Sankhyā : The Indian Journal of Statistics
Special Issue on Bayesian Analysis
1998, Volume 60, Series A, Pt. 3, pp. 405-417

ISSUES IN BAYESIAN LOSS ROBUSTNESS

By JACINTO MARTÍN
Universidad Politecnica de Madrid, Madrid
DAVID RÍOS INSUA
Universidad Rey Juan Carlos, Madrid
and
FABRIZIO RUGGERI
CNR-IAMI, Milano

SUMMARY. Bayesian robustness has concentrated mainly on checking sensitivity of a Bayesian analysis with respect to changes in the prior or the model. We deal here with several issues concerning robust Bayesian analysis with respect to the loss function. Stemming from foundational results, we suggest that a main computational objective would be the obtainment of the set of nondominated alternatives. We then discuss a number of structure questions concerning the nondominated set, mainly its existence and relation with the set of Bayes alternatives and discuss procedures to compute the nondominated set. Since this set may be too big to reach a final decision, we mention some problems concerning gathering additional information.

1. **Introduction**

Our concern here is with Bayesian robustness, see Berger (1994), which stems from an appreciation of the potential difficulties in assessing the inputs necessary to conduct a Bayesian analysis and/or the need to check the impact of those inputs on the conclusions of that analysis.

Previous work on robust Bayesian analysis has concentrated mainly on inference problems. As a consequence, efforts have centered on studies of the local and/or global behaviour of a predictive or posterior probability or expected loss, when the prior and/or model varies in a certain class. Checking the influence of the loss function on the conclusions of an analysis is broadly recognised as important, but not thoroughly studied, see various discussions following Berger (1994). The main reason there suggested refers to the difficulty of assessing

AMS (1991) *subject classification.* Primary 62E10; secondary 60E05.
Key words and phrases. Robust Bayesian analysis, classes of loss functions, nondominated alternative, Bayes alternative, $\mathcal{L}$-minimax.

loss functions in applied settings, because of limited information or lack of time, which would explain why statisticians have concentrated on using a fairly limited set of stylised loss functions, including quadratic, absolute value or 0-1 loss functions. In turn, the appropriate attitude in this case would be that since loss functions are difficult to assess we should try to work with a class $\mathcal{L}$ of loss functions.

Similar issues arise in related contexts. For example, Clyde and Chaloner (1995) describe multiobjective design problems where each objective is associated with a utility function, studying the computation of nondominated designs. Frequentists have also worried about the difficulties of obtaining loss functions and, accordingly, worked with classes of them, with emphasis on the concept of $\mathcal{L}$-admissibility, as a minimal requirement for the acceptability of an estimator, see Brown (1975) or Hwang (1985). A popular topic in economic analysis, see Levy (1992), has been stochastic dominance, which refers to choice under risk when the utility function belongs to a given class.

Recently, several authors have paid attention to loss robustness, see e.g. Makov (1994) or Dey *et al.* (1995). Most work in the area has concentrated on extensions of typical sensitivity to the prior studies to the case in which there is imprecision in the loss function. For example, one could compute the range of the posterior loss of an alternative, when the loss function ranges in a class. As we shall argue in Section 2, these straightforward extensions demand some care. On the other hand, by appealing to the foundations of robust Bayesian analysis, we suggest a shift on the emphasis in Bayesian loss robustness computations: various foundations, see e.g. Nau *et al.* (1997), suggest ordering alternatives in a Pareto sense, according to a class of expected losses taken with respect to a class of probability distributions and a class of loss functions. As a consequence, robust Bayesian analysis should devote some effort to the problem of computing nondominated alternatives.

In Section 3, we study a number of structural properties of the nondominated set, mainly its existence and relations with the set of Bayes alternatives. We then study the computation of the nondominated set. We then discuss issues concerning the size of the nondominated set and provide relations between nondominated and $\mathcal{L}$-minimax alternatives, ending up with comments.

2. Basic Notation and Concepts

We start within the standard decision theoretical framework, see e.g. Berger (1985). We have to choose among a set $\mathcal{A}$ of alternatives a. We model uncertainty about states $\theta \in \Theta$ with a prior distribution $P(\theta)$, which, in the presence of information x provided by an experiment with likelihood $l(x|\theta)$, is updated to the posterior $P(\theta|x)$. When we choose alternative a and state θ happens, we obtain a consequence $c(a, \theta) \in \mathcal{C}$, the space of consequences. In many statistical problems, it is assumed that $c(a,\theta) = a - \theta$. We evaluate consequences with a loss function $L(a,\theta)$, and alternatives according to their posterior expected loss $T(L,a) = \int L(a,\theta) dP(\theta|x)$. We suggest as optimal the alternative of minimum posterior expected loss, which we designate Bayes alternative (for loss function L).

We shall depart from this framework assuming that we have a class $\mathcal{L}$ of loss functions, instead of a single one. This could be due to our inability to model exactly the preferences of a decision maker, e.g. because of limited information or lack of time, a desire to check the impact of the loss function over the conclusions of the analysis, or model consensus among several decision makers. We could be tempted to extend standard robust Bayesian procedures to our setting in a straightforward manner. For example, we could compute, for a given alternative, the range of its expected loss when the loss ranges in the class. Large ranges would suggest lack of robustness.

However, these extensions demand some care. Suppose for example that the class of losses includes functions of the type $L(a,\theta) + k$, for $k \in [k_0, k_1]$, which happens e.g. in the so called *universal class*, see Hwang (1985). Note that all those loss functions are strategically equivalent, therefore leading to the same optimal decision, hence having a robust problem. Should we insist in computing the range of posterior expected losses, we would find it to be greater than $k_1 - k_0$, which, if large, may suggest that we lack robustness. A similar problem raises if we consider losses of the form $kL(a,\theta)$, when k ranges in an interval. Related concerns refer to the debate decision robustness vs. loss robustness, see Srinivasan and Kadane, in discussion to Berger (1994). If we insist on using the range of the posterior expected loss as a sensitivity measure, we should then be careful in defining the class of loss functions: in order to avoid this problem, we could fix the loss of the worst consequence to be the upper bound for all functions in the class, and the loss of the best one to be the lower bound for all functions. However, with this condition, we could not use typical loss functions like the quadratic or the absolute value.

One way forward in this problem is to go back to foundations of robust Bayesian analysis, see e.g. Nau *et al.* (1997) or Ríos Insua and Martín (1994a). We are specifically interested in problems in which preferences are modelled with a class $\mathcal{L}$ of functions L and beliefs with a distribution P. Within this incompleteness in preferences context, foundational results suggest that preferences among alternatives will follow a Pareto order with respect to the class of expected losses. To wit, we shall find alternative b at most as preferred as alternative a if, and only if, the posterior expected loss of b is greater than or equal to the posterior expected loss of a, for each loss function L in $\mathcal{L}$. In symbols, we shall write

$$b \preceq a \iff T(L,a) \leq T(L,b), \ \forall L \in \mathcal{L}$$

When $T(L,a) \leq T(L,b)$, $\forall L \in \mathcal{L}$, with strict inequality for one loss function in the class, it will be $b \preceq a$ and $\neg(a \preceq b)$, and we shall write $b \prec a$, saying that a dominates b. Clearly, we would discard alternative b from the analysis. Therefore, the natural solution concept in this context is that of nondominated alternative.

DEFINITION 1. *$a \in \mathcal{A}$ is nondominated if there is no other alternative $b \in \mathcal{A}$ such that $a \prec b$.*

Note that the classical concept of admissibility is based as well on inequalities between performance measures relative to alternatives; in our case, we use posterior expected losses, for various loss functions, whereas in the case of admissibility, we use the risk function, for various priors.

We believe that a computational question of interest in Bayesian loss robustness analysis should be the calculation of the set of nondominated alternatives, much as we are interested in the existence of a Bayes alternative in a conventional Bayesian analysis or, more generally, optimal alternatives in an optimisation problem. The computation of the nondominated set allows us to discard definitely inferior alternatives, those that are dominated, hence reducing the set of alternatives on which the Decision Maker should focus attention.

3. Nondominated Sets: Structural Properties

We discuss first some structural properties of the nondominated set, namely its existence and relations with the set of Bayes alternatives.

3.1 *Existence of nondominated alternatives.* A first issue to consider is when do nondominated alternatives exist. Simple examples show that this set may be empty. Suppose e.g. that $\mathcal{A} = (0,1] \times (0,1]$, with a generic alternative defined by $a = (a_1, a_2)$. Suppose also that $\mathcal{L} = \{L_1, L_2\}$, with $L_1(a,\theta) = a_1$ and $L_2(a,\theta) = a_2$, with a generic state space Θ. Then, whatever the prior and the model are, which in fact are irrelevant in this case, for any alternative a there is another alternative a' such that $a \prec a'$, so that the nondominated set is empty.

The following results show important statistical problems in which the nondominated set is nonempty. The first one refers to problems in which the set of alternatives is finite, a relevant example being that of multiple hypothesis testing, see e.g. Berger (1985). The proof follows immediately from the transitivity of the dominance property.

PROPOSITION 1. *If the set $\mathcal{A}$ of alternatives is finite, the nondominated set is not empty.*

A more general result of similar proof is

PROPOSITION 2. *If for one loss function in the class $\mathcal{L}$ the set of Bayes alternatives is nonempty and finite, then there exists, at least, one nondominated alternative.*

By repeated application of Proposition 2 we may find a representative set of nondominated alternatives. In some cases, the process ends in a finite number of steps. An important case is that in which the class is finitely generated, that is there is a set $\{L_1, L_2, ..., L_n\} \subset \mathcal{L}$ such that $\forall L \in \mathcal{L}$, $L = \sum_{i=1}^n w_i L_i$, $\sum w_i = 1$, $w_i \geq 0$. An example is that of multiobjective designs, in which we have several loss functions and we find nondominated designs looking for designs which are optimal for convex combinations of the original loss functions, see Clyde and Chaloner (1996). We assume an appropriate topology for Proposition 3.

PROPOSITION 3. *If the set $\mathcal{A}$ of alternatives is compact and the class $\mathcal{L}$ is generated by a finite number of loss functions continuous in $\mathcal{A}$, uniformly with respect to θ, then the nondominated set is not empty.*

PROOF. Let $L_1, ..., L_n$ be the loss functions generating $\mathcal{L}$. Fix L_1. Since L_1 is continuous in $\mathcal{A}$, uniformly in θ, $T(L_1, .)$ will also be continuous in $\mathcal{A}$. To wit, for each ϵ there is δ such that if $|b - a| < \delta$, $|L_1(b, \theta) - L_1(a, \theta)| < \epsilon, \forall \theta$. Then, $|T(L_1, a) - T(L_1, b)| \leq \int |L_1(b, \theta) - L_1(a, \theta)| d\pi(\theta|x) \leq \epsilon$. Since $T(L_1, .)$ is continuous in $\mathcal{A}$, which is compact, the corresponding set of minimisers, i.e. the set B_1 of Bayes actions, will be compact, as may be easily seen. If B_1 is finite the result follows immediately from Proposition 2, one of the alternatives in B_1 being nondominated. If not, we may find a compact set $B_2 \subset B_1$ that minimises the expected loss under L_2 in B_1. If the set is finite then one of the alternatives in B_2 is nondominated. Otherwise, repeat with L_3 the procedure, eventually, reaching B_n, which will be a set of nondominated alternatives. □

Similar results may be obtained, e.g. if $L_1(a, \theta)$ is continuous in a for each θ and the density of $P(\theta|x)$ with respect to the underlying measure is bounded above.

3.2 *Nondominated and Bayes alternatives.* Recall that, from a foundational point of view, in cases of imprecision in the loss function we should look for nondominated alternatives. When there is precision in the loss, nondominated and Bayes alternatives coincide, and we would like to relate them in general cases. A simple case is that in which there is a unique alternative a which is Bayes for all $L \in \mathcal{L}$. Then it is the unique nondominated alternative.

The following example shows that the sets of nondominated and Bayes alternatives can be different.

EXAMPLE 1. Let

$$\mathcal{A} = \left\{(1, x),\ x \in \left(\frac{1}{2}, 1\right)\right\} \cup \left\{(x, 0),\ x \in \left(0, \frac{1}{2}\right)\right\} \cup \left\{\left(\frac{3}{4}, \frac{1}{4}\right)\right\},$$

and $\mathcal{L} = \{L_1, L_2\}$, with $L_1(a, \theta) = -a_1$ and $L_2(a, \theta) = a_2$. Then, $(3/4, 1/4)$ is the unique nondominated alternative. The Bayes alternatives for L_1 are $\{(1, x),\ x \in (1/2, 1)\}$. The Bayes alternatives for L_2 are $\{(x, 0),\ x \in (0, 1/2)\}$. □

An important result relating the sets of Bayes and nondominated alternatives is given by the following corollary of Proposition 1.

COROLLARY 1. *If any loss function in the class $\mathcal{L}$ has a unique Bayes alternative, then the set of Bayes alternatives is contained in the set of nondominated ones.*

We will see later (Proposition 7) that the two sets coincide for a quite large class of loss functions. Note, though, that both sets can be different even if there is a unique Bayes solution for each loss function in the class.

EXAMPLE 2. Let $\mathcal{A} = \{(a_1, a_2) : a_1^2 + a_2^2 \geq 1, 0 \leq a_1, a_2 \leq 1\}$. Suppose also that $\mathcal{L} = \{L_1, L_2\}$, with $L_1(a, \theta) = a_1$ and $L_2(a, \theta) = a_2$, with a generic state space Θ. Then, whatever the prior and the model are, the set of nondominated alternatives is $\mathcal{A} = \{(a_1, a_2) : a_1^2 + a_2^2 = 1, 0 \leq a_1, a_2 \leq 1\}$, whereas $(1, 0)$ is Bayes for L_1 and $(0, 1)$ is Bayes for L_2. □

Under appropriate conditions both concepts are intimately related. For example, following Proposition 2, if the set of Bayes alternatives is finite for a loss function, one of those alternatives is nondominated. As a consequence of Corollary 1, if there is a unique Bayes alternative for $L \in \mathcal{L}$, then it is nondominated. Reciprocal results are not true. For example, let

$$\mathcal{A} = \left\{(1, x),\ x \in \left(\frac{1}{2}, 1\right)\right\} \cup \left\{(x, 0),\ x \in \left(0, \frac{1}{2}\right)\right\} \cup \left\{\left(\frac{3}{4}, \frac{1}{4}\right)\right\},$$

If $\mathcal{L} = \{L_1, L_2\}$ with $L_1(a, \theta) = a_1$ and $L_2 = -a_2$, then $(3/4, 1/4)$ is the unique nondominated alternative and there is no Bayes one.

Under conditions similar to those of Proposition 3, and with basically the same proof, we have

PROPOSITION 4. *Suppose $\mathcal{L}$ is generated by a finite number of continuous, uniformly in θ, loss functions in $\mathcal{A}$, which is a compact set. Then, for any L_i there is one nondominated alternative that is Bayes for L_i.*

The provision of more general results, either concerning existence of nondominated alternatives or the relation between nondominated and Bayes alternatives, requires the adoption of abstract topological conditions much as it is done in multiobjective optimisation, see e.g. Yu (1986). However, it turns out to be much more interesting to consider specific statistical examples, as we illustrate in the next subsection.

3.3 *An example: Set estimation.* We consider set estimation under the loss function

$$L_s(\theta, C) = s(vol(C)) - I(\theta \in C),$$

where C is a subset, s is an increasing function and $I(\cdot)$ is the indicator function. Casella, Hwang and Robert (1994) studied admissibility of estimators under such loss functions.

For simplicity, we shall assume that $P(\theta|x)$ gives positive measure to any set with positive Lebesgue measure. We shall assume also that the density function $p(\theta|x)$, corresponding to $P(\theta|x)$, is not constant on any subset with positive Lebesgue measure. This allows for the existence of HPD subsets, $\{\theta : p(\theta|x) \geq k\}, k > 0$, with any arbitrary volume. It can be easily shown that the expected loss of an alternative C is given by

$$\int_\Theta L(\theta, C)dP(\theta|x) = s(vol(C)) - P(C|x).$$

We shall consider a class $\mathcal{L}$ of loss functions L_s depending on s, $\mathcal{L} = \{L_s : s_L(t) \leq s(t) \leq s_U(t), \forall t\}$, where s_L and s_U are given increasing functions. In such case, we have:

PROPOSITION 5. *The nondominated alternatives are necessarily HPD sets.*

PROOF. Suppose there exists a nondominated alternative B which is not a HPD set. Let A be a HPD set with the same volume as B; such A exists because of the previous assumptions on the density $p(\theta|x)$. There exist two subsets $I \subseteq B \cap A^C$ and $E \subseteq A \cap B^C$, with the same volume such that $\inf_{\theta \in E} P(\theta|x) > \sup_{\theta \in I} P(\theta|x)$. Therefore, it follows that $P(A|x) > P(B|x)$. Moreover, for every $L_s \in \mathcal{L}$, or equivalently, for every s satisfying the constraints,

$$s(vol(A)) - P(A|x) - (s(vol(B)) - P(B|x)) = P(B|x) - P(A|x) < 0$$

so that B is dominated by A. □

It is not true, however, that any HPD set is a nondominated alternative, as shown in the following example.

EXAMPLE 3. Let $P(\theta|x)$ be a Beta distribution $\mathcal{B}(2,1)$, with density $P(\theta|x) = 2\theta$, corresponding, e.g., to a Bernoulli model under a uniform prior. The intervals $A_y = [y, 1]$ are the HPD sets with $P(A_y|x) = 1 - y^2$ and volume $1 - y$. The corresponding expected loss is $s(1-y) - 1 + y^2$.

Consider the class $\mathcal{L}$ of all loss functions L_s with $s(t)$ within the band defined by the functions $s_L(t) = t^2$ and $s_U(t) = t$. We will show that the HPD sets A_y, with $y > \dfrac{1+\sqrt{1/2}}{2}$, are dominated by $A_{1/2}$. In fact, for every s such that $s_L \leq s \leq s_U$, we have

$$\int_\Theta L_s(\theta, A_{1/2})dP(\theta|x) - \int_\Theta L_s(\theta, A_x)dP(\theta|x) = -1/4 + 2y - 2y^2,$$

which is negative for those y's. □

Finally, we can see that nondominated alternatives are not always Bayes.

EXAMPLE 4. Consider $\mathcal{L} = \{L_1, L_2\}$, where the losses L_1 and L_2 correspond to the size function $s_1(t) = t^3$ and $s_2(t) = t$, respectively. The Bayes alternatives are the HPD sets A_y with y equal to $y_1 = (4-\sqrt{7})/3$ and $y_2 = 1/2$, respectively. It can be easily shown that any HPD set A_y, with $y \in (y_1, y_2)$ is a nondominated alternative, despite not being a Bayes one. □

4. Nondominated Set: Computations

We turn now to the issue of computing the nondominated set. We shall study an important case in which computations may be performed exactly. In most cases, however, we shall have to turn to procedures to approximate the nondominated set. For that reason a scheme is provided in Martin *et al.* (1997), which requires procedures to check dominance between alternatives.

The case of interest is that of bands of convex loss functions. Let $\lambda(t)$ be a function on $\mathbb{R}$ which is positive (negative, null) if and only if $t > 0$ ($t < 0, t = 0$, respectively). The function $\Lambda(a,\theta) = \int_0^{\theta-a} \lambda(t)dt$ defines a loss function such that $\Lambda(a,\theta) = \Lambda(\theta - a)$ and $\Lambda^{'}(t) = \lambda(t)$, for all real t. Note that widely used loss functions may be obtained for appropriate choices of λ functions: for example, the squared error loss $\Lambda(a,\theta) = (\theta - a)^2$ is given by $\lambda(t) = 2t$; the absolute error loss, $\Lambda(a,\theta) = |\theta - a|$ is obtained for $\lambda(t) = t/|t|$ (with $\lambda(0) = 0$); the LINEX loss function, see Varian (1974), $\Lambda(a,\theta) = e^{\gamma(a-\theta)} - \gamma(a-\theta) - 1$ is obtained for $\lambda(t) = \gamma(1 - e^{-\gamma t})$. Note also that symmetric loss functions can be obtained if $\lambda(-t) = -\lambda(t)$ for all real t. Moreover, we can obtain a strictly convex loss function by considering $\lambda(t)$ such that $\lambda^{'}(t) > 0$ for all real $t \neq 0$. Hence, we have an extremely flexible definition of a loss function.

Given λ, we may perturb it by changing the values of $\lambda(t)$ in some intervals. Among possible changes of λ, we shall consider those within a band. Consider two functions $v(t)$ and $u(t)$ defined on $\mathbb{R}$ which are positive (negative, null) if and only if $t > 0$ ($t < 0, t = 0$) and such that $v(t) \leq u(t), \forall t \in \mathbb{R}$. Let V and U be their associated loss functions. We will consider the class

$$\mathcal{L} = \{\Lambda : v(t) \leq \lambda(t) \leq u(t),\ \forall t \in \mathbb{R}\},$$

and call it *band of convex loss functions class.* We shall compute the set of nondominated alternatives for that class with $\mathcal{A} = \Theta = \mathbb{R}$. We provide first some preliminary results.

PROPOSITION 6.

$$\sup_{\Lambda \in \mathcal{L}} (T(\Lambda, a) - T(\Lambda, b)) = \begin{cases} T(U,a) - T(U,b) & a < b \\ T(V,a) - T(V,b) & a > b \end{cases}$$

PROOF. Let $F(\theta|x)$ be the distribution function corresponding to $P(\theta|x)$. The result follows from

$$\int_{\mathbb{R}} \Lambda(a,\theta)dP(\theta|x) - \int_{\mathbb{R}} \Lambda(b,\theta)dP(\theta|x) = \int_{\mathbb{R}} \left\{ \int_{\theta-b}^{\theta-a} \lambda(t)dt \right\} dP(\theta|x)$$

$$= \int_{\mathbb{R}} \{F(b+t|x) - F(a+t|x)\}\, \lambda(t)dt$$

and $F(b+t|x)-F(a+t|x)$ being either nonnegative or nonpositive for all real t. □

Let a_V and a_U be the (not necessarily unique) Bayes alternatives under losses V and U, respectively.

Lemma 1. *Let $P(\theta|x)$ be such that $P(A|x) > 0$ for all measurable subsets A with non-null Lebesgue measure. Suppose $v(t) \leq u(t)$ for any real t, with $v(t) < u(t)$ on some interval I. It follows that $a_V \leq a_U$.*

Proof. We prove the result by contradiction. Suppose there exist two Bayes alternatives a_V and a_U such that $a_V > a_U$. As the alternatives are Bayes, it follows that

$$\int_{\mathbb{R}} [V(a_V,\theta) - V(a_U,\theta)]dP(\theta|x) \leq 0 \text{ and } \int_{\mathbb{R}} [U(a_V,\theta) - U(a_U,\theta)]dP(\theta|x) \geq 0.$$

Such conditions are equivalent to

$$\int_{\mathbb{R}} \left\{ \int_{\theta - a_U}^{\theta - a_V} v(t)dt \right\} dP(\theta|x) \leq 0 \text{ and } \int_{\mathbb{R}} \left\{ \int_{\theta - a_U}^{\theta - a_V} u(t)dt \right\} dP(\theta|x) \geq 0.$$

Combining both conditions, we have

$$\int_{\mathbb{R}} \left\{ \int_{\theta - a_U}^{\theta - a_V} [u(t) - v(t)]dt \right\} dP(\theta|x) \geq 0. \qquad \ldots (1)$$

Since $u(t) - v(t)$ is nonnegative for any real t and strictly positive on an interval I (possibly $\mathbb{R}$), and $\theta - a_U > \theta - a_V$ for all θ, the inner integral in (1) is strictly negative for all θ such that $I \cap [\theta - a_V, \theta - a_U] \neq \emptyset$. Such θ's belong to a measurable set A, with $P(A|x) > 0$, therefore the integral in (1) is strictly negative, contradicting the assumption $a_V > a_U$. □

For Proposition 7, we assume that V and U are strictly convex loss functions so that there exists a unique Bayes alternative for each of them. This result is relevant since it shows that Bayes and nondominated alternatives coincide for a quite large class of loss functions.

Proposition 7. *Let a_V and a_U be the Bayes alternatives corresponding to losses V and U obtained, respectively, from v and u. Suppose $v'(t) > 0$ and $u'(t) > 0$ for all real $t \neq 0$. The interval $[a_V, a_U]$ is the set of all Bayes alternatives, which coincides with the set of nondominated alternatives.*

Proof. Any alternative $a < a_V$ is dominated by a_V since, for any $\Lambda \in \mathcal{L}$, it holds that

$$\int_{\mathbb{R}} \Lambda(a,\theta)dP(\theta|x) - \int_{\mathbb{R}} \Lambda(a_V,\theta)dP(\theta|x) = \int_{\mathbb{R}} \left\{ \int_{\theta - a_V}^{\theta - a} \lambda(t)dt \right\} dP(\theta|x)$$

$$\geq \int_{\mathbb{R}} \left\{ \int_{\theta-a_V}^{\theta-a} v(t)dt \right\} dP(\theta|x) = \int_{\mathbb{R}} V(a,\theta)dP(\theta|x) - \int_{\mathbb{R}} V(a_V,\theta)dP(\theta|x) > 0.$$

Similarly, any alternative $a > a_U$ is dominated by a_U, so that the set of nondominated alternatives is contained in $[a_V, a_U]$.

Because of Lemma 1, the set of Bayes alternatives is contained in $[a_V, a_U]$. Consider the function $\lambda_\epsilon(t) = (1-\epsilon)v(t)+\epsilon u(t)$, $\epsilon \in [0,1]$, and the corresponding loss function $\Lambda_\epsilon \in \mathcal{L}$. It is easy to see that the Bayes alternative a_ϵ is unique and a_ϵ is a nondecreasing function of ϵ, again as a consequence of Lemma 1.

Suppose there exists $\tilde{\epsilon} \in (0,1)$ and $\tilde{a}$ such that $\lim_{\epsilon\to\tilde{\epsilon}^-} a_\epsilon < \tilde{a} < a_{\tilde{\epsilon}}$. Because of the continuity of $\Lambda_\epsilon(a,\theta)$ in both ϵ and a, it follows that there is $\eta > 0$ such that $a_{\tilde{\epsilon}-\eta} < \tilde{a}$ and

$$\begin{aligned}
\int_{\mathbb{R}} \Lambda_{\tilde{\epsilon}}(a_{\tilde{\epsilon}},\theta)dP(\theta|x) &< \int_{\mathbb{R}} \Lambda_{\tilde{\epsilon}}(a_{\tilde{\epsilon}-\eta},\theta)dP(\theta|x)\\
\Longrightarrow \int_{\mathbb{R}} \Lambda_{\tilde{\epsilon}}(\tilde{a},\theta)dP(\theta|x) &< \int_{\mathbb{R}} \Lambda_{\tilde{\epsilon}}(a_{\tilde{\epsilon}-\eta},\theta)dP(\theta|x)\\
\Longrightarrow \int_{\mathbb{R}} \Lambda_{\tilde{\epsilon}-\eta}(\tilde{a},\theta)dP(\theta|x) &< \int_{\mathbb{R}} \Lambda_{\tilde{\epsilon}-\eta}(a_{\tilde{\epsilon}-\eta},\theta)dP(\theta|x),
\end{aligned}$$

which is impossible. Therefore, a_ϵ is a continuous function of ϵ and any alternative in $[a_L, a_U]$ is Bayes and nondominated, because of Proposition 2. □

Example 5. Suppose that $P(\theta|x)$ is $\mathcal{N}(0,1)$, and $\mathcal{L}$ is given by

$$v(t) = \begin{cases} 3t & t < 0 \\ t & t \geq 0 \end{cases}$$

and

$$u(t) = \begin{cases} t & t < 0 \\ 3t & t \geq 0 \end{cases}$$

Such a class contains the squared error loss (for $\lambda(t) = 2t$). Besides, $V(t)$ and $U(t)$ are the loss functions in $\mathcal{L}$ which penalises the most, respectively, the negative and the positive values of t. It can be shown that the interval $[-.3989, .3989]$ is the set of nondominated (and Bayes) alternatives. △

In most cases, however, we shall not be able to compute exactly the nondominated set, and we shall need a scheme to approximate it. One such scheme may be seen in Riós Insua and Martin (1994b).

5. The Size of the Nondominated Set

In some cases, nondominance is a very powerful concept leading to a unique nondominated alternative. Here is a non-trivial case. Suppose that $P(\theta|x)$ is a symmetric, unimodal distribution such that $P(A|x) > 0$ for any measurable subset A. Without loss of generality, assume that the mode is 0. Let $\mathcal{L}$ be the class of all convex, symmetric loss functions which are not constant in a subset with positive posterior probability.

PROPOSITION 8. *Given P and $\mathcal{L}$ as above, the mode 0 is the Bayes alternative for any $L \in \mathcal{L}$ and the unique nondominated alternative.*

PROOF. For any $a \in \Theta$ and for any $L \in \mathcal{L}$, it follows that

$$\begin{aligned}
&\int_{\mathbb{R}} [L(0,\theta) - L(a,\theta)] dP(\theta|x) \\
&= \int_0^\infty [L(\theta) - L(\theta - a)] dP(\theta|x) + \int_0^\infty [L(-\theta) - L(-\theta - a)] dP(-\theta|x) \\
&= \int_0^\infty [2L(\theta) - L(\theta - a) - L(\theta + a)] dP(\theta|x) \\
&< 0,
\end{aligned}$$

the inequality following from the convexity of L, which is nonconstant on a subset with positive probability. The result follows immediately. □

However, in most cases the nondominated set will be too big to reach a final decision. Note that as a byproduct of the procedures to compute or approximate the nondominated set, we obtain estimates on the differences in posterior expected losses among nondominated alternatives. If these were not large, we would conclude that these perform not too differently in terms of their posterior expected loss and, basically, we would not loose too much by recommending any of those alternatives.

One possibility would be to elicit additional information from the decision maker and further constrain the class. Clearly, in this case the set of nondominated alternatives will be smaller and we could hope that this iterative process would converge until the nondominated set is small enough to reach a final decision. Martin and Riós Insua (1997) provide ideas for aiding in eliciting additional information when robustness lacks.

Alternatively, it is conceivable that we might not be able to elicit additional information. In such cases, see e.g. Makov (1994), we might appeal to *ad hoc* concepts like $\mathcal{L}$-minimax alternatives, as a way to pick a nondominated alternative, as we show in the following result. Recall that $a_M \in \mathcal{A}$ would be $\mathcal{L}$-minimax if $\max_{L\in\mathcal{L}} T(L, a_M) = \min_{a\in\mathcal{A}} \max_{L\in\mathcal{L}} T(L, a)$.

PROPOSITION 9. *If the set of $\mathcal{L}$-minimax alternatives is finite, one of them is nondominated.*

Suppose that a is $\mathcal{L}$-minimax and dominated by another alternative a'. It is easy to see that a' is also $\mathcal{L}$-minimax. The result follows because the set is finite. □

As a corollary, if there is a unique $\mathcal{L}$-minimax alternative, it is nondominated. Similar results may be obtained, for other *ad hoc* concepts like $\mathcal{L}$-minimax regret alternatives.

6. Discussions

We have described the case in which imprecision in preferences is modelled with a class of loss functions. Rather than undertaking a straightforward extension of robust Bayesian analysis and analysing the local or global behaviour of a posterior or predictive expectation, we have gone back to foundations of robust Bayesian analysis and concluded that a main computational issue in this field should be the computation of the nondominated set.

We have explored several questions in that direction, mainly existence of nondominated alternatives, relations between Bayes and nondominated alternatives, relations between $\mathcal{L}$-minimax and nondominated alternatives, and provided ways to compute the nondominated set.

There are many other classes of loss functions for which the results here stated should be computed, and we view this paper as a first one in that direction. Some of the classes will be parallel to those used in sensitivity to the prior studies, see Berger (1994). Lindley (1976) introduces other important classes. Finally, another source of classes of loss functions is the stochastic dominance literature, see Levy (1992) for a review, which concentrates mainly on loss functions whose derivatives of increasing order alternate in sign. Martin et al (1997) provide additional ideas.

Acknowledgments. This work has been supported by grants from CNR, NATO, CICYT and CAM. Discussions with Carlos Conde are gratefully acknowledged.

References

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

———— (1994), An overview of robust Bayesian analysis (with discussion), *Test*, **3**, 5-124.

Brown, L. (1975), Estimation with incompletely specified loss functions, *Jour. Amer. Stat. Assoc.*, **70**, 417-427.

Casella, G., Hwang, G. and Robert, C. (1994), Loss functions for set estimation. In *Statistical Decision Theory and Related Topics V*, (Gupta, S. and Berger, J. eds.), 237-251, Springer, New York.

Clyde, M. and Chaloner, K. (1996), The equivalence of constrained and weighted designs in multiple objective design problems, *Jour. Amer. Stat. Assoc.*, **91**, 1236-1244.

Dey, D., Lou, K. and Bose, S. (1995), A Bayesian approach to loss robustness, *Technical Report*, University of Connecticut, Storrs.

Hwang, J.T. (1985), Universal domination and Stochastic domination: Estimation simultaneously under a broad class of Loss functions, *Ann. Statist.,* **13,** 295-314.

Levy, H. (1992), Stochastic dominance and expected utility analysis: survey and analysis, *Management Science,* **38**, 555-593.

Lindley, D. (1976), A class of utility functions, *Ann. Statist.,* **4**, 1-10.

Makov, U. (1994), Some Aspects of Bayesian loss-robustness, *Jour. Stat. Plan. Infer.* **38**, 359-370.

Martin, J. and Riós Insua, D. (1997), Local sensitivity analysis in Bayesian Decision Theory. In *Bayesian Robustness* (Berger, J., Betrò, B., Moreno, E., Pericchi, L.R., Ruggeri, F., Salinetti, G. and Wasserman, L. eds), Lecture Notes IMS, 29, IMS, Hayward.

Martin, J., Riós Insua, D. and Ruggeri, F. (1997), Issues in Bayesian loss robustness, *Tech. Rep.*, ISDS, Duke University.

Nau, R., Ríos Insua, D. and Martín, J. (1997), The Shape of Incomplete Preferences and the Foundation of Robust Bayesian Statistics, *Tech. Rep.* Madrid Technical University.

Ríos Insua, D. and Martín, J. (1994a), On the foundations of robust decision making. In *Decision Theory and Decision Analysis: Trends and Challenges*, (Ríos, S. ed.), Kluwer A. P., Amsterdam.

Ríos Insua, D. and Martín, J. (1994b), Robustness issues under imprecise beliefs and preferences, *J. Stat. Plann. Infer.,* **48**, 383-389.

Varian, H. R. (1974), A Bayesian approach to real state assessment. In Studies in Bayesian Econometrics and Statistics in Honor of L.J.Savage (Fienberg, S. and Zellner, A. eds.), North-Holland, Amsterdam.

Yu, P.L. (1986), *Multiple Criteria Decision Making*, Plenum Press.

Jacinto Martín
Dept. Inteligencia Artificial
Facultad de Informatica
Univ. Politecnica de Madrid
Campus de Montegancedo
28660, Boadilla, Spain
e-mail : jrmartin@fi.upm.es

David Ríos Insua
Escuela de Ingenieria
Universidad Rey Juan Carlos
Independencia 12
28931, Mostoles, Spain
e-mail : d.rios@escet.urjc.es

Fabrizio Ruggeri
Istituto per le Applicazioni della
Matematica e dell'Informatica
Consiglio Nazionale delle Ricerche
Via Ampére 56
I-20131 Milano, Italy
e-mail : fabrizio@iami.mi.cnr.it