RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **96** (1), 2002, pp. 55–61
Estadística e Investigación Operativa / Statistics and Operations Research

A note on an open problem in the foundations of statistics

Alejandro García del Amo and David Ríos Insua

Abstract. We study a problem in the foundations of statistics, that of modelling imprecision in preferences and beliefs within the expected utility framework. This problem is at the core of sensitivity analysis in Bayesian statistics. We provide an answer in the form of state-dependent expected utilities.

Una nota sobre un problema abierto en fundamentos de Estadística

Resumen. Estudiamos un problema abierto en los fundamentos de la estadística, el de modelizar imprecisión en preferencias y creencias dentro del marco de la utilidad esperada. Este problema fundamenta el análisis de sensibilidad en estadística bayesiana. Proporcionamos una solución en términos de utilidades esperadas dependientes del estado.

1. An open problem in the foundations of statistics

Various axiomatic frameworks, including those of Von Neumann-Morgenstern, Savage, DeGroot and Anscombe and Aumann, see e.g. French and Ríos Insua (2000) for a review, essentially lead to the (subjective) expected utility model, which is at the core of Bayesian Statistics and Decision Theory and much of current research in Economics. That various paths lead essentially to the same model is reassuring and, in a sense, one of the main strengths of the Bayesian approach. At the same time, the axiomatic foundations facilitate a critical analysis of the theory, by putting into question various axioms and possibly obtaining alternative models, see Anand (1987) for a discussion.

Following that argument, in this note we explore foundations for a model of beliefs based on a class of probability distributions, of preferences based on a class of utility functions and leading to the comparison of alternatives through Pareto-type inequalities of their expected utilities. As explained in Ríos Insua and Ruggeri (2000), these would provide the foundations of much of recent research in robust Bayesian analysis, and is a problem yet to be solved. The essential issue is what are the consequences of dropping the completeness axiom, and what other conditions are needed to achieve such representation.

Since its inception, see e.g. Von Neumann and Morgenstern (1944), Savage (1954) and Aumann (1962), many authors have dwelt on the issue of complete judgements because of the difficulty of assessing beliefs and preferences precisely, specially in presence of several DM's. Indeed, many authors including Girón and Ríos (1980), Ríos Insua (1990, 1992), Nau (1995), Seidenfeld et al (1995), Shapley and Baucells (1998) or Dubra et al (2001), have dwelt on the problem providing partial answers (for example, by considering precision in beliefs but not in preferences as in Dubra et al, or the other way round, as in Girón and Ríos).

Presentado por Sixto Ríos García.
Recibido: 30 de Abril de 2002. *Aceptado*: 8 de Mayo de 2002.
Palabras clave / Keywords: 62C10,46E40.
Mathematics Subject Classifications: Expected utility, Incompleteness, Quasi Orders, Young measures.

The few attempts on the general open problem, see e.g. Ríos Insua and Martin (1994), have used too strong conditions.

In this note, based on conditions in Ríos Insua and Criado (2000), we deal with problems of decision making under uncertainty to obtain a novel state-dependent expected utility representation and suggest ways to handle this general problem which, we believe, would close a long-lived issue in the foundations of statistics. We shall use here Young measures, introduced by Young (1937; 1942) as a tool to deal with certain problems arising in variational calculus that do not admit classical solutions because approximate solutions show typically rapid oscillations. They are also known as parametrized measures, relaxed controls, measurable families of probabilities, transition probabilities or stochastic kernels.

2. Basic structure and problem

We consider decision making problems under uncertainty, within Anscombe and Aumann (1963) framework. In such framework, we assume that preferences are established among Anscombe Aumann acts, which are functions from a set of states into the set of probability measures over a space of consequences. We shall cast them in a somewhat more abstract way, by adopting a functional analytic view of the problem. Specifically, we shall view Anscombe Aumann acts as Young measures. More details on the necessary theoretical background can be found in, e.g., Aliprantis and Border (1999).

The basic elements in our problem are a set S of states of nature, which we shall assume is a compact subset of $\mathbb{R}^n$ endowed with Lebesgue measure μ_S, over $\mathcal{F}_S$; the $\sigma-$algebra of Lebesgue measurable subsets of S, and X, a compact subset of $\mathbb{R}^m$, which will be the space of consequences. We shall consider the set $\mathcal{P}(X)$ of all probability measures on the Borel sets $\mathcal{B}_X$ of X endowed with the topology of convergence in distribution (or weak* topology). A decision maker establishes his preferences $\preceq$ when comparing alternatives, which in our context will be functions f from S into $\mathcal{P}(X)$. Intuitively, if the state s finally holds, we shall obtain the probability measure f_s over X.

We shall assume some mathematical structure over those alternatives. To wit, we shall consider the set $Y(S;X)$ of Young measures, which consists of all $f \in L^{\infty}_{w^*}(S;ca(X))$ such that $f_s \in \mathcal{P}(X)$ a.e. in S, where $ca(X)$ is the space of all signed measures of bounded variation on $\mathcal{B}_X$ and $L^{\infty}_{w^*}(S;ca(X))$ designates the Banach space of all (equivalence classes of) weak*-measurable functions $f : S \to ca(X)$ such that

$$\|f\|_\infty = \text{ ess sup } \{\|f_s\| : s \in S\} < \infty.$$

Note that this can be identified with the topological dual of the Banach space $L^1(S;C(X))$ of all (equivalence classes of) measurable functions $g : S \to C(X)$ such that

$$\|g\|_1 = \int_S \|g(s)\| \, ds < \infty$$

(Cembranos and Mendoza, 1997), where $C(X)$ is the set of continuous functions over X. Equivalently, a function $f : S \to \mathcal{P}(X)$ is a Young measure if $s \to \int_S u \, df_s$ is measurable for every $u \in C(X)$. We will assume that $Y(S;X)$ is endowed with the relativization $\sigma(Y(S;X), L^1(S;C(X)))$ of the weak* topology $\sigma(L^{\infty}_{w^*}(S;ca(X)), L^1(S;C(X)))$.

The general aim of Bayesian foundations in this context would be to obtain the representation

$$f \preceq g \Longleftrightarrow \int_S \left[\int_X u(x) df_s(x)\right] v(s) \, ds \le \int_S \left[\int_X u(x) dg_s(x)\right] v(s) \, ds$$

u being a utility function and v a density function. Conditions for this representation are well-known in the Bayesian literature, see French and Ríos Insua (2000) for further information. They all require that $\preceq$ is a weak order (i.e, it is transitive and complete). However, in many applications we need to deal with incomplete relations, suggesting a representation of the type

$$f \preceq g \Longleftrightarrow \left(\int_S \left[\int_X u(x) df_s(x)\right] v(s) \, ds \le \int_S \left[\int_X u(x) dg_s(x)\right] v(s) \, ds, \forall u \in U, v \in V\right)$$

where U is a set of utility functions and V is a set of density functions.

We aim at providing such representation here. The type of argument we shall use is to relate a closed convex cone with the preference order which will be a quasi order, then use a Hahn-Banach type of representation and, finally, explore the consequences of such representation.

Recall that a reflexive and transitive (binary) relation $\preceq$ on a set A is called a quasi-order on A. A subset C in a vector space L is a wedge if $C + C \subset C$ and $\lambda C \subset C$ for all $\lambda \geq 0$. Note that a wedge C in L determines a quasi-order $\preceq$ by

$$x \preceq y \text{ if } y - x \in C.$$

Moreover, this relation is compatible (with the vector structure of L) in the sense that $x \preceq y$ implies $x + z \preceq y + z$ and $\lambda x \preceq \lambda y$ for all $z \in L$ and $\lambda > 0$. Conversely, if $\preceq$ is a reflexive, transitive and compatible relation on L, and if we define $C = \{x \in L : x \succeq 0\}$, then C is a wedge in L, and $\preceq$ is exactly the quasi-order induced by C.

If E is a locally convex Hausdorff space and C is a wedge in E, then it is easy to see that the set $C' = \{x' \in E' : \langle x, x' \rangle \geq 0 \text{ for all } x \in C\}$ is a wedge in E'.

The following result will be essential in our approach. The first two equivalences may be seen in Namioka (1957, Corollary 4.2), the other two equivalences being simple.

Lemma 1 *Let E be a locally convex Hausdorff space and let $\preceq$ be the quasi-order induced by a wedge C in E. The following statements are equivalent.*

1. *C is closed.*
2. *If $x \in E$, then $x \succeq 0$ if and only if $\langle x, x' \rangle \geq 0$ for all $x' \in C'$.*
3. *If $x, y \in E$, then $x \preceq y$ if and only if $\langle x, x' \rangle \leq \langle y, x' \rangle$ for all $x' \in C'$.*
4. *There exists a subset $\mathcal{W}$ of E' such that $x \preceq y$ if and only if $\langle x, w \rangle \leq \langle y, w \rangle$ for all $w \in \mathcal{W}$.*

We shall also need the following result about Young measures (Roubíček, 1997, Corollary 3.1.7).

Theorem 1 *The set of all Young measures $Y(S; X)$ is a metrizable convex weak* compact set.*

Finally, if $u \in C(S)$ and $v \in C(X)$, we write $u \otimes v$ for the function $s \to u \otimes v(s) = u(s)v$, from $S \to C(X)$. Moreover $C(S) \otimes C(X)$ is the collection of functions $\sum_{i=1}^{j} u_i \otimes v_i$, where $j \in \mathbb{N}$, $u_i \in C(S)$, and $v_i \in C(X)$. Since $Y(S; X)$ is a weak* compact set, it follows from the denseness of $C(S) \otimes C(X)$ in $L^1(S; C(X))$ (Warga, 1972, Theorem I.5.25) that

$$\sigma(Y(S; X), L^1(S; C(X))) = \sigma(Y(S; X), C(S) \otimes C(X))$$

(Wilansky, 1978, Lemma 9-5-1).

3. A representation with state-dependent expected utilities

In this section, we provide a representation in terms of state-dependent expected utilities. We identify a class of functions that allows us to represent a quasi order within the set of Young measures. Those functions may be expressed as expected utilities, with utilities viewed as a weighted combination of utility functions, the weights depending on states. Specifically, we have the following

Theorem 2 *The three conditions*

1. *$(Y(S; X), \preceq)$ is a quasi-order.*
2. *(Independence Axiom.) For any $f, g, h \in Y(S; X)$ and any $\lambda \in (0, 1)$, $f \preceq g$ implies $\lambda f + (1-\lambda) h \preceq \lambda g + (1 - \lambda) h$.*

3. *(Continuity Axiom.) If* (f_k) *and* (g_k) *are convergent sequences in* $Y(S;X)$ *such that* $f_k \preceq g_k$ *for each* k*, then* $\lim f_k \preceq \lim g_k$.

are equivalent to the existence of a class $\mathcal{W}$ *of functions* $w(s,x) = \sum_{i=1}^{j} u_i(s)v_i(x)$*, where* $u_i \in C(S)$*, and* $v_i \in C(X)$*, such that*

$$f \preceq g \Longleftrightarrow \int_S \left[\int_X w(s,x)df_s(x)\right] ds \leq \int_S \left[\int_X w(s,x)dg_s(x)\right] ds \text{ for all } w \in \mathcal{W}.$$

PROOF. Let us show the non trivial implication. To this end, consider the set

$$C = \{\lambda(g-f) : \lambda > 0 \text{ and } f \preceq g\}.$$

It is easy to see that C is a wedge. From Lemma 1, in order to complete the proof, it suffices to show that C is closed in $L^{\infty}_{w^*}(S;ca(X))$, which we assume to be endowed with the weak* topology. The proof is presented in steps. The first two have been noted recently by Shapley and Baucells (1998) (on an arbitrary mixture space); see also Dubra et al. (2001). First note that the continuity axiom allows strengthen the independence axiom to

- *Claim 1. For any* $f,g,h \in Y(S;X)$ *and any* $\lambda \in (0,1)$, $f \preceq g$ *if and only if* $\lambda f + (1-\lambda)h \preceq \lambda g + (1-\lambda)h$.

As a consequence, we have

- *Claim 2. If* $f,g \in Y(S;X)$*, then* $f \preceq g$ *if and only if* $g - f \in C$.

We next consider two distinguishing properties of $\mathcal{P}(X)$. Noting that $0 \in \mathcal{P}(X) - \mathcal{P}(X)$, the first one is immediate from the convexity of $\mathcal{P}(X) - \mathcal{P}(X)$.

- *Claim 3. If* $p,q \in \mathcal{P}(X)$ *and* $0 < \lambda < 1$*, then there exist two probability measures* r *and* s *satisfying* $\lambda(q-p) = s - r$.

At this point, we introduce some further notation. If μ is a signed measure on $\mathcal{B}_X$, the positive and negative variations of μ are the measures on $\mathcal{B}_X$ defined by the formulae

$$\mu^+(A) = \sup\{\mu(B) : B \in \mathcal{B}_X \text{ and } B \subset A\}$$

and

$$\mu^-(A) = \sup\{-\mu(B) : B \in \mathcal{B}_X \text{ and } B \subset A\}$$

for $A \in \mathcal{B}_X$. Jordan decomposition theorem shows that $\mu = \mu^+ - \mu^-$. As $\|(q-p)^+\| = \|(q-p)^-\|$, we get

- *Claim 4. If* $p,q \in \mathcal{P}(X)$*, then there exist two probability measures* r *and* s *such that* $q - p = \|(q-p)^+\|(s-r)$.

This result can be generalized as follows.

- *Claim 5. Let* $f,g \in Y(S;X)$*. Then, there exist* $r,s \in Y(S;X)$ *such that* $g - f = \|g-f\|(s-r)$*. Moreover, if* $f \neq g$, $f \preceq g$ *if and only if* $r \preceq s$.

The validity of the second claim follows immediately from claim 2. On the other hand, for $f = g$ the conclusion of the first claim is obvious. Let us, therefore, assume that $f \neq g$.

For each k there exists a partition $\{S_k^1, ..., S_k^{J(k)}\}$ of S such that

$$h_k = \sum_{j=1}^{J(k)} h_k^j \chi_{S_k^j} \to h \quad \text{in} \quad L_{w^*}^\infty(S; ca(X))$$

for all $h \in Y(S;X)$, where $h_k^j \in \mathcal{P}(X)$ is the functional on $C(X)$ defined by

$$h_k^j(v) = \frac{1}{|S_k^j|} \int_{S_k^j} < v, f_s > ds.$$

For details, see Roubíček (1997, Theorems 3.1.3 and 3.1.6). In particular, this implies

$$g_k - f_k \to g - f \ \text{ in } \ L_{w^*}^\infty(S; ca(X)).$$

Now, we may easily verify that

$$\|g_k - f_k\| = \max_j \|g_k^j - f_k^j\| \leq \|g - f\|.$$

By claim 4, there exist $\tilde{r}_k^j, \tilde{s}_k^j \in \mathcal{P}(X)$ such that

$$g_k^j - f_k^j = \|(g_k^j - f_k^j)^+\|(\tilde{s}_k^j - \tilde{r}_k^j).$$

In fact, since

$$g_k^j - f_k^j = \|g - f\| \frac{\|(g_k^j - f_k^j)^+\|}{\|g - f\|}(\tilde{s}_k^j - \tilde{r}_k^j) \text{ and } \frac{\|(g_k^j - f_k^j)^+\|}{\|g - f\|} \in [0,1],$$

we can use claim 3 to obtain two probability measures r_k^j and s_k^j such that

$$g_k^j - f_k^j = \|g - f\|(s_k^j - r_k^j).$$

From this last equality, we infer that

$$g_k - f_k = \|g - f\|(s_k - r_k)$$

with

$$r_k = \sum_{j=1}^{J(k)} r_k^j \chi_{S_k^j} \text{ and } s_k = \sum_{j=1}^{J(k)} s_k^j \chi_{S_k^j}$$

in $Y(S;X)$. This allows us to use Theorem 1 to obtain $r, s \in Y(S;X)$ and subsequences (r_{k_i}) and (s_{k_i}) such that $r_{k_i} \longrightarrow r$ and $s_{k_i} \to s$ in $L_{w^*}^\infty(S; ca(X))$. But then,

$$g - f = \|g - f\|(s - r),$$

and the proof of the claim is finished.

- *Claim 6. C is closed in $L_{w^*}^\infty(S; ca(X))$.*

Since $C(X)$ is separable, it is only necessary to verify that C is sequentially closed in $L^{\infty}_{w^*}(S; ca(X))$ (Wilansky, 1978, p. 192): If $(f_k), (g_k)$ are sequences in $Y(S;X)$ such that $f_k \preceq g_k$ and if (λ_k) is a sequence in $(0, \infty)$ such that $\lambda_k(g_k - f_k) \to h$ in $L^{\infty}_{w^*}(S; ca(X))$, then we have to verify that $h \in C$.

By the uniform boundedness principle, the fact that $\lambda_k(g_k - f_k) \to h$ in $L^{\infty}_{w^*}(S; ca(X))$ implies that $(\lambda_k(g_k - f_k))$ is strongly bounded, that is, there exists some $M > 0$ such that $\|\lambda_k(g_k - f_k)\| \leq M$ for each k. Next, from claim 5 it follows that we can choose $\gamma_k > 0$ and $r_k, s_k \in Y(S;X)$ with $r_k \preceq s_k$ and $\|s_k - r_k\| = 1$ such that $g_k - f_k = \gamma_k(s_k - r_k)$. Thus

$$\|\lambda_k(g_k - f_k)\| = \|\lambda_k \gamma_k (s_k - r_k)\| = \lambda_k \gamma_k \leq M$$

for each k, so we can assume (by passing to a subsequence if necessary) that $\lambda_k \gamma_k \to \lambda$. On the other hand, by Theorem 1, there are $r, s \in Y(S;X)$ such that (by passing to a subsequence if necessary) $r_k \to r$ and $s_k \to s$ in $L^{\infty}_{w^*}(S; ca(X))$. Hence $h = \lambda(s - r)$ with $r \preceq s$ (use the continuity axiom), and the proof is finished. ■

4. Discussion

We have provided a solution to an open problem in the foundations of statistics in terms of state dependent expected utilities. We should check whether similar results may be provided for state independent utilities (or, at least, some substantial part of the set of expected utilities). We could also explore cases in which there is imprecision in beliefs, but not in preferences, and the other way round. We have based our analysis in Anscombe Aumann framework. Savage framework should be also explored. Finally, updating of the involved class of probabilities in terms of Bayes type of rules should be studied. This will be the subject of further work.

Acknowledgement. Work supported by grants from CICYT, CAM, Castilla and Leon Government and URJC.

References

[1] Aliprantis, C. D. and Border, K. C. (1999). *Infinite-dimensional analysis. A hitchhiker's guide*. Second edition. Springer-Verlag, Berlin.

[2] Anand, P. (1987). Are the preference axioms really rational? *Theory and Decision*, **23**, 189–214.

[3] Anscombe, F. and Aumann, R. (1963). A definition of subjective probability. *Ann. Statist.*, **34**, 199–205.

[4] Aumann, R. (1962). Utility theory without the completeness axiom, *Econometrica*, **30**, 445–462.

[5] Cembranos, P. and Mendoza, J. (1997). *Banach spaces of vector-valued functions*. Lecture Notes in Mathematics, **1676**. Springer-Verlag, Berlin.

[6] Dubra, J., Maccheroni, F. and Ok, E. (2001). Expected utility theory without the completeness axiom. *Cowles Foundation Discussion Paper, 1294*.

[7] French, S. and Ríos Insua, D. (2000). *Statistical Decision Theory*, Arnold, London.

[8] Girón, F.J. and Ríos, S. (1980). Quasi Bayesian behaviour: a more realistic approach to decision making? In *Bayesian Statistics*. J. M. Bernardo, M. H. De Groot, D. Lindley and A.F.M. Smith, eds., Valencia University Press, Valencia.

[9] Namioka, I. (1957). *Partially ordered linear topological spaces*. Mem. Amer. Math. Soc. **24**.

[10] Nau, R. (1995). The shape of incomplete preferences. *Technical Report*, Duke University.

[11] Ríos Insua, D. (1990). *Sensitivity Analysis in Multiobjective Decision Making*, Springer-Verlag, Berlin.

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

[13] Ríos Insua, D. and Criado, R. (2000). Topics on the foundations of robust Bayesian analysis, in *Robust Bayesian Analysis*, D. Ríos Insua and F. Ruggeri eds., Springer.

[14] Ríos Insua, D. and Martin, J. (1994). On the foundations of robust decision making. In *Decision Theory and Decision Analysis: Trends and Challenges*, S. Ríos eds. Kluwer Academic Publishers, Cambridge.

[15] Ríos Insua, D. and Ruggeri, F. (2000). *Robust Bayesian Analysis*, Springer-Verlag, New York.

[16] Roubíček, T. (1997). *Relaxation in optimization theory and variational calculus*, Walter de Gruyter and Co., Berlin.

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

[18] Seidenfeld, T., Schervish, M. and Kadane, J. (1995). A representation of partially ordered preferences. *Ann. Statist.*, **23**, 2168–2217.

[19] Shapley, L.S. and Baucells, M. (1998). Multiperson utility, *UCLA Working paper 779*.

[20] Valadier, M. (1990). Young measures. In *Methods of nonconvex analysis*. Lectures given at the First 1989 C.I.M.E. Session held in Varenna, June 15–23, 1989. A. Cellina, eds., Lecture Notes in Mathematics, **1446**, Springer-Verlag, Berlin.

[21] Von Neumann, J. and Morgenstern, O. (1944). *The Theory of Games and Economic Behaviour*, Princeton Univ. Press.

[22] Warga, J. (1972). *Optimal control of differential and functional equations*, Academic Press, New York-London.

[23] Wilansky, A. (1978). *Modern methods in topological vector spaces*, McGraw-Hill International Book Co., New York.

[24] Young, L.C. (1937). Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *C. R. Soc. Sci. Varsovie*, **30**, 212–234.

[25] Young, L.C. (1942). Generalized surfaces in the calculus of variations I and II. *Ann. of Math.*, **43** (2), 84–103 and 530–544.

A. García del Amo and D. Ríos Insua
Departamento de Ciencias Experimentales e Ingeniería
Universidad Rey Juan Carlos
Madrid. Spain