Study of the Probabilistic Information of a Random Set*

ENRIQUE MIRANDA
University of Oviedo, Spain

INÉS COUSO
University of Oviedo, Spain

PEDRO GIL
University of Oviedo, Spain

Abstract

Given a random set coming from the imprecise observation of a random variable, we study how to model the information about the distribution of this random variable. Specifically, we investigate whether the information given by the upper and lower probabilities induced by the random set is equivalent to the one given by the class of the distributions of the measurable selections; together with sufficient conditions for this, we also give examples showing that they are not equivalent in all cases.

Keywords

random sets, upper and lower probabilities, measurable selections, Choquet integral

1 Introduction

Random sets have been successfully applied in such different fields as economy ([11]) or stochastic geometry ([14]), and they have been studied under different interpretations, like the behavioral ([19]) or the evidential one ([7]). In this paper, we will interpret a random set as the result of the imprecise observation of a random variable ([13]). Under this interpretation, our information about the probability distribution of the random variable is given by the class of distributions of the *measurable selections* of the random set. This class of distributions is a subset of the class of probability measures bounded between the *upper* and *lower* probabilities ([7]) of the random set. These functions satisfy Walley's axioms of

*This paper has been partially funded by FEDER-MCYT, grant number BFM2001-3515, and Fundación Banco Herrero.

coherence ([21]), and are moreover ∞-alternating and ∞-monotone, respectively ([20]).

Although working with the upper and lower probabilities leads to a number of mathematical simplifications ([20, 21]), the information they provide is in general more imprecise than the one given by the set of distributions of the measurable selections ([16, 18]). In this paper, we will investigate under which conditions these two models are equivalent. The results we obtain will show when it is advisable to model our information through the upper and lower probabilities and when this produces a loss of precision.

In Section 2, we introduce some concepts and notations that we will use in the rest of the paper, and recall some previous works on the subject. In Section 3, we investigate the information that the upper and lower probabilities give about the distribution of the original random variable, and about the value of this distribution on an arbitrary set. Finally, in Section 4 we give some additional comments and remarks.

2 Preliminary concepts

We will consider a probability space $(\Omega, \mathcal{A}, P)$, a measurable space $(X, \mathcal{A}')$ and a multi-valued mapping $\Gamma : \Omega \to \mathcal{P}(X)$. If X is a topological space, we will denote β_X its Borel σ-field. A topological space is said to be *Polish* when it is separable and complete for some compatible metric d, and it is called *Souslin* if it is the bijective image of a Polish space. The multi-valued mapping will be called open (resp. complete, closed, compact) if $\Gamma(\omega)$ is an open (resp. complete, closed, compact) subset of X for every $\omega \in \Omega$.

Formally, a random set is a multi-valued mapping satisfying some measurability condition. There are different conditions, such as the weak, the strong, or the graph-measurability ([12]). Most of them are based on the notion of upper and lower inverse:

Definition 1 *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ be a measurable space and $\Gamma : \Omega \to \mathcal{P}(X)$ a multi-valued mapping. Given $A \in \mathcal{A}'$, its* **upper inverse** *is $\Gamma^*(A) = \{\omega \in \Omega \mid \Gamma(\omega) \cap A \neq \emptyset\}$, and its* **lower inverse** *is $\Gamma_*(A) = \{\omega \in \Omega \mid \emptyset \neq \Gamma(\omega) \subseteq A\}$.*

When there is no possible confusion about the multi-valued mapping we are working with, we will use the notation $A^* := \Gamma^*(A)$ and $A_* := \Gamma_*(A)$. By a **random set** we will mean throughout a strongly measurable multi-valued mapping. The strong measurability is necessary for the upper and lower probabilities of the random set to be defined on $\mathcal{A}'$.

Definition 2 *A multi-valued mapping is called* **strongly measurable** *if $A^* \in \mathcal{A}$ $\forall A \in \mathcal{A}'$.*

Note that $A_* = X^* \cap ((A^c)^*)^c\ \forall A \in \mathcal{A}'$, whence if Γ is strongly measurable, we also have $A_* \in \mathcal{A}\ \forall A \in \mathcal{A}'$. The concepts of upper and lower probabilities induced by a random set were introduced by Dempster in [7]:

Definition 3 *Given a random set* $\Gamma : \Omega \to \mathcal{P}(X)$*, the* **upper probability** *of* $A \in \mathcal{A}'$ *is* $P^*_\Gamma(A) = \frac{P(A^*)}{P(X^*)}$*, and its* **lower probability** *is* $P_{*\Gamma}(A) = \frac{P(A_*)}{P(X^*)}$.

When there is no ambiguity about which random set is inducing the upper and lower probability, we will denote $P^* := P^*_\Gamma$ and $P_* := P_{*\Gamma}$.

As we said in the introduction, we will regard a random set as the result of the imprecise observation of a random variable $U_0 : \Omega \to X$ (which we will call *original* random variable), in the sense that for every ω in the initial space all we know about $U_0(\omega)$ is that it belongs to the set $\Gamma(\omega)$. As a consequence, $\Gamma(\omega)$ will be assumed to be non-empty for every ω, and hence $P^*(A) = P(A^*)$ and $P_*(A) = P(A_*)$ for all $A \in \mathcal{A}'$. The upper and lower probabilities induced by a random set are conjugate functions, and they are moreover ∞-alternating and ∞-monotone capacities, respectively ([20]). This means in particular that they satisfy Walley's axioms of coherence ([21]).

If Γ is the imprecise observation of U_0, all we know about this variable is that it belongs to the class of **measurable selections** (or **selectors**) of Γ,

$$S(\Gamma) := \{U : \Omega \to X \text{ measurable} \mid U(\omega) \in \Gamma(\omega)\ \forall \omega\}.$$

The probability distribution of U_0 belongs to

$$P(\Gamma) := \{P_U \mid U \in S(\Gamma)\},$$

and our information about $P_{U_0}(A)$ is given by the set of values

$$P(\Gamma)(A) := \{P_U(A) \mid U \in S(\Gamma)\}.$$

There are two other classes of probabilities that may be useful in some situations. The first one is

$$\Delta(\Gamma) := \{Q \text{ probability} \mid Q(A) \in P(\Gamma)(A)\ \forall A \in \mathcal{A}'\}.$$

This is the set of distributions whose values are compatible with the information given by the random set. It is clear that $P(\Gamma) \subseteq \Delta(\Gamma)$. If they coincide, the information about the distribution of the original random variable is equivalent to the information about the values it takes. On the other hand, we can also consider the class

$$M(P^*) := \{Q \text{ probability} \mid Q(A) \leq P^*(A)\ \forall A \in \mathcal{A}'\}$$

of distributions dominated by P^*, or **credal set** generated by P^*. This class is convex and easier to handle in practice than $P(\Gamma)$. Using the inequalities $P_*(A) \leq$

$P_U(A) \leq P^*(A)$, valid for any $U \in S(\Gamma)$, $A \in \mathcal{A}'$, we deduce that $\Delta(\Gamma) \subseteq M(P^*)$. We see then that $P(\Gamma) \subseteq \Delta(\Gamma) \subseteq M(P^*)$. As we showed in [16], both inclusions can be strict, and in some cases the use of the upper and lower probabilities can produce a loss of precision, which in turn can cause some misjudgements. It is therefore interesting to see in which cases it is reasonable to use P^* and P_*.

Although the class of the distributions of the selectors of a random set ([1, 9]) and the upper probability it induces ([14, 20]) have been thoroughly studied in the literature, the connection between them has not received much attention. It was investigated for the case of X finite in [16], and for some particular infinite spaces in [3, 6, 10, 15, 18]. Our goal in this paper is to somewhat fill this gap. Specifically, we will study two different problems:

- First, we will investigate the relationship between $\Delta(\Gamma)$ and $M(P^*)$, which tells us if the upper and the lower probabilities are informative enough about the value $P_{U_0}(A)$ for some arbitrary $A \in \mathcal{A}'$.
- Then, we will study when $P(\Gamma) = M(P^*)$, i.e., under which conditions the upper probability keeps all the information about P_{U_0}.

3 Study of the probabilistic models for P_{U_0}

3.1 $P^*(A), P_*(A)$ as a model for $P_{U_0}(A)$

Let us start investigating the relationship between $\Delta(\Gamma)$ and $M(P^*)$. As we mentioned before, $\Delta(\Gamma)$ models the information that Γ gives about the probability values of the elements in $\mathcal{A}'$. Therefore, by investigating its equality with $M(P^*)$ we will see whether P^* and P_* are informative enough about the 'true' probability of an arbitrary set A. This is formally stated in the following proposition.

Proposition 1 *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ a measurable space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. Then,*

$$\Delta(\Gamma) = M(P^*) \Leftrightarrow P(\Gamma)(A) = [P_*(A), P^*(A)] \; \forall A \in \mathcal{A}'.$$

Let us consider then some arbitrary $A \in \mathcal{A}'$, and let us study the relationship between $P(\Gamma)(A)$ and $[P_*(A), P^*(A)]$. It is clear that the latter is a superset of the former. In order to give conditions for the equality, we must see if the maximum and minimum values of $P(\Gamma)(A)$ coincide with $P^*(A)$ and $P_*(A)$, and also if $P(\Gamma)(A)$ is convex.

This problem was studied in [18]. We showed there that $P(\Gamma)(A)$ has a maximum and a minimum value (it is indeed a closed subset of $[0,1]$), and that these values do not coincide in all cases with $P^*(A), P_*(A)$, even in the non-trivial case of $S(\Gamma) \neq \emptyset$. Moreover, $P(\Gamma)(A)$ is not convex in general. The following theorem gives sufficient conditions for the equalities $P^*(A) = \max P(\Gamma)(A)$ and $P_*(A) = \min P(\Gamma)(A)$. It generalizes previous results from [6].

Theorem 1 *[18] Consider* $(\Omega, \mathcal{A}, P)$ *a probability space,* (X, τ) *a topological space and* $\Gamma : \Omega \to \mathcal{P}(X)$ *a random set. Under any of the following conditions:*

1. Ω *is complete,* X *is Souslin and* $Gr(\Gamma) \in \mathcal{A} \otimes \beta_X$,
2. X *is a separable metric space and* Γ *is compact,*
3. X *is a separable metric space and* Γ *is open,*
4. X *is a Polish space and* Γ *is closed,*
5. X *is a* σ*-compact metric space and* Γ *is closed,*

we have $P^*(A) = \max P(\Gamma)(A)$ *and* $P_*(A) = \min P(\Gamma)(A)$ $\forall A \in \beta_X$. *Moreover, if*

6. X *is a separable metric space and* Γ *is complete,*

then $P^*(A) = \max P(\Gamma)(A), P_*(A) = \min P(\Gamma)(A)$ $\forall A \in Q(\{B_n\}_n)$, *where* $\{B_n\}_n = \{B(x_i; q_j) \mid i \in \mathbb{N}, q_j \in \mathbb{Q}\}$ *is a countable basis of* $\tau(d)$ *associated to a countable dense set* $\{x_n\}_n$ *and* $Q(\{B_n\}_n)$ *is the field generated by* $\{B_n\}_n$.

This theorem gives sufficient conditions for the equalities $P^* = \max P(\Gamma)$ and $P_* = \min P(\Gamma)$. The coherence of P^* implies ([21]) that it is the upper envelope of the set of the finitely additive probabilities it dominates. We have proven that, under conditions (1) to (5) from Theorem 1, it is indeed the upper envelope of the class of *countably* additive probabilities induced by the selectors. A similar (symmetrical) remark can be made for P_*.

Let us remark in passing that results established in Theorem 1 guarantee the existence of a selector of Γ whose distribution coincides with P^* on a finite chain. Indeed, in [5] Couso showed that the equality $P^*(A) = \sup P(\Gamma)(A)$ $\forall A \in \mathcal{A}'$ implies the equality between the Choquet integral of a bounded random variable respect to the upper probability of a random set ([8]) and the supremum of class of the integrals respect to the distributions of the measurable selections. This allows us to deduce the following result, which generalizes theorem 1 from [3].

Theorem 2 *Let* $\Gamma : \Omega \to \mathcal{P}(X)$ *be a random set and* $V : X \to \mathbb{R}$ *a bounded random variable. Under any of the conditions (1) to (5) from the previous theorem,* $(C)\int V dP^* = \sup\{\int V dP_U \mid U \in S(\Gamma)\}$ *and* $(C)\int V dP_* = \inf\{\int V dP_U \mid U \in S(\Gamma)\}$.

On the other hand, we have already remarked that the equality between $\Delta(\Gamma)$ and $M(P^*)$ relies on the equalities $P^*(A) = \max P(\Gamma)(A)$ and on the convexity of $P(\Gamma)(A)$ for every $A \in \mathcal{A}'$. Concerning the latter, we have proven the following:

Proposition 2 *[18] Let* $\Gamma : \Omega \to \mathcal{P}(X)$ *be a random set, and consider* $A \in \mathcal{A}'$. *Let* $U_1, U_2 \in S(\Gamma)$ *satisfy* $P_{U_1}(A) = \max P(\Gamma)(A)$, $P_{U_2}(A) = \min P(\Gamma)(A)$. *Then,*

$$P(\Gamma)(A) \text{ is convex} \Leftrightarrow U_1^{-1}(A) \setminus U_2^{-1}(A) \text{ is not an atom } ^1.$$

In particular, $P(\Gamma)(A)$ is convex $\forall A \in \mathcal{A}'$ if the initial space is non-atomic; this condition holds for instance if we have some additional information stating that P_{U_0} is continuous. Nevertheless, the non-atomicity of $(\Omega, \mathcal{A}, P)$ is not necessary for $P(\Gamma)(A)$ to be convex, as we showed in [16]. If we join Theorem 1 and Proposition 2, we derive the following corollary:

Corollary 1 *Let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set satisfying any of the conditions (1) to (5) from Theorem 1. If $A^* \setminus A_*$ is not an atom for any $A \in \beta_X$, $\Delta(\Gamma) = M(P^*)$.*

3.2 P^*, P_* as a model for P_{U_0}

Let us study now the equality between $P(\Gamma)$ and $M(P^*)$, which tells whether the upper probability keeps all the available information about the distribution of the original random variable, P_{U_0}. The class of the distributions of the selectors has been studied for some types of random sets (see for instance [1, 9, 10]). However, its relationship with the credal set generated by the upper probability has not been investigated in detail. In [16], we studied this problem for the case of X finite, and in [15] the attention was focused on random intervals. On the other hand, Castaldo and Marinacci proved in [3] a result for compact random sets on Polish spaces.

The equality between $\Delta(\Gamma)$ and $M(P^*)$ does not guarantee that $P(\Gamma) = M(P^*)$, and neither does the equality between $P(\Gamma)$ and $\Delta(\Gamma)$ ([16]). Then, a possible approach for our problem would be determining sufficient conditions for $P(\Gamma) = \Delta(\Gamma)$, and join them with the ones stated in Corollary 1. Unfortunately, it does not seem easy (except in trivial situations) to characterize this last equality. We are going to show that a reasoning based on the extreme points of $M(P^*)$ will be more fruitful in our context: it allows us to easily characterize the equality between $P(\Gamma)$ and $M(P^*)$ in the finite case, and we can use this to derive some results for the case of X separable metric. When X is finite, the extreme points of $M(P^*)$ are in correspondence with the permutations on X, in the following manner[2]:

Theorem 3 *[4] Consider $X = \{x_1, \ldots, x_n\}$ finite and μ a 2-alternating capacity on $\mathcal{P}(X)$. For any $\pi \in S^n$, define a probability Q_π on $\mathcal{P}(X)$ satisfying*

$$Q_\pi(\{x_{\pi(1)}, \ldots, x_{\pi(j)}\}) = \mu(\{x_{\pi(1)}, \ldots, x_{\pi(j)}\}) \forall j = 1, \ldots, n.$$

Then, $Ext(M(\mu)) = \{Q_\pi \mid \pi \in S^n\}$ and $M(\mu) = Conv(\{Q_\pi \mid \pi \in S^n\})$.

---

[1] By this we mean that for every $\alpha \in (0,1)$ there is some measurable $B \subseteq U_1^{-1}(A) \setminus U_2^{-1}(A)$ with $P(B) = \alpha P(U_1^{-1}(A) \setminus U_2^{-1}(A))$.

[2] This theorem is an extension, for 2-alternating capacities, of a result given by Dempster ([7]) for random sets on finite spaces.

We can see ([16]) that given X finite and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set, it is $Ext(M(P^*)) \subseteq P(\Gamma)$, and as a consequence $P(\Gamma) = M(P^*) \Leftrightarrow P(\Gamma)$ is convex. This equivalence does not hold for the case of X infinite, as the following example shows:

Example 1 (sketch) *Let $\Gamma : (0,1) \to \mathcal{P}([0,1])$ be defined between the probability space $((0,1), \beta_{(0,1)}, \lambda_{(0,1)})$ and the measurable space $([0,1], \beta_{[0,1]})$ by $\Gamma(\omega) = (0,\omega)\ \forall\omega \in (0,1)$. It is easy to see that this mapping is strongly measurable.*

- *Given $U \in S(\Gamma)$, it can be checked that $P_U(\{0\}) = 0, P_U([0,x]) \geq x\ \forall x$, and $\lambda_{(0,1)}(\{x \in (0,1) \mid P_U([0,x]) = x\}) = 0$.*
- *Conversely, consider a probability measure $Q : \beta_{[0,1]} \to [0,1]$ satisfying the three previous properties. This implies that it also satisfies $Q([0,x)) \geq x$ and $Q([0,x)) > x$ for all but a null subset of $(0,1)$, that we will denote N_Q. The quantile function U of Q is a measurable mapping satisfying $P_U = Q, U(\omega) \in \Gamma(\omega)\ \forall\omega \notin N_Q$. We can modify U on N_Q without affecting its measurability so that all its values are included in those of Γ, whence we deduce that $Q \in P(\Gamma)$.*
- *We deduce that $P(\Gamma)$ is the class of probability measures with $Q(\{0\}) = 0, Q([0,x]) \geq x\ \forall x$ and $Q([0,x]) > x$ for all but a null subset of $[0,1]$, and we can easily check that this class is convex.*
- *The Lebesgue measure $\lambda_{[0,1]}$ on $\beta_{[0,1]}$ satisfies $\lambda_{[0,1]}(A) \leq P^*(A)\ \forall A \in \beta_{[0,1]}$; hence, it belongs to $M(P^*)$, and clearly it does not satisfy $\lambda_{[0,1]}([0,x]) > x$ with probability 1. As a consequence, $P(\Gamma) \subsetneq M(P^*)$.*

In [17], we investigated the form of the extreme points of $M(\mu)$ for the case of μ 2-alternating and upper continuous, and for (X,d) a separable metric space. The idea in that paper was to approximate a distribution $Q : \beta_X \to [0,1]$ by distributions coinciding with Q on some finite fields. We will use a similar reasoning in our next theorem, where we consider the upper probability P^* induced by a random set (and hence not necessarily upper continuous). We will work in this paper with the topology of the weak convergence, whose main properties can be found in [2]. Together with the well-known Portmanteau's theorem, we will also use the following result:

Proposition 3 *[2] Let (X,d) be a separable metric space, and consider a class $\mathcal{U} \subseteq \beta_X$ such that (i) it is closed under finite intersections and (ii) every open set is a finite or countable union of elements from $\mathcal{U}$. Let $\{P_n\}_n, P$ be a family of probability measures on β_X such that $P_n(A) \to P(A)\ \forall A \in \mathcal{U}$. Then, the sequence $\{P_n\}_n$ converges weakly to P.*

Let $\{x_n\}_n$ be a countable set dense on (X,d), and define $\{B_n\}_n := \{B(x_i;q_j) \mid i \in \mathbb{N}, q_j \in \mathbb{Q}\}$ a countable basis of $\tau(d)$. Let us denote $Q(\{B_n\}_n)$ the field generated by $\{B_n\}_n$, Q_n the field generated by $\{B_1,\ldots,B_n\}$. Then, $Q(\{B_n\}_n) = \cup_n Q_n$, and it can easily be checked that $Q(\{B_n\}_n)$ satisfies the hypotheses (i) and (ii) stated in the previous proposition. Any element of Q_n is a (finite and disjoint) union of elements from $\mathcal{D}_n := \{C_1 \cap C_2 \cap \cdots \cap C_n \mid C_i \in \{B_i, B_i^c\} \forall i : 1, \ldots, n\}$. Let us denote this class $\mathcal{D}_n := \{E_1^n, \ldots, E_{k_n}^n\}$.

Theorem 4 *Let $(\Omega, \mathcal{A}, P)$ be a probability space, (X, d) a separable metric space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set such that $P^*(A) = \max P(\Gamma)(A)\ \forall A \in Q(\{B_n\}_n)$. Then,*

1. $\overline{M(P^*)} = \overline{Conv(P(\Gamma))}$.
2. $\overline{P(\Gamma)} = \overline{M(P^*)} \Leftrightarrow \overline{P(\Gamma)}$ *is convex.*

Proof.

1. It is clear that $\overline{Conv(P(\Gamma))} \subseteq \overline{M(P^*)}$. Conversely, consider $Q_1 \in M(P^*)$, and fix $n \in \mathbb{N}$. Consider the finite measurable space $(\mathcal{D}_n, \mathcal{P}(\mathcal{D}_n))$, and let us define the multi-valued mapping

$$\begin{aligned} \Gamma_n : \Omega &\to \mathcal{P}(\mathcal{D}_n) \\ \omega &\hookrightarrow \{E_i^n \mid \Gamma(\omega) \cap E_i^n \neq \emptyset\}. \end{aligned}$$

   - Given $I \subseteq \{1, \ldots, k_n\}, \Gamma_n^*(\{E_i^n\}_{i \in I}) = \{\omega \mid \exists i \in I, E_i^n \in \Gamma_n(\omega)\} = \{\omega \mid \exists i \in I, \Gamma(\omega) \cap E_i^n \neq \emptyset\} = \Gamma^*(\cup_{i \in I} E_i^n) \in \mathcal{A} \Rightarrow \Gamma_n$ is strongly measurable.
   - Define a probability measure $Q : \mathcal{P}(\mathcal{D}_n) \to [0,1]$ by $Q(\{E_i^n\}) = Q_1(E_i^n)\ \forall i$. Then, given $I \subseteq \{1, \ldots, k_n\}$,

$$Q(\{E_i^n\}_{i \in I}) = Q_1(\cup_{i \in I} E_i^n) \leq P_\Gamma^*(\cup_{i \in I} E_i^n) = P_{\Gamma_n}^*(\{E_i^n\}_{i \in I}),$$

   whence $Q \in M(P_{\Gamma_n}^*)$.

   Now, from Theorem 3 $M(P_{\Gamma_n}^*) = Conv(\{Q_\pi \mid \pi \in S^{k_n}\})$, where the probability measure $Q_\pi : \mathcal{P}(\mathcal{D}_n) \to [0,1]$ is defined by $Q_\pi(\{E_{\pi(1)}^n, \ldots, E_{\pi(j)}^n\}) = P_{\Gamma_n}^*(\{E_{\pi(1)}^n, \ldots, E_{\pi(j)}^n\}) = P_\Gamma^*(\cup_{i=1}^j E_{\pi(j)}^n)\ \forall j = 1, \ldots, k_n$.

   For any of these extreme points, there is some $P_\pi \in P(\Gamma)$ with $P_\pi(E_j^n) = Q_\pi(\{E_j^n\})\ \forall j = 1, \ldots, k_n$: it suffices to take into account that, as we have seen in Theorem 2, we can approximate P_Γ^* on a finite chain. As a consequence, for the probability $Q \in Conv(\{Q_\pi \mid \pi \in S^n\})$ defined above, there is some $P_n \in Conv(P(\Gamma))$ such that $P_n(E_j^n) = Q(\{E_j^n\}) = Q_1(E_j^n)\ \forall j = 1, \ldots, k_n$. The sequence $\{P_n\}_n \subseteq Conv(P(\Gamma))$ satisfies $P_n(A) \to Q_1(A)$ for all $A \in Q(\{B_n\}_n)$. Applying Proposition 3, we conclude that $\{P_n\}_n$ converges weakly to Q_1, whence $M(P^*) \subseteq \overline{Conv(P(\Gamma))}$ and we deduce the desired equality.

2. For the direct implication, it suffices to see that $\overline{M(P^*)}$ is convex. Consider $P_1, P_2 \in \overline{M(P^*)}, \alpha \in (0,1)$; then, there are $\{P_n^1\}_n, \{P_n^2\}_n \subset M(P^*)$ converging weakly to P_1, P_2, respectively. Let $A \in \beta_X$ be a $(\alpha P_1 + (1-\alpha)P_2)$-continuity set. It is $0 = (\alpha P_1 + (1-\alpha)P_2)(\delta(A))^3 = \alpha P_1(\delta(A)) + (1-\alpha)P_2(\delta(A))$, and therefore A is also a P_1, P_2-continuity set. From Portmanteau's theorem (see for instance [2]), $P_n^1(A) \to P_1(A)$ and $P_n^2(A) \to P_2(A)$, whence $(\alpha P_n^1 + (1-\alpha)P_n^2)(A) \to (\alpha P_1 + (1-\alpha)P_2)(A)$, and again using Portmanteau's theorem we deduce that the sequence $\{\alpha P_n^1 + (1-\alpha)P_n^2\}_n \subset M(P^*)$ converges weakly to $\alpha P_1 + (1-\alpha)P_2$. Hence, this probability belongs to $\overline{M(P^*)}$.

   For the converse implication, assume that $\overline{P(\Gamma)}$ is convex. Then, applying the first point of this theorem, it is

$$\overline{M(P^*)} = \overline{Conv(P(\Gamma))} \subseteq \overline{Conv(\overline{P(\Gamma)})} = \overline{\overline{P(\Gamma)}} = \overline{P(\Gamma)} \Rightarrow \overline{P(\Gamma)} = \overline{M(P^*)}.$$

□

The second part of this theorem extends a result mentioned before for the finite case (it can be checked that in that case both $P(\Gamma)$ and $M(P^*)$ are closed). We deduce that a way to determine conditions for the equality $\overline{M(P^*)} = \overline{P(\Gamma)}$ is to give sufficient conditions for the convexity of $\overline{P(\Gamma)}$. We have done this in our next theorem. It uses the following supporting result.

Lemma 1 *Let $(\Omega, \mathcal{A}, P)$ be a non-atomic probability space, (X,d) a separable metric space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. Then, the class of probabilities $\mathcal{H}_n := \{Q : \mathcal{P}(\mathcal{D}_n) \to [0,1]$ probability $\mid \exists Q' \in P(\Gamma)$ such that $Q(\{E_i^n\}) = Q'(E_i^n)\ \forall E_i^n \in \mathcal{D}_n\}$ is convex for every n.*

Proof. Fix $n \in \mathbb{N}$, and consider $P_1, P_2 \in \mathcal{H}_n, \alpha \in (0,1)$. Then, there exist $U_1, U_2 \in S(\Gamma)$ with $P_{U_1}(E_i^n) = P_1(\{E_i^n\}), P_{U_2}(E_i^n) = P_2(\{E_i^n\})\ \forall i = 1, \ldots, k_n$. Let us consider the measurable partition of Ω given by $\{C_{ij} \mid i,j = 1, \ldots, k_n\}$ with $C_{ij} = U_1^{-1}(E_i^n) \cap U_2^{-1}(E_j^n)$; from the non-atomicity of $(\Omega, \mathcal{A}, P)$, there is, for every i,j, some measurable $D_{ij} \subseteq C_{ij}$ such that $P(D_{ij}) = \alpha P(C_{ij})$. Define $C = \cup_{i,j} C_{ij}$, and

$$U := U_1 I_C + U_2 I_{C^c}.$$

Then, U is a measurable combination of selectors of Γ, whence $U \in S(\Gamma)$. Moreover,

[3] $\delta(A)$ denotes here the boundary of the set A.

$$
\begin{aligned}
P_U(E_l^n) = P(U_1^{-1}(E_l^n) \cap C) + P(U_2^{-1}(E_l^n) \cap C^c) \\
&= \sum_{i=1}^{k_n} P(D_{li}) + \sum_{j=1}^{k_n} (P(C_{jl}) - P(D_{jl})) \\
&= \sum_{i=1}^{k_n} \alpha P(C_{li}) + \sum_{j=1}^{k_n} (1-\alpha) P(C_{jl}) \\
&= \alpha P_{U_1}(E_l^n) + (1-\alpha) P_{U_2}(E_l^n) \ \forall l = 1, \ldots, k_n,
\end{aligned}
$$

and we deduce that $\alpha P_1 + (1-\alpha)P_2 \in \mathcal{H}_n$. □

Theorem 5 *Let $(\Omega, \mathcal{A}, P)$ be a non-atomic probability space, (X,d) a separable metric space and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. Then, $\overline{P(\Gamma)}$ is convex.*

Proof. Let us show first that $Conv(P(\Gamma)) \subseteq \overline{P(\Gamma)}$. Consider $P_1, P_2 \in P(\Gamma), \alpha \in (0,1)$. Applying the previous lemma, for every n there is $Q_n \in P(\Gamma)$ with $Q_n(A) = (\alpha P_1 + (1-\alpha)P_2)(A)\ \forall A \in \mathcal{Q}_n$, where $\mathcal{Q}_n$ is the field generated by $\{B_1, \ldots, B_n\}$. Now, applying Proposition 3 we deduce that $\{Q_n\}_n$ converges weakly to $\alpha P_1 + (1-\alpha)P_2$ and this probability belongs to $\overline{P(\Gamma)}$. From this, we deduce in particular the equality $\overline{Conv(P(\Gamma))} = \overline{P(\Gamma)}$. The first set in this equality is the closure of a convex set of probabilities defined on a separable metric space. Following the course of reasoning from the proof of point 2 from Theorem 4, we can deduce that $\overline{Conv(P(\Gamma))}$ (and hence $\overline{P(\Gamma)}$) is convex. □

A similar proof would allow us to deduce that $\overline{\Delta(\Gamma)}$ is convex when $(\Omega, \mathcal{A}, P)$ is non-atomic and (X,d) separable. Now, using Theorems 1, 4 and 5, we derive the following result:

Corollary 2 *Let $(\Omega, \mathcal{A}, P)$ be a probability space, (X,d) be a separable metric space, and $\Gamma : \Omega \to \mathcal{P}(X)$ a random set. Under any of the following conditions:*

1. *Γ is open,*
2. *Γ is complete,*
3. *X is σ-compact and Γ is closed,*

$\overline{M(P^*)} = \overline{Conv(P(\Gamma))}$. If in addition $(\Omega, \mathcal{A}, P)$ is non-atomic, then $\overline{M(P^*)} = \overline{P(\Gamma)}$.

Proof. The first part follows from Theorem 1 and the first point of Theorem 4. For the second part, it suffices to apply the second point of Theorem 4 and Theorem 5. □

This corollary extends results from [3, 16], and tells us that under fairly general conditions, the upper probability can be used to model the available information without producing a (big) loss of precision. It also extends some results from [10]: it is proven there that given two closed random sets Γ_1, Γ_2 on a separable Banach space, the equality between $P^*_{\Gamma_1}$ and $P^*_{\Gamma_2}$ implies that $\overline{Conv(P(\Gamma_1))}$ is equal to $\overline{Conv(P(\Gamma_2))}$. Similar results can be seen in [1, 9]. We have proven that it is indeed $P^*_{\Gamma_1} = P^*_{\Gamma_2} \Rightarrow \overline{Conv(P(\Gamma_1))} = \overline{Conv(P(\Gamma_2))} = \overline{M(P^*_{\Gamma_1})} = \overline{M(P^*_{\Gamma_2})}$, and only requiring Γ_i to be complete on a separable metric space $\forall i = 1,2$. On the other hand, we deduce that under the hypotheses of the second part of Corollary 2, if $P(\Gamma)$ is weakly closed, it is also convex, and $M(P^*)$ is closed. The converses are not true in general. The following example shows that $P(\Gamma)$ is not necessarily closed when $M(P^*)$ is closed:

Example 2 *[15] Consider $\Gamma : [0,1] \to \mathcal{P}([0,1])$ defined by $\Gamma(\omega) = [-\omega, \omega]\ \forall \omega \in [0,1]$. Then, it can be proven that $M(P^*)$ is closed (indeed, this holds for any compact random set on a Polish space). However, given the selectors $A, B \in S(\Gamma)$ defined by $A(\omega) = -\omega, B(\omega) = \omega\ \forall \omega$, it can be checked that $\frac{P_A + P_B}{2} \notin P(\Gamma)$. This shows that $P(\Gamma)$ is not convex. As a consequence, it is not closed either: if it were, it would be $P(\Gamma) = \overline{P(\Gamma)} = \overline{M(P^*)} = M(P^*)$ convex, a contradiction.*

On the other hand, Example 1 shows that $P(\Gamma)$ is not closed either if it is convex. Indeed, that implication does not hold even if $P(\Gamma) = M(P^*)$:

Example 3 *Consider $(\Omega, \mathcal{A}, P) = ((0,1), \beta_{(0,1)}, \lambda_{(0,1)})$ a non-atomic probability space, and let $\Gamma : \Omega \to \mathcal{P}(\mathbb{R})$ be constant on $(0,1)$. Then, $M(P^*) = \{Q : \beta_{\mathbb{R}} \to [0,1]$ probability $\mid Q((0,1)) = 1\}$. Given a probability measure $Q \in M(P^*)$, its quantile function $U : (0,1) \to \mathbb{R}$ is a selector of Γ and satisfies $P_U = Q$, whence $P(\Gamma) = M(P^*)$ convex. However, the sequence of degenerate probability measures on $\frac{1}{n}$, $\{\delta_{\frac{1}{n}}\}_n \subseteq P(\Gamma)$, converges weakly to $\delta_0 \notin P(\Gamma)$. Hence, this set is not closed.*

4 Conclusions and open problems

In this paper, we have compared the different models of the probabilistic information given by a random set, when this random set is the imprecise observation of a random variable. We have considered three different sets of probability measures, and through them we have investigated whether an imprecise probability model in terms of the upper and lower probabilities is useful in this context.

The results we have established allow us to conclude that under fairly general conditions, the upper and lower probabilities induced by a random set can be used to summarize the information on the distribution of the original random variable without a substantial loss of precision. Nevertheless, there are still a number of particular cases of random sets which are worth a detailed study. We would also like to study the topological structure of $P(\Gamma)$ and $M(P^*)$ under other than the topology of the weak convergence, and derive other sufficient conditions for the equalities $\Delta(\Gamma) = M(P^*)$ and $\overline{P(\Gamma)} = \overline{M(P^*)}$. Finally, it would also be interesting (though we are not very optimistic in this respect) to obtain sufficient conditions for the equality $P(\Gamma) = M(P^*)$ in terms of the images of the random set, as it was done in [15] for the particular case of random intervals.

References

[1] Z. Arstein and S. Hart. Law of large numbers for random sets and allocation processes. *Mathematics of Operations Research*, 6:485–492, 1981.

[2] P. Billingsley. *Convergence of probability measures*. Wiley, New York, 1968.

[3] A. Castaldo and M. Marinacci. Random correspondences as bundles of random variables. In *Proceedings of the 2nd ISIPTA Conference*, Ithaca (New York), 2001.

[4] A. Chateauneuf and J.-Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17:263–283, 1989.

[5] I. Couso. *Teoría de la Probabilidad para datos imprecisos. Algunos aspectos*. Ph.D. Thesis, University of Oviedo, 1999. In Spanish.

[6] I. Couso, S. Montes and P. Gil. Second order possibility measure induced by a fuzzy random variable. In C. Bertoluzza, M. A. Gil and D. A. Ralescu, editors, *Statistical modeling, analysis and management of fuzzy data*. Springer, 2002.

[7] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

[8] D. Denneberg. *Non-additive measure and integral*. Kluwer Academic Publishers, 1994.

[9] S. Hart and E. Kohlberg. Equally distributed correspondences. *Journal of Mathematical Economics*, 1:167–674, 1974.

[10] C. Hess. The distribution of unbounded random sets and the multivalued strong law of large numbers in nonreflexive banach spaces. *Journal of Convex Analysis*, 6:163–182, 1999.

[11] W. Hildenbrand. *Core and Equilibria of a Large Economy*. Princeton University Press, Princeton, 1974.

[12] C. J. Himmelberg. Measurable relations. *Fundamenta Mathematicae*, 87:53–72, 1975.

[13] R. Kruse and K. D. Meyer. *Statistics with vague data*. D. Reidel Publishing Company, Dordretch, 1987.

[14] G. Matheron. *Random sets and integral geometry*. Wiley, New York, 1975.

[15] E. Miranda. Estudio de la información probabilística de los intervalos aleatorios. In *Proceedings of the 27th SEIO conference*, Lérida (Spain), 2003. In Spanish.

[16] E. Miranda, I. Couso and P. Gil. Upper probabilities and selectors of random sets. In P. Grzegorzewski, O. Hryniewicz and M. A. Gil, editors, *Advances in soft computing*. Physica-Verlag, 2002.

[17] E. Miranda, I. Couso and P. Gil. Extreme points of credal sets generated by 2-alternating capacities. *International Journal of Approximate Reasoning*, 33:95-115, 2003.

[18] E. Miranda, I. Couso and P. Gil. Upper probabilities attainable by distributions of measurable selections, 2002, submitted.

[19] E. Miranda, G. de Cooman and I. Couso. Imprecise probabilities induced by multi-valued mappings. In *Proceedings of the 9th IPMU Conference*, Annecy (France), 2002.

[20] H. T. Nguyen. On random sets and belief functions. *Journal of Mathematical Analysis and Applications*, 63:531–542, 1978.

[21] P. Walley. *Statistical reasoning with imprecise probabilities*. Chapman and Hall, London, 1991.

Enrique Miranda is with the Department of Statistics and O. R. of the University of Oviedo, C-Calvo Sotelo s/n, 33007, Oviedo. E-mail:emiranda@correo.uniovi.es

Inés Couso is with the Department of Statistics and O. R. of the University of Oviedo, E. S. de Marina Civil, Ctra. de Villaviciosa s/n 33203, Gijón. E-mail:couso@pinon.ccu.uniovi.es

Pedro Gil is with the Department of Statistics and O. R. of the University of Oviedo, C-Calvo Sotelo s/n, 33007, Oviedo. E-mail:pedro@pinon.ccu.uniovi.es