Ann. Inst. Statist. Math.
Vol. 45, No. 1, 187–199 (1993)

RANGE OF THE POSTERIOR PROBABILITY OF AN INTERVAL FOR PRIORS WITH UNIMODALITY PRESERVING CONTAMINATIONS

S. SIVAGANESAN

Department of Mathematical Sciences, University of Cincinnati, Old Chemistry Building (ML 25), Cincinnati, OH 45221-0025, U.S.A.

(Received February 21, 1991; revised February 27, 1992)

Abstract. Range of the posterior probability of an interval over the ϵ-contamination class $\Gamma = \{\pi = (1-\epsilon)\pi_0 + \epsilon q : q \in Q\}$ is derived. Here, π_0 is the elicited prior which is assumed unimodal, ϵ is the amount of uncertainty in π_0, and Q is the set of all probability densities q for which $\pi = (1-\epsilon)\pi_0 + \epsilon q$ is unimodal with the same mode as that of π_0. We show that the sup (resp. inf) of the posterior probability of an interval is attained by a prior which is equal to $(1-\epsilon)\pi_0$ except in one interval (resp. two disjoint intervals) where it is constant.

Key words and phrases: Posterior probability, unimodality preserving contaminations, Bayesian robustness.

1. Introduction

1.1 *Background*

Robustness or sensitivity of Bayesian methods with respect to small changes in an elicited prior distribution has been an active research area in the past few years. A brief review, and references where motivation and details on recent developments can be seen are given in Subsection 1.3.

We consider a situation where we observe $X \sim f(x \mid \theta)$, and are interested in the (scalar) parameter θ. In Bayesian analysis one is required to elicit the prior information about θ in terms of a single prior distribution, π_0. This prior distribution is typically chosen so that it has a (mathematically) convenient functional form and conforms to certain prior features that are easy to elicit. It therefore becomes essential to study the sensitivity of any posterior measure of interest to small changes in the elicited prior π_0. The general approach that has recently been taken is the following. A class Γ, consisting of priors that are 'close' to π_0 and have the same easy-to-elicit features, is first constructed. Then, the sensitivity of a posterior measure of interest is investigated by computing its range as the prior varies over Γ. The idea being that when this range is 'small', one can be satisfied that the posterior measure is insensitive to plausible changes in the elicited prior.

1.2 *Statement of the problem*

Here, we assume that π_0 is unimodal with (unique) mode θ_0, and consider the ϵ-contamination class of priors given by

$$(1.1)\quad \Gamma = \{\pi = (1-\epsilon)\pi_0 + \epsilon q : q \text{ is a prob. density such that } \pi \text{ is unimodal with mode } \theta_0 \text{ and } \pi(\theta_0) \le h_0\}.$$

Here, $0 < \epsilon < 1$ can be thought of as the amount of elicitation 'error' or the uncertainty in π_0, and $h_0 \geq (1-\epsilon)\pi_0(\theta_0)$ is a constant. The bound h_0 for $\pi(\theta_0)$ is used to avoid the concentration of all contamination mass at θ_0. The value of h_0 may be chosen so as to satisfy an elicited upper bound on the prior probability of an appropriate interval containing θ_0. Some other choices are given in Berger and Berliner (1986). (Also, specifying a very large value for h_0 will effectively remove this restriction.) Moreover, it turns out that posterior answers are, in general, not particularly sensitive to the choice of h_0. The above class was first studied in Berger and Berliner (1986) where the ML-II prior for this class was derived, and later in Sivaganesan (1989), where the range of the posterior mean was derived. When one believes that the prior distribution is close to π_0 and that it is unimodal with mode θ_0, the above class is a conceptually appealing way of expressing the uncertainty in π_0. For further discussions and motivation for choosing this class, see the articles indicated above.

Clearly, the posterior probability of an interval is a quantity of interest in many situations. In this paper, we show how one can compute the range of the posterior probability of an interval as the prior varies over the class Γ, as given in (1.1).

We have organized the paper as follows. In Section 2, we define certain subclasses of Γ and state certain assumptions that will be used throughout the paper. In Section 3, we give the main results concerning the calculation of the sup and inf of the posterior probability of an interval, and provide some illustrative examples. Finally, in Section 4, we provide the proofs of the main results.

1.3 *History*

The notion that the prior information, usually being vague, can only be expressed by a class of prior distributions (rather than a single prior) was first expressed by Good, e.g., see Good (1983). This was then re-vitalized in Berger (1984), which has an extensive and illuminating discussion on this notion and lays the basic foundation for the "robust Bayesian view". For more on this, also see Walley (1990) and the references therein.

Various types of classes of priors have been proposed and studied. Among them are the ϵ-contamination class, the density band class, and the class specified by quantiles. The ϵ-contamination class was first considered in this context by Huber (1973) where the range of the posterior probability is considered for the class of arbitrary contaminations. Number of variations of this class have later been studied, among others, by Sivaganesan and Berger (1989), Sivaganesan (1989, 1990), Moreno and Cano (1988) and Lavine *et al.* (1991). The density band class was first studied in DeRobertis and Hartigan (1981), and was later extended by Lavine (1991*a*, 1991*b*), DasGupta and Studden (1991), DasGupta (1991) and Bose (1990).

The class specified by quantiles has also received much interest. This was first studied in Berliner and Goel (1986) where the problem of finding the range of the posterior probability of certain intervals was considered. Different versions of this class were later studied by Berger and O'Hagen (1988), O'Hagen and Berger (1988), Moreno and Cano (1988) and Sivaganesan (1990). Among other types of classes that have been studied are the parametric classes, e.g., see Polasek (1985), and the classes specified by distribution bands, e.g., see Basu and DasGupta (1990).

2. Preliminaries

Here, we define two sub-classes, Γ_U and Γ_L, of Γ which we will use in finding the range of the posterior probability. The class Γ_U consists of priors each of which is equal to $(1-\epsilon)\pi_0$ except in some interval B where it is a constant. In Section 3, we will show that the sup of the posterior probability of an interval C is attained by a prior $\pi \in \Gamma_U$ for which $B \subseteq C$. The class Γ_L will be defined likewise, except that it will consist of those distributions that are constant in two disjoint intervals, and, we will show in Section 3 that the inf of the posterior probability of C is attained by a member of Γ_L. Now we define these two classes more explicitly. For simplicity in these definitions, and in the forthcoming calculations, we will assume that π_0 is continuous in its support.

2.1 *Class* Γ_U

Let $\tilde{\pi} \in \Gamma$ be of the form

$$\tilde{\pi}(\theta) = \begin{cases} K & \theta \in B \\ (1-\epsilon)\pi_0(\theta_0) & \theta \notin B, \end{cases} \tag{2.1}$$

for some open interval B and an appropriate constant K. Here, the length of the interval B (given an end point) and the value of K are implicitly defined to satisfy the requirement that $\tilde{\pi} \in \Gamma$. Specifically, K is equal to the value of $(1-\epsilon)\pi_0(\cdot)$ at the end point of B closer to θ_0 when $\theta_0 \notin B$, and when $\theta_0 \in B$, it is equal to h for some $(1-\epsilon)\pi_0(\theta_0) \leq h \leq h_0$. And, the length of the interval B satisfies the condition

$$\int_B [\tilde{\pi}(\theta) - (1-\epsilon)\pi_0(\theta)]d\theta = \epsilon. \tag{2.2}$$

Thus, $\tilde{\pi}$ will take one of the following four forms determined by B.

(i) When $B = (s, t(s))$ with $s > \theta_0$,

$$\tilde{\pi}(\theta) = \begin{cases} (1-\epsilon)\pi_0(s) & \theta \in B \\ (1-\epsilon)\pi_0(\theta) & \theta \notin B, \end{cases}$$

where $t(s)$, for $s > \theta_0$, is determined by (2.2).

(ii) When $B = (s(t), t)$ with $t < \theta_0$,

$$\tilde{\pi}(\theta) = \begin{cases} (1-\epsilon)\pi_0(t) & \theta \in B \\ (1-\epsilon)\pi_0(\theta) & \theta \notin B, \end{cases}$$

where $s(t)$, for $t < \theta_0$, is determined by (2.2).

(iii) When $B = (\theta_0, t(h))$ for $t(h) > \theta_0$ (or $= (t(h), \theta_0)$ for $t(h) > \theta_0$),

$$\tilde{\pi}(\theta) = \begin{cases} h & \theta \in B \\ (1-\epsilon)\pi_0(\theta) & \theta \notin B, \end{cases}$$

where $(1-\epsilon)\pi_0(\theta_0) \le h \le h_0$, and $t(h)$ satisfies (2.2).

(iv) When $B = (s, t)$ for $s < \theta_0 < t$,

$$\tilde{\pi}(\theta) = \begin{cases} h_0 & \theta \in B \\ (1-\epsilon)\pi_0(\theta) & \theta \notin B, \end{cases}$$

where the values of s and t are chosen to satisfy (2.2).

Now, we define $\Gamma_U \subseteq \Gamma$ by

$$\Gamma_U = \{\pi \in \Gamma : \pi \text{ is of the form } \tilde{\pi}\}. \tag{2.3}$$

2.2 *Class* Γ_L

Here, we consider priors $\bar{\pi} \in \Gamma$ which are of the form

$$\bar{\pi}(\theta) = \begin{cases} c_1 & \theta \in B_1 \\ c_2 & \theta \in B_2 \\ (1-\epsilon)\pi_0(\theta) & \text{elsewhere.} \end{cases}$$

Here, $B_i = (a_i, b_i)$ $(i = 1, 2)$, are two disjoint intervals, and c_i $(i = 1, 2)$ are given by

$$c_i = \begin{cases} (1-\epsilon)\pi_0(a_i) & \text{if } a_i > \theta_0 \\ (1-\epsilon)\pi_0(b_i) & \text{if } b_i < \theta_0 \\ h & \text{if } a_i \le \theta_0 \le b_i, \end{cases}$$

for some $(1-\epsilon)\pi_0(\theta_0) \le h \le h_0$.

Observe that the values of the end points a_i, b_i $(i = 1, 2)$ together must satisfy the requirement that $\bar{\pi}$ defined above be in Γ. Specifically, one must have,

$$\sum_{i=1}^{2} c_i(b_i - a_i) - \int_{B_1 U B_2} (1-\epsilon)\pi_0(\theta) d\theta = \epsilon. \tag{2.4}$$

Note that the contamination that yields $\bar{\pi}$ has its mass, ϵ, concentrated over two disjoint intervals B_1 and B_2 in such a way that $\bar{\pi}$ is constant over each of B_1 and B_2 (while remaining unimodal with mode θ_0). Here, we allow the possibility of one interval being empty (or having zero mass for the contamination)—resulting in a $\bar{\pi}$ having the same form as $\tilde{\pi}$ of Subsection 2.1. Note that $\bar{\pi}$, like $\tilde{\pi}$, also has the conceptually simple form of being constant over a certain range, and equal to $(1-\epsilon)\pi_0(\theta)$ elsewhere. Now, we define $\Gamma_L \subseteq \Gamma$ by

$$\Gamma_L = \{\pi \in \Gamma : \pi \text{ is of the form } \bar{\pi}\}. \tag{2.5}$$

2.3 *Notation and assumptions*

Throughout the rest of the article we will be concerned with an interval $C = (a, b)$ in the real line, and assume that the length of this interval is large enough to contain an interval B, as defined in Subsection 2.1 (see equation (2.2)). This assumption is made for ease of presentation; the other case can be treated using similar lines of argument.

We will use the notation $l(\theta)$ to denote the likelihood function (for the data x at hand), and assume that it is unimodal with unique mode, denoted by $\hat{\theta}$. We will also use $P^{\pi}(C)$ to denote the posterior probability of C with respect to prior π, i.e.,

$$P^{\pi}(C) = \frac{\int_C l(\theta)\pi(\theta)d\theta}{\int l(\theta)\pi(\theta)d\theta}, \tag{2.6}$$

and the symbols $\overline{P}(C)$ and $\underline{P}(C)$, respectively, to denote the sup and inf of $P^{\pi}(C)$ taken over $\pi \in \Gamma$ (for Γ as in (1.1)), i.e.,

$$\overline{P}(C) = \sup_{\pi\in\Gamma} P^{\pi}(C) \quad \text{and} \quad \underline{P}(C) = \inf_{\pi\in\Gamma} P^{\pi}(C). \tag{2.7}$$

3. Statement of results and examples

In this section, we give the main results which can be used to calculate the sup and inf of the posterior probability of an interval. We then give an illustrative example with normal prior and normal likelihood.

3.1 *Calculation of* $\overline{P}(C)$

In the following theorem we show that $\overline{P}(C)$ can be obtained by maximizing $P^{\pi}(C)$ over the subclass Γ_U. The proof is given in Section 4.

THEOREM 3.1. *Let $\overline{P}(C)$ and Γ_U be as in* (2.3) *and* (2.7), *and suppose $l(\theta)$ is unimodal. Then,*

$$\overline{P}(C) = \sup_{\pi\in\Gamma_U} P^{\pi}(C).$$

The calculation of $\overline{P}(C)$ is greatly simplified since the maximization over Γ_U can be easily done due to the simplicity of the form of $\pi \in \Gamma_U$ (see Subsection 2.1). This is apparent from the following expressions for $P^{\pi}(C)$. Note that, for $\pi \in \Gamma_U$,

$$P^{\pi}(C) = \frac{A_0 + \int_B I_C(\theta)l(\theta)[K-(1-\epsilon)\pi_0(\theta)]d\theta}{A + \int_B l(\theta)[K-(1-\epsilon)\pi_0(\theta)]d\theta}, \tag{3.1}$$

where $A = (1-\epsilon)m(x \mid \pi_0)/\epsilon$ and $A_0 = AP^{\pi_0}(C)$. For instance, if $B = (s, t(s))$ for some $s > \theta_0$ (see Subsection 2.1), $P^\pi(C)$ can be written as

$$\frac{A_0 + (1-\epsilon)\int_s^{t(s)} I_C(\theta)l(\theta)[\pi_0(s) - \pi_0(\theta)]d\theta}{A + (1-\epsilon)\int_s^{t(s)} l(\theta)[\pi_0(s) - \pi_0(\theta)]d\theta}.$$

The advantage here is that we can write $P^\pi(C)$, for $\pi \in \Gamma_U$, in terms of an end point of the interval B. Thus, the calculation is essentially reduced to maximizing a function of one variable. In the following remarks, which are easy to verify, we give bounds on the end points of the interval $B = (s, t)$ for which the sup is attained. These can be useful in the calculation of $\overline{P}(C)$.

Remarks.

1. When $\theta_0 \in C = (a, b)$, it suffices to consider $a < s < t < b$.
2. When $\theta_0 < a$, it is sufficient to consider $\theta_0 \leq s < t <\leq b$.
3. Similarly, when $\theta_0 > b$, it is sufficient to consider $a < s < t < \theta_0$.

3.2 *Calculation of* $\underline{P}(C)$

It turns out that in order to calculate $\underline{P}(C)$, one only needs to minimize $P^\pi(C)$ over the sub-class Γ_L of Γ. We state this result in the following theorem whose proof is given in Section 4.

THEOREM 3.2. *Let* Γ_L *and* $\underline{P}(C)$ *be as in* (2.5) *and* (2.7), *and suppose that* $l(\theta)$ *is unimodal. Then,*

$$\underline{P}(C) = \inf_{\pi \in \Gamma_L} P^\pi(C).$$

The above result means that the prior $\pi \in \Gamma$ which minimizes $P^\pi(C)$ is of the form $\bar{\pi}$, as given in Subsection 2.2. That is, it is equal to $(1-\epsilon)\pi_0(\cdot)$ except in two (disjoint) intervals (B_1 and B_2) where it is constant.

These intervals will always be on the opposite sides of C and, in most cases, each will either be contiguous with C or partly overlapping C. Thus, for $\bar{\pi} \in \Gamma_L$, we have

$$P^{\bar{\pi}}(C) = \frac{A_0 + \sum_{i=1}^2 \int_{B_i} I_C(\theta)l(\theta)[c_i - (1-\epsilon)\pi_0(\theta)]d\theta}{A + \sum_{i=1}^2 \int_{B_i} l(\theta)[c_i - (1-\epsilon)\pi_0(\theta)]d\theta}, \tag{3.2}$$

where c_i $(i = 1, 2)$ is the constant value of $\bar{\pi}$ over the interval $B_i = (a_i, b_i)$ (see Subsection 2.2). Thus, for $\bar{\pi} \in \Gamma_L$, $P^{\bar{\pi}}(C)$ can be expressed as a function of the four end points of the intervals B_1 and B_2. Note, however, that these end points are subject to the constraint (2.4). Calculation of $\underline{P}(C)$ is thus reduced to minimizing a function of three variables subject to bounds.

Very often, however, one can make more specific statements about the lengths and locations of the intervals B_1 and B_2 in a way that can be very useful in further simplifying the calculation of $\underline{P}(C)$. We outline these in the following remarks, where we let B_1 be the interval on the left of C.

Remarks.

(1) When $\sup_{\theta \leq a} l(\theta)$ is 'much smaller' (or 'much larger') than $\sup_{\theta \geq b} l(\theta)$, B_1 (or B_2) will be empty. For instance, if $B_2 = (s, t)$ minimizes $P^{\pi}(C)$ over $\pi \in \Gamma_L$ for which $B_1 = \phi$ and, $\inf_{B_2} l(\theta) > \sup_{\theta \leq a} l(\theta)$, then this inf is the same as the overall inf $\underline{P}(C)$.

(2) When θ_0 and $\hat{\theta}$ are both in C, it can be easily verified that B_1 and B_2, when both non-empty, will be overlapping C and satisfy $l(a_1) = l(b_2)$.

(3) When $\theta_0 \notin C$ and is smaller (or larger) than $\hat{\theta}$, it can be easily seen that B_1 (or B_2) will be contiguous with C, i.e., b_1 (or a_2) will be equal to a (or b).

Example 1. Consider the canonical normal example where $X \mid \theta \sim N(\theta, 1)$ and $\pi_0 \equiv N(0, 2)$, and let $\epsilon = 0.1$. Suppose now that $x = 1.0$. Then, the 95% π_0-HPD credible region is given by $C = (-.93, 2.27)$. Calculation yields $\overline{P}(C) = .951$ and $\underline{P}(C) = .945$. The sup is attained by $\pi \in \Gamma_U$, for which $B = (0, 1.81)$, and the inf is attained for $\pi \in \Gamma_L$ for which $B_1 = (-1.01, -.79)$ and $B_2 = (1.57, 3.01)$.

When $x = 3.0$ in the above, the 95% π_0-HPD credible region is $C = (.40, 3.60)$. In this case $\overline{P}(C)$ is .963, and is attained by $\pi \in \Gamma_U$ for which $B = (1.80, 3.43)$. Also, minimization of (3.2) with $B_1 = \phi$ yields an inf value of .861 when $B_2 = (2.48, 3.60)$. Hence, using Remark (1) above, we have $\underline{P}(C) = 0.861$.

Example 2. (Berger (1985), p. 147) An intelligence test is given to a child, and the result X is $N(\theta, 100)$, where θ is distributed according to the prior $\pi_0 \equiv N(100, 225)$. It is desired to test $H_0 : \theta \leq 100$ versus $H_1 : \theta > 100$. Suppose that the prior uncertainty is expressed by a class of the form Γ with $\epsilon = 0.1$, and an additional restriction that the median is 100 for each π. This additional restriction can be easily accommodated in the calculations with simple modifications. These calculations yield a range of (.89, .91) for the posterior probability of H_1 (or, (.10, .12) for the Bayes factor) when $x = 115$. Similarly, when $x = 125$, the range for the posterior probability is found to be (.98, .99).

Example 3. Suppose that the life time of an electronic component has an exponential distribution with mean θ (in units of 100 hrs), and that the prior information about θ is elicited by the distribution $\pi_0 \equiv$ Inverse Gamma(9, .01). A sample of 5 components tested has a total life time of 65. Suppose that we are interested in the posterior probability that $\theta > 8$, when the uncertainty in the prior π_0 is expressed by a class of the form Γ (with $\theta_0 = 10$). Calculations using the results of Theorems 3.1 and 3.2 yield a range of (.91, .96) for the posterior probability. The sup is attained by a prior in Γ_U for which $B = (10.6, 14.6)$, and the inf is attained for a prior in Γ_L for which $B_1 = (6.2, 9.3)$ and $B_2 = \phi$.

The choice of h_0 had either little or no effect on the above answers. The extent of the influence of h_0 largely depends on the function for which the posterior expectation is considered. When this function is an indicator function of a set, as in this paper, the effect would be minimal as opposed to those functions which have much more 'variability' (at θ_0).

4. Proofs of the theorems

In the following theorem, we re-state the main result of Theorem 4.2 of Berger and Berliner (1986) using the class Γ_U.

THEOREM 4.1. *When $l(\theta)$ is unimodal,*

$$\sup_{\pi\in\Gamma}\int l(\theta)\pi(\theta)d\theta = \sup_{\tilde{\pi}\in\tilde{\Gamma}}\int l(\theta)\tilde{\pi}(\theta)d\theta.$$

We now give some lemmas that will aid in the proof of the main results.

LEMMA 4.1. *Suppose π_0' is a (sub-)probability density with support (c,d), and is decreasing. For $0<\epsilon'<1$ and $h_0'\geq\pi_0'(c)$,*

$$\Gamma' = \{\pi=\pi_0'+\epsilon' q : q \text{ is a prob. density with support } (c,\infty), \\ \pi(c)\leq h_0' \text{ and } \pi \text{ is decreasing}\}.$$

Also, let $\pi'\in\Gamma'$ maximizes $\int_c^d l(\theta)\pi(\theta)d\theta$ over Γ'. Then, if $l(\theta)$ is unimodal π' is either of the form

$$\pi'(\theta)=\begin{cases}\pi_0'(s) & c<s<\theta<t\\ \pi_0'(\theta) & \theta\notin(s,t),\end{cases} \tag{4.1}$$

where s and t are such that $\pi'\in\Gamma'$, or

$$\pi'(\theta)=\begin{cases}h & c<\theta<t\\ \pi_0'(\theta) & \theta\geq t,\end{cases} \tag{4.2}$$

for some $-\epsilon'<\eta<\epsilon'$ and $t>c$.

PROOF. Follows from Theorem 4.1.

LEMMA 4.2. *Suppose π_0' is a (sub-)probability density with support (c,d), and is unimodal with mode $\theta_0\in(c,d)$. For $0<\epsilon'<1$, let*

$$\Gamma' = \{\pi=\pi_0'+\epsilon' q : q \text{ is a prob. density with support } (c,\infty), \\ \pi(\theta_0)\leq h_0' \text{ and } \pi \text{ is unimodal with mode } \theta_0\}.$$

Also, let $\pi'\in\Gamma'$ maximizes $\int_c^d l(\theta)\pi(\theta)d\theta$ over Γ'. Then, for $c<d\leq\infty$ and $l(\theta)$ decreasing, π' is of the form

$$\pi'(\theta)=\begin{cases}h & c<\theta<c'\\ \pi_0'(\theta) & \theta\geq c',\end{cases} \tag{4.3}$$

where, $h = \sup\{\pi(c) : \pi \in \Gamma'\}$, *and* c' *is chosen to satisfy* $\pi' \in \Gamma'$. *Moreover, when* $-\infty \le c < d$, *and* $l(\theta)$ *is increasing,* π' *is of the form*

$$\pi'(\theta) = \begin{cases} h & d' < \theta < d \\ \pi_0'(\theta) & \theta \le d'. \end{cases} \tag{4.4}$$

Here, $h = \sup\{\pi(d) : \pi \in \Gamma'\}$, *and* d' *is chosen so that* $\pi' \in \Gamma'$.

PROOF. For a given $\pi \in \Gamma'$, there exists c_1 such that $c \le c_1 \le c'$, $\pi'(\theta) \ge \pi(\theta)$ in (c, c_1), and $\pi'(\theta) \le \pi(\theta)$ in (c_1, ∞). It is now easy to show that $\int_c^d l(\theta)\pi(\theta) \le \int_c^d l(\theta)\pi'(\theta)d\theta$, proving (4.3). (4.4) follows similarly.

We observe that it is easy to verify the existence of $\pi \in \Gamma$ which attains $\overline{P}(C)$, let this be $\hat{\pi}$. Similarly we will use π^* to denote that $\pi \in \Gamma$ which attains $\underline{P}(C)$. That $\hat{\pi}$, $\pi^* \in \Gamma$, and are unique is easy to verify under mild conditions. We will, however, omit the details for simplicity.

LEMMA 4.3. (i) *When* $C = (a, b) \subseteq (\theta_0, \infty)$, $\hat{\pi}(\theta)$ *for* $\theta \notin C$ *is of the form*

$$\hat{\pi}(\theta) = \begin{cases} \hat{\pi}(a) & a_1 < \theta < a \\ (1-\epsilon)\pi_0(\theta) & \theta \le a_1 \text{ or } \ge b. \end{cases} \tag{4.5}$$

Here, a_1 *is defined as that number, when it exists, for which* $(1-\epsilon)\pi_0(a_1) = \hat{\pi}(a)$. *Else,* $a_1 = \theta_0$.

(ii) *When* $\theta_0 \in C = (a, b)$, $\hat{\pi}(\theta)$ *for* $\theta \notin C$ *is given by*

$$\hat{\pi}(\theta) = (1-\epsilon)\pi_0(\theta).$$

PROOF. Let $\gamma = \overline{P}(C)$. Then, using (2.6), we have $\int (I_C(\theta) - \gamma)l(\theta)\pi(\theta)d\theta \le 0$ for all $\pi \in \Gamma$. Now, using the above and the fact that $\int (I_C(\theta) - \gamma)l(\theta)\hat{\pi}(\theta)d\theta = 0$, we obtain

$$\int (I_C(\theta) - \gamma)l(\theta)(\pi(\theta) - \hat{\pi}(\theta))d\theta \le 0 \quad \text{for all } \pi \in \Gamma. \tag{4.6}$$

For $n > 0$ sufficiently large, let $\pi_n \in \Gamma$ be defined by

$$\pi_n(\theta) = \begin{cases} (1-\epsilon)\pi_0(\theta) & \theta < a_1,\ b < \theta < n \\ \hat{\pi}(\theta) & a \le \theta \le b \\ \hat{\pi}(a) & a_1 \le \theta < a. \end{cases}$$

Here a_1 is as defined in the statement of the lemma, and $\pi_n(\theta)$, for $\theta \ge b + n$, is defined so that $\pi_n \in \Gamma$. Now, clearly $\pi_n = \hat{\pi}$ on (a, b), and $\pi_n \le \hat{\pi}$ on $(-\infty, n)\backslash(a, b)$. Hence, using (4.6), $\int (I_C(\theta) - \gamma)l(\theta)(\pi_n(\theta) - \hat{\pi}(\theta))d\theta \le 0$ for all $n > 0$, which gives

$$\int_{(-\infty,n)\backslash(a,b)} (-\gamma)l(\theta)(\pi_n(\theta) - \hat{\pi}(\theta))d\theta + \int_{(n,\infty)} (-\gamma)l(\theta)(\pi_n(\theta) - \hat{\pi}(\theta))d\theta \le 0.$$

Now, letting A_n and B_n, respectively, be the first and second terms in the above inequality, we have $A_n \geq 0$ is increasing and $B_n \to 0$ as $n \to \infty$. But, we have $A_n + B_n \leq 0$. This gives, $\lim_{n\to\infty} A_n = 0$, and hence $A_n = 0$ for all $n > 0$. Thus, $\hat{\pi}(\theta) = \pi_n(\theta)$ on $(-\infty, n)\backslash(a,b)$ for all $n > 0$. This proves part (i) of the lemma. The proof of part (ii) is similar, and is omitted.

LEMMA 4.4. *Let $C = (a,b) \subseteq (\theta_0, \infty)$. Then, for $\theta \in C$, π^* is of the form*

$$\pi^*(\theta) = \max\{(1-\epsilon)\pi_0(\theta), \pi^*(b)\}.$$

The proof is similar to that of Lemma 4.3, and hence is omitted.

PROOF OF THEOREM 3.1. Our goal is to show that $\hat{\pi} \in \Gamma_U$. The proof is complicated by the need to consider various cases. For brevity and readability, we give the proof for the case where $C = (a,b) \subseteq (\theta_0, \infty)$ and $l(\theta)$ is modal in C. The proof for the other cases are similar, and are therefore omitted.

Now, using (i) of Lemma 4.3, we obtain that $\hat{\pi}$ satisfies (4.5), i.e., $\hat{\pi}(\theta) = (1-\epsilon)\pi_0(\theta)$ for $\theta \geq b$ and for $\theta \leq a_1$. Moreover, it is clear that $\hat{\pi}$ maximizes $\int_a^b l(\theta)\pi(\theta)d\theta$ over all $\pi \in \Gamma$ for which $\pi = \hat{\pi}$ on C', the complement of C. Thus, letting

$$\pi'(\cdot) = (1-\epsilon)\pi_0(\cdot)I_C(\cdot) \quad \text{and} \quad \epsilon' = \int_a^b (\hat{\pi}(\theta) - (1-\epsilon)\pi_0(\theta))d\theta,$$

we have (using Lemma 4.1) that $\hat{\pi}$ in C is either of the form (4.1) or (4.2), with c replaced by a. Now, let $\hat{\pi}$ in C be of the form (4.1). Then, we must have

$$\hat{\pi}(a) = (1-\epsilon)\pi_0(a), \qquad \text{i.e.,} \qquad a_1 = a \text{ in (4.5)}. \tag{4.7}$$

To show this, suppose that $\hat{\pi}(a) > (1-\epsilon)\pi_0(a)$, and let

$$\pi_1(\theta) = \begin{cases} (1-\epsilon)\pi_0(\theta) & \text{for } \theta \notin C \\ \hat{\pi}(\theta) & \text{for } \theta \in C. \end{cases}$$

Then,

$$\begin{aligned} 0 &= \int (I_C(\theta) - \gamma)l(\theta)\hat{\pi}(\theta)d\theta \\ &= \int_{a_1}^a (-\gamma)l(\theta)\hat{\pi}(\theta)d\theta + \int_a^b (1-\gamma)l(\theta)\hat{\pi}(\theta)d\theta + \int_{(a_1,b)'} (-\gamma)(1-\epsilon)l(\theta)\pi_0(\theta)d\theta \\ &< \int_{a_1}^a (-\gamma)l(\theta)\pi_1(\theta)d\theta + \int_a^b (1-\gamma)l(\theta)\pi_1(\theta)d\theta + \int_{(a_1,b)'} (1-\gamma)l(\theta)\pi_1(\theta)d\theta \\ &= \int (I_C(\theta) - \gamma)l(\theta)\pi_1(\theta)d\theta. \end{aligned}$$

Thus, $P^{\pi_1}(C) > \gamma$, which is a contradiction since, as is easy to verify, π_1 is either in Γ, or is the limit of a sequence of π's in Γ. This proves (4.7). Thus, when $\hat{\pi}$ is of the form (4.1), we get, using (4.5) and (4.7), that $\hat{\pi} \in \Gamma_U$. Now, let $\hat{\pi}$ be of the form (4.2). Then, using (4.5), we have

$$\hat{\pi}(\theta) = \begin{cases} h & \text{for } a < \theta < t \\ (1-\epsilon)\pi_0(a_1) & \text{for } a_1 < \theta < a \\ (1-\epsilon)\pi_0(\theta) & \text{for } \theta < a_1,\, \theta > t, \end{cases}$$

for some $t \in (a, b)$ and $h \leq (1-\epsilon)\pi_0(a_1)$. Now, if $h < (1-\epsilon)\pi_0(a_1)$, one can, as before, construct a π_1 such that $P^{\pi_1}(C) > \gamma$. Thus $h = (1-\epsilon)\pi_0(a_1)$. Hence,

$$\hat{\pi}(\theta) = \begin{cases} (1-\epsilon)\pi_0(a_1) & \text{for } a_1 \leq \theta \leq t \\ (1-\epsilon)\pi_0(\theta) & \text{otherwise}, \end{cases}$$

concluding the proof that $\hat{\pi} \in \Gamma_U$. □

PROOF OF THEOREM 3.2. This proof also requires the consideration of various cases. To keep the presentation simple and easy to read, we consider only the case where $C = (a, b) \subseteq (\theta_0, \infty)$ and $l(\theta)$ is modal in C. The proofs for the other cases follow similar lines of argument. Our goal is to show that $\pi^* \in \Gamma_L$. From Lemma 4.4, we have,

$$\pi^*(\theta) = \max\{(1-\epsilon)\pi_0(\theta), \pi^*(b)\} \quad \text{for } \theta \in C. \tag{4.8}$$

Now, it is easy to see that π^* maximizes $\int_b^\infty l(\theta)\pi(\theta)d\theta$ among all $\pi \in \Gamma$ such that $\pi(\theta) = \pi^*(\theta)$ in $(-\infty, b)$. Hence, using (4.3) of Lemma 4.2 (note that $l(\theta)$ is decreasing in (b, ∞)), we have

$$\pi^*(\theta) = \begin{cases} \pi^*(b) & \text{for } b \leq \theta \leq b' \\ (1-\epsilon)\pi_0(\theta) & \text{for } \theta > b', \end{cases} \tag{4.9}$$

for suitable $b' \geq b$. Similarly, we have that π^* maximizes $\int_{-\infty}^a l(\theta)\pi(\theta)d\theta$ among all $\pi \in \Gamma$ for which $\pi(\theta) = \pi^*(\theta)$ for $\theta \in (a, \infty)$. Hence, from (4.4), we have

$$\pi^*(\theta) = \begin{cases} h & \text{for } a' \leq \theta \leq a \\ (1-\epsilon)\pi_0(\theta) & \text{for } \theta < a', \end{cases} \tag{4.10}$$

for some $h \geq (1-\epsilon)\pi_0(a)$ and $a' \leq a$. Thus, combining (4.8), (4.9) and (4.10), we have

$$\pi^*(\theta) = \begin{cases} \pi^*(a) & \text{for } a' \leq \theta \leq a \\ \pi^*(b) & \text{for } b_1 \leq \theta \leq b' \\ (1-\epsilon)\pi_0(\theta) & \text{otherwise}, \end{cases}$$

for some suitable constants $a' \leq a$, $b_1 < b$ and $b' > b$ so that $\pi^* \in \Gamma$. Hence, $\pi^* \in \Gamma_L$, which completes the proof.

Discussion. We have shown how the range of the posterior probability of an interval can be found for priors with unimodality preserving contaminations. The prior that attains the sup has the conceptually simple form of being equal to

$(1-\epsilon)\pi_0$ except in an interval where it is constant. This also was the form of both the ML-II prior, derived by Berger and Berliner (1986), and the prior attaining the sup (inf) of the posterior mean as shown in Sivaganesan (1989). Interestingly, the prior that attains the inf (of the posterior probability of an interval) also has a conceptually simple, yet (possibly) different, form of being equal to $(1-\epsilon)\pi_0$ except in two disjoint intervals where it is constant. These ranges can be of interest in evaluating robustness with respect to prior in interval estimation and in testing hypotheses.

Acknowledgement

The author would like to express his sincere gratitude to an anonymous referee for the very valuable comments.

References

Basu, S. and DasGupta, A. (1990). Bayesian analysis under distribution bands, Tech. Report 90-48, Department of Statistics, Purdue University, Indiana.

Berger, J. (1984). The robust Bayesian viewpoint (with discussion), *Robustness of Bayesian Analysis* (ed. J. Kadane), 63–144, North-Holland, Amsterdam.

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

Berger, J. (1990). Robust Bayesian analysis: sensitivity to the prior, *J. Statist. Plann. Inference*, **25**, 303–328.

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

Berger, J. and O'Hagan, A. (1988). Ranges of posterior probabilities for unimodal priors with specified quantiles, *Bayesian Statistics III* (eds. J. M. Bernardo, M. H. Degroot, D. V. Lindley, and A. F. M. Smith), Oxford University Press.

Berliner, L. M. and Goel, P. (1986). Incorporating partial prior information: range of posterior probabilities, Tech. Report 357, Department of Statistics, Ohio State University.

Bose, S. (1990). Bayesian robustness with shape constrained priors and mixture priors, Ph.D. Thesis, Purdue University, Indiana.

DasGupta, A. (1991). Diameter and volume minimizing confidence sets in Bayes and classical problems, *Ann. Statist.*, **19**, 1225–1243.

DasGupta, A. and Studden, W. (1991). Robust Bayesian experimental designs in normal linear models, *Ann. Statist.*, **19**, 1244–1256.

DeRobertis, L. and Hartigan, J. A. (1981). Bayesian inference using intervals of measures, *Ann. Statist.*, **9**, 235–244.

Good, I. J. (1983). *Good Thinking: The Foundations of Probability and Its Applications*, University of Minnesota Press, Minneapolis.

Huber, P. J. (1973). The use of Choquet capacities in statistics, *Proc. 39th Session ISI*, **45**, 181–188.

Lavine, M. (1991*a*). Sensitivity in Bayesian statistics: the prior and likelihood, *J. Amer. Statist. Assoc.*, **86**, 396–399.

Lavine, M. (1991*b*). An approach to robust Bayesian analysis with multidimensional problems, *J. Amer. Statist. Assoc.*, **86**, 400–403.

Lavine, M., Wasserman, L. and Wolpert, R. (1991). Bayesian inference with specified prior marginals, *J. Amer. Statist. Assoc.*, **86**, 964–971.

Moreno, E. and Cano, J. A. (1988). Robust Bayesian analysis with ϵ-contaminations partially known, Tech. Report, University of Granada, Spain.

O'Hagan, A. and Berger, J. (1988). Ranges of posterior probabilities for quasiunimodal priors with specified quantiles, *J. Amer. Statist. Assoc.*, **83**, 503–508.

Polasek, W. (1985). Sensitivity analysis for general and hierarchical linear regression models, *Bayesian Inference and Decision Techniques with Applications* (eds. P. K. Goel and A. Zellner), North-Holland, Amsterdam.

Sivaganesan, S. (1989). Sensitivity of posterior mean to unimodality preserving contaminations, *Statist. Decisions*, **7**, 77–93.

Sivaganesan, S. (1991). Sensitivity of some posterior summaries when prior is unimodal with specified quantiles, *Canad. J. Statist.*, **19**, 57–66.

Sivaganesan, S. and Berger, J. (1989). Ranges of posterior measures for priors with unimodal contaminations, *Ann. Statist.*, **17**, 868–889.

Walley, P. (1990). *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, London.