensional space, i.e., S_m is a set of the form $\{\underset{\sim}{X} : a_m \leq ||\underset{\sim}{X} - \underset{\sim}{\mu}|| \leq b_m\}$. We give qualitative descriptions of the behavior of the set S_m as m increases. In particular, we show that the set S_m converges to the empty set as $m \to \infty$; these results are given in section 2. Notice versions of some of these results for the one dimensional case were proved in DasGupta (1988); also see Fan and Berger (1989) for related results.

Apart from the multidimensional t priors, other natural priors in this case are general scale mixtures of normal priors (the t prior is one such). In section 3, we give a short treatment of the problem for general scale mixtures. In particular, we give a theorem describing a condition on the mixing distribution required for the posterior to be starunimodal for every $\underset{\sim}{x}$. A few examples are given to illustrate the theorem and to help qualitatively explain the message of this theorem. As stated before, a major consequence of the starunimodality of the posterior is a rather sharp lower bound on the probability of a starshaped set C containing the mode $\underset{\sim}{\nu} = \underset{\sim}{\nu}(\underset{\sim}{x})$. Specifically, suppose $\underset{\sim}{x}$ and the prior for $\underset{\sim}{\theta}$ are such that the posterior for $\underset{\sim}{\theta}$ for this specific $\underset{\sim}{x}$ is starunimodal with mode at $\underset{\sim}{\nu}$. Then it turns out that for $r > 0$,

$$P(\underset{\sim}{\theta} \varepsilon C/\underset{\sim}{x}) \geq \tau_C(r), \text{ where}$$

$\tau_C(r)$ is the rth order posterior moment of a suitable functional of $\underset{\sim}{\theta}$ depending on the set C (this will be made more precise in the body of the paper). By choosing C large enough, one can make $\tau_C(r) \geq 1 - \alpha$ (usually exactly equal to $1 - \alpha$) where $0 < \alpha < 1$ is specified. C then is a guaranteed credible set of posterior probability at least $1 - \alpha$. This technique works for any $r > 0$. We demonstrate that typically there exists a moderate value of r for which the probability bound is the sharpest and the set C resulting from this $r > 0$ is seen to be an efficient credible set for $\underset{\sim}{\theta}$ in terms of size of the set. Notice we have an enormous amount of flexibility in choosing the shape of the credible set because the technique works for any starshaped set (in particular spherical, elliptical and rectangular regions). We do not claim that our sets are HPD regions for this problem. Indeed, the motivation derives from the fact that it is extremely difficult to find the HPD sets here. These results are given in Section 4. In Section 5, we make some concluding remarks.

2. Unimodality of the posterior for spherically symmetric priors. In this section we establish appropriate unimodality properties of the posterior for $\underset{\sim}{\theta}$ when $\underset{\sim}{\theta}$ has a spherically symmetric prior. Recall that unimodality of the posterior is crucial for obtaining efficient credible sets. As stated in the introduction, there is no unique concept of unimodality in multidimensions. We will, for the most part, consider star unimodality. This is because it is a natural unimodality property and furthermore suffices for the subsequent construction of credible sets. Logconcavity or strong unimodality, on the other hand, is interesting on its own merit. For the sake of completeness we first define star unimodality and then prove an equivalence theorem stating the equivalence of logconcavity of the posterior of $\underset{\sim}{\theta}$ for all $\underset{\sim}{X}$ and star unimodality of the posterior of $\underset{\sim}{\theta}$ for all $\underset{\sim}{X}$. Aside from the theoretical importance, this result has practical implications in the sense that logconcavity is typically much

easier to verify than star unimodality in the problems we are considering. The theorem enables us to verify starunimodality for all $\underset{\sim}{X}$ by verifying logconcavity for all $\underset{\sim}{X}$ instead.

Definition: Let $\underset{\sim}{X} \sim F$ be an absolutely continuous random variable on $\mathbb{R}^p$ with density $f(\underset{\sim}{X})$. We say $\underset{\sim}{X}$ (or equivalently f) is starunimodal about zero if $f(t\underset{\sim}{X}) \leq f(s\underset{\sim}{X})$ for all $0 < t \leq s < \infty$ and all $\underset{\sim}{X}$. $\underset{\sim}{X}$ is said to be starunimodal about $\underset{\sim}{\nu}$ if $\underset{\sim}{X} - \underset{\sim}{\nu}$ is starunimodal about zero. $\underset{\sim}{\nu}$ will be called the mode of $\underset{\sim}{X}$.

Remark Notice that if f is differentiable, then starunimodality is equivalent to $\frac{d}{dt} f(t\underset{\sim}{X}) \leq 0$ for all $t > 0$ and all $\underset{\sim}{X}$.

Theorem 2.1 Let $\underset{\sim}{X} \sim N(\underset{\sim}{\theta}, \sigma^2 I)$ and let $\underset{\sim}{\theta} \sim \pi\frac{(||\theta - \underset{\sim}{\mu}||^2)}{\tau^2}$. Assume $\underset{\sim}{\mu}, \sigma^2, \tau^2$ are known. Suppose $\pi(\cdot)$ is twice differentiable and decreasing. Then the posterior distribution of $\underset{\sim}{\theta}$ is logconcave for all $\underset{\sim}{X}$ if and only if it is starunimodal for all $\underset{\sim}{X}$.

Proof: Clearly we only need prove the if part. We will, without loss, assume $\underset{\sim}{\mu} = \underset{\sim}{0}$ and further let $\sigma^2 = \tau^2 = 1$. The case of general σ^2 and τ^2 is exactly similar. Let $\pi(\underset{\sim}{\theta}|\underset{\sim}{X})$ denote the posterior density of $\underset{\sim}{\theta}$. Clearly, $-\log \pi(\underset{\sim}{\theta}|\underset{\sim}{X})$ is proportional to $\phi_1(||\theta||^2) + \frac{1}{2}||\theta - X||^2$, where $\phi_1 = -\log \pi$. Suppose now $\pi(\underset{\sim}{\theta}|\underset{\sim}{X})$ is starunimodal for all $\underset{\sim}{X}$ with mode at $\underset{\sim}{\nu} = \underset{\sim}{\nu}(\underset{\sim}{X})$. $\therefore$ $\underset{\sim}{Z} = \underset{\sim}{\theta} - \underset{\sim}{\nu}$ is starunimodal for all $\underset{\sim}{X}$ with mode at $\underset{\sim}{0}$. It then follows that $\forall\, t > 0,\ \forall\, \underset{\sim}{Z},\ \forall\, \underset{\sim}{X}$,

$$\frac{d}{dt}\left\{\Phi_1(||t\underset{\sim}{Z} + \underset{\sim}{\nu}||^2) + \frac{1}{2}||t\underset{\sim}{Z} + \underset{\sim}{\nu} - \underset{\sim}{X}||^2\right\} \geq 0$$

$$\iff 2\phi_1'(||t\underset{\sim}{Z} + \underset{\sim}{\nu}||^2)(t||\underset{\sim}{Z}||^2 + \underset{\sim}{\nu}'\underset{\sim}{Z}) + (t||\underset{\sim}{Z}||^2 + \underset{\sim}{Z}'(\underset{\sim}{\nu} - \underset{\sim}{X})) \geq 0 \ \ \forall\, t > 0,\ \ \forall\, \underset{\sim}{Z},\ \ \forall\, \underset{\sim}{X}. \quad (2.1)$$

Letting $t \to 0$ in (2.1), we have

$$2\phi_1'(||\underset{\sim}{\nu}||^2) \cdot \underset{\sim}{\nu}'\underset{\sim}{Z} + \underset{\sim}{Z}'(\underset{\sim}{\nu} - \underset{\sim}{X}) \geq 0 \ \ \forall\, \underset{\sim}{Z},\ \ \forall\, \underset{\sim}{X}$$
$$\iff (2\phi_1'(||\underset{\sim}{\nu}||^2) \cdot \underset{\sim}{\nu} + \underset{\sim}{\nu} - \underset{\sim}{X})'\underset{\sim}{Z} \geq 0 \ \ \forall\, \underset{\sim}{Z},\ \ \forall\, \underset{\sim}{X}, \quad (2.2)$$

from which it immediately follows

$$(2\phi_1'(||\underset{\sim}{\nu}||^2) + 1) \cdot \underset{\sim}{\nu} = \underset{\sim}{X}, \quad (2.3)$$

i.e., $\underset{\sim}{\nu} = a\underset{\sim}{X}$ for suitable a. Notice 'a' may (and will, usually) depend on $\underset{\sim}{X}$. Given $\underset{\sim}{X} \neq \underset{\sim}{0}$, 'a' can be found from the equation

$$a(2\phi_1'(a^2||\underset{\sim}{X}||^2) + 1) - 1 = 0; \tag{2.4}$$

it is easy to check that $0 < a < 1$ and that $a = a(||\underset{\sim}{X}||)$ is continuous in $||\underset{\sim}{X}||$. Substituting $a\underset{\sim}{X}$ for $\underset{\sim}{\nu}$ and $\underset{\sim}{Z} = \underset{\sim}{X}$ in (2.1), one then has

$$2\phi_1'((t+a)^2\underset{\sim}{X}'\underset{\sim}{X})(t+a) + t + a - 1 \geq 0 \;\; \forall\, t, \;\; \forall\, \underset{\sim}{X}. \tag{2.5}$$

Multiplying both sides of (2.5) by $||\underset{\sim}{X}||$, writing ω for $(t+a)||\underset{\sim}{X}||$, ω_0 for $a||\underset{\sim}{X}||$ and letting $f(\omega) = 2\phi_1'(\omega^2)\omega + \omega$, one then obtains that given $||\underset{\sim}{X}|| > 0$, there exists $\omega_0 > 0$ such that

$$f(\omega_0) = ||\underset{\sim}{X}||,$$

$$\text{and } f(\omega) > ||\underset{\sim}{X}|| \;\; \forall\, \omega > \omega_0.$$

Since f must then be increasing, it follows that $\pi(\underset{\sim}{\theta}|\underset{\sim}{X})$ is logconcave for all $\underset{\sim}{X}$. This is because direct computations give that the Hessian matrix of $-\log \pi(\underset{\sim}{\theta}|\underset{\sim}{X}) = (1+2\phi_1'(||\underset{\sim}{\theta}||^2))\cdot I + 4\phi_1''(||\underset{\sim}{\theta}||^2)\underset{\sim}{\theta}\underset{\sim}{\theta}'$ which is nonnegative definite if $1+2\phi_1'(||\underset{\sim}{\theta}||^2)+4||\underset{\sim}{\theta}||^2\phi_1''(||\underset{\sim}{\theta}||^2) \geq 0$. This last inequality, however, follows if $f(\omega)$ is an increasing function.

<u>Corollary 1</u> Let $\underset{\sim}{X} \sim N(\underset{\sim}{\theta}, \sigma^2 I)$ and $\underset{\sim}{\theta} \sim t(m, \underset{\sim}{\mu}, \tau^2 I)$; then the posterior of $\underset{\sim}{\theta}$ given $\underset{\sim}{X}$ is starunimodal for all $\underset{\sim}{X}$ if and only if $\frac{m\tau^2}{\sigma^2} > \frac{m+p}{8}$.

Proof: We give the proof here for the case $\sigma^2 = \tau^2 = 1$ (the proof for the general case is essentially the same). Clearly, we can assume $\underset{\sim}{\mu} = \underset{\sim}{0}$. From Theorem 2.1, the posterior is starunimodal for all $\underset{\sim}{X}$ if and only if it is logconcave for all $\underset{\sim}{X}$. The condition for logconcavity is $1 + 2\phi_1'(||\underset{\sim}{\theta}||^2) + 4||\underset{\sim}{\theta}||^2\phi_1''(||\underset{\sim}{\theta}||^2) \geq 0$ for all $\underset{\sim}{\theta}$ which reduces to $m > \frac{p}{7}$ on computation, as required.

The above corollary implies that if $\underset{\sim}{\theta} \sim t(m, \underset{\sim}{\mu}, \tau^2 I)$ and if $\frac{m\tau^2}{\sigma^2} \leq \frac{m+p}{8}$, then there exist appropriate $\underset{\sim}{X}$ such that the posterior of $\underset{\sim}{\theta}$ given $\underset{\sim}{X}$ is not starunimodal. The next theorem characterizes the set of $\underset{\sim}{X}$ for which the posterior of $\underset{\sim}{\theta}$ given $\underset{\sim}{X}$ is starunimodal for any given specific values of m, σ^2, τ^2. We take $\underset{\sim}{\mu} = \underset{\sim}{0}$; if $\underset{\sim}{\mu} \neq \underset{\sim}{0}$, all assertions hold with $\underset{\sim}{X} - \underset{\sim}{\mu}$ in place of $\underset{\sim}{X}$. For the sake of brevity, we will denote $\frac{\underset{\sim}{\theta}}{\sigma} = \underset{\sim}{Z}$, $\frac{\underset{\sim}{X}}{\sigma} = \underset{\sim}{X}_0$, $\frac{\sigma^2}{m\tau^2} = \alpha, m + p = \beta$, and $\gamma = \frac{1}{\alpha} = \frac{m\tau^2}{\sigma^2}$. Ignoring constants of proportionality, the posterior density $\pi(\underset{\sim}{\theta}|\underset{\sim}{X})$ equals

$$\pi(\underset{\sim}{\theta}|\underset{\sim}{X}) = e^{-\frac{1}{2\sigma^2}}(\underset{\sim}{\theta} - \underset{\sim}{X})'(\underset{\sim}{\theta} - \underset{\sim}{X}) \cdot \frac{1}{\left(1 + \frac{\underset{\sim}{\theta}'\underset{\sim}{\theta}}{m\tau^2}\right)^{\frac{m+p}{2}}} .$$

$$\therefore -\log \pi(\underset{\sim}{\theta}|\underset{\sim}{X}) = (\underset{\sim}{Z} - \underset{\sim}{X}_0)'(\underset{\sim}{Z} - \underset{\sim}{X}_0) + \frac{m+p}{2}\log(1 + \alpha \underset{\sim}{Z}'\underset{\sim}{Z}). \tag{2.6}$$

If the posterior of $\underset{\sim}{\theta}$ is starunimodal about, say, $\underset{\sim}{\nu}(\underset{\sim}{X}) = \underset{\sim}{\nu}$, then using the argument of Theorem (2.1) it now follows using (2.6) that

$$\underset{\sim}{\nu} = a\underset{\sim}{X}_0,$$

where a solves

$$g(a) = a - 1 + \frac{\alpha\beta a}{1 + \alpha a^2 \underset{\sim}{X}_0'\underset{\sim}{X}_0} = 0. \tag{2.7}$$

$$\iff \alpha \underset{\sim}{X}_0'\underset{\sim}{X}_0 a^3 - \alpha \underset{\sim}{X}_0'\underset{\sim}{X}_0 a^2 + (1 + \alpha\beta)a - 1 = 0 \tag{2.8}$$

Substituting $\frac{1}{\underset{\sim}{X}_0'\underset{\sim}{X}_0} = y$, (2.8) reduces to

$$h(a) = a^3 - a^2 + (\beta + \gamma)ay - \gamma y = 0. \tag{2.9}$$

Notice $h(a)$ always has a root in the open interval (0, 1) and cannot have any roots outside of [0, 1]. If the posterior of $\underset{\sim}{\theta}$ given $\underset{\sim}{X}$ is starunimodal, then $h(a)$ has only one root. Since $h(a)$ is a cubic in a, it is wellknown (see, e.g., Press et al, page 146) that

$$\text{letting } Q = \frac{1 - 3(\beta + \gamma)y}{9}$$
$$\text{and } R = \frac{-2 + 9(\beta + \gamma)y - 27\gamma y}{54},$$

$h(a)$ has only one real root if and only if

$$Q^3 - R^2 \leq 0$$

which, on lengthy but straightforward algebra reduces to

$$4(\beta + \gamma)^3 y^2 - (\beta^2 - 8\gamma^2 + 20\beta\gamma)y + 4\gamma \geq 0 \tag{2.10}$$

For future reference, we will define $S_{1,m} = \{\underset{\sim}{X} : (2.10) \text{ holds}\}$; recall here $\underset{\sim}{y} = \frac{1}{||\underset{\sim}{X}_0||^2}$. If $h(a)$ has a unique root in the open interval (0, 1), say a^*, then the posterior, if it is starunimodal, will be starunimodal with mode at $a^*\underset{\sim}{X}_0$ (the unique root a^* can be obtained in a closed form; see Press et al (1986)). Equivalently, $\underset{\sim}{Z} = \underset{\sim}{\theta} - a^*\underset{\sim}{X}_0$ will be starunimodal with mode at $\underset{\sim}{0}$. If $f(\underset{\sim}{z}|\underset{\sim}{X})$ denotes the posterior density of $\underset{\sim}{z} = \underset{\sim}{\theta} - a^*\underset{\sim}{X}_0$, then starunimodality of f is equivalent to

$$\frac{d}{dt} f(t\underset{\sim}{z}|\underset{\sim}{X}) \leq 0 \quad \forall\, t > 0, \quad \forall\, \underset{\sim}{z},$$

which, on direct computation, reduces to

$$\begin{aligned}
&\{t||\underset{\sim}{Z}||^2 + (a^* - 1)\underset{\sim}{X}_0'\underset{\sim}{z}\}\{1 + \alpha(t\underset{\sim}{z} + a^*\underset{\sim}{X}_0)'(t\underset{\sim}{z} + a^*\underset{\sim}{X}_0)\} \\
&\quad + \alpha\beta(t||\underset{\sim}{z}||^2 + a^*\underset{\sim}{X}_0'\underset{\sim}{z}) \geq 0 \quad \forall\, t > 0, \quad \forall\, \underset{\sim}{z} \\
\Longleftrightarrow\ & \alpha||\underset{\sim}{z}||^4 t^3 + [2\alpha a^*\underset{\sim}{X}_0'\underset{\sim}{z}||\underset{\sim}{z}||^2 + \alpha(a^* - 1)\underset{\sim}{X}_0'\underset{\sim}{z}||\underset{\sim}{z}||^2]t^2 \\
&\quad + [(1 + \alpha {a^*}^2 \underset{\sim}{X}_0'\underset{\sim}{X}_0)||\underset{\sim}{z}||^2 + 2\alpha a^*(a^* - 1)(\underset{\sim}{X}_0'\underset{\sim}{Z})^2 + \alpha\beta||\underset{\sim}{z}||^2]t \\
&\quad + \{(a^* - 1)(1 + \alpha {a^*}^2 \underset{\sim}{X}_0'\underset{\sim}{X}_0) + \alpha\beta a^*\}\underset{\sim}{X}_0'\underset{\sim}{z} \geq 0 \quad \forall\, t > 0, \quad \forall\, \underset{\sim}{z} \qquad (2.11)
\end{aligned}$$

Because a^* solves $h(a) = 0$ (or equivalently, solves (2.7)), (2.11) reduces to the quadratic inequality

$$\begin{aligned}
&\alpha||\underset{\sim}{z}||^4 t^2 + ||\underset{\sim}{z}||^2 \underset{\sim}{X}_0'\underset{\sim}{z}(2\alpha a^* + \alpha(a^* - 1))t \qquad (2.12) \\
&\quad + [||\underset{\sim}{z}||^2(1 + \alpha\beta + \alpha {a^*}^2 \underset{\sim}{X}_0'\underset{\sim}{X}_0) - 2\alpha a^*(1 - a^*)(\underset{\sim}{X}_0'\underset{\sim}{z})^2] \geq 0, \quad \forall\, t > 0, \quad \forall\, \underset{\sim}{z}.
\end{aligned}$$

For $\underset{\sim}{z} \neq \underset{\sim}{0}$, the roots of (2.12) are given by

$$t = \frac{-(3\alpha a^* - \alpha)\underset{\sim}{X}_0'\underset{\sim}{z} \pm \sqrt{(\alpha - 3\alpha a^*)^2(\underset{\sim}{X}_0'\underset{\sim}{z})^2 - 4\alpha[||\underset{\sim}{z}||^2(1 + \alpha\beta + \alpha {a^*}^2 \underset{\sim}{X}_0'\underset{\sim}{X}_0) - 2\alpha a^*(1 - a^*)(\underset{\sim}{X}_0'\underset{\sim}{z})^2]}}{2\alpha||\underset{\sim}{z}||^2} \qquad (2.13)$$

We claim that (2.12) ≥ 0 for $\forall\, t > 0$, $\forall\, \underset{\sim}{z}$ if and only if the discriminant is < 0 $\forall\, \underset{\sim}{z}$. This is because if the discriminant is ≥ 0 for some $\underset{\sim}{z}$, say $\underset{\sim}{z}_0$, then for one of $\underset{\sim}{z} = \underset{\sim}{z}_0$ and $\underset{\sim}{z} = -\underset{\sim}{z}_0$, at least one of the two roots in (2.13) is strictly positive which in turn implies that there must exist $t > 0$ for which (2.12) is strictly negative for this particular $\underset{\sim}{z}$, a contradiction.

The condition that the discriminant is strictly negative for all $\underset{\sim}{z}$ is

$$\begin{aligned}
&(\alpha - 3\alpha a^*)^2(\underset{\sim}{X}_0'\underset{\sim}{z})^2 + 8\alpha^2 a^*(1-a^*)(\underset{\sim}{X}_0'\underset{\sim}{z})^2.\\
&< 4\alpha(1+\alpha\beta+\alpha a^{*^2}\underset{\sim}{X}_0'\underset{\sim}{X}_0)||\underset{\sim}{z}||^2 \quad \forall\, \underset{\sim}{z}\\
\Longleftrightarrow &(1+a^*)^2\alpha^2(\underset{\sim}{X}_0'\underset{\sim}{z})^2 < 4\alpha(1+\alpha\beta+\alpha a^{*^2}\underset{\sim}{X}_0'\underset{\sim}{X}_0)||\underset{\sim}{z}||^2 \quad \forall\, \underset{\sim}{z}\\
\Longleftrightarrow &(1+a^*)^2||\underset{\sim}{X}_0||^2 < 4(\beta+\gamma+a^{*^2}||\underset{\sim}{X}_0||^2)\\
\Longleftrightarrow &(-3a^{*^2}+2a^*+1)||\underset{\sim}{X}_0||^2 < 4(\beta+\gamma) \qquad (2.14)
\end{aligned}$$

On the other hand (2.7) implies that

$$||\underset{\sim}{X}_0||^2 = \frac{(1+\alpha\beta)a^*-1}{\alpha a^{*2}(1-a^*)} \qquad (2.15)$$

Substituting (2.15) for $||\underset{\sim}{X}_0||^2$, (2.14) reduces to

$$(3a^{*^2}-2a^*-1)[(1+\alpha\beta)a^*-1] + 4(1+\alpha\beta)a^{*^2}(1-a^*) > 0. \qquad (2.16)$$

Factorizing $3a^{*^2}-2a^*-1$ as $(3a^*+1)(a^*-1)$, (2.16) reduces to

$$a^{*^2}(1+\alpha\beta) + a^*(2-\alpha\beta) + 1 > 0. \qquad (2.17)$$

Again, for future reference, we will define

$$S_{2,m} = \{\underset{\sim}{X}\varepsilon S_{1,m} : (2.17) \text{ holds}\};$$

here a^* is the unique root of $h(a) = 0$.

Since the posterior is starunimodal if and only if (2.10) and (2.17) both hold, we have the following theorem:

<u>Theorem 2.2</u> Let $\underset{\sim}{X} \sim N(\underset{\sim}{\theta}, \sigma^2 I)$ and $\underset{\sim}{\theta} \sim t(m, \underset{\sim}{\mu}, \tau^2 I)$. Then the posterior of $\underset{\sim}{\theta}$ given $\underset{\sim}{X}$ is starunimodal if and only if $\underset{\sim}{X}\varepsilon S_{2,m}$, where $\gamma = \frac{m\tau^2}{\sigma^2}, \beta = m+p, \gamma = \frac{1}{\alpha}$, and $y = \frac{\sigma^2}{||\underset{\sim}{X}||^2}$.

Remark: Recall that for $\gamma > \frac{1}{8}\beta$, the posterior is starunimodal for all $\underset{\sim}{X}$. Indeed, under this condition, an involved argument does imply that $S_{1,m} = S_{2,m} = \mathbb{R}^p$. In DasGupta (1988), it was proved that if $X \sim N(\theta, \sigma^2)$ and $\theta \sim t(m, \mu, \tau^2)$, then the posterior of θ given

X is unimodal for all X if and only if $\frac{m\tau^2}{\sigma^2} > \frac{m+1}{8}$; otherwise the posterior is unimodal for $|X - \mu| \leq a$ and $|X - \mu| \geq b$ for suitable a and b. In other words, the posterior fails to be unimodal for $a < |X - \mu| < b$. We will now give a similar result for the p-dimensional case. First we need the following proposition.

<u>Proposition 2.3</u>. Suppose $\underset{\sim}{X}_1, \underset{\sim}{X}_2$ are such that $h(a)$ has only one root for $\underset{\sim}{X} = \underset{\sim}{X}_i, i = 1, 2$. Let $a^*(\underset{\sim}{X}_i)$ denote the root of $h(a)$ for $\underset{\sim}{X} = \underset{\sim}{X}_i$. Suppose $||\underset{\sim}{X}_1|| < ||\underset{\sim}{X}_2||$. Then $a^*(||\underset{\sim}{X}_1||) \leq a^*(||\underset{\sim}{X}_2||)$.

Proof: For any fixed $\underset{\sim}{X}$, writing y for $\frac{1}{||\underset{\sim}{X}_0||^2}$, by (2.7), $a^*(||\underset{\sim}{X}||)$ satisfies

$$a^*(y) - 1 + \frac{\alpha\beta y a^*(y)}{\alpha a^{*2}(y) + y} = 0. \tag{2.18}$$

Denote $\frac{1}{||\underset{\sim}{X}_1||^2} = y_1, \frac{1}{||\underset{\sim}{X}_2||^2} = y_2$; note $y_2 < y_1$. Then we have,

$$\begin{aligned} a^*(y_1) - 1 + \frac{\alpha\beta y_1 a^*(y_1)}{\alpha a^{*2}(y_1) + y_1} &= 0 \\ \Longrightarrow a^*(y_1) - 1 + \frac{\alpha\beta a^*(y_1) y_2}{\alpha a^{*2}(y_1) + y_2} &< 0 \end{aligned} \tag{2.19}$$

(since $\frac{y}{c+y}$ is increasing as a function of y for positive c). We claim (2.19) implies $a^*(y_2) > a^*(y_1)$, which will prove the proposition. For if $a^*(y_2) \leq a^*(y_1)$, then for $y = y_2$, the function $h(a)$ is zero at $a = a^*(y_2)$, negative at $a = a^*(y_1)$ and positive at $a = 1$. By continuity of $h(a)$, there must therefore be another root of $h(a)$ between $a^*(y_1)$ and 1, a contradiction to the assumption that $h(a)$ has a unique root for $\underset{\sim}{X} = \underset{\sim}{X}_2$.

We now prove the following theorem.

<u>Theorem 2.4</u> $S^c_{2,m}$ is a set of the form $\{\underset{\sim}{X} : a_m \leq ||\underset{\sim}{X} - \underset{\sim}{\mu}|| \leq b_m\}$ for suitable constants a_m, b_m depending on m, σ^2, τ^2.

Proof: Assume without loss of generality that $\underset{\sim}{\mu} = \underset{\sim}{0}$. By definition, $S_{2,m}$ is the set of all $\underset{\sim}{X}$ such that (2.10) and (2.17) both hold. Plainly, the set of $\underset{\sim}{X}$ for which (2.10) holds is a set of the form

$$\{\underset{\sim}{X} | ||\underset{\sim}{X}|| \geq a_{1,m} \text{ or } ||\underset{\sim}{X}|| \leq b_{1,m}\}. \tag{2.20}$$

Also, by Proposition 2.3 one gets that the set of $\underset{\sim}{X}$ for which (2.17) holds is a set of the form

$$\{\underset{\sim}{X} : ||\underset{\sim}{X}|| \geq a_{2,m} \text{ or } ||\underset{\sim}{X}|| \leq b_{2,m}\}. \tag{2.21}$$

Combining (2.20) and (2.21), it then follows that, in general, $S^c_{2,m}$ is a set of the form

$$\{\underset{\sim}{X} : c_{1,m} \leq ||\underset{\sim}{X}|| \leq d_{1,m}\} \cup \{\underset{\sim}{X} : c_{2,m} \leq ||\underset{\sim}{X}|| \leq d_{2,m}\}. \tag{2.22}$$

It is easy to directly verify, however, that $S^c_{2,m}$ can be a union of two distinct spherical bands only for the two configurations $b_{2,m} \leq a_{2,m} < b_{1,m} \leq a_{1,m}$ or $b_{1,m} \leq a_{1,m} < b_{2,m} \leq a_{2,m}$. We will now prove that none of these two configurations can arise. Notice it is only necessary to consider the case $\beta \geq 8\gamma$, because otherwise the posterior is star-unimodal for all $\underset{\sim}{X}$.

By direct computation,

$$a^2_{1,m} = \frac{8(\beta+\gamma)^3}{\beta^2 - 8\gamma^2 + 20\beta\gamma - \sqrt{\beta(\beta-8\gamma)^3}}$$

and

$$b^2_{1,m} = \frac{8(\beta+\gamma)^3}{\beta^2 - 8\gamma^2 + 20\beta\gamma - \sqrt{\beta(\beta-8\gamma)^3}}. \tag{2.23}$$

On simplification, (2.23) reduces to

$$a^2_{1,m} = \frac{\beta^2 - 8\gamma + 20\beta\gamma + \sqrt{\beta(\beta-8\gamma)^3}}{8\gamma}$$

and

$$b^2_{1,m} = \frac{\beta^2 - 8\gamma^2 + 20\beta\gamma - \sqrt{\beta(\beta-8\gamma)^3}}{8\gamma}. \tag{2.24}$$

Also, note that (2.17) holds if and only if

$$a^* \geq \frac{\beta - 2\gamma + \sqrt{\beta(\beta-8\gamma)}}{2(\beta+\gamma)} \qquad \text{or} \qquad a^* \leq \frac{\beta - 2\gamma - \sqrt{\beta(\beta-8\gamma)}}{2(\beta+\gamma)}.$$

Combining this with (2.15), one therefore gets that (2.17) holds if and only if

$$||\underset{\sim}{X}||^2 \geq \frac{4(\beta+\gamma)^3(\beta - 4\gamma + \sqrt{\beta(\beta-8\gamma)})}{[\beta - 2\gamma + \sqrt{\beta(\beta-8\gamma)}]^2[\beta + 4\gamma - \sqrt{\beta(\beta-8\gamma)}]} = a^2_{2,m}$$

or

$$||\underset{\sim}{X}||^2 \le \frac{4(\beta+\gamma)^3(\beta-4\gamma-\sqrt{\beta(\beta-8\gamma)})}{[\beta-2\gamma-\sqrt{\beta(\beta-8\gamma)}]^2[\beta+4\gamma+\sqrt{\beta(\beta-8\gamma)}]} = b_{2,m}^2. \tag{2.25}$$

Again, algebra reduces (2.25) to

$$a_{2,m}^2 = \frac{(\beta+\gamma)^2[(\beta+\sqrt{\beta(\beta-8\gamma)})^2-16\gamma^2]}{4\gamma[\beta+\sqrt{\beta(\beta-8\gamma)}-2\gamma]^2}$$

and

$$b_{2,m}^2 = \frac{(\beta+\gamma)^2[(\beta-\sqrt{\beta(\beta-8\gamma)})^2-16\gamma^2]}{4\gamma[\beta-\sqrt{\beta(\beta-8\gamma)}-2\gamma]^2}. \tag{2.26}$$

Straightforward algebra now gives that

$$\begin{aligned} &a_{1,m}^2 - b_{2,m}^2 \\ &= \frac{[\beta^2-8\gamma^2+20\beta\gamma+\sqrt{\beta(\beta-8\gamma)^3}][\beta-2\gamma-\sqrt{\beta(\beta-8\gamma)}]^2 - 2(\beta+\gamma)^2[(\beta-\sqrt{\beta(\beta-8\gamma)})^2-16\gamma^2]}{8\gamma[\beta-\sqrt{\beta(\beta-8\gamma)}-2\gamma]^2} \end{aligned} \tag{2.27}$$

The numerator of (2.27) simplifies to

$$\frac{16(\beta-8\gamma)\gamma^2(\beta+\gamma)^2}{(\beta^2-6\beta\gamma+2\gamma^2)\sqrt{\beta(\beta-8\gamma)}+\beta(\beta-8\gamma)(\beta-2\gamma)}, \tag{2.28}$$

which is nonnegative since $\beta-8\gamma$, $\beta-2\gamma$ and $\beta^2-6\beta\gamma+2\gamma^2$ are each nonnegative. This proves that $a_{1,m} \ge b_{2,m}$. Similar arguments give that $a_{2,m} \ge b_{1,m}$. Hence $S_{2,m}^c$, i.e., the set of $\underset{\sim}{X}$ for which the posterior of $\underset{\sim}{\theta}$ is not starunimodal, is a single spherical band.

3. Starunimodality and logconcavity for general mixture normal priors. In Section 2, we derived a necessary and sufficient condition for the posterior of $\underset{\sim}{\theta}$ to be starunimodal for all $\underset{\sim}{X}$ (or equivalently, logconcave for all $\underset{\sim}{X}$), when $\underset{\sim}{\theta}$ has a spherically symmetric t prior. In this section, we give an analogous necessary and sufficient condition when the prior for $\underset{\sim}{\theta}$ is a scale mixture of normals of the form

$$\pi(\underset{\sim}{\theta}'\underset{\sim}{\theta}) = \int \frac{e^{\frac{-\underset{\sim}{\theta}'\underset{\sim}{\theta}}{2\eta}}}{\eta^{p/2}} dG(\eta), \tag{3.1}$$

where G is a probability measure on $(0,\infty)$. Note $\pi(\cdot)$ is not a probability density because it has not been normalized; however, this will naturally have no bearing on any of the results derived in this section.

Theorem 3.1 Let $\underset{\sim}{X} \sim N(\underset{\sim}{\theta}, I)$ and let $\underset{\sim}{\theta}$ have a spherically symmetric prior density $\pi(\underset{\sim}{\theta}'\underset{\sim}{\theta})$ given by (3.1). Then the posterior of $\underset{\sim}{\theta}$ given $\underset{\sim}{X}$ is starunimodal for all $\underset{\sim}{X}$ if and only if

$$2\lambda \ \mathrm{Var}_\lambda(u) \leq 1 + E_\lambda(u) \text{ for all } \lambda,$$

where $\lambda \geq 0$ denotes $\frac{\underset{\sim}{\theta}'\underset{\sim}{\theta}}{2}$ and u is distributed as $p(u|\lambda) = e^{-\lambda u} u^{\frac{p}{2}} dH(u) / \int e^{-\lambda u} u^{\frac{p}{2}} dH(u)$ where H is the probability distribution of $\frac{1}{\eta}$ induced by the distribution G for η in (3.1).

Remark: For the case when $\underset{\sim}{X} \sim N(\underset{\sim}{\theta}, \sigma^2 I)$, Theorem 3.1 holds with $\lambda = \frac{\underset{\sim}{\theta}'\underset{\sim}{\theta}}{2}$ replaced by $\underset{\sim}{\theta}'\underset{\sim}{\theta}/2\sigma^2$.

Before we give a proof of Theorem 3.1, we will give one short example illustrating the use of Theorem 3.1.

Example 1. Let $\underset{\sim}{\theta}$ have the t prior with density proportional to $\frac{1}{\left(1+\frac{\underset{\sim}{\theta}'\underset{\sim}{\theta}}{m\tau^2}\right)^{\frac{m+p}{2}}}$. In this case, it is well known that G is absolutely continuous with density g proportional to $e^{-\frac{\gamma}{2\eta}} \cdot (\frac{1}{\eta})^{\frac{\beta-p}{2}+1}$, where as before $\gamma = \frac{m\tau^2}{\sigma^2} = m\tau^2$ and $\beta = m+p$. Consequently, H is also absolutely continuous with density h proportional to $e^{-\frac{\gamma}{2}u} u^{\frac{\beta-p}{2}-1}$. Direct computation now yields that under the density $e^{-\lambda u} u^{\frac{p}{2}} h(u) / \int e^{-\lambda u} u^{\frac{p}{2}} h(u) du$,

$$E_\lambda(u) = \frac{\beta}{2(\lambda + \frac{\gamma}{2})}$$
$$\text{and } \mathrm{Var}_\lambda(u) = \frac{\beta}{2(\lambda + \frac{\gamma}{2})^2}.$$

Thus, the posterior is starunimodal for all $\underset{\sim}{X}$ if and only if

$$\frac{\lambda\beta}{(\lambda + \frac{\gamma}{2})^2} - \frac{\beta}{2(\lambda + \frac{\gamma}{2})} - 1 \leq 0 \qquad \forall\, \lambda > 0$$

which, on multiplying by $(\lambda + \frac{\gamma}{2})^2$ on both sides reduces to the quadratic inequality

$$\lambda^2 + (\gamma - \frac{\beta}{2})\lambda + \frac{\gamma^2}{4} + \frac{\gamma\beta}{4} \geq 0 \qquad \forall\, \lambda > 0. \tag{3.2}$$

It is now easy to argue that (3.2) holds if and only if the discriminant $(\gamma - \frac{\beta}{2})^2 - (\gamma^2 + \gamma\beta)$ is negative, which is equivalent to $\beta < 8\gamma$, i.e., $m\tau^2 > \frac{m+p}{8}$, precisely the condition obtained earlier in Corollary 1.

Proof of Theorem 3.1: It follows from the proof of Theorem 2.1 that the posterior is starunimodal for all $\underset{\sim}{X}$ if and only if $1 + 2\phi_1'(||\underset{\sim}{\theta}||^2) + 4||\underset{\sim}{\theta}||^2\phi_1''(||\underset{\sim}{\theta}||^2) \geq 0$ for all $\underset{\sim}{\theta}$, where $\phi_1(||\underset{\sim}{\theta}||^2) = -\log \pi(||\underset{\sim}{\theta}||^2)$. Denoting $||\underset{\sim}{\theta}||^2$ by s,

$$\pi(s) = \int e^{-\frac{s}{2}u} u^{\frac{p}{2}} dH(u);$$

therefore, by direct computation,

$$1 + 2\phi_1'(s) + 4s\phi_1''(s) \geq 0$$

$$\Longleftrightarrow 1 + \frac{\int e^{-\frac{s}{2}u} u^{\frac{p}{2}+1} dH(u)}{\int e^{-\frac{s}{2}u} u^{\frac{p}{2}} dH(u)} - s\left\{\frac{\int e^{-\frac{s}{2}u} u^{\frac{p}{2}+2} dH(u)}{\int e^{-\frac{s}{2}u} u^{\frac{p}{2}} dH(u)} - \left(\frac{\int e^{-\frac{s}{2}u} u^{\frac{p}{2}+1} dH(u)}{\int e^{-\frac{s}{2}u} u^{\frac{p}{2}} dH(u)}\right)^2\right\} \geq 0,$$

which is the desired result on using $\lambda = \frac{s}{2}$.

Remark: The main message of Theorem 3.1 seems to be that if the variance of u is small compared to its mean under $p(u|\lambda)$, then the posterior is starunimodal for all $\underset{\sim}{X}$. If the mixing distribution G has this property itself, then it will often carry over to $p(u|\lambda)$. For example, if $\pi(\cdot)$ is a normal density (i.e., G is a point mass), then the posterior is clearly starunimodal for every $\underset{\sim}{X}$.

4. <u>Construction of credible regions and their efficiency</u>. In this section, we exploit the properties of starunimodality derived in Section 2 to construct credible regions for $\underset{\sim}{\theta}$ under t priors. We first need the following definition and a lemma.

<u>Definition</u> Let $S \subseteq \mathbb{R}^p$ be any starshaped set, starshaped about $\underset{\sim}{0}$. The Minkowski functional of the set S is defined as

$$\pi_S(\underset{\sim}{v}) = \inf\{a > 0 : \underset{\sim}{v} \varepsilon aS\}.$$

<u>Lemma 4.1</u> Let $\underset{\sim}{V}$ have a starunimodal distribution on $\mathbb{R}^p$ with mode at $\underset{\sim}{0}$. Let $S \subseteq \mathbb{R}^p$ be any starshaped set containing $\underset{\sim}{0}$. Let $a > 0$ be any fixed number and let $r > 0$ be such that $E(W^r) < \infty$ where $W^{\frac{1}{p}}$ denotes the Minkowski functional of the set S. Then

$$P(\underset{\sim}{V} \varepsilon aS) \geq 1 - \left(\frac{r}{r+1}\right)^r \cdot \frac{E(W^r)}{a^{pr}}.$$

Proof: First note that since $\underset{\sim}{V}$ is starunimodal with mode at $\underset{\sim}{0}$, $\underset{\sim}{V}$ admits the representation

$$\underset{\sim}{V} = U^{\frac{1}{p}} \underset{\sim}{Z},$$

where $U \sim u[0,1]$, and $U, \underset{\sim}{Z}$ are independent (see page 40, Dharmadhikari and Joag-Dev (1988)). Also, it follows from the definition of $\pi_S(\underset{\sim}{V})$ and the fact that S is starshaped that $\pi_S(\underset{\sim}{V})$ is homogeneous of degree 1, i.e., for $\alpha > 0, \pi_S(\alpha \underset{\sim}{v}) = \alpha \pi_S(\underset{\sim}{v})$.

$\therefore$ with probability 1,

$$\begin{aligned} \pi_S(\underset{\sim}{V}) &= \pi_S(U^{\frac{1}{p}} \underset{\sim}{Z}) = U^{\frac{1}{p}} \pi_S(\underset{\sim}{Z}) \\ \Longrightarrow W &= (\pi_S(\underset{\sim}{V}))^p = U(\pi_S(\underset{\sim}{Z}))^p. \end{aligned} \tag{4.1}$$

Since $U, \underset{\sim}{Z}$ are independent, so are U and $(\pi_S(\underset{\sim}{Z}))^p$ and therefore, by Khintchine (1938) (also see page 6, Dharmadhikari and Joag-Dev (1988)), W is a scalar-valued unimodal random variable with mode at 0. Since for any $a > 0$,

$$\pi_S(\underset{\sim}{V}) < a \Longrightarrow \underset{\sim}{V} \varepsilon aS,$$

we now have,

$$\begin{aligned} P(\underset{\sim}{V} \varepsilon aS) &\geq P(\pi_S(\underset{\sim}{V}) < a) \\ &= P(W < a^p) \\ &\geq 1 - \left(\frac{r}{r+1}\right)^r \frac{E(W^r)}{a^{pr}} \end{aligned}$$

by Theorem 1.11 in Dharmadhikari and Joag-Dev (1988). This proves the theorem.

<u>Corollary 2</u> For any fixed $0 < \alpha < 1$, let

$$\underline{a} = \inf_{r>0} \left(\frac{r}{(r+1)\alpha^{1/r}} ||W||_r \right)^{1/p},$$

where $||W||_r = (E(W^r))^{1/r}$. Then,

$$P(\underset{\sim}{V} \varepsilon \underline{a} S) \geq 1 - \alpha.$$

Proof: Given any $r > 0, 0 < \alpha < 1$, if we let

$$a = \left(\frac{r}{(r+1)\alpha^{1/r}} ||W||_r \right)^{1/p} \tag{4.2}$$

then from Theorem 4.1 we have,

$$P(\underset{\sim}{V} \varepsilon aS) \geq 1 - \alpha, \tag{4.3}$$

from which the corollary follows by using the Dominated Convergence Theorem on noting that $aS \downarrow \underline{a}S$ as $a \downarrow \underline{a}$.

We are now in a position to state the principal result of this section.

<u>Theorem 4.2</u> Let $\underset{\sim}{X} \sim N(\underset{\sim}{\theta}, \sigma^2 I)$ and $\underset{\sim}{\theta} \sim t(m, \underset{\sim}{0}, \tau^2 I)$. Given m, σ^2, τ^2, suppose $\underset{\sim}{X} \varepsilon S_{2,m}$ where $S_{2,m}$ is defined in (2.16). Let $\underset{\sim}{\nu} = \underset{\sim}{\nu}(\underset{\sim}{X}) = a^* \underset{\sim}{X}$ be the mode of the posterior distribution of $\underset{\sim}{\theta}$ given $\underset{\sim}{X}$ where a^* is the unique root of (2.9). Then, for any starshaped set S, starshaped about $\underset{\sim}{0}$, and any $0 < \alpha < 1$,

$$P[\underset{\sim}{\theta} - \underset{\sim}{\nu} \varepsilon \underline{a} S | \underset{\sim}{X}] \geq 1 - \alpha,$$

where $\underline{a}$ is as in Corollary 2, with $\underset{\sim}{V} = \underset{\sim}{\theta} - \underset{\sim}{\nu}$.

Proof: Use Corollary 2 and the fact that for $\underset{\sim}{X} \varepsilon S_{2,m}$, the posterior of $\underset{\sim}{\theta}$ is starunimodal with mode at $\underset{\sim}{\nu} = a^* \underset{\sim}{X}$.

Remark: Since S is any starshaped set containing $\underset{\sim}{0}$, Theorem 4.2 provides a great amount of flexibility in obtaining Bayesian credible regions for $\underset{\sim}{\theta}$ with a guaranteed posterior probability of $1 - \alpha$. However, the choice of S must necessarily be dictated by the convenience in computing $||W||_r$. It turns out that for natural starshaped sets, the Minkowski functional $\pi_S(\cdot)$ is easy to calculate in a closed form. This is very useful for later calculating $||W||_r$. For example, if S is an ellipsoid

$$S = \{\underset{\sim}{\theta} : \underset{\sim}{\theta}' A^{-1} \underset{\sim}{\theta} \leq 1\},$$

then $\pi_S(\underset{\sim}{\theta}) = \sqrt{\underset{\sim}{\theta}' A^{-1} \underset{\sim}{\theta}}$. If, on the other hand, S is a symmetric hyperrectangle

$$S = \{\underset{\sim}{\theta} : |\theta_i| \leq b_i\},$$

then $\pi_S(\underset{\sim}{\theta}) = \max\limits_{1 \leq i \leq p} \frac{|\theta_i|}{b_i}$.

Such simple closed form expressions for $\pi_S(\cdot)$ make it easier to compute $||W||_r$. Another concern here is that the use of Chebyshev-type inequalities in Lemma 4.1 can make the credible regions of Theorem 4.2 inefficient in the sense that it may be possible to reduce the size of the credible region drastically without making the posterior probability less than $1-\alpha$. We show substantial evidence below that by using $\underline{a}$ instead of 'a' given by (4.2) with a fixed r (such as $r = 1$ or 2 say), we can make the credible regions quite efficient. We therefore recommend that as a general practice, a few different values of r (say integral values $r = 1, 2, \ldots$ etc.) should be tried to find out which r gives a small value of 'a'. Our methods then produce reasonably efficient credible regions under robust flat tailed priors with a guaranteed posterior probability of $1-\alpha$ and it is important to note that in the process we do not need to compute anything more than a few posterior moments $||W||_r$ for a few values of $r > 0$. The posterior moments can be evaluated by simulation from the posterior distribution; that is what we do below. But, alternatively, one can often evaluate the posterior moments $E(W^r)$ by doing a series of only one dimensional numerical integrals provided the set S is a nice starshaped set, such as an ellipsoid or a hyperrectangle. We comment here that $\underline{a}$ will usually correspond to a fractional $r > 0$. The simulation methods can be applied equally well by first choosing a near optimum integral value of r and then taking a grid of fractional values near this integral r. This will result in greater efficiency.

Example We take $\underset{\sim}{X} \sim N(\underset{\sim}{\theta}, \sigma^2 I)$ and $\theta \sim t(m, \underset{\sim}{0}, \tau^2 I)$. Two different values of p are considered, $p = 2, 3$, and m is taken to be 6. All calculations given below are for $\alpha = .05$ and for elliptical credible regions $S_a = \{\underset{\sim}{\theta} : (\underset{\sim}{\theta} - \underset{\sim}{\nu})' A^{-1} (\underset{\sim}{\theta} - \underset{\sim}{\nu}) \leq a^2\}$, where A is taken to be the posterior covariance matrix. See Berger (1985), Tierney and Kadane (1984), Johnstone and Velleman (1984) etc for methods to approximate (posterior) moments. We show below the minimum value of 'a' for $r = 1, 2, 3, 4, 5$. This is denoted by a_0, and the r to which it corresponds is denoted by r_0. Thus the credible region with guaranteed posterior probability of .95 we recommend is

$$S_0 = \{\underset{\sim}{\theta} = (\underset{\sim}{\theta} - \underset{\sim}{\nu})' A^{-1} (\underset{\sim}{\theta} - \underset{\sim}{\nu}) \leq a_0^2\}.$$

Efficiency is calculated as follows: the ellipsoid S_0 is continuously shrunk by decreasing 'a' further from a_0 until for the simulated sample a posterior probability of .95 is attained. The value of 'a' for which this happens is called $a_{\min}$ and the efficiency is measured by $\frac{a_{\min}}{a_0}$. Thus, for example, $\frac{a_{\min}}{a_0} = .8$ will mean that the length of each axis of S_0 can be reduced 20 percent without lowering the posterior probability below .95. For the sake of completeness, we also give the actual posterior probability of S_0 (notice .95 is a lower bound). This is denoted by P_0. Different values of σ^2, τ^2 and $\underset{\sim}{X}'\underset{\sim}{X}$ are tried. We choose $\underset{\sim}{X}'\underset{\sim}{X} = \eta + \delta v$, where η, v stand for the marginal mean and standard deviation of $\underset{\sim}{X}'\underset{\sim}{X}$ for a specific combination of m, σ^2 and τ^2 and δ is allowed to vary. Simulation sizes vary between 500 and 5,000.

Table 1: Table of efficiency of credible regions

p	σ^2	τ^2	δ	r_0	a_0	P_0	$a_{\min}$	$a_{\min}/a_0$
2	.25	1	.1	2	3.34565	.98100	2.67582	.79979
3	.25	1	-.5	1	5.47449	.99700	2.79106	.50982
2	1	1	.1	3	2.98927	.98400	2.44745	.81874
2	1	1	.5	3	3.22387	.98500	2.71662	.84266
2	1	1	1	2	3.33021	.99200	2.55734	.76792
2	1	1	2	2	4.31434	.99900	2.55147	.59139
3	1	1	.1	3	3.61290	.99200	2.79821	.77451
3	1	1	.5	2	4.27755	.99600	2.86880	.67066
2	4	1	.1	3	2.96171	.98800	2.52982	.85418

In general, the credible regions appear to be more efficient for $p = 2$ and for smaller values of $|\delta|$, i.e., when $\underset{\sim}{X}'\underset{\sim}{X}$ is close to its marginal mean. The credible regions seem to be the least efficient when $\frac{\sigma^2}{\tau^2}$ is small and p is large and $|\delta|$ is also large. However, more computation is needed before something definite can be asserted with complete confidence.

5. Concluding remarks. Shape behaviors of the posterior are always theoretically interesting. In this article we describe how knowledge of appropriate forms of unimodality is also practically useful in deriving reasonably efficient credible sets with a guaranteed posterior probability. We have dealt with only the symmetric cases here, i.e., the variance-covariances matrices of $\underset{\sim}{X}$ as well as $\underset{\sim}{\theta}$ are assumed to be proportional to the identity matrix. Many of our results and methods should be applicable to the case of general known covariance matrices as well, although the posterior mode will not, in such cases, be on the line joining the data and the prior mean. Point estimation of $\underset{\sim}{\theta}$ using the poste-

rior mode is also of interest. Risk behavior of the posterior mode and comparison to the posterior mean are both important issues. A convenience in using the mode as opposed to the mean is that the mean needs to be approximated by numerical methods but the mode, as we describe in the paper, has a closed form expression. This allows further study of its properties. These issues will be considered elsewhere. We hope that our results on estimation of a normal mean with flat tailed priors will be useful in robust estimation problems.

Acknowledgement: Herman Rubin actively helped with the simulation. Jim Berger and Bill Studden made many useful suggestions. They all have our deepest appreciation.

References

1. Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York.
2. DasGupta, A. (1988). Volume and diameter minimizing confidence sets in Bayes and classical problems (tentatively accepted by the *Ann. Stat.*).
3. Dawid, A.P. (1973). Posterior expectations for large observations, *Biometrika*, **60**, 664–666.
4. Dharmadhikari, S. and Joag-Dev, K. (1988). *Unimodality, convexity, and applications.* Academic Press, New York.
5. Dickey, J.M. (1976). Approximate posterior distributions. *J. Amer. Statist. Assoc.*, **71**, 680–689.
6. Fan, T.H. and Berger, J.O. (1989). *Behavior of the posterior distribution and inferences for a normal mean with t prior distributions.* Technical Report #89–26C, Department of Statistics, Purdue University.
7. Hill, B. (1980). *Robust analysis of the random model and weighted least squares regression. In Evaluation of Econometric Models.* Academic Press, New York.
8. Johnstone, I. and Velleman, P. (1985). Efficiency scores, variance decompositions, and Monte Carlo swindles. *J. Amer. Statist. Assoc.*, **80**, 851–862.
9. Khintchine, A.Y. (1938). On unimodal distributions. *Izv. Nauchno-Isled. Inst. Mod. Mech. Tomsk. Gos. Univ.*, **2**, 1–7.
10. Meeden, G. and Isaacson, D. (1977). Approximate behavior of the posterior distribution for a large observation. *Ann. Statist.*, **5**, 899–908.
11. Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1986). *Numerical Recipes, the art of scientific computing*, Cambridge University Press, New York.

12. Tierney, L. and Kadane, J.B. (1986). Accurate approximations for posterior moments and marginals. *J. Amer. Statist. Assoc.*, **81**