Bayesian Robustness with Quantile Loss Functions*

J.P. ARIAS
Universidad de Extremadura, Spain

J. HERNÁNDEZ
Universidad de Extremadura, Spain

J. MARTÍN
Universidad de Extremadura, Spain

A. SUÁREZ
Universidad de Cádiz, Spain

Abstract

Bayes decision problems require subjective elicitation of the inputs: beliefs and preferences. Sometimes, elicitation methods may not perfectly represent the Decision Maker's judgements. Several foundations propose to overlay this problem using robust approaches. In these models, beliefs are modelled by a class of probability distributions and preferences by a class of loss functions. Thus, the solution concept is the set of non-dominated alternatives. In this paper we focus on the computation of the efficient set when the preferences are modelled by a class of convex loss functions, specifically the quantile loss functions. We illustrate the idea with examples and introduce the use of stochastic dominance in this context.

Keywords

Bayesian robustness, non-dominated alternatives, Bayes alternatives, quantile loss functions, stochastic orders, quantile class of prior distributions

1 Introduction

Robust Bayesian analysis arises to avoid demanding an excessively precision in the decision maker's judgements concerning his beliefs and preferences. Thus, the

*This work has been supported in part by a grant of Junta de Extremadura IPR00A075. We thank the referees for their fruitful comments

imprecision in preferences leads to a class of loss functions while the imprecision in beliefs is modelled by a class of prior probability distributions which would be actualized via Bayes Theorem. For some interesting revisions on Bayesian Robustness axiomatic systems see e.g. Ríos Insua and Martín [13], Nau [11], Seidenfeld et al [18] and Weber [19].

In summary, using a class Γ of prior distributions over the set of states Θ and a class $\mathcal{L}$ of loss functions, given $a, b \in \mathcal{A}$, set of alternatives, we say that $b \preceq a$ if and only if

$$T(a, L, \pi) \leq T(b, L, \pi),\ \forall \pi \in \Gamma, \forall L \in \mathcal{L},$$

where $T(a, L, \pi)$ is the posterior expected loss for the action a, L is the loss function, π is the prior and $\preceq$ is the preference relationship between alternatives:

$$T(a, L, \pi) = \frac{\int_\Theta L(a, \theta) l(\theta) d\pi(\theta)}{\int_\Theta l(\theta) d\pi(\theta)},$$

$l(\theta)$ being the likelihood for an experiment x.

This model is similar to a multicriteria optimization problem. The optimal solution is the one that minimizes $T(\cdot, L, \pi)$ for every pair $\pi \in \Gamma, L \in \mathcal{L}$. Unfortunately, in general, that optimal solution does not exist. Thus, the non-dominated set is taken as an starting point. Any dominated alternative must be discarded. See Coello [6] for an excellent discussion on multiobjective optimization. We say that a dominates b if and only if $a \prec b$, (that is, $a \preceq b$ and $\neg(b \preceq a)$). A non-dominated alternative a is such that there is no other alternative b which dominates a. Arias [1] and Arias and Martín [2] provide theoretic results about the existence of such a set and its relationship with the set of Bayes alternatives. Martín and Arias [8] provide a method based on comparing pairs to approximate the non-dominated set. Some references for Bayesian sensitivity are Berger [4], Ríos Insua and Ruggeri [14] and Ríos et al [15].

We study the calculus of the non-dominated set for problems in which the imprecision in preferences is modelled by quantile loss functions. We give general results that we will particularize for classes of quantile prior distributions, see Moreno and Cano [9]. Since we are interested in Bayesian inference, we will consider $\mathcal{A} = \mathbb{R}$ although the results will be easily applicable when $\mathcal{A}$ is an interval of $\mathbb{R}$.

We organize this work as follows. We begin with some results concerning convex loss functions and their implications in the calculus of the non-dominated set. Secondly, we particularize for quantile loss functions, indicating the relationship with the Bayes alternatives in this case. We also consider a quantile class for prior distributions giving some results and an example. Third part of the paper is dedicated to various stochastic orders, only those that hold for the posterior distributions once the priors have been ordered, and how they can be used in order to calculate the non-dominated set.

2 Bayesian Robustness with convex loss functions

We will denote $\mathcal{L}_C$ the class of all convex loss functions in $\mathcal{A}$. Every loss function $L \in \mathcal{L}_C$, verifies for all θ, $a, b \in \mathcal{A}$, and $\lambda \in [0,1]$ that

$$L(\lambda a + (1-\lambda)b, \theta) \leq \lambda L(a,\theta) + (1-\lambda)L(b,\theta). \tag{1}$$

A first useful result, easy to prove, is:

Lemma 1 *Let Γ be any class of prior distributions and $\mathcal{L}_C$ the class of convex loss functions. The function $T(\cdot, L, \pi) : \mathcal{A} \longrightarrow \mathbb{R}$ is convex for every pair $(L,\pi) \in \mathcal{L}_C \times \Gamma$.*

A well known result is that every convex function is continuous in the interior, see Roberts and Varberg [16]. Then considering the set of alternatives, $\mathbb{R}$, the function $T(\cdot, L, \pi)$ is continuous in $\mathbb{R}$ and if it exists the set of Bayes alternatives, this will be a closed interval in $\mathbb{R}$. In the case that, for some pair $(L,\pi) \in \mathcal{L}_C \times \Gamma$ the set of Bayes alternatives is empty, the function $T(\cdot, L, \pi)$ will be increasing or decreasing in $\mathbb{R}$ (strictly increasing or strictly decreasing if the functions are strictly convex).

If the set of Bayes alternatives is not empty, then the function $T(\cdot, L, \pi)$ is strictly decreasing in $(-\infty, a_{(L,\pi)})$ and strictly increasing in $(a^{(L,\pi)}, +\infty)$, being

$$\begin{aligned} a_{(L,\pi)} &= \min_{a \in B_{(L,\pi)}} a, \text{ and} \\ a^{(L,\pi)} &= \max_{a \in B_{(L,\pi)}} a. \end{aligned}$$

Note that the alternatives $a_{(L,\pi)}$ and $a^{(L,\pi)}$ are also Bayes for (L,π).

An immediate result is that the set of non-dominated alternatives is included in the closed interval $[\mu_*, \mu^*]$, being μ_* and μ^*, respectively, the infimum and the supremum of the Bayes alternatives, that is,

$$\begin{aligned} \mu_* &= \inf_{(L,\pi) \in \mathcal{L} \times \Gamma} a_{(L,\pi)}, \text{ and} \\ \mu^* &= \sup_{(L,\pi) \in \mathcal{L} \times \Gamma} a^{(L,\pi)}. \end{aligned}$$

In the Bayesian literature the range of this interval is considered as the robustness measure of the problem, see Berger [4]. However, if we are interested in calculating exactly the set of non-dominated alternatives, we can give a more accurate approximation using the following result due to Arias et al [3]:

Theorem 1 *Let $\mathcal{L} \subseteq \mathcal{L}_C$ be a family of convex loss functions, Γ a class of prior probability distributions, so that, for every pair $(L,\pi) \in \mathcal{L} \times \Gamma$, the set of Bayes* *alternatives* $B_{(L,\pi)}$ *is not empty and let*

$$\begin{aligned} a_* &= \inf_{(L,\pi)\in\mathcal{L}\times\Gamma} a^{(L,\pi)}. \\ a^* &= \sup_{(L,\pi)\in\mathcal{L}\times\Gamma} a_{(L,\pi)}. \end{aligned}$$

We have

1. *If* $a_* < a^*$*, then* $(a_*, a^*) \subseteq \mathcal{ND}(\mathcal{A}) \subseteq [a_*, a^*]$.
2. *If* $a_* \geq a^*$*, then* $\mathcal{ND}(\mathcal{A}) = [a^*, a_*]$.

In order to study the robustness of the problem, it is not necessary to determine whether the alternatives a_* and a^* are dominated or not. Nevertheless, it is interesting to calculate the efficient set in an accurate way. In this paper we will see inference problems modelled by particular classes of loss functions and prior distributions in which we can assure that the extremes of the interval a_* and/or a^* are non-dominated alternatives. If the set of Bayes alternatives is empty for some pair $(L,\pi) \in \mathcal{L} \times \Gamma$ then the result is valid considering $a_* = -\infty$ (when $T(\cdot, L, \pi)$ is increasing) or $a^* = \infty$ (decreasing).

3 Quantile loss functions

Let us consider the case where preferences are modelled by quantile loss functions. A particular case of this type is the absolute value loss function. The class of quantile loss functions is defined as

$$\mathcal{L} = \{L_p : L_p(a,\theta) = |a-\theta| - a(2p-1),\ p \in [0,1]\} \quad (1).$$

Functions equivalent to these have been used in Economy, such as the ones studied by Geweke [7]

$$L(a,\theta) = c_1(a-\theta)I_{(-\infty,a]}(\theta) + c_2(\theta-a)I_{(a,+\infty)}(\theta) \quad (2)$$

where I is the indicator function. Bayes alternatives for this type of function are the quantiles of order $c_2/(c_1+c_2)$ (If $c_1 = c_2$, it coincides with the median) Thus, they are asymmetric functions with different weights on the positive and negative errors. Next example shows the use of this type of functions.

Example 1 *Noortwijt and Gelder [12] studied the Bayes estimators of the optimal dyke height under asymmetric linear loss function. Let us suppose we have to decide the height of the dykes to prevent flooding. The height of the dyke h will be the decision variable and* $h_0 = 3.25$ *the initial height at the moment when the decision has to be taken. Inundation will occur as soon as the sea water level* *exceeds the hight of the dyke. We assume that the maximal sea levels per year* X_i $i = 1,\ldots,n$ *are conditionally independent, exponentially distributed, with a known location parameter* $x_0 = 1.96$ *meters and an unknown parameter* λ *with expected value 0.33 meters. Therefore the likelihood function is*

$$l(x|\lambda) = \prod_{i=1}^{n} f(x_i,\lambda) = \prod_{i=1}^{n} \frac{1}{\lambda} exp\{-\frac{x_i - x_0}{\lambda}\}.$$

The prior density of λ *is assumed to be an inverted gamma distribution with scale parameter* $\mu > 0$ *and shape parameter* $\nu > 0$

$$I_g(\lambda,\nu,\mu) = [\mu^{\nu}/\Gamma(\nu)]\lambda^{-(\nu+1)} exp\{-\mu/\lambda\} \quad \lambda > 0.$$

The loss function is (2) with $c_1 = 5.37 \cdot 10^7$ *and* $c_2 = 1.94 \cdot 10^7$. △

This type of loss function have also been used in Forecast Theory, see Capistrán [5] and references therein.

We will use the functions defined in (1) as they only depend on a single parameter. Quantile loss functions are convex in $\mathcal{A}$. The posterior expected loss is

$$T(a,L_p,\pi) = D_{\theta|x}(a) - a(2p-1),$$

and their Bayes alternatives are the quantiles of order p of the posterior distributions, since

$$T'(a,L_p,\pi) = 2F_{\theta|x}(a) - 2p,$$

for every point a where the distribution function is continuous.

Let us recall that it is called quantile of order p of a random variable X, the value $Q_X(p)$ such that

$$P[X \leq Q_X(p)] \geq p \ \text{ and}$$
$$P[X \geq Q_X(p)] \geq 1-p.$$

As it happened with the absolute value loss function, when using quantile loss functions, the posterior distribution quantiles may not be unique.

Based on theorem 1 we have the following result, when there is precision in DM's beliefs.

Proposition 1 *Let* $\mathcal{L}$ *be the class of quantile loss functions*

$$\mathcal{L} = \{L_p : L_p(a,\theta) = |a-\theta| - a(2p-1),\ p \in [p_0,p_1]\}$$

and π *a prior distribution so that the posterior distribution quantiles are unique, then*

$$\mathcal{ND}_{\pi}(\mathcal{A}) = [Q_{\pi}(p_0), Q_{\pi}(p_1)],$$

where $Q_{\pi}(p_0)$ *and* $Q_{\pi}(p_1)$ *are, respectively, the quantiles of order* p_0 *and* p_1 *of the posterior distribution of* π.

This result can be generalized in the case that we also have imprecision in the decision maker's beliefs.

Proposition 2 *Let $\mathcal{L}$ be the class of quantile loss functions*

$$\mathcal{L} = \{L_p : L_p(a,\theta) = |a-\theta| - a(2p-1),\ p \in [p_0, p_1]\},$$

a class of distributions

$$\Gamma = \{\pi : \pi(\theta|x) \text{ with posterior quantiles } Q_\pi(p) \text{ unique},\ p \in [0,1]\}$$

and the values

$$a_* = \mu_* = \inf_{\pi\in\Gamma} Q_\pi(p_0) \text{ and}$$
$$a^* = \mu^* = \sup_{\pi\in\Gamma} Q_\pi(p_1),$$

then

$$(\mu_*, \mu^*) \subseteq \mathcal{ND}(\mathcal{A}) \subseteq [\mu_*, \mu^*].$$

If posterior quantiles are not unique we must appeal to theorem 1 with

$$a_* = \inf_{\pi\in\Gamma}\{\sup Q_\pi(p_0)\}$$

$$a_* = \sup_{\pi\in\Gamma}\{\inf Q_\pi(p_1)\}$$

and

$$(a_*, a^*) \subseteq \mathcal{ND}(\mathcal{A}) \subseteq [a_*, a^*].$$

In general it can not be assured that μ_* or μ^* are non-dominated alternatives as we illustrate with the following example.

Example 2 *Let us consider the class of absolute value loss functions and a class of discrete posterior distributions with probability distribution:*

$(n \in \mathbb{N})$

$$\pi_n(\theta) = \begin{cases} \dfrac{2n+3}{4(n+1)} & \text{if } \theta = \dfrac{1}{n}, \\ \\ \dfrac{2n+1}{4(n+1)} & \text{if } \theta = 1. \end{cases}$$

The Bayes alternative for each distribution π_n would be its posterior median $1/n$ and the set of non-dominated alternatives is the interval $(0,1]$. The alternative 0 is dominated by the alternative 1, since for every $n \in \mathbb{N}$

$$T(0, L, \pi_n) = \frac{2n^2+3n+3}{4n(n+1)} > \frac{2n^2+n-3}{4n(n+1)} = T(1, L, \pi_n).$$

△

Obviously if μ_* and μ^* are unique Bayes alternatives, then they are also non-dominated alternatives.

Note that, if having precision in beliefs, then the range of the non-dominated set is the range between the quantile p_0 and the quantile p_1. Sometimes the non-dominated set can be the same as the (HPD), as in next example. However, so that this happens the elicitation of the class should depend on the posterior distribution.

Example 3 *Let us consider the class of loss functions*

$$\mathcal{L} = \{L_p : L_p(a,\theta) = |a-\theta| - a(2p-1),\ p \in [0,0.8]\}$$

and a Pareto prior distribution with parameters α *and* β*.*

We take a sample $\{X_1,\ldots,X_n\}$ *of a population which is distributed following an uniform distribution with mean* $\theta/2$*. Therefore, the posterior distribution is*

$$P(\alpha+n, \max\left(\beta, X_{(n)}\right)),\ \textit{with}\ X_{(n)} = \max\{X_1,\ldots,X_n\}$$

Then, the posterior quantiles of π *are*

$$Q_\pi(p) = \frac{1}{\sqrt[\alpha+n]{p}} \max\left(\beta, X_{(n)}\right).$$

Thus, the set of non-dominated alternatives would be the closed interval

$$ND(\mathcal{A}) = \left[\max\left(\beta, X_{(n)}\right), Q_\pi(0.8)\right].$$

This means than the non-dominated set coincides with the confidence interval HPD for the parameter θ *at a confidence level of* 80% *.*

The table 1 shows the non-dominated set when $\alpha = 2$ *and* $\beta = 5$ *for various samples.*

n	$X_{(n)}$	$ND(A)$	range
10	3.2	[5, 5.0938]	0.0938
50	4.5	[5, 5.0215]	0.0215
100	4.8	[5, 5.011]	0.011
500	5.9	[5.9, 5.9026]	0.0026
1000	4.7	[5, 5.0011]	0.0011

Table 1: Non-dominated set for the example 3.

△

3.1 Relationship with the non-dominated set

An important question is the relationship between the Bayes set and the non-dominated set. It is easy to prove that, in general, they are different, see Arias *et al* [1]. In this case, we have the following result:

Proposition 3 *Let $\mathcal{L}$ be the class of quantile loss functions with $p \in [p_0, p_1]$. If the class Γ of prior distributions is convex and $Q_\pi(p)$ are unique for every $\pi \in \Gamma$ and $p \in (p_0, p_1)$, then the set of non-dominated alternatives is the closure of the set of Bayes actions, and the interiors of both sets are the same.*

Proof.

If the set of prior distributions is convex, then the set of posterior distributions is convex as well, see Arias et al[1]. As the quantiles are unique for any π, all bayes actions are non-dominated. So, we only have to prove that given a and b bayes actions for (π, L_1) and(π, L_2), $\alpha a + (1-\alpha)b$ is also bayes for $\alpha \in (0,1)$. Consider $a = Q_{\pi_1}(p')$ and $b = Q_{\pi_2}(p'')$ with $p', p'' \in [p_0, p_1]$ and $a < b$ $p' < p''$.

Then

$$\int_{-\infty}^{\alpha a+(1-\alpha)b} \pi_1(\theta|x)d\theta \quad > p'$$

$$\int_{-\infty}^{\alpha a+(1-\alpha)b} \pi_2(\theta|x)d\theta \quad < p''$$

so, there is β such that

$$\beta \int_{-\infty}^{\alpha a+(1-\alpha)b} \pi_1(\theta|x)d\theta + (1-\beta)\int_{-\infty}^{\alpha a+(1-\alpha)b} \pi_2(\theta|x)d\theta \quad = p \in (p', p'')$$

Then $\alpha a + (1-\alpha)b$ is the $Q_\pi(p)$ with $\pi(\cdot|x) = \beta\pi_1(\cdot|x) + (1-\beta)\pi_2(\cdot|x)$

□

4 Quantile prior distributions

We now consider some classes of prior distributions to model imprecision in beliefs. Let $A_i = \lfloor \theta_i, \theta_{i+1} \rceil$ [1] be a partition of the parameter space and the class of prior distributions:

$$\Gamma_Q = \{\pi : \pi(A_i) = q_i, \quad i = 1 \dots n, \quad q_i \geq 0 \quad \forall i \quad \sum_i q_i = 1\}$$

This is a particular case of the quantile class, see Moreno and Cano [9] and Moreno and Pericchi [10] and Martín and Ríos Insua [13] among others.

A well known result states that the suprema and the infima of functionals over Γ are attained for discrete distributions. So, we have

Lemma 2 *We have*

$$\max_{\pi_d \in \Gamma_d} \pi(\bigcup_{j=1}^{i} A_i|x) = \frac{\sum_{j=1}^{i} \max_{\theta \in A_j} l(x|\theta)p_j}{\sum_{j=1}^{i} \max_{\theta \in A_j} l(x|\theta)p_j + \sum_{k=i+1}^{n} \min_{\theta \in A_k} l(x|\theta)p_k}$$

[1] by $\lfloor a, b \rceil$ we denote any type of interval in $\mathbb{R}$

Let us denote r_i the second term in (8). Lemma 2 lead us to the following iterative scheme to calculate μ_*

$r_0 = 0 \quad i = 0$
while $r_i < p_0 \quad i = i + 1$
  compute r_i

Let A_k be the first interval for which $r_k \geq p_0$. We define now

$$r_k(\theta) = \frac{\sum_{j=1}^{i-1} \max_{\lambda \in A_j} l(x|\lambda) p_j + l(x|\theta) p_k}{\sum_{j=1}^{i-1} \max_{\lambda \in A_j} l(x|\lambda) p_j + l(x|\theta) p_k + \sum_{j=k+1}^{n} \min_{\lambda \in A_j} l(x|\lambda) p_j} \quad \forall \theta \in A_k$$

then

$$\mu_* = \inf\{\theta \in A_k : r_k(\theta) \geq p_0\}$$

By Theorem 1, for the calculus of a_*, we will distinguish two cases, if the last inequality is strict then $\mu_* = a_*$ which is the only quantile of order p. Otherwise, there are several quantiles. For the calculus of a_* we proceed as follows. In A_k we will search a point $a > \mu_*$ for which the inequality is strict. If such point exists then $a_* = \inf\{a \in A_k : r_k(a) > p_0\}$. If there is no point $a > \mu_*$ in A_k such that this is verified then $a_* = \inf\{a \in A_{k+1} : r_{k+1}(a) > p_0\}$ and so on.

Example 4 *The decision maker considers that negative errors are more important than positive, so he uses the class of loss functions:*

$$\mathcal{L} = \{L_p : L_p(a, \theta) = |a - \theta| - a(2p - 1),\ p \in [0.45, 0.48]\}.$$

For believes representation he adopts a quantile class with quantiles given in Table 2.

A_i	$[-\infty, -16.44)$	$[-16.44, -5.22)$	$[-5.24, -2.53)$	$[-2.53, -1.25)$	$[-1.25, 0)$
p_i	0.05	0.25	0.1	0.05	0.05
A_i	$[0, 1.25)$	$[1.25, 2.53)$	$[2.53, 5.24)$	$[5.24, 16.44)$	$[16.44, \infty)$
p_i	0.05	0.05	0.1	0.25	0.05

Table 2: Probabilities for some intervals

Applying the proposed iterative scheme to obtain $\inf Q_\pi(0.45)$ *we get the* r_i *values showed table 3. Therefore* $\min_{\pi \in \Gamma} Q_\pi(0.45) \in A_4 = [-2.533, -1.257]$ *and solving* $\min\{\theta \in A_i$ *such that* $r_k(\theta) \geq 0.45\}$ *we obtain* $\inf Q_\pi(0.45) = -2.533$. *Moreover,* $r_k(-2.5333) > 0.45$ *so* $a_* = -2.533$.

We apply the equivalent algorithm for $sup_{\pi \in \Gamma} Q_\pi(0.48)$ *obtaining 2.533. Then* $\mathcal{ND}(\mathcal{A}) = [-2.533, 2.533]$

i	1	2	3	4	5	6	7	8	9	10
r_i	0.0001	0.28	0.44	0.54	0.65	0.76	0.87	0.99	1	1

Table 3: r_i values

The class of quantiles Γ_Q can be generalized considering bounds over the sets A_i obtaining the class

$$\Gamma_{QG} = \{\pi : \alpha_i \leq \pi(A_i) \leq \beta_i, \quad 0 \leq \alpha_i \leq \beta_i \leq 1\}$$

In this case the proposed scheme can be modified taking into account that

$$\Gamma_{QG} = \bigcup_{\alpha \leq p \leq \beta} \{\pi : \pi(A_i) = p_i \quad \sum_i pi = 1\}$$

where $\alpha = (\alpha_1, \dots, \alpha_n)$ $p = (p_1, \dots, p_n)$ $\beta = (\beta_1, \dots, \beta_n)$ and $\alpha \leq p \leq \beta$ indicates $\alpha_i \leq p_i \leq \beta_i \quad i = 1, \dots, n$.

Thus, to calculate r_k it is enough to consider sequently the linear problems:

$$\max \sum_{i=1}^{k} \max_{\theta_i \in A_i} l(x|\theta_i) p_i$$

$$s.a. \quad \sum_{i=1}^{n} p_i = 1$$

$$\alpha_i \leq p_i \leq \beta_i \quad i = 1, \dots, n$$

with optimum $p_1^*, \dots, p_k^*$ and

$$\min \sum_{i=k+1}^{n} \min_{\theta_i \in A_i} l(x|\theta_i) p_i$$

$$s.a. \sum_{i=1}^{k} p_i^* + \sum_{i=k+1}^{n} p_i = 1$$

$$\alpha_i \leq p_i \leq \beta_i \quad i = 1, \dots, n$$

with optimum $p_{k+1}^*, \dots, p_n^*$ and we replace in the algorithm p_i for p_i^*. A similar modification give us the values μ_* and a_*.

A natural extention of the quantile class for continuous parameters is the class

$$\Gamma_{LU} = \{\pi : L(A) \leq \pi(A) \leq U(A), \quad \forall A \in \beta\}$$

where β is a σ-field on the state set Θ. This is the class studied, among others, by Moreno and Pericchi [10] who provide the following result for posterior probabilities of sets in β.

Theorem 2 *Let A be an arbitrary set in* β. *Suppose that* $l(x|\theta)$, *L and U satisfy* $U([l(x|\theta) = z]) = L([l(x|\theta) = z]) = 0$ *for any* $z \geq 0$ *where* $[l(x|\theta) = z] = \{\theta : l(x|\theta) = z\}$. *Then, we have*

$$\text{(i)} \quad \text{if} \quad U(A) + L(A^c) > 1 \quad \text{then} \sup_{\pi \in \Gamma_{LU}} P^{\pi}(A|x) = P^{\pi_0}(A|x)$$

$$\text{where} \quad \pi_0(d\theta) = U(d\theta) I_{A \bigcap [l(x|\theta) \geq z_A]}(\theta) + L(d\theta) I_{A \bigcap [l(x|\theta) < z_A] \bigcap A^c}(\theta)$$

$$z_A \quad \text{being such that} \quad \pi_o(\Theta) = 1$$

$$\text{(ii)} \quad \text{if} \quad U(A) + L(A^c) = 1 \quad \text{then} \sup_{\pi \in \Gamma_{LU}} P^{\pi}(A|x) = P^{\pi_0}(A|x)$$

$$\text{where} \quad \pi_0(d\theta) = U(d\theta) I_A(\theta) + L(d\theta) I_{A^c}(\theta)$$

$$\text{(iii)} \quad \text{if} \quad U(A) + L(A^c) < 1 \quad \text{then} \sup_{\pi \in \Gamma_{LU}} P^{\pi}(A|x) = P^{\pi_0}(A|x)$$

$$\text{where} \quad \pi_o(d\theta) = U(d\theta) I_{A \bigcup A^c \bigcap [f(x|\theta) < z_A]}(\theta) + L(d\theta) I_{A^c \bigcap [f(x|\theta) \geq z_A]}(\theta)$$

$$z_A \quad \text{being such that} \quad \pi_o(\Theta) = 1$$

This resuls allow us to compute $\mathcal{ND}(\mathcal{A})$.

Theorem 3 *Let be the class* $\Gamma_{LU} = \{\pi : L(A) \leq \pi(A) \leq U(A), \quad \forall A \in \beta\}$ *and* $\mathcal{L} = \{L_p : L_p(a,\theta) = |a - \theta| + a(2p-1),\ p \in [p_0, p_1]\}$ *and* $l(x|\theta)$, *L and U satisfy* $U([l(x|\theta) = z]) + L([l(x|\theta) = z]) = 0 \quad \forall z \geq 0$ *then*

$$\mu_* = \inf\{\theta \in \Theta : \sup_{\pi \in \Gamma_{LU}} P^{\pi}(-\infty, \theta) \geq p_0\},$$

where P^{π} *denotes the posterior probability.*

Proof. Let $\theta_* = \inf\{\theta \in \Theta : \sup_{\pi \in \Gamma_{LU}} P^{\pi}(-\infty,\theta) \geq p_0\}$. If $\theta < \theta_*$ then $P^{\pi}(-\infty,\theta) < p_0 \quad \forall \pi$. Then by Theorem 1, θ is a dominated alternative. Moreover, if θ_* satisfies $\sup_{\pi \in \Gamma_{LU}} \{P^{\pi}(-\infty,\theta_*) \geq p_0\}$ then there is $\pi_* \in \Gamma_{LU}$ such that $\theta_* \in Q_{\pi_*}(p_0)$. In other case, by previous Theorem 2 there is a sequence of values θ_n such that is exists $\pi_{\theta_n} \in \Gamma_{LU}, \quad \theta_n \in Q_{\pi_{\theta_n}}(p_0)$ with $\theta_n \to \theta_*$ so $\theta_* = \mu_*$. □

Theorem 4 *If L and U verify that* $L(A) > 0 \quad U(A) > 0 \quad \forall A$ *with* $\mu(A) > 0$ *then* $\mu_* = a_*$.

Proof. It is easy to prove that π_0 of theorem 2 has unique quantiles. □

These results can be applied using searching methods based on simulation schemes.

5 Stochastic order applied to the calculus of the non-dominated set

The relationship between a prior distribution $\pi(\theta)$ and its corresponding posterior distribution $\pi(\theta|x)$, through Bayes Theorem, is not simple in the sense that, properties in the prior distribution not always hold for the posterior distribution.

In this context, starting from the class Γ of prior distributions which models the decision maker's uncertainty, it would be greatly useful if one could be able to establish order relationships between the posterior distributions from the order relationships already known among the prior distributions. In other words, given two distributions $\pi_1(\theta)$ and $\pi_2(\theta)$ belonging to the class Γ such that $\pi_1(\theta) \prec \pi_2(\theta)$, where $\prec$ is an order relationship between both distributions, it would be of great interest that this relationship remained in the posterior distributions, this is, that it is verified $\pi_1(\theta|x) \prec \pi_2(\theta|x)$. We find the ideal tool for this study in the general theory of stochastic orders. We introduce a brief summary about the concept of stochastic order between distribution functions and the definitions of various orders. The applications of such orders notably simplifies the calculus of the non-dominated set as we will show.

Let Γ be a family of distribution functions over which a binary relationship, which is a partial order, has been defined "$\prec$". Each time we assess the relationship $F \prec G$, we will extend this order to the random variables $X \prec Y$, where F and G are the distribution functions of X and Y respectively.

Definition 1 *The random variable X is said to be stochastically smaller than the random variable Y, we will denote $F \prec_{st} G$, if $F(x) \geq G(x)$ for every x belonging to $\mathbb{R}$ being F and G the corresponding distribution functions.*

This is the most common order in the stochastic distribution theory. If two random variables are stochastically ordered, this implies that all their location parameters are also ordered. Let us remember that, in many examples in decision theory, the Bayes alternatives are the location parameters of the posterior distributions. Besides, it is immediate, from the definition, that the stochastic order between two variables implies the order between their respective quantiles.

Definition 2 *Given X and Y two continuous random variables with density functions f and g respectively, we will say that X is smaller in likelihood ratio than Y, we will denote $X \prec_{lr} Y$, if*

$$\frac{g(t)}{f(t)} \quad \text{is increasing over the union of the supports of } X \text{ and } Y,$$

where $a/0$ is consider ∞ every time that a is greater than zero.

Given two random variables X and Y it is verified that

$$X \prec_{lr} Y \quad \Rightarrow \quad X \prec_{st} Y$$

see Shaked and Shantikumar [17] and Whitt [20].

As indicated in the beginning of this section, the relationship between the density function of the prior distribution and the posterior distribution is not easily treatable from a mathematic point of view; although, this relationship is more intuitive when we study properties associated to the quotient of two prior distributions. Due to the form of the posterior density function, it is not difficult to translate these properties to the quotient of the respective posterior distributions. In this way we give the following two propositions, easy to prove, but of great use as we show later with various examples.

Proposition 4 *Let $\pi_1(\theta)$ and $\pi_2(\theta)$ be two prior distributions for an unknown parameter of interest. Let $\pi_1(\theta|x)$ and $\pi_2(\theta|x)$ be the respective posterior distributions of the parameter once the sampling experiment information has been incorporated. Then if $\pi_1(\theta) \prec_{lr} \pi_2(\theta)$ it is verified that $\pi_1(\theta|x) \prec_{lr} \pi_2(\theta|x)$. Particularly, it is verified that $\pi_1(\theta|x) \prec_{st} \pi_2(\theta|x)$.*

Example 5 *Let us consider a decision problem where the decision maker's beliefs are modelled by a parametric class of Pareto distributions with unknown parameter α*

$$\Gamma = \{\pi \sim \mathcal{P}(\alpha,\beta) : \alpha \in [\alpha_1, \alpha_2], \alpha_1, \beta > 0\}$$

and the preferences are modelled by the class of quantile loss functions

$$\mathcal{L} = \{L_p : L_p(a,\theta) = |a-\theta| - a(2p-1),\ p \in [p_1, p_2]\}.$$

The class of Pareto distributions can be ordered in the sense of likelihood ratio, since, for any two distributions $\mathcal{P}(\alpha_1,\beta)$, $\mathcal{P}(\alpha_2,\beta)$,

$$\frac{\pi_2(\theta)}{\pi_1(\theta)} = \frac{\alpha_2}{\alpha_1}\left(\frac{\beta}{\theta}\right)^{\alpha_2-\alpha_1}$$

is an increasing function in $[\beta, +\infty)$, as long as $\alpha_1 > \alpha_2$. Then, it is stochastically ordered and, therefore, all the location parameters are ordered, in particular, the quantiles are ordered. If we take a sample of size n of a population that is distributed according to an uniform distribution of mean $\theta/2$, we have that, by proposition 4, the non-dominated set coincides with the closed interval

$$\left[Q_{\mathcal{P}(\alpha_2,\beta)}(p_1), Q_{\mathcal{P}(\alpha_1,\beta)}(p_2)\right].$$

Table 4 shows the non-dominated set for some samples and supposing that $\beta = 4, \alpha_1 = 2, \alpha_2 = 4, p_1 = \frac{1}{3}$ and $p_2 = \frac{1}{3}$.

References

[1] Arias. J.P., El conjunto no dominado en la teoría de la decisión bayesiana robusta, *Ph.D. Thesis*, Universidad de Extremadura, 2002.

n	$X_{(n)}$	$ND(A)$	range
4	3.2	$[4.208, 4.489]$	0.2819
50	4.5	$[4.534, 4.560]$	0.0264
100	4.8	$[4.819, 4.833]$	0.0139
500	5.9	$[5.905, 5.908]$	0.0034
1000	4.7	$[4.702, 4.703]$	0.0013
10000	6.1	$[6.1002, 6.1004]$	0.00017

Table 4: Non-dominated set for the example 5.

[2] Arias, J.P. and Martín, J., Uncertainty in beliefs and preferences: Conditions for optimal alternatives, *Annals of Mathematics and Artificial Intelligence*, 35, 1-4, 3-10, 2002.

[3] Arias, J.P., Martín, J.R., Suárez, A., Bayesian Robustness with convex loss functions, *Technical Paper*, Universidad de Extremadura, avaliable at jparias@unex.es, 2003.

[4] Berger, J., An overview of robust Bayesian analysis (with discussion), *Test*, 3, 5-124, 1994.

[5] Capistrán Carmona, C., A review of Forecast Theory using generalized loss functions, *Working paper*, University of California, San Diego, 2003.

[6] Coello Coello, C. A., Van Veldhuizen, D. A. and Lamont, G. B., *Evolutionary Algorithms for Solving Multi-Objetive Problems*, Kluber A.P., 2002.

[7] Geweke, J., A brief digression on ancillarity and nuisance parameters, *Lecture Notes in Economics, Applies Econometrics II* 8-205, 1997.

[8] Martín, J. and Arias, J.P., Computing the efficient set in Bayesian decision problems, *Robust Bayesian Analysis*, eds. D. Ríos Insua and F. Ruggeri, Lectures Notes in Statistics, vol 152, Springer, 2000.

[9] Moreno, E. and Cano, J.A., Robust Bayesian analysis for ε-contaminations partially known, *Journal Royal Statistical Society B* 53, 143-155, 1991.

[10] Moreno E. and Pericchi R., Bands of probability measures: a robust bayesian analysis, *Bayesian Statistics 4*, Oxford O.P.707-713 (eds. Bernardo *et al*) , 1992.

[11] Nau, Robert F., The shape of incomplete preferences *Working Paper,* Fuqua School of Bussines, Duke University, 1996.

[12] Van Noortwijk, J. and van Gelder, P., Bayesian estimation of quantiles for the purpose of flood prevention. B.L. Edge, ed., *Proceedings of the 26th International Conference on Coastal Engineering*, Copenhagen, Denmark, 3529-3541, 1998.

[13] Ríos Insua, D. and Martín J., On the foundations of robust Decision Making *Decision Theory and Decision Analysis: Trends and Challenges* Rios, S., Kluwer A.P., 1994.

[14] Ríos Insua, D. and Ruggeri, F., *Robust Bayesian Analysis,* Springer, 2000.

[15] Ríos Insua, D., Ruggeri, F. and Martín, J., Bayesian sensitivity analysis: a review, *Sensitivity Analysis*, (eds. Saltelli *et al*), New York: Wiley, 2000.

[16] Roberts, A.W. and Varberg, A.D., *Convex Functions*, Academic Press. New York, 1973.

[17] Shaked, M. and Shanthikumar, J.G., *Stochastic orders and their applications,* Academic Press, New York, 1994.

[18] Seidenfeld, T., Shervish, M.J., and Kadane, J., A representation of partially ordered preferences, *Annals of Statistics,* 23, 2168-2217, 1995.

[19] Weber, M., Decision Making with incomplete information, *European Journal of Operations Research*, 28, 4-57, 1987.

[20] Whitt, W. Uniform conditional variability ordering of probability distributions, *Journal of Applied Probability,* 22, 619-633, 1985.

Pablo Arias belongs to the Department of Mathematics, Universidad de Extremadura, cáceres, 10071, Spain. E-mail: jparias@unex.es

Javier Hernández belongs to the Department of Mathematics, Universidad de Extremadura, Badajoz, 06006, Spain. E-mail: javierhs@materiales.unex.es

Jacinto Martín belongs to the Department of Mathematics, Universidad de Extremadura, Cáceres, 10071, Spain. E-mail: jrmartin@unex.es

Alfonso Suárez belongs to the Department of Statistics, Universidad de Cádiz, Spain E-mail: alfonso.suarez@uca.es