RANGES OF POSTERIOR PROBABILITIES FOR QUASI-UNIMODAL PRIORS WITH SPECIFIED QUANTILES

by

Anthony O'Hagan
University of Warwick

James O. Berger
Purdue University

Technical Report #87-11

Department of Statistics
Purdue University

August 1987

Authors' Footnote. Anthony O'Hagan is Senior Lecturer in the Department of Statistics, University of Warwick, Coventry CV4 7AL, England, and James Berger is the Richard M. Brumfield Distinguished Professor of Statistics at Purdue University, West Lafayette, IN 47907. The authors are grateful to Mr. Kun-Liang Lu for performing the numerical work, to the National Science Foundation (U.S.) Grant DMS–8401996 and the Science and Engineering Research Council (U.K.) for support of the research, and to referees for helpful suggestions.

◇

RANGES OF POSTERIOR PROBABILITIES FOR QUASI-UNIMODAL PRIORS WITH SPECIFIED QUANTILES

by

Anthony O'Hagan
University of Warwick

James O. Berger
Purdue University

ABSTRACT

Suppose several quantiles of the prior distribution for θ are specified or, equivalently, the prior probabilities of a partitioning collection of intervals $\{I_i\}$ are given. Suppose, in addition, that the prior distribution is assumed to be unimodal. Rather than selecting a single prior distribution to perform a Bayesian analysis, it is of interest to consider the class of all prior distributions compatible with these inputs. For this class and unimodal likelihood functions, the ranges of the posterior probabilities of the I_i, and of the posterior c.d.f. at the specified prior quantiles, were determined in Berger and O'Hagan (1987). Unfortunately, calculations with this class can be difficult. Here a similar, but much easier to analyze, class of quasi-unimodal prior distributions is considered, and compared with other classes.

Key Words: Prior quantiles; Quasi-unimodal prior; Bayesian robustness; Ranges of Posterior probabilities.

Authors' Footnote. Anthony O'Hagan is Senior Lecturer in the Department of Statistics, University of Warwick, Coventry CV4 7AL, England, and James Berger is the Richard M. Brumfield Distinguished Professor of Statistics at Purdue University, West Lafayette, IN 47907. The authors are grateful to Mr. Kun-Liang Lu for performing the numerical work, to the National Science Foundation (U.S.) Grant DMS–8401996 and the Science and Engineering Research Council (U.K.) for support of the research, and to referees for helpful suggestions.

1. INTRODUCTION

In any practical Bayesian analysis a prior distribution for a continuous parameter cannot be specified in complete detail. To do so would imply infinitely many prior probability judgements. Instead, only a few judgements are actually made. The prior specification is then usually completed by assuming a reasonable, and preferably tractable, form of distribution which fits the judgements that have been made. Furthermore, the judgements that are required are often not prior probabilities but complex functions of them, such as prior means and variances. We consider here a scenario of prior specification which does not require the fitting of a specific distribution to the prior judgements. The user assigns his prior probabilities for the parameter lying in each of the intervals $I_1, I_2, \ldots, I_m$, which are contiguous and partition the real line; this amounts to asserting $m-1$ points of the prior c.d.f. of the parameter. We then assume only that the prior distribution π lies in some set Π of distributions all of which agree with the stated prior probabilities over the intervals $\{I_i\}$. The posterior distribution π^* then lies in a corresponding set Π^*, and we consider what bounds are thereby implied for relevant posterior probabilities.

Example 1. In Martz and Waller (1982), Example 5.1 supposes that two engineers are concerned with the mean life θ of a proposed new industrial engine. The two engineers, A and B, quantify their beliefs about θ in terms of the probabilities given in Table 1 for being in specified intervals. Note that A has substantially more precise beliefs than does B.

Table 1. Specified Prior Probabilities of Intervals

i	Interval I_i	$p_i^A = Pr(\theta \in I_i \mid A)$	$p_i^B = Pr(\theta \in I_i \mid B)$
1	[0, 1000)	0.01	0.15
2	[1000, 2000)	0.04	0.15
3	[2000, 3000)	0.20	0.20
4	[3000, 4000)	0.50	0.20
5	[4000, 5000)	0.15	0.15
6	[5000, ∞)	0.10	0.15

Either of these probability specifications determines a class of prior distributions π, namely

$$\Pi_0 = \{\pi : p_i = Pr(\theta \in I_i) = \int_{I_i} \pi(d\theta) \text{ for all } i\}. \tag{1.1}$$

Berliner and Goel (1986) determine the ranges of the posterior probabilities of the I_i when Π_0 is the assumed class of priors. Earlier, DeRobertis (1978) had considered the related class of priors with $Pr(\theta \in I_i) \geq \gamma_i$ for all i. For both this class and Π_0 the ranges of the posterior probabilities are often quite large because the classes include discrete distributions concentrated at "least favorable" configurations. The engineers in Example 1 might well deny that such discrete priors are plausible reflections of their prior beliefs for a continuous parameter, and might insist that they actually have smooth prior densities. Indeed, it would not be uncommon to encounter the belief that Π is actually unimodal, leading to a class such as

$$\Pi_2 = \{\text{unimodal } \pi : p_i = Pr(\theta \in I_i) \text{ for all i}\}. \tag{1.2}$$

Use of this more realistic class can sharply reduce the variability in posterior answers, as will be demonstrated in Section 3.

Although maximizations and minimizations over Π_2 can be reduced to low dimensional numerical optimization, the algorithm we developed for doing so is extremely complex. (This algorithm is discussed in Berger and O'Hagan (1987).) Here we consider a class, Π_1, of priors which is similar to Π_2, but is much simpler to analyze. Evidence is given that Π_2 can be replaced by this broader class, with little degradation in the answers. Section 2 presents basic notation that will be needed, while Section 3 illustrates the type of answers obtained, through several examples. Section 4 presents the algorithm for analysis with Π_1.

Previous work on finding ranges of posterior quantities for classes of priors mainly dealt with conjugate priors (e.g. Leamer (1978, 1982) and Polasek (1985)). Huber (1973) was the first to explicitly consider a large "nonparametric" class of priors. He determined the range of the posterior probability of a set when Π is an ε-contamination class of priors having the form $\pi = (1-\varepsilon)\pi_0 + \varepsilon q$; here π_0 is a single elicited prior, ε reflects the uncertainty in π_0, and q is a "contamination". Huber considered the case where all contaminations (even discrete) are allowed. Berger and Berliner (1986), Sivaganesan (1986a, 1986b), and Sivaganesan and Berger (1986) considered a variety of generalizations, to different classes of contaminations (e.g. unimodal) and different posterior criteria (e.g. the posterior mean and variance). DeRobertis and Hartigan (1981) considered a large class of priors specified by a type of upper and lower envelope on the prior density, and also find ranges of posterior quantities of interest. Each of these classes is plausible as a model of prior uncertainty. Classes such as Π_0 and Π_2 perhaps have the advantage of being the simplest to understand and elicit. Other work dealing with similar classes of priors includes Bierlien (1967), Kudō (1967), West (1979), Manski (1981), Lambert and Duncan (1981), Cano, Hernández, and Moreno (1985), and Lehn and Rummel (1987). Related analyses for testing situations can be found in Edwards, Lindman, and Savage (1963), Berger and Sellke (1987), Berger and Delampady (1987), Casella and Berger (1987), and Delampady (1986). Other related works include Kadane and Chuang (1978), Wolfenson and Fine (1982), Berger (1984, 1985, 1987), and Walley (1986). These latter works of Berger and Walley also include general review and history of the subject.

2. NOTATION AND THE FORMAL PROBLEM

Prior information is to be stated for an unknown, continuous parameter $\theta \in [a_0, a_m]$ by giving

$$p_i = Pr(\theta \in I_i) = Pr(a_{i-1} \leq \theta \leq a_i)$$

for $i = 1, 2, \ldots, m$. The intervals partition the parameter space $[a_0, a_m]$ and their endpoints $a_1, a_2, \ldots, a_{m-1}$ are arbitrary, possibly even specified by the user. Infinite parameter spaces are included by $a_0 = -\infty$ and/or $a_m = \infty$. It is assumed that there is an underlying prior density $\pi(\theta)$ on $[a_0, a_m]$ constrained by

$$\int_{a_{a-1}}^{a_i} \pi(\theta) d\theta = p_i \qquad (i = 1, 2, \ldots, m). \tag{2.1}$$

Data are obtained, yielding a likelihood function $\ell(\theta)$. We will assume that $\ell(\theta)$ is unimodal, with mode θ_0. For an arbitrary prior density π, the posterior density π^* is, by Bayes theorem,

$$\pi^*(\theta) = \pi(\theta)\ell(\theta) / \int_{a_0}^{a_m} \pi(t)\ell(t) dt. \tag{2.2}$$

Of interest is some set

$$C = \cup_{i \in \Omega} I_i, \tag{2.3}$$

where Ω is some subset of the indices $\{1, 2, \ldots, m\}$. We will seek bounds on $Pr^*(C)$, the posterior probability of the set C. The two cases of most common interest will be $C = I_i$

and $C = \cup_{i \leq n} I_i$; for the latter case, $Pr^*(C)$ is the posterior c.d.f. evaluated at a_i. Sets, C, more general than (2.3) can be considered (see Section 4.2), as could quantities such as the posterior mean, but the analyses then become messier.

For a given C and class of priors, Π, we seek the range of the posterior probability of C as π ranges over Π. Specifically, we will calculate

$$\overline{P}^*_\Pi(C) = \sup_{\pi \in \Pi} Pr^*(C), \tag{2.4}$$

$$\underline{P}^*_\Pi(C) = \inf_{\pi \in \Pi} Pr^*(C). \tag{2.5}$$

We assume that the prior is asserted to be unimodal. This is possible, of course, only when the specified p_i are compatible with unimodality. This will be the case when the constants

$$q_i = p_i/(a_i - a_{i-1}), \qquad i = 1, 2, \ldots, m \tag{2.6}$$

(defined to be zero if the denominator is infinite) satisfy

$$q_1 \leq q_2 \leq \ldots \leq q_k \geq q_{k+1} \geq q_{k+2} \geq \cdots \geq q_m \tag{2.7}$$

for some k. Note that q_i is the uniform density on I_i which has mass p_i. We will henceforth assume that (2.7) holds, and that the prior mode is known to be in I_k.

In addition to Π_0 and Π_2 defined in (1.1) and (1.2), we consider the quasi-unimodal class Π_1 defined as follows. First, let

$$\pi_i(\theta) = \pi(\theta)\chi_{I_i}(\theta), \quad i = 1, 2, \ldots, m, \tag{2.8}$$

where χ_I denotes the indicator function on the set I. Then $\pi = \sum_i \pi_i$, and π_i isolates that part of π which is in I_i. Now unimodality of π implies that

$$\begin{aligned} \pi_i(a_i) &\leq \pi_{i+1}(a_i), \quad i = 1, 2, \ldots, k-1 \\ \pi_i(a_i) &\geq \pi_{i+1}(a_i), \quad i = k+1, k+2, \ldots, m. \end{aligned} \tag{2.9}$$

These constraints at the boundaries "tie together" the intervals when dealing with Π_2, making numerical optimization complex. We shall define Π_1 by slightly relaxing these constraints at the boundary.

Consider, for instance, the interval I_2 when $k > 3$. Unimodality implies that π_2 is increasing. Furthermore, since π_1 must also be increasing it is certainly the case that $\pi_1(a_1) \geq q_1$. Similarly $\pi_3(a_2) \leq q_3$. This suggests replacing the constraints in (2.9) by the weaker constraints

$$\pi_2(a_1) \geq q_1, \quad \pi_2(a_2) \leq q_3.$$

This simplification will greatly reduce the complexity of the numerical optimization, and will be seen to have only a minor effect on the answers. Table 2 results from applying this reasoning to each interval.

Table 2. Bounds on π_i Imposed by Unimodality

Interval number	Conditions
$1 \leq i \leq k-2$	π_i increasing, $\pi_i(a_{i-1}) \geq q_{i-1}$, $\pi_i(a_i) \leq q_{i+1}$
$i = k-1$	π_i increasing, $\pi_i(a_{i-1}) \geq q_{i-1}$
$i = k$	π_k unimodal, $\pi_k(a_{k-1}) \geq q_{k-1}$, $\pi_k(a_k) \geq q_{k+1}$
$i = k+1$	π_i decreasing, $\pi_i(a_i) \geq q_{i+1}$
$k+2 \leq i \leq m$	π_i decreasing, $\pi_i(a_{i-1}) \leq q_{i-1}$, $\pi_i(a_i) \geq q_{i+1}$

Let $\Pi_{(i)}$ consist of all π_i satisfying the conditions for interval i in Table 2, and also satisfying the probability condition (2.1). Define the *quasi-unimodal* class to be

$$\Pi_1 = \{\pi : \quad \pi_i \in \Pi_{(i)}, \quad i = 1, 2, \ldots, m\}. \tag{2.10}$$

Note that the truly unimodal class is

$$\Pi_2 = \{\pi : \quad \pi \in \Pi_1 \text{ and (2.9) holds}\}.$$

Note also that $\Pi_2 \subset \Pi_1 \subset \Pi_0$, so that the posterior ranges $(\underline{P}^*_\Pi(C), \overline{P}^*_\Pi(C))$ will be nested in reverse order for these classes.

3. EXAMPLES

We delay the discussion of the algorithm for analysis with Π_1 until after the presentation of several illustrative examples. In these examples we calculate the range of the posterior probabilities of the intervals I_i, and also of the c.d.f. evaluated at the a_i. The format used for each example is to present the range of the posterior probability of the relevant set C as the interval $(\underline{P}^*_\Pi(C), \overline{P}^*_\Pi(C))$. For each example, the classes Π_0 and Π_2 defined in Section 1, and the class Π_1 defined in Section 2, will be considered.

Example 1 (continued). Data becomes available in the form of two independent life-times which are exponentially distributed with mean θ. The observed life-times are 2000 and 2500 hours, leading to a likelihood function

$$\ell(\theta) = \theta^{-2} \exp(-4500/\theta).$$

Tables 3 and 4 present the ranges of the posterior probabilities of the intervals I_i for engineers A and B, respectively. Tables 5 and 6 present the ranges of the posterior c.d.f.s evaluated at the a_i, for A and B respectively. For engineer B we make the natural assumption that the prior mode is specified to be $a_k = 3000$ (see Table 1) in calculations with Π_1 and Π_2. For engineer A, the probabilities in Table 1 lead to no natural restriction on the prior mode, other than that it be in the interval $[3000, 4000)$. The calculations with Π_1 and Π_2 allow an arbitrary mode in this interval.

Table 3. Posterior Ranges for $C = I_i$, Engineer A

I_i	p_i	Π_0	Π_1	Π_2
[0,1000)	0.01	(0,0.006)	(0.001,0.004)	(0.001,0.004)
[1000,2000)	0.04	(0.019,0.057)	(0.037,0.049)	(0.037,0.049)
[2000,3000)	0.20	(0.214,0.291)	(0.222,0.260)	(0.225,0.260)
[3000,4000)	0.50	(0.476,0.613)	(0.512,0.584)	(0.517,0.584)
[4000,5000)	0.15	(0.106,0.164)	(0.121,0.147)	(0.121,0.147)
[5000,∞)	0.10	(0,0.083)	(0,0.071)	(0,0.071)

Table 4. Posterior Ranges for $C = I_i$, Engineer B

I_i	p_i	Π_0	Π_1	Π_2
[0,1000)	0.15	(0,0.111)	(0.020,0.023)	(0.020,0.023)
[1000,2000)	0.15	(0.088,0.255)	(0.171,0.197)	(0.172,0.197)
[2000,3000)	0.20	(0.235,0.391)	(0.282,0.327)	(0.283,0.327)
[3000,4000)	0.20	(0.197,0.349)	(0.247,0.288)	(0.248,0.288)
[4000,5000)	0.15	(0.125,0.233)	(0.149,0.175)	(0.149,0.175)
[5000,∞)	0.15	(0,0.146)	(0,0.121)	(0,0.121)

Table 5. Posterior Ranges for $C = [0, a_i]$, Engineer A

a_i	Π_0	Π_1	Π_2
1000	(0,0.006)	(0.001,0.004)	(0.001,0.004)
2000	(0.0194,0.062)	(0.038,0.053)	(0.039,0.050)
3000	(0.241,0.341)	(0.262,0.310)	(0.265,0.308)
4000	(0.769,0.886)	(0.794,0.871)	(0.800,0.870)
5000	(0.917,1)	(0.929,1)	(0.929,1)

Table 6. Posterior Ranges for $C = [0, a_i]$, Engineer B

a_i	Π_0	Π_1	Π_2
1000	(0,0.111)	(0.020,0.023)	(0.020,0.023)
2000	(0.096,0.327)	(0.191,0.221)	(0.192,0.221)
3000	(0.388,0.623)	(0.474,0.547)	(0.476,0.547)
4000	(0.659,0.860)	(0.725,0.830)	(0.728,0.830)
5000	(0.854,1)	(0.879,1)	(0.879,1)

Note first that Π_1 and Π_2 yield usefully small ranges of posterior probabilities, in all cases. For instance, if engineer A is willing to assume unimodality as well as the given p_i, then he knows that his posterior probability that $\theta \in [3000, 4000)$ lies between 0.517 and 0.584, while his posterior probability that $\theta \leq 2000$ lies between 0.039 and 0.050. For engineer B, the corresponding ranges are 0.248 to 0.288, and 0.192 to 0.221. These ranges are small enough that the engineers can probably make decisions on this basis, obviating the need for more detailed prior specification.

Note also that Π_1 and Π_2 tend to yield very similar answers, so that the relaxation from unimodality to quasi-unimodality does not seem to matter greatly. On the other hand, Π_0 yields substantially broader intervals (typically 2 to 4 times larger than Π_2, say), indicating that imposing unimodality or quasi-unimodality has a pronounced effect.

A secondary point of interest is the very small interval of posterior probabilities that is obtained for interval I_1 of engineer B when Π_1 and Π_2 is used. The reason serves as a warning about casual assumption of the unimodality constraint. It is easy to see that, when two adjacent intervals have equal q_i (as do I_1 and I_2 for engineer B), then any unimodal prior must have its mode in one of the intervals or be uniform over those intervals. In Table 4 the mode could only be between 2000 and 4000, so all priors in Π_1 and Π_2 are uniform over I_1 and I_2. Thus there may be little variation in the prior (over Π_i) under the unimodality assumption if certain of the adjacent q_i are nearly equal (the central intervals excepted).

Example 2. As a second example, we illustrate the methodology on a standard type of Bayesian example. Suppose subjective elicitation yields the following intervals and corresponding prior probabilities, p_i, for a normal mean θ.

Table 7. Intervals and Prior Probabilities: Normal Example

I_i	$(-\infty,-2)$	(-2,-1)	(-1,0)	(0,1)	(1,2)	$(2,\infty)$
p_i	0.08	0.16	0.26	0.26	0.16	0.08

A "textbook" Bayesian analysis would be to notice that the p_i are a good match to a $\mathcal{N}(0,2)$ (normal, with mean 0 and variance 2) prior distribution. Suppose now that $x = 1.5$ is observed from a $\mathcal{N}(\theta,1)$ experiment. Then usual conjugate prior Bayesian theory would be employed, resulting in a $\mathcal{N}(1,2/3)$ posterior distribution. The resulting posterior probabilities of the I_i are listed in Table 8 as p_i^*.

As an indication of the robustness of the p_i^* to the prior normality assumption, we can calculate the ranges of the posterior probabilities of the I_i for the various classes of priors we are considering. These results are given in Table 8.

Table 8. Posterior Ranges for $C = I_i$, Normal Example

I_i	p_i^*	Π_0	Π_1	Π_2
$(-\infty,-2)$	.0001	(0,0.001)	(0,0.0002)	(0,0.0002)
$(-2,-1)$	.007	(0.001,0.029)	(0.006,0.011)	(0.006,0.011)
$(-1,0)$	.103	(0.024,0.272)	(0.095,0.166)	(0.095,0.166)
$(0,1)$	.390	(0.208,0.600)	(0.320,0.447)	(0.322,0.447)
$(1,2)$	.390	(0.265,0.625)	(0.355,0.475)	(0.357,0.473)
$(2,\infty)$	.110	(0,0.229)	(0,0.156)	(0,0.156)

The p_i^* are reasonably robust, except possibly for p_6^*. Also of interest is the now very dramatic difference between the Π_0 ranges and the ranges for the unimodality classes; their sizes differ by roughly a factor of 4. This provides further evidence of the value of incorporating the unimodality assumption (if subjectively warranted). Of course, these are but two examples, and situations can be constructed where there is little difference between the results for Π_0 and Π_1, but our general experience in looking at a variety of examples is that incorporation of unimodality typically has a substantial effect.

One final comment: the degree of robustness in situations such as Example 2 will typically depend strongly on the data x. In particular, as x gets extreme, so that the likelihood and the prior clash, substantially less robustness will be observed (cf. Berger and Berliner (1986) and Sivaganesan and Berger (1986)).

4. OPTIMIZING OVER THE QUASI-UNIMODAL CLASS

Here we describe the algorithm for maximizing or minimizing $Pr^*(C)$ over $\pi \in \Pi_1$. Note that

$$
\begin{aligned}
Pr^*(C) &= \sum_{i\in\Omega} \omega_i \Big/ \sum_{i=1}^{m} \omega_i \\
&= \left[1 + \frac{\sum_{i\notin\Omega} \omega_i}{\sum_{i\in\Omega} \omega_i}\right]^{-1}, \qquad (4.1)
\end{aligned}
$$

where

$$\omega_i = \int_{a_{i-1}}^{a_i} \pi_i(\theta)\ell(\theta)d\theta.$$

Hence

$$\overline{P}^*_\Pi(C) = \left[1 + \frac{\sum_{i\notin\Omega} \underline{\omega}_i}{\sum_{i\in\Omega} \overline{\omega}_i}\right]^{-1}, \qquad (4.2)$$

$$\underline{P}^*_\Pi(C) = \left[1 + \frac{\sum_{i\notin\Omega} \overline{\omega}_i}{\sum_{i\in\Omega} \underline{\omega}_i}\right]^{-1}, \qquad (4.3)$$

where

$$\overline{\omega}_i = \sup_{\pi_i \in \Pi_{(i)}} \omega_i, \quad \underline{\omega}_i = \inf_{\pi_i \in \Pi_{(i)}} \omega_i.$$

The reduction to independent optimizations over each interval is the great simplification that results from use of Π_1.

For $i \neq k-1$, k or $k+1$, $\overline{\omega}_i$ and $\underline{\omega}_i$ are achieved at π_i which are either uniform (equal to q_i) or single step functions (with step heights of q_{i-1} and q_{i+1}). This is established by the arguments used in the Appendix of Berger and O'Hagan (1987). For $i = k-1$ or $i = k+1$, π_i can be unbounded above, and it is easily seen that $\overline{\omega}_i$ (or $\underline{\omega}_i$) are achieved at step functions which degenerate into a uniform segment plus a point mass at a_{k-1} or a_k, respectively. The interval I_k is another special case, and can yield a point mass at any point in the interval. Following through with arguments similar to those in the Appendix of Berger and O'Hagan (1987), and defining $L(x,y) = \int_x^y \ell(\theta)d\theta$, one finds that in each interval the optimizing prior is one of the following seven types.

Uniform (denoted by U); On I_i,

$$\pi_i(\theta) = q_i. \tag{4.4}$$

Full step (denoted by S): On I_i,

$$\pi_i(\theta) = \begin{cases} q_{i-1} & \text{if } a_{i-1} \le \theta \le s_i \\ q_{i+1} & \text{if } s_i < \theta \le a_i, \end{cases} \tag{4.5}$$

where

$$s_i = (q_{i+1}a_i - q_{i-1}a_{i-1} - p_i)/(q_{i+1} - q_{i-1}). \tag{4.6}$$

Limited step right (denoted by V): On I_i

$$\pi_i(\theta) = \begin{cases} q_{i-1} & \text{if } a_{i-1} \le \theta \le \sigma_i^* \\ f_i^* & \text{if } v_i^* < \theta \le a_i, \end{cases} \tag{4.7}$$

where

$$f_i^* = \min(f_i, q_{i+1}), \tag{4.8}$$

$$f_i = \{p_i - q_{i-1}(v_i - a_{i-1})\}/(a_i - v_i), \tag{4.9}$$

$$v_i^* = \min(v_i, s_i), \tag{4.10}$$

and v_i is the solution in (a_{i-1}, θ_0) to the equation

$$(a_i - v_i)\ell(v_i) = L(v_i, a_i). \tag{4.11}$$

Limited step left (denoted by W): On I_i

$$\pi_i(\theta) = \begin{cases} g_i^* & \text{if } a_{i-1} \le \theta \le w_i^* \\ q_{i+1} & \text{if } w_i^* < \theta \le a_i, \end{cases} \tag{4.12}$$

where

$$g_i^* = \min(g_i, q_{i-1}) \tag{4.13}$$

$$g_i = \{p_i - q_{i+1}(a_i - w_i)\}/(w_i - a_{i-1}), \tag{4.14}$$

$$w_i^* = \max(w_i, s_i) \qquad 4.15)$$

and w_i is the solution in (θ_0, a_i) to the equation

$$(w_i - a_{i-1})\ell(w_i) = L(a_{i-1}, w_i). \tag{4.16}$$

Point mass right (denoted by R): On I_i

$$\pi_i(\theta) = q_{i-1}, \tag{4.17}$$

plus a point mass at $\theta = a_i$ with probability

$$r_i = p_i - q_{i-1}(a_i - a_{i-1}). \tag{4.18}$$

Point mass left (denoted by L): On I_i

$$\pi_i(\theta) = q_{i+1}, \tag{4.19}$$

plus a point mass at $\theta = a_{i-1}$ with probability

$$l_i = p_i - q_{i+1}(a_i - a_{i-1}). \tag{4.20}$$

Point mass center (denoted by C): On I_i

$$\pi_i(\theta) = \begin{cases} q_{i-1} & \text{if } a_{i-1} \leq \theta < \theta_0 \\ q_{i+1} & \text{if } \theta_0 < \theta \leq a_i, \end{cases} \tag{4.21}$$

plus a point mass at $\theta = \theta_0$ with probability

$$c_i = p_i - q_{i-1}(\theta_0 - a_{i-1}) - q_{i+1}(a_i - \theta_0). \tag{4.22}$$

Table 9 below identifies which of the seven types is optimal for each interval. A "Max" interval is one for which $\overline{\omega}_i$ is sought; a "Min" interval is one for which $\underline{\omega}_i$ is sought (see (4.2) and (4.3)). Also, "likelihood form" refers to whether $\ell(\theta)$ is increasing, decreasing, or both (called modal) in the interval. In the modal interval (note that there can be only one), it is necessary to distinguish between three cases, depending on a comparison of the values of the likelihood at the endpoints with the average of the likelihood over the interval.

Table 9. Optimizing Interval Classification

Likelihood Form	Class	Intervals: I_1 to I_{k-2}	I_{k-1}	I_k	I_{k+1}	I_{k+2} to I_m
Increasing	Max	T	R	R	U	U
	Min	U	U	L	L	T
Decreasing	Max	U	U	L	L	T
	Min	T	R	R	U	U
Modal:						
(i) $L(a_{i-1}, a_i) \leq \ell(a_{i-1})(a_i - a_{i-1})$	Max	U	U	C	W	W
	Min	W	R	R	U	U
(ii) $L(a_{i-1}, a_i) \leq \ell(a_i)(a_i - a_{i-1})$	Max	V	V	C	U	U
	Min	U	U	L	L	V
(iii) otherwise	Max	V	V	C	W	W
	Min	W	R	*	L	V

The only case not given by Table 9 is for interval I_k when the likelihood is modal in I_k, both $\ell(a_{i-1})$ and $\ell(a_i)$ are less than $L(a_{i-1}, a_i)/(a_i - a_{i-1})$, and the interval type is "Min". This is shown in Table 9 as a "*" because here the optimal π_k is either type R or type L, but depends on a complicated criterion. The optimum is type L if

$$(q_k - q_{k-1})\ell(a_k) + (q_{k+1} - q_k)\ell(a_{k-1})$$
$$> (q_{k+1} - q_{k-1})L(a_{k-1}, a_k)/(a_k - a_{k-1});$$

otherwise it is type R.

Table 9 together with equations (4.4) to (4.22) explicitly define the optimizing $\pi \in \Pi_1$, and the corresponding posterior density is defined by (2.2). The $\overline{\omega}_i$ and $\underline{\omega}_i$ in (4.2) and (4.3) can even be given explicitely, according to the type of π_i, as follows:

Type U : $\omega_i = q_i L(a_{i-1}, a_i)$;

Type S : $\omega_i = q_{i-1} L(a_{i-1}, s_i) + q_{i+1} L(s_i, a_i)$;

Type V : $\omega_i = q_{i-1} L(a_{i-1}, v_i^*) + f_i^* L(v_i^*, a_i)$;

Type W : $\omega_i = g_i^* L(a_{i-1}, w_i^*) + q_{i+1} L(\omega_i^*, a_i)$;

Type R : $\omega_i = q_{i-1} L(a_{i-1}, a_i) + r_i \ell(a_i)$;

Type L : $\omega_i = q_{i+1} L(a_{i-1}, a_i) + l_i \ell(a_{i-1})$;

Type C : $\omega_i = q_{i-1} L(a_{i-1}, \theta_0) + q_{i+1} L(\theta_0, a_i) + c_i \ell(\theta_0)$.

This explicit solution is much easier to compute than the solution for Π_2 in Berger and O'Hagan (1987). Since $\Pi_1 \supset \Pi_2$, the bounds it provides on $Pr^*(C)$ will in general be wider than those for Π_2, but the numerical examples presented in Section 3 indicated that the difference is typically small. Indeed, the optimal $\pi \in \Pi_2$ may actually lie in Π_1, in which case the bounds will be the same.

4.2 Analysis for Arbitrary Intervals C

If $C = [c_1, c_2]$ with c_1 or c_2 (or both) being unequal to any a_i, a simple modification of the algorithm in Section 4.1 allows for calculation of $\overline{P}^*(C)$ and $\underline{P}^*(C)$: simply create a new partition by including c_1 and c_2 with the $\{a_i\}$.

The only hitch is that, if say I_i is separated into $(a_{i-1}, c_j]$ and $(c_j, a_i]$, one must determine how the mass p_i, allocated to I_i, is to be divided up between the two new intervals. The answer is conceptually clear: give as much mass as possible, subject to maintaining the prescribed unimodality constraints (i.e., ordering of the q_i), to that interval in C (not in C) if $\overline{P}^*(C)$ ($\underline{P}^*(C)$) is desired.

Clearly, this method will generalize to quite arbitrary sets C.

REFERENCES

BERGER, J. (1984), "The Robust Bayesian Viewpoint (With Discussion)", in *Robustness of Bayesian Analyses*, ed. J. Kadane, Amsterdam: North-Holland.

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

BERGER, J. (1987), "Robust Bayesian Analysis: Sensitivity to the Prior," Technical Report #87-10, Department of Statistics, Purdue University.

BERGER, J. and BERLINER, L.M. (1986), "Robust Bayes And Empirical Bayes Analysis With ε-Contaminated Priors", *Annals of Statistics*, **14**, 461-486.

BERGER, J. and DELAMPADY, M. (1987), "Testing Precise Hypotheses," To appear in *Statistical Science.*

BERGER, J. and O'HAGAN, A. (1987), "Ranges of Posterior Probabilities For Unimodal Priors With Specified Quantiles," To appear in *Bayesian Statistics III*, eds. J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith.

BERGER, J., and SELLKE, T. (1987), "Testing a Point Null Hypothesis: The Irreconcilability of Significance Levels and Evidence", *Journal of the American Statistical Association* **82**, 112–122.

BERLINER, L.M., and GOEL, P. (1986), "Incorporating Partial Prior Information: Ranges of Posterior Probabilities", Technical Report No. 357, Department of Statistics, Ohio State University.

BIERLEIN, D. (1967), "Zur Einbeziehung der Erfahrung in Spieltheoretische Modelle," *Z. Oper. Res. Verfahren III*, 29–54.

CANO, J.A., HERNÁNDEZ, A., and MORENO, E. (1985), "Posterior Measure Under Partial Prior Information," *Statistica* **2**, 219–230.

CASELLA, G., and BERGER, R. (1987), "Reconciling Bayesian And Frequentist Evidence in the One-Sided Testing Problem", *Journal of the American Statistical Association*, **82**, 106–111.

DELAMPADY, M. (1986), "Testing a Precise Hypothesis: Interpreting P-Values From a Robust Bayesian Perspective", *Ph.D. Thesis*, Purdue University, West Lafayette.

DE ROBERTIS, L. (1978), "The Use of Partial Prior Knowledge In Bayesian Inference", *Ph.D. Thesis*, Yale University, New Haven.

DE ROBERTIS, L., and HARTIGAN, J.A. (1981), "Bayesian Inference Using Intervals of Measures", *Annals of Statistics*, **1**, 235-244.

EDWARDS, W., LINDMAN, H., and SAVAGE, L.J. (1963), "Bayesian Statistical Inference for Psychological Research," *Psychological Review*, **70**, 193–242.

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

HUBER, P.J. (1973), "The Use of Choquet Capacities in Statistics", *Bulletin of the International Statistical Institute*, **45**, 181-191.

KADANE, J.B., and CHUANG, D.T. (1978), "Stable Decision Problems", *Annals of Statistics*, **6**, 1095-1110.

KUDŌ, H. (1967), "On Partial Prior Information and the Property of Parametric Sufficiency," *Proceedings of Fifth Berkeley Symposium on Statistics and Probability*, **1**, Berkeley: University of California Press.

LAMBERT, D. and DUNCAN, G. T. (1986), "Single-Parameter Inference Based on Partial Prior Information," *Canadian Journal of Statistics* **14**, 297–305.

LEAMER, E.E. (1978), *Specification Searches*, New York: Wiley.

LEAMER, E.E. (1982), "Sets of Posterior Means With Bounded Variance Prior", *Econometrica*, **50**, 725-736.

LEHN, J. and RUMMEL, F. (1987), "Gamma minimax estimation of a binomial probability under squared error loss," *Statistics and Decisions 5*.

MANSKI, C.F. (1981), "Learning and Decision Making When Subjective Probabilities Have Subjective Domains," *Annals of Statistics*, **9**, 59–65.

MARTZ, H.F., and WALLER, R.A. (1982), *Bayesian Reliability Analysis*, New York: Wiley.

POLASEK, W. (1985), "Sensitivity Analysis for General and Hierarchical Linear Regression Models", in *Bayesian Inference and Decision Techniques with Applications*, eds. P.K. Goel and A. Zellner, Amsterdam: North-Holland.

SIVAGANESAN, S. (1986a), "Robust Bayes Analysis with Arbitrary Contaminations", Technical Report No. SMU-DS-TR-198, Department of Statistical Science, Southern Methodist University.

SIVAGANESAN, S. (1986b), "Sensitivity Of The Posterior Mean To Unimodality Preserving Contaminations", Technical Report No. SMU-DS-TR-199, Department of Statistical Science, Southern Methodist University.

SIVAGANESAN, S., and BERGER, J. (1986), "Ranges Of Posterior Measures For Priors With Unimodal Contaminations", Technical Report #86-41, Department of Statistics, Purdue University.

WALLEY, P. (1986), *Rationality and Vagueness*, Manuscript in preparation.

WEST, S. (1979), "Upper and Lower Probability Inferences for the Logistic Function," *Annals of Statistics*, **7**, 490–413.

WOLFENSON, M., and FINE, T.L. (1982), "Bayes-Like Decision Making With Upper And Lower Probabilities", *Journal of the American Statistical Association*, **77**, 80-88.