ON THE FOUNDATIONS OF ROBUST BAYESIAN STATISTICS

by

David Ríos Insua and Jacinto Martín

David Ríos Insua
Purdue University
Universidad Politécnica de Madrid

Jacinto Martín
Universidad Politécnica de Madrid

Technical Report #94-35C

Department of Statistics
Purdue University

December 1994

△

ON THE FOUNDATIONS OF ROBUST BAYESIAN STATISTICS

BY

DAVID RÍOS INSUA AND JACINTO MARTÍN

Decision Analysis Group

Universidad Politécnica de Madrid

Abstract

Since its inception, the foundations of Bayesian Statistics and Decision Theory have been criticised on several grounds, specially due to the excessive precision demanded to the judgmental inputs to a Bayesian analysis. This has led to several models allowing for some incompleteness in those inputs. This paper provides a unifying perspective on this problem, giving foundations for decision making under risk, when there is imprecision in the decision maker's preferences, and decision making under uncertainty, when there is imprecision in the decision maker's beliefs and preferences.

(KEYWORDS: Statistics, Decision Theory, Bayesian Analysis, Robustness, Classes of priors, Classes of utilities)

Robust Bayesian Analysis is an increasingly popular topic in Bayesian Statistics, Berger (1994). It stems from the observation that Bayesian analysis places excessive demands to the judgmental inputs (beliefs and preferences) of a Decision Maker. Indeed, since its inception, this has been the main criticism to the foundations of Bayesian Decision Theory and Inference, Savage

(1954). Even several criticisms of these foundations on different grounds, see e.g. Schmeidler (1989), have as starting point the incompleteness of a Decision Maker's judgments. The acknowledgment of this incompleteness has also led to work in areas such as stochastic dominance (Levy, 1992), sensitivity analysis (Rios Insua, 1990), error modeling (Lindley, Tversky and Brown, 1979) and alternative models of Decision Making and Inference (Nau, 1992; Gilboa and Schmeidler, 1993).

There is substantial work providing foundations for some of these models, in limited contexts, since they assume that preferences are precise, but not beliefs, or the other way round. For example, Aumann (1962) provides foundations leading to models of incomplete preferences by classes of utility functions in finite consequence spaces; Giron and Rios (1980) provide results leading to models of incomplete beliefs by classes of probability distributions, when preferences over consequences are precise.

This paper provides a unifying and more general perspective on this problem. The general theme is that appropriate results from functional analysis together with convenient separability conditions may be used to provide foundations for the models suggested above. Section 1 discusses in several directions some basic results in functional analysis. Section 2 deals with decision making under risk; we show how to model imprecision in preferences by means of a class of utility functions, in a very general context. Section 3 deals with decision making under uncertainty, within Anscombe and Aumann (1963) framework. First, we give expected utility representations with classes of state dependent utilities, and then a representation with classes of state independent utilities. We address also issues of modeling and updating beliefs.

Seidenfeld, Schervish and Kadane (1992) (SSK from now on) and Nau (1994) provide related approaches to the problem. They obtain a representation of preferences in terms of a class of expected utilities similar to our Theorem 2. Our setting is more general in that we allow for continuous state and consequence spaces. Then, SSK go on to prove that preferences are represented by a class of probability/utility pairs, where the utilities are almost state-independent, in the presence of two axioms relating to state-independent utility. Nau shows that adding the sure-thing principle allows to represent one of the expected utilities as a probability/utility pair; in order to get a representation with state independent utilities, an additional condition is required. Our main result is a representation in terms of a class of probabilities and a class of state independent utilities arbitrarily paired, obtained with the aid of a separability condition. Our representation is less general but more convenient computationally. We are also able to deal with the continuous case. It also fits better within general approaches to robust Bayesian analysis commonly proposed, see e.g. Berger (1994).

As usual, given a relation $\preceq$ in a set X, $\prec$ and $\sim$ will designate, respectively, the associated strict preference and indifference relations.

1 Modeling quasi orders in linear spaces

Most of the results we shall deal with may be framed into the following lemma. We shall discuss it in detail, specially questions concerning boundedness, uniqueness and connections with related results. We sketch the proof.

For a full one, see Rios Insua (1992).

Lemma 1 *Let X be a convex set in a normed real space Y and $\preceq$ a binary relation on it. Then, the following three conditions*

A1. $(X, \preceq)$ *is a quasi order (transitive and reflexive).*

A2. *For* $\alpha \in (0,1), x, y, z \in X$, $x \preceq y \iff \alpha x + (1-\alpha) z \preceq \alpha y + (1-\alpha) z$.

A3. *For* $x, y, z, r \in X$, $(\alpha x + (1-\alpha) y \preceq \alpha z + (1-\alpha) r, \forall \alpha \in (0,1]) \Longrightarrow y \preceq r$.

are equivalent to the existence of a set W of continuous linear functions w on Y such that, $\forall x, y \in X$,

$$x \preceq y \iff (w(x) \leq w(y),\ \forall w \in W).$$

Proof. Define $E = \{z : z = y - x,\ x, y \in X,\ x \preceq y\}$, $S = \{z : z = \sum_{i=1}^{n} \alpha_i z_i,\ z_i \in E,\ \alpha_i \geq 0\}$. The closure of S, $cl(S)$, is a closed, convex cone (in a normed real space). There is a family W of continuous linear functions w on Y such that $z \in cl(S) \iff (w(z) \geq 0,\ \forall w \in W)$. Prove now that for $x, y \in X$, $y - x \in cl(E) \iff x \preceq y$. Consequently, for $x, y \in X$, $x \preceq y \iff (w(x) \leq w(y),\ \forall w \in W)$.

The converse is immediate. □

Note that the result is a modification of Von Neumann-Morgestern conditions (Fishburn, 1970), in that we do not require the relation to be complete and we adopt a different continuity condition (A3). Indeed, we easily get:

Corollary 1 *Under the conditions of Lemma 1, if, in addition,* $(X, \preceq)$ *is complete, there is a function* w *on* X *such that* $\forall x, y \in X, \alpha \in (0,1)$,

$$\begin{aligned} x \preceq y &\iff w(x) \leq w(y), \\ w(\alpha x + (1-\alpha)y) &= \alpha w(x) + (1-\alpha) w(y). \end{aligned}$$

In some applications, we are interested in bounded functionals. Condition A4 in Corollary 2 is a standard assumption on the existence of best and worst elements. SSK (1992) derive it from basic principles.

Corollary 2 *Under the conditions of Lemma 1, if, in addition,*

X *is compact, or,*

A4. There are $x_*, x^* \in X$ *such that* $x_* \preceq x \preceq x^*$, $\forall x \in X$,

then the linear functions w *are bounded in* X.

In the first case, the w's are continuous in a compact set X, therefore they are bounded. Under A4, it is $w(x_*) \leq w(x) \leq w(x^*)$, $\forall x \in X, \forall w \in W$. □

It is also interesting to determine uniqueness conditions of W. By construction, W is a maximal set. There may be subsets of W representing the same order. The following result provides relations between classes of functionals representing $\preceq$. Its proof is simple.

Lemma 2 *Suppose a family* W *of continuous linear functions represents* $(X, \preceq)$, *as in Lemma 1. Then, if* W' *is another such family,* $\forall w' \in W'$, *there are* $\lambda_1, ..., \lambda_n > 0$, $w_1, \ldots, w_n \in W$ *and* β *such that*

$$w' = \sum_{i=1}^{n} \lambda_i w_i + \beta.$$

Note also that if W' is a family of linear functions representing $(X, \preceq)$ as above, then $[W']$ and $gen(W')$ will generate $(X, \preceq)$, where $[W']$ and $gen(W')$ are, respectively, the convex hull and the set of generators of W'.

We conclude this section with two other results leading to the conclusions of Lemma 1. Their proof is very similar. SSK use a result resembling Lemma 3, with a very different proof, to show the existence of W, in a specific context. Giron and Rios (1980) and Walley (1991) use, essentially, Lemma 4, in a particular context, to model incompleteness in beliefs by a class of probability distributions.

Lemma 3 *Let X be a convex set in a normed real space Y and $\preceq$ a binary relation on it. Then, conditions A1, A2 and*

A3'. For $\{x_n\}, \{y_n\}, x, y \in X$ such that $x_n \rightarrow x$, $y_n \rightarrow y$ and $x_n \preceq y_n$, $\forall n$, it is $x \preceq y$.

are equivalent to the existence of a set W of continuous linear functions w on Y such that, $\forall x, y \in X$,

$$x \preceq y \iff (w(x) \leq w(y), \; \forall w \in W).$$

Lemma 4 *Let X be a convex, symmetric with respect to $\mathbf{0}$, set in a normed real space Y and $\preceq$ a binary relation on it. Then, conditions A1, A2 and*

A3". For $\{x_n\}, x, y, z \in X$ such that $x_n \rightarrow x$ and $y \preceq x_n \preceq z$, $\forall n$, it is $y \preceq x \preceq z$.

are equivalent to the existence of a set W of continuous linear functions w on Y such that, $\forall x, y \in X$,

$$x \preceq y \iff (w(x) \leq w(y), \; \forall w \in W).$$

Lemmas 1, 3 and 4 differ in the continuity condition, which, in turn, is different to Von Neumann and Morgestern's. These continuity conditions are essentially equivalent if conditions A1, A2 hold and the relation is complete. This is not the case when completeness does not hold. Note, however, that A3' implies A3 and A3".

2 Decision making under risk

Our first application will be to decision making problems under risk. In Rios Insua (1992), we provide an expected utility representation via a class of utility functions, in the simple case. See also related results in Aumann (1964) and Fishburn (1970, 1982). We shall study here the general case adopting some boundedness conditions, and a proof based on distribution functions rather than lotteries (or probability distributions over the space $\mathcal{C}$ of consequences).

Theorem 1 *Let $\mathcal{C} \subset \mathbf{R}^n$, $\mathbf{F}$ be a convex set of distribution functions whose support is compact and contained in $\mathcal{C}$, and includes the set of degenerate distribution functions. Let $\preceq_*$ be a binary relation in $\mathbf{F}$. Then,*

U1. $(\mathbf{F}, \preceq_*)$ *is a quasi order.*

U2. For $\alpha \in (0,1), F, G, H \in \mathbf{F}$,

$$F \preceq_* G \iff \alpha F + (1-\alpha)H \preceq_* \alpha G + (1-\alpha)H.$$

U3. For $F, G, H, L \in \mathbf{F}$, $(\alpha F+(1-\alpha)G \preceq_* \alpha H+(1-\alpha)L, \forall \alpha \in (0,1]) \Longrightarrow G \preceq_* L.$

are equivalent to the existence of a class $\mathcal{U}$ of real functions u on $\mathcal{C}$ such that

$$F \preceq_* G \iff \left(\int u dF \leq \int u dG,\ \forall u \in \mathcal{U}\right).$$

Proof. We are under the conditions of Lemma 1: **F** is a convex subset of the normed space $\mathcal{F}$ of real fuctions on $\mathcal{C}$, the norm being

$$\|f\| = \sup_{c \in \mathcal{C}} |f(c)|,$$

and A1, A2, A3 hold. Therefore, there is a family W of continuous functions w such that

$$\begin{aligned} F \preceq_* G &\iff (w(F) \leq w(G), \forall w \in W), \\ w(\alpha F + (1-\alpha)G) &= \alpha w(F) + (1-\alpha) w(G), \end{aligned}$$

for $F, G \in \mathbf{F}, \alpha \in (0,1)$. We shall prove that $w(F) = \int u dF$ for some u.

Let us fix $w \in W$. For $c \in \mathcal{C}$, define $u(c) = w(F_c)$, where F_c is the distribution function degenerate at c. Then, $w(F_c) = \int u dF_c$. Let F be the distribution function associated with a simple probability distribution on $\mathcal{C}$. Then, $F = \sum_{i=1}^n p(c_i) F_{c_i}$, where $p(c_i)$ is the probability at c_i. Due to linearity of w,

$$w(F) = \sum_{i=1}^n p(c_i) w(F_{c_i}) = \sum_{i=1}^n p(c_i) u(c_i) = \int u dF.$$

Note now that u is locally bounded. This is an immediate consequence of the fact that if $\alpha_n \to 1$ and $c_n \to c$, since $(1-\alpha_n) F_{c_n} + \alpha_n F_c \to F_c$, we shall have $(1-\alpha_n) u(c_n) + \alpha_n u(c) \to u(c)$. Suppose that F has compact support $K \subset \mathcal{C}$. Then u is bounded in K.

Assume that F is continuous. For a given n, and $i = 1, \ldots, 2^n$, let $A_{in} = \{c : (i-1)/2^n < F(c) \leq i/2^n\}$ and $c_{in} \in \arg\inf_{c \in cl(A_{in})} u(c)$. Let $F_n =$

$\sum_{i=1}^{2^n}(1/2^n)F_{c_{in}}$. Then, by continuity, $w(F_n)\rightarrow w(F)$. Moreover, $w(F_n) = \sum_{i=1}^{2^n} u(c_{in})/2^n = \int u_n dF$, with $u_n(c) = \sum_{i=1}^{2^n} u(c_{in})I_{A_{in}}(c)$, where $I_{A_{in}}$ is the indicator function of A_{in}. $\{u_n\}$ is a sequence of simple functions such that $u_n \uparrow u$. Therefore, $\int u_n dF \rightarrow \int udF$, and $w(F) = \int udF$.

When F has a finite number of discrete jumps, we may write it as $F = \alpha F_1 + (1-\alpha)F_2$, where F_1 is simple, F_2 is continuous and $\alpha \in [0,1]$. Then,

$$w(F) = \alpha w(F_1) + (1-\alpha)w(F_2) = \alpha \int udF_1 + (1-\alpha)\int udF_2$$

$$= \int ud(\alpha F_1 + (1-\alpha)F_2) = \int udF.$$

When F has a countable number of discrete jumps, a convergence argument follows easily. □

The same proof applies in these two cases:

Corollary 3 *The result in Theorem 1 holds when* $\mathbf{F}$ *is the set of all distribution functions over* $\mathcal{C}$ *if, in addition,*

$\mathcal{C}$ *is compact, or*

U4. There are $F_*, F^* \in \mathbf{F}$ *such that* $F_* \preceq_* F \preceq_* F^*$, $\forall F \in \mathbf{F}$.

In what follows, whenever required, we shall write $p \preceq_* q$ when $F_p \preceq_* F_q$ and $w(p) = w(F_p)$, with F_p, F_q the distribution functions associated with lotteries p and q, respectively.

3 Decision making under uncertainty

Our second application is to decision making problems under uncertainty, under Anscombe and Aumann (1963) framework. In Rios Insua (1992), we deal with this problem in the simple case. Here we shall deal with the general case, using weaker conditions.

Our first result will be a representation in terms of a family of cardinal utility functions. The basic elements will be a space S of states; a space $\mathcal{C}$ of consequences, endowed with an algebra $\mathcal{B}$; the set $\mathcal{P}$ of probability distributions over $\mathcal{C}$; the set $\mathcal{G}$ of Anscombe-Aumann acts (or functions from S into $\mathcal{P}$, so that $\mathcal{G} = \mathcal{P}^S$); and a binary relation $\preceq$ in $\mathcal{G}$ modeling preferences. Given $p \in \mathcal{P}$, we define the constant act $p \in \mathcal{G}$, such that $p(s) = p$, $\forall s \in S$. In such a way, we assume that $\preceq$ on $\mathcal{G}$ induces a relation $\preceq$ on $\mathcal{P}$, in a natural way, that is $p \preceq q$ if $pI_S \preceq qI_S$.

Theorem 2 *Let S be a set, $\mathcal{C} \subset \mathbf{R}^n$, $\mathcal{B}$ an algebra over $\mathcal{C}$, and $\mathcal{P}$ the set of probability distributions over $(\mathcal{C}, \mathcal{B})$. Let $\preceq$ be a binary relation in $\mathcal{G} = \mathcal{P}^S$. Then, the three conditions*

V1. $(\mathcal{G}, \preceq)$ *is a quasi order.*

V2. For $\alpha \in (0,1), f, g, h \in \mathcal{G}$, $f \preceq g \iff \alpha f + (1-\alpha)h \preceq \alpha g + (1-\alpha)h$.

V3. For $f, g, h, l \in \mathcal{G}$, $(\alpha f + (1-\alpha)g \preceq \alpha h + (1-\alpha)l, \forall \alpha \in (0,1]) \Longrightarrow g \preceq l$.

are equivalent to the existence of a class V of functions v on $\mathcal{G}$ such that

$$f \preceq g \iff (v(f) \leq v(g), \forall v \in V),$$

$$v(\alpha f + (1-\alpha)g) = \alpha v(f) + (1-\alpha)v(g),$$

for $f, g \in \mathcal{G}, \alpha \in (0,1)$.

Proof. $\mathcal{G}$ is a convex subset of the normed space $\mathcal{H} = \{g : S \to \mathcal{X}\}$, where $\mathcal{X}$ is the set of additive real functions in $(\mathcal{C}, \mathcal{B})$, endowed with the norm $\|g\| = \sup_{s \in S} \|g(s)\|$ and $\|g(s)\| = \sup_{C \in \mathcal{B}} |g(s)(C)|$. The result follows from Lemma 1. □

We study now expected utility representations. First, we consider state dependent utilities, assuming there is an underlying measure μ, and under the additional condition $V4$. The result is partly based on Fishburn (1970, thm 13.1), following a suggestion in SSK. Note, however, that we deal with the more general case of incomplete preferences and infinite state spaces.

Theorem 3 *Under the conditions of Theorem 2, suppose that, in addition, $\mathcal{A}$ is a σ-algebra on S and*

V4. There is a nonatomic finite measure μ on $\mathcal{A}$ such that, for $f, g \in \mathcal{G}$

$$\mu(A) = 0,\ f = g \text{ on } A^c \Longrightarrow f \sim g.$$

Then, there is a class $\mathcal{L}$ of real functions l on $S \times \mathcal{P}$ such that

$$f \preceq g \iff \left(\int l(s, f(s))d\mu(s) \le \int l(s, g(s))d\mu(s), \forall l \in \mathcal{L}\right),$$

whenever f, g are simple Anscombe-Aumann acts.

Proof. $\mathcal{G}$ is a convex subset of the space $\mathcal{H} = \{g : S \to \mathcal{X}\}$, where $\mathcal{X}$ is the set of additive real functions in $(\mathcal{C}, \mathcal{B})$. Consider the equivalence relation $\approx$

on $\mathcal{H}$

$$f \approx g \iff f = g \text{ on } A^c, \text{ with } \mu(A) = 0.$$

$\mathcal{H}/\!\approx$ is a normed space with norm $\|\mathbf{g}\| = \int_S \|g(s)\| d\mu(s)$, with $\mathbf{g} \in \mathcal{H}/\approx$, $g \in \mathbf{g}$ and $\|g(s)\| = \sup_{C \in \mathcal{B}} |g(s)(C)|$. Moreover, $\mathcal{G}/\!\approx$ is a convex subset of $\mathcal{H}/\!\approx$, and the binary relation $\preceq'$ in $\mathcal{G}/\!\approx$

$$\mathbf{f} \preceq' \mathbf{g} \iff \exists\, f \in \mathbf{f},\ g \in \mathbf{g} \text{ such that } f \preceq g,$$

is well-defined and satisfies conditions A1, A2 and A3 of Lemma 1. Therefore, there is a class V of linear continuous functions v on $\mathcal{G}/\!\approx$ such that $\mathbf{f} \preceq' \mathbf{g} \iff (v(\mathbf{f}) \leq v(\mathbf{g}),\ \forall\ v \in V)$. For $f \in \mathcal{G}$, define $v(f) = v(\mathbf{f})$, with $f \in \mathbf{f}$. Then, $f \preceq g \iff (v(f) \leq v(g),\ \forall\ v \in V)$. Take a constant act $p \in \mathcal{G}$ and define a function v' as follows: $v'(f) = v(f) - v(p)$, for each $v \in V$. We have $f \preceq g \iff (v'(f) \leq v'(g), \forall v \in V)$, $v'(p) = 0$ and $v'(\alpha f + (1-\alpha)g) = \alpha v'(f) + (1-\alpha)v'(g)$, for $\alpha \in (0,1)$, $f, g \in \mathcal{G}$.

Given $A \in \mathcal{A}$, $q \in \mathcal{P}$ define $u(A,q) = v'(qI_A + pI_{A^c})$. We prove now that $u(.,q)$ is a σ-additive function in $\mathcal{A}$, which is absolutely continuous with respect to μ. Prove first that $u(.,q)$ is additive in $\mathcal{A}$. Choose $A, B \in \mathcal{A}$ and disjoint. Then,

$$u(A,q) + u(B,q) = v(qI_A + pI_{A^c}) + v(qI_B + pI_{B^c}) - 2v(p) =$$

$$v(qI_{A\cup B} + pI_{(A\cup B)^c}) - v(p) = u(A \cup B, q).$$

Prove now that $u(.,q)$ is σ-additive. Let $\{A_i\}_{i=1}^{\infty}$ be a collection of mutually disjoint sets in $\mathcal{A}$. Let $B_n = \bigcup_{i=1}^{n} A_i$, $A = \bigcup_{i=1}^{\infty} A_i$. Then,

$$qI_{B_n} + pI_{B_n^c} \overset{n\to\infty}{\longrightarrow} qI_A + pI_{A^c},$$

since

$$\|qI_{B_n} + pI_{B_n^c} - (qI_A + pI_{A^c})\| = \|(p-q)I_{A\setminus B_n}\| =$$

$$\int \|(p-q)I_{A\setminus B_n}(s)\| d\mu(s) = \|p-q\| \int_{A\setminus B_n} d\mu(s) \leq \mu(\bigcup_{i=n}^{\infty} A_i)$$

Since the functions v' are continuous, we have $u(B_n, q) \overset{n\to\infty}{\longrightarrow} u(A, q)$. Due to additivity, $u(B_n, q) = \sum_{i=1}^{n} u(A_i, q)$. Therefore, $\sum_{i=1}^{\infty} u(A_i, q) = u(\bigcup_{i=1}^{\infty} A_i, q)$. μ-absolute continuity of $u(., q)$ follows immediately.

Then, by Radon-Nikodym theorem, see e.g. Ash (1971),

$$u(A, q) = \int_A l(s, q) d\mu(s),$$

for some function $l(., q)$ in S.

Let $f : S \to \mathcal{P}$ be a simple Anscombe-Aumann act, i.e., $f = \sum_{i=1}^{n} p_i I_{A_i}$, with $p_i \in \mathcal{P}$ and the sets A_i forming a measurable partition of S. Then,

$$f + (n-1)p = \sum_{i=1}^{n} (p_i I_{A_i} + pI_{A_i^c}),$$

and

$$(1/n)v(f) + ((n-1)/n)v(p) - v(p) = (1/n)v(\sum_{i=1}^{n}(p_i I_{A_i} + pI_{A_i^c})) - v(p),$$

that is

$$v'(f) = \sum_{i=1}^{n} v'(p_i I_{A_i} + pI_{A_i^c})$$

$$= \sum_{i=1}^{n} u(A_i, p_i) = \sum_{i=1}^{n} \int_{A_i} l(s, p_i) d\mu(s) = \int l(s, f(s)) d\mu(s).$$

□

The extension of the results to, say, continuous acts, require the adoption of some boundedness condition, as follows:

Proposition 1 *Under the conditions of Theorem 3, suppose that, in addition:*

There are constant acts f_*, f^* *such that* $f_* \preceq f \preceq f^*$, $\forall f \in \mathcal{G}$.

Then, if S *is compact, the conclusion of Theorem 3 holds also for continuous Anscombe-Aumann acts.*

Proof. Before dealing with the proof, we shall provide some additional facts about functions $u(A, q)$. Assume $q_n \to q$; then, since $q_n I_A + p I_{A^c} \to q I_A + p I_{A^c}$, we have $u(A, q_n) \to u(A, q)$, since v is continuous. Therefore, $\int_A l(s, q_n) d\mu(s) \to \int_A l(s, q) d\mu(s)$. Since this holds $\forall A \in \mathcal{A}$, we conclude that $l(., q_n) \xrightarrow{\mu} l(., q)$, in measure μ.

Let $f : S \to \mathcal{P}$ be continuous. For each s, n define $A_n(s) = \{t : \|f(t) - f(s)\| < 1/n\}$. These sets form an open cover of S: then, there is a finite subcover $\{A_{n1}, \ldots, A_{nm_n}\}$. Let $B_{n1} = A_{n1}$, $B_{ni} = A_{ni} \setminus \bigcup_{j=1}^{i-1} B_{nj}$, $i = 1, \ldots, m_n$. Choose $s_{ni} \in cl(B_{ni})$ arbitrarily and define $f_n = \sum_{i=1}^{n} f(s_{ni}) I_{B_{ni}}$. Then, $f_n \to f$, so that, by continuity, $v'(f_n) \to v'(f)$. By Theorem 3, $v'(f_n) = \int l(s, f_n(s)) d\mu(s)$. Now, since $f_n(s) \to f(s), \forall s$, uniformly, $l(s, f_n(s)) \to l(s, f(s))$ in measure.

Then, taking into account the boundedness condition, we may appeal to an extended dominated convergence theorem so that

$$\int l(s, f_n(s)) d\mu(s) \to \int l(s, f(s)) d\mu(s).$$

Consequently,

$$v'(f) = \int l(s, f(s))d\mu(s)$$

□

Combining the two previous results, we may provide a state-dependent utility representation for a wide class of Anscombe-Aumann acts.

Corollary 4 *Under the conditions of Theorem 3 and the boundedness condition of Proposition 1, the result holds for Anscombe-Aumann acts which are continuous in a compact set and simple outside, and for convex combinations of those acts.*

We look now for subjective expected utility representations. First, we show how to model incomplete beliefs by means of a class of probability distributions. These are updated applying Bayes' theorem to each of the distributions in the class. Our first result models beliefs through a class of finitely additive probability distributions, with the help of a weak version of the sure-thing principle (V6) and a nontriviality condition (V5). We may view (V6) also as a monotonicity or dominance condition.

Theorem 4 *Under the conditions of Theorem 2, suppose that, in addition, $\mathcal{A}$ is an algebra on S and there are constant acts p, q such that*

V5. $p \prec q$.

V6. For $A, B \in \mathcal{A}$, $A \subset B$, $pI_{A^c} + qI_A \preceq pI_{B^c} + qI_B$.

Then, if we define,

$$A \preceq_\ell B \iff pI_{A^c} + qI_A \preceq pI_{B^c} + qI_B,$$

for $A, B \in \mathcal{A}$, there is a set P of (finitely additive) probability distributions r such that

$$A \preceq_\ell B \iff (r(A) \leq r(B),\ \forall r \in P).$$

Proof. Under the conditions of Theorem 2, there is a family V of additive functions v such that

$$A \preceq_\ell B \iff (v(pI_{A^c} + qI_A) \leq v(pI_{B^c} + qI_B),\ \forall v \in V).$$

Let V' be the set of functions such that $v(p) < v(q)$, which is nonempty, due to V5. Define V'' as follows; for each $v \in V'$, define $v' \in V''$ according to

$$v'(.) = \frac{v(.) - v(p)}{v(q) - v(p)}.$$

We have, therefore, $A \preceq_\ell B \iff (v'(pI_{A^c}+qI_A) \leq v'(pI_{B^c}+qI_B), \forall v' \in V'')$. Define, for each $v' \in V''$,

$$r_{v'}(A) = v'(pI_{A^c} + qI_A),$$

and $P = \{r :\ r = r_{v'},$ for some $v' \in V''\}$. Then,

$$A \preceq_\ell B \iff (r(A) \leq r(B),\ \forall r \in P).$$

Moreover, $\forall r \in P$, and $A, B \in \mathcal{A}$ which are disjoint,

$$r(S) = v'(q) = 1.$$

$$0 = v'(p) \leq v'(pI_{A^c} + qI_A) = r(A).$$

$$r(A \cup B) = v'(pI_{(A \cup B)^c} + qI_{A \cup B}) + v'(p) =$$

$$v'(pI_{A^c} + qI_A) + v'(pI_{B^c} + qI_B) = r(A) + r(B).$$

□

A possible criticism is that $\preceq_\ell$ depends on p, q. Ways of overcoming this include

- Add as an axiom a definition of $\preceq_\ell$ independent of p, q such as

  $$pI_{A^c} + qI_A \preceq pI_{B^c} + qI_B \iff p'I_{A^c} + q'I_A \preceq p'I_{B^c} + q'I_B,$$

  whenever $p \prec q$, $p' \prec q'$, similar to Savage's (1954) fourth axiom.

- Admit that, due to incompleteness in judgments, the above condition might not hold and modify the definition of $\preceq_\ell$ as follows:

  $$A \preceq_\ell B \iff (pI_{A^c} + qI_A \preceq pI_{B^c} + qI_B, \forall p, q : p \prec q).$$

  This second alternative requires slight modifications in the proof of Theorem 4.

- Replace V6 with a stronger dominance condition such as

  $$f(s) \preceq g(s), \forall s \Longrightarrow f \preceq g.$$

  Schmeidler (1989) shows that, in the presence of completeness, this condition is equivalent to the sure-thing principle. See our discussion below on this principle.

Since information evolves in time, we need a procedure to update beliefs. By imposing an additional condition, we may do it applying Bayes' formula to each distribution in the class. $B|D$ will represent event B, given event D.

Proposition 2 *Under the conditions of Theorem 4, let* $D \in \mathcal{A}$ *such that* $\emptyset \prec_\ell D$. *Suppose that, in addition,*

V7. For $B, C \in \mathcal{A}$, $B|D \preceq_\ell C|D \iff B \cap D \preceq_\ell C \cap D$.

Then, there is a family P *of probability distributions* r*, such that*

$$B \preceq_\ell C \iff (r(B) \leq r(C), \forall r \in P),$$
$$B|D \preceq_\ell C|D \iff (r(B|D) \leq r(C|D), \forall r \in P : r(D) > 0).$$

Proof. Build P as in Theorem 4. Then,

$$B|D \preceq_\ell C|D \iff B \cap D \preceq_\ell C \cap D \iff (r(B \cap D) \leq r(C \cap D), \forall r \in P)$$

$$\iff (r(B \cap D)/r(D) \leq r(C \cap D)/r(D), \forall r \in P : r(D) > 0) \iff$$

$$(r(B|D) \leq r(C|D), \forall r \in P : r(D) > 0).$$

□

Clearly, those $r \in P$ such that $r(D) = 0$ are irrelevant in the case above, since $r(B \cap D) = r(C \cap D) = 0$. Also, $\{r \in P : r(D) > 0\} \neq \emptyset$, since $\emptyset \prec_\ell D$.

We might wish to have a class of σ-additive probability distributions representing $\preceq_\ell$. Proposition 3 deals with such question. Its proof follows combining ideas from Theorems 3 and 4.

Proposition 3 *Under the conditions of Theorem 2, suppose that $\mathcal{A}$ is a σ-algebra on S and $(\mathcal{G}, \preceq)$ verifies V4, V5 and V6. Then, the probability distributions r in P modeling $\preceq_\ell$ are σ-additive, and absolutely continuous with respect to μ.*

Once solved the issue of modeling beliefs, we may go back to subjective expected utility representations. Note first that the representation in Theorem 3 uses the underlying measure μ for each v'. Under the conditions of Proposition 3, we have a natural family of probability distributions modeling beliefs and we may attain subjective (state dependent) expected utilities.

Corollary 5 *Under the conditions of Theorem 4 and Proposition 3, there is a class $\mathcal{W}$ of pairs of probability distributions r and state dependent utilities $u(.,.)$ such that*

$$f \preceq g \iff \left(\int u(s, f(s))dr(s) \leq \int u(s, g(s))dr(s), \forall (r, u) \in \mathcal{W}\right),$$

for acts f, g as in Theorem 3.

Proof. Let p, q be as in Theorem 4. Suppose that $v(p) < v(q)$. Define r as in Theorem 4; we have $r(A) = \int_A m(s)d\mu(s)$. If $m(s) \neq 0$, define $u(s, f(s)) = l(s, f(s))/m(s)$. If $m(s) = 0$, define $u(s, f(s)) = l(s, f(s))$. If $v(p) = v(q)$, define $r = \mu/\mu(S)$, $u = l$. □

We may consider whether adding a sure-thing principle to axioms leads to state-independent subjective expected utility representations. Proposition 4 below provides a partial answer to our enquiry, in the sense that expected utilities are represented as convex combinations of state independent expected utilities. Actually, Nau (1994) proves that, in the finite case, one of the expected utilities in $gen(V)$, see Theorem 2, is state independent and that to attain separability of all generators we have to add a stronger condition. SSK (1992) provide an almost state-independent representation, in the presence of a sure-thing principle and a stochastic dominance condition.

For this result, we assume finite set of states and space of consequences. We need first a definition of a non-null state.

Definition 1 *State $s \in S$ is non-null if*

$$p \prec pI_{S\setminus\{s\}} + qI_{\{s\}}$$

whenever $p \prec q$.

Proposition 4 *Assume the conditions of Theorem 2, with finite S and $\mathcal{C}$, the nontriviality condition V5 and a sure thing principle like:*

> *If $h \in \mathcal{G}, p, q \in \mathcal{P}$ are such that*
>
> $$hI_{S\setminus\{s\}} + pI_{\{s\}} \preceq hI_{S\setminus\{s\}} + qI_{\{s\}}$$
>
> *for s non-null, then*
>
> $$hI_{S\setminus\{s'\}} + pI_{\{s'\}} \preceq hI_{S\setminus\{s'\}} + qI_{\{s'\}}$$
>
> *for any other s' non-null.*

hold. Then, there is a class of expected utilities representing the order and those expected utilities can be represented as convex combinations of state independent subjective expected utilities.

Proof. Under the conditions of Theorem 2, and S being finite, a standard argument, see Kreps (1988), leads to the state-dependent expected utility representation

$$f \preceq g \iff (\sum_{s\in S}\sum_{c} u^v(s,c)f(s)(c) \leq \sum_{s\in S}\sum_{c} u^v(s,c)g(s)(c), \forall v \in gen(V)),$$

where $f(s)(c)$ is the probability of obtaining the consequence c under state s for act f, and $u^v(s,c)$ is the utility of consequence c under state s, associated with expected utility v. Assume that s, s' are non-null. Then, for $h \in \mathcal{G}$

$$(\sum_{c} u^v(s,c)p(c) \leq \sum_{c} u^v(s,c)q(c), \forall v \in gen(V)) \iff$$

$$hI_{S\setminus\{s\}} + pI_{\{s\}} \preceq hI_{S\setminus\{s\}} + qI_{\{s\}} \iff$$

$$hI_{S\setminus\{s'\}} + pI_{\{s'\}} \preceq hI_{S\setminus\{s'\}} + qI_{\{s'\}} \iff$$

$$(\sum_{c} u^v(s',c)p(c) \leq \sum_{c} u^v(s',c)q(c), \forall v \in gen(V)).$$

Thus, first and last inequalities define quasiorders in $\mathcal{P}$, $\preceq_s$ and $\preceq_{s'}$, which coincide. We appeal now to Lemma 2 and its ensuing discussion. Fix a non-null state s_0. Then, each $u^v(s,.)$ may be written

$$u^v(s,c) = \sum_{i=1}^{m_v} a_s^{v,i} u^{v,i}(s_0,c) + b_s^v.$$

Therefore, we may write, for $v \in gen(V)$, after some algebra,

$$v(f) = \sum_{s\in S}(\sum_{i=1}^{m_v} a_s^{v,i}(\sum_{c} u^{v,i}(s_0,c)f(s)(c))) + \sum_{s\in S} b_s^v,$$

so that

$$f \preceq g \iff$$

$$(\sum_{s\in S}\sum_{i=1}^{m_v} a_s^{v,i}\sum_c u^{v,i}(s_0,c)f(s)(c) \leq \sum_{s\in S}\sum_{i=1}^{m_v} a_s^{v,i}\sum_c u^{v,i}(s_0,c)g(s)(c), \forall v \in gen(V)).$$

Define $u^{v,i}(s_0,c) = u^{v,i}(c)$, $\lambda_i^v = \sum_{s\in S} a_s^{v,i}$ and $\mu_v^i(s) = a_s^{v,i}/\lambda_i^v$, since we may assume that $\lambda_i^v > 0$. Then, interchanging the summations, we get

$$f \preceq g \iff (\forall v \in gen(V),$$

$$\sum_{i=1}^{m_v}\lambda_i^v \sum_{s\in S}\mu_v^i(s)\sum_c u^{v,i}(c)f(s)(c) \leq \sum_{i=1}^{m_v}\lambda_i^v \sum_{s\in S}\mu_v^i(s)\sum_c u^{v,i}(c)g(s)(c)).$$

If required, we may normalise the weights λ_i^v. □

Nau (1994) shows that, for one $v \in gen(V)$, we may assume that $m_v = 1$, but this does not happen in general.

We shall provide a state independent representation assuming a strong separability condition (V8) below. We use the concept of probability distribution $x(f,r)$ induced by a probability distribution r over $(S,\mathcal{A})$ and an act $f \in \mathcal{G}$ defined by

$$x(f,r)(C) = \int f(s)(C)dr(s),$$

with $C \in \mathcal{B}$, see Fishburn (1982). Essentially, we convert the decision making problem under uncertainty into one under risk, by appealing to the class of probability distributions deduced from the preferences, as in Theorem 4. Since there is imprecision about beliefs, we associate with each act f under uncertainty the corresponding set of acts $\{x(f,r)\}_{r\in P}$ under risk. A minimal requirement for consistency of preferences under risk and under uncertainty seems, therefore, (V8) below. Note also that, in decision analysis and statistical decision theory, it is often the case that beliefs and preferences are assessed separately. Therefore, the previous model is adequate for that case. It also corresponds to general approaches to robust Bayesian analysis outlined in Berger (1994). Also, note that for a constant act $f = pI_S$, $x(f,r) = p, \forall r \in \mathcal{P}$, so that (V8) introduces no inconsistency in this case. Similarly, when $f = pI_{A^c} + qI_A$, $x(f,r) = (1 - r(A))p + r(A)q$, so, again (V8) introduces no inconsistency with $\preceq_\ell$. Moreover, for an Anscombe Aumann act f such that $f(s)$ is degenerate for every state, also called pure horse lottery, we have $x(f,r) = r \circ f^{-1}$ and (V8) is completely natural. We should mention that (V8) is redundant when preferences are complete, assuming the sure-thing principle. Nau's (1994) example suggest that this is not the case in the incomplete case.

Theorem 5 *Let* $\mathbf{G} \subset \mathcal{G}$ *such that its elements are convex combinations of functions which are continuous in a compact set and simple outside it. Suppose that* $(\mathbf{G}, \preceq)$ *satisfies the conditions of Theorem 4 and let* P *be the corresponding family of probability distributions. Suppose that, in addition, for* $f, g \in \mathbf{G}$

V8. $f \preceq g \iff (x(f,r) \preceq x(g,r), \forall r \in P)$.

Then, there is a set W *of functions* w *such that*

$$f \preceq g \iff (\int w(f(s))dr(s) \leq \int w(g(s))dr(s), \forall r \in P, \forall w \in W).$$

Proof. Let W be as in Theorem 1, with $\preceq_*$ the restriction of $\preceq$ to the set of constant acts in $\mathcal{G}$. We have $f \preceq g \iff (w(x(f,r)) \leq w(x(g,r)), \forall r \in P, \forall w \in W)$. Suppose that f is a simple Anscombe-Aumann act, i.e., $f =$

$\sum_{i=1}^{n} p_i I_{A_i}$. Then,

$$x(f,r) = \sum_{i=1}^{n} r(A_i) p_i$$

and

$$w(x(f,r)) = w(\sum_{i=1}^{n} r(A_i)p_i) = \sum_{i=1}^{n} r(A_i)w(p_i) = \int w(f(s))dr(s).$$

Let f be continuous in $K \subset S$ (compact) and simple in K^c. For $x \in K$, define

$$A_n(x) = \{y \in K : \|f(x) - f(y)\| < 1/n\}.$$

These sets form an open cover of K: there is a finite subcover $A_n(x_{n1}), \ldots,$ $A_n(x_{nm_n})$. Let $B_{n1} = A_n(x_{n1})$ and $B_{ni} = A_n(x_{ni}) \setminus \bigcup_{j=1}^{i-1} B_{nj}$. Define

$$f_n = \sum_{i=1}^{m_n} f(y_i) I_{B_{ni}} + f I_{K^c},$$

where y_i is an arbitrary element of B_{ni}. Then, $f_n \rightarrow f$, $x(f_n, r) \rightarrow x(f,r)$, and $w(x(f_n,r)) \rightarrow w(x(f,r))$. Moreover,

$$w(x(f_n,r)) = \int w(f_n(s))dr(s).$$

We apply now the dominated convergence theorem. First, due to continuity, $w(f_n(s)) \rightarrow w(f(s)), \forall s$. Moreover, w is linear and continuous, therefore it is bounded, and since $\|f_n(s)\| \leq 1$, $\forall n, s$, the function $g(s) = \|w\|$ is a dominating integrable function. Thus,

$$\int w(f_n(s))dr(s) \rightarrow \int w(f(s))dr(s).$$

Therefore,

$$w(f) = \int w(f(s))dr(s)$$

Similarly, for convex combinations of these functions. □

Of course, as in Theorem 1, we may associate to $w \in W$ the corresponding $u \in \mathcal{U}$ so that

$$\int w(f(s))dr(s) = \int_S \int_C u df(s) dr.$$

We conclude this section with two corollaries:

Corollary 6 *Under the conditions of Theorem 5, if the restriction of $\preceq$ to the set of constant acts is complete, there is w such that*

$$f \preceq g \iff \left(\int w(f(s))dr(s) \leq \int w(g(s))dr(s), \forall r \in P\right)$$

Apply Corollary 1, Theorem 1 and Theorem 5. □

Essentially, this is the type of representation in Giron and Rios (1980) and Walley (1991), and the type of model used in conventional robust Bayesian analysis, see Berger (1994).

Corollary 7 *Under the conditions of Theorem 5, if $\preceq_\ell$ is complete, there is r such that*

$$f \preceq g \iff \left(\int w(f(s))dr(s) \leq \int w(g(s))dr(s), \forall w \in W\right)$$

Apply Corollary 1, Theorem 4 and Theorem 5. □

This is, essentially, the type of model used in stochastic dominance problems, see Levy (1992).

4 Conclusions

We have provided axiomatic foundations to model incompleteness in a DM's judgments by means of a class of utility functions and a class of probability distributions, arbitrarily paired. We have analysed the cases of decision making under risk and decision making under uncertainty. In this latter case, we have considered both state-dependent and state-independent utility models. With our results, we unify and support many streams of recent research, specially in the areas of robustness and sensitivity studies in Bayesian Statistics and Decision Theory. We believe that a main consequence of our approach is that there should be a shift from 'conventional' robust Bayesian computations to the type of computations discussed in Rios Insua and Martin (1994).

Of course many issues remain unsolved. For example, it would be interesting to provide similar foundations in Savage's framework.

Acknowledgements This research has been financed by grants from DGICYT, UPM and MEC-British Council and completed while the first author visited Purdue University. The authors are grateful to many discussions with Bob Nau and Simon French.

References

[1] Anscombe, F., Aumann, R. (1963) A definition of subjective probability, *Ann. Math. Stat.*, 34, 199-205.

[2] Aumann, R. (1962) Utility theory without the completeness axiom, *Econometrica*, 30, 445-462; (1964), 32, 210-212.

[3] Ash, R. (1971) *Real Analysis and Probability*, Academic Press.

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

[5] Fishburn, P.C. (1970) *Utility Theory for Decision Making*, Wiley.

[6] Fishburn, P.C. (1982) *The Foundations of Expected Utility*, D. Reidel.

[7] Gilboa, I., Schmeidler, D. (1993) Updating ambiguous beliefs, *J. Econ. Theory*, 59, 33-44.

[8] Giron, F.J. and Rios, S. (1980) Quasi Bayesian behaviour: A more realistic approach to decision making?, in Bernardo, De Groot, Lindley, Smith (eds) *Bayesian Statistics*, Valencia U.P.

[9] Kreps, D. (1988) *Notes on the Theory of Choice*, Westview.

[10] Levy, H. (1992) Expected utility analysis and stochastic dominance: survey and analysis, *Mgt. Sci.*, 38, 555-593.

[11] Lindley, D., Tversky, A., Brown, R. (1979) On the reconciliation of probability assessments, *JRSS, A*, 142, 146-180.

[12] Nau, R. (1992) Indeterminate probabilities on finite sets, *Annals of Statistics*, 20, 1737-1767.

[13] Nau, R. (1994) On the shape of incomplete preferences, *Wkg. Paper, Fuqua School of Business*, Duke Univ.

[14] Rios Insua, D. (1990) *Sensitivity Analysis in Multiobjective Decision Making*, Springer Verlag.

[15] Rios Insua, D. (1992) The foundations of robust decision making: the simple case, *Test*, 1, 69-78.

[16] Rios Insua, D. and Martin, J. (1994) Robustness issues under imprecise beliefs and preferences, *Jour. Stat. Plan. Inf.*, 40, 383-389.

[17] Savage, L. J. (1954) *The Foundations of Statistics*, Wiley.

[18] Schmeidler, D. (1989) Subjective probability and expected utility without additivity, *Econometrica*, 37, 571-587.

[19] Seidenfeld, T., Schervish, M., Kadane, J. (1992) A representation of partially ordered preferences, *Res. Rep. Dept. of Philosophy*, Carnegie Mellon (presented at 43 NBER-NSF Meeting).

[20] Walley, P. (1991) *Statistical Reasoning with Imprecise Probabilities*, Chapman Hall.