Random sets and belief functions

Enrique Miranda
University of Oviedo
mirandaenrique@uniovi.es

3rd Belief School, September 2015

Outline

- Random sets on finite spaces.
- Representation by belief functions.
- Particular cases.
- Connection with measurable selections.
- Extensions to the infinite case.

Random sets on finite spaces

Consider a probability space $(\Omega, \mathcal{A}, P)$ and a finite space X. A random set is a map

$$\Gamma : \Omega \to \mathcal{P}(X)$$

satisfying the following measurability condition:

$$A^* := \{\omega : \Gamma(\omega) \cap A \neq \emptyset\} \in \mathcal{A} \ \forall A \subseteq X.$$

We shall assume throughout that $\Gamma(\omega) \neq \emptyset$ for every ω (which will be equivalent to dealing with normalized belief functions).

Measurability

The above condition is called strong measurability, and in this case, where X is finite, is equivalent to each of the following conditions:

1. $A_* := \{\omega : \Gamma(\omega) \subseteq A\} \in \mathcal{A} \ \forall A \subseteq X$.
2. $\{\omega : \Gamma(\omega) = A\} \in \mathcal{A} \ \forall A \subseteq X$.

It reduces to the usual measurability condition of random variables when $\Gamma(\omega)$ is a singleton for every ω.

A_*, A^* are called the lower and upper inverses of A, respectively.

Example

Consider
$\Omega = \{1,2,3\}, \mathcal{A} = \{\{1,2\},\{3\},\Omega,\emptyset\}, P(\{1,2\}) = \frac{2}{3}, P(\{3\}) = \frac{1}{3}$
and $\Gamma : \Omega \to \mathcal{P}(\{1,2,3\})$ given by

$$\Gamma(1) = \{1,2\}, \Gamma(2) = \{2,3\}, \Gamma(3) = \{1,3\}.$$

Given $A = \{1\}$, it holds that $A^* = \{\omega : 1 \in \Gamma(\omega)\} = \{1,3\} \notin \mathcal{A}$.
Thus, Γ is NOT a random set.

Upper and lower probabilities of a random set

Given a set $A \subseteq X$, Dempster defined its upper and lower probabilities by

$$P^*(A) := P(A^*) = P(\{\omega : \Gamma(\omega) \cap A \neq \emptyset\})$$

and

$$P_*(A) := P(A_*) = P(\{\omega : \Gamma(\omega) \subseteq A\}).$$

It holds that $P_*(A) \leq P^*(A)\ \forall A \subseteq X$. They are moreover conjugate functions: $P^*(A) = 1 - P_*(A^c)\ \forall A \subseteq X$. As we shall see, there is a connection with belief functions.

Example

Consider the same multi-valued mapping as before, but now with $\mathcal{A} = \mathcal{P}(\Omega)$ and $P(\{1\}) = P(\{2\}) = P(\{3\}) = \frac{1}{3}$, so that Γ is a random set.

Given $A = \{1\}$, it holds that

- $A^* = \{1, 3\} \Rightarrow P^*(A) = \frac{2}{3}$.
- $A_* = \{\omega : \Gamma(\omega) = \{1\}\} = \emptyset \Rightarrow P_*(A) = 0$.

Particular case: random variables

When Γ is single-valued, then for every $A \subseteq X$ it holds that

$$A_* = A^* = \Gamma^{-1}(A).$$

Thus, the measurability condition is the usual measurability condition, and the lower and upper probabilities coincide with the probability measure induced by the random variable.

Exercise

Consider $\Omega = \{1, 2, 3\}$ with the probability measure $P(\{1\}) = 0.3, P(\{2\}) = 0.5, P(\{3\}) = 0.2$ and the random set $\Gamma : \Omega \to \mathcal{P}(\{1, 2, 3\})$ given by

$$\Gamma(1) = \{1, 2\}, \Gamma(2) = \{2, 3\}, \Gamma(3) = \{1, 2, 3\}.$$

Determine the upper and lower probabilities of the sets $A = \{1\}, B = \{1, 2\}$ and $C = \{2, 3\}$.

Basics of belief functions (again)

Given a finite space X, a belief function or ∞-monotone Choquet capacity on $\mathcal{P}(X)$ is a function $Bel : \mathcal{P}(X) \to [0,1]$ such that for every natural number n and every family $\{A_1, \dots, A_n\}$ of subsets of X, it holds that

$$Bel(A_1 \cup \dots A_n) \geq \sum_{\emptyset \neq I \subseteq \{1,\dots,n\}} (-1)^{|I|+1} Bel(\cap_{i \in I} A_i).$$

Basic probability assignment

From Shafer, a function $m : \mathcal{P}(X) \to [0,1]$ is called a basic probability assignment when it satisfies $m(\emptyset) = 0$ and $\sum_{A \subseteq X} m(A) = 1$.

- Given a basic probability assignment m, the function $Bel : \mathcal{P}(X) \to [0,1]$ given by

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

  is a belief function.
- If Bel is a belief function, the map $m : \mathcal{P}(X) \to [0,1]$ given by

$$m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} Bel(B)$$

  is a basic probability assignment.

Plausibility functions

The conjugate Pl of a belief function is called a plausibility function. It is related to the same basic probability assignment via the formula

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B).$$

Moreover, this correspondence between belief functions, basic probability assignments and plausibility measures is one-to-one. m is called the Möbius inverse of Bel.

Focal elements

Given a belief function *Bel* with Möbius inverse m, a subset A of X is called a focal element of m when $m(A) \neq 0$. In particular, the focal elements of a belief function are those sets for which $m(A) > 0$.

The focal elements are useful when working with a lower probability. In this sense, in game theory we have the so-called k-additive measures, which are those whose focal elements have cardinality smaller or equal than k.

Belief functions and random sets (Nguyen, 1978)

Let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set. Then its lower probability P_* is a belief function, and its upper probability P^* is the conjugate plausibility function.

The Möbius inverse of P_* is given by

$$m(A) = P(\{\omega : \Gamma(\omega) = A\}).$$

Thus, the focal elements of P_* are the subsets A of X for which $P(\Gamma^{-1}(A)) > 0$.

Example

Let $(\Omega, \mathcal{A}, P) = ([0,1], \beta_{[0,1]}, \lambda_{[0,1]})$, $X = \{1,2,3\}$ and $\Gamma : \Omega \to \mathcal{P}(X)$ given by

$$\Gamma(\omega) = \begin{cases} \{1,2\} & \text{if } \omega < 0.3 \\ \{3\} & \text{if } \omega = 0.3 \\ \{1,2,3\} & \text{if } \omega \in (0.3, 0.5] \\ \{2,3\} & \text{if } \omega > 0.5 \end{cases}$$

Then P_* is the belief function with focal elements $m(\{1,2\}) = 0.3, m(\{1,2,3\}) = 0.2, m(\{2,3\}) = 0.5$. As a consequence, we obtain for instance $P_*(\{1,3\}) = 0$, $P_*(\{1,2\}) = 0.3$, $P^*(\{2,3\}) = 1 = P^*(\{1,2\})$.

From belief functions to random sets

We have seen that any random set induces a belief function. Conversely, any belief function *Bel* can be obtained as the lower probability P_* of a random set: this result is called Choquet's theorem, and we say that the random set is the source of Bel.

To see this, consider an arbitrary order among the focal elements of m, $A_1 \prec A_2 \prec \cdots \prec A_n$ and define $\Gamma : [0,1) \to \mathcal{P}(X)$ by

$$\Gamma(\omega) = A_i \text{ if } \omega \in [a_{i-1}, a_i),$$

where $a_{-1} = 0$, $a_i = \sum_{1 \leq j \leq i} m(A_j)$.

Thus, the two models (random sets and belief functions) are equally expressive.

Non-uniqueness

Note that two different random sets may have the same lower probability P_*: if we have the basic probability assignment m with focal elements $\{A_1, \ldots, A_n\}$, we could also consider $\Omega = \{1, \ldots, n\}, \mathcal{A} = \mathcal{P}(\Omega)$, with $P(\{i\}) = m(A_i)$ and let

$$\Gamma : \Omega \to \mathcal{P}(X)$$
$$i \hookrightarrow A_i.$$

Example

If we consider the belief function *Bel* on $\mathcal{P}(\{1,2,3,4\})$ with basic probability assignment

$$m(\{1,2,3\}) = 0.2 = m(\{1\}), m(\{2,3\}) = 0.1 = m(\{4\}), m(\{3,4\}) = 0.4,$$

then we can consider $\Omega = \{1,2,3,4,5\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, P the probability measure determined by

$$P(1) = P(2) = 0.2, P(3) = P(4) = 0.1, P(5) = 0.4$$

and the random set Γ given by

$$\Gamma(1) = \{1,2,3\}, \Gamma(2) = \{1\}, \Gamma(3) = \{2,3\}, \Gamma(4) = \{4\}, \Gamma(5) = \{3,4\}.$$

Then the lower probability of this random set is *Bel*.

Exercise

Consider the belief function *Bel* on $\mathcal{P}(\{1,2,3\})$ given by

A	{1}	{2}	{3}	{1,2}	{1,3}	{2,3}	{1,2,3}
Bel(A)	0.1	0.2	0	0.5	0.3	0.4	1

Determine a random set having *Bel* as its lower probability.

Example: vacuous belief functions

The case where we have the least amount of information corresponds to the basic probability assignment $m(X) = 1$, $m(A) = 0$ for every $A \subsetneq X$. The corresponding belief and plausibility functions are

$$Bel(A) = 0 \ \forall A \neq X,$$

$$Pl(A) = 1 \ \forall A \neq \emptyset.$$

These are called vacuous, and model the most imprecise situation.

The corresponding random set would be $\Gamma : \Omega \to \mathcal{P}(X)$ given by $\Gamma(\omega) = X$ for all ω.

Example: probability measures

Let $Bel : \mathcal{P}(X) \to [0, 1]$ be a belief function, and let Pl be its conjugate plausibility function. The following are equivalent:

1. Bel is a probability measure.
2. The focal elements of μ are singletons.
3. $Bel = Pl$.
4. $Bel(A) + Bel(A^c) = 1$ for every $A \subseteq X$.

For them, the associated random set Γ satisfies $|\Gamma(\omega)| = 1 \; \forall \omega$, and becomes thus a random variable. We say that the belief function is Bayesian.

Particular cases: possibility measures

Given X finite, an upper probability $\overline{P} : \mathcal{P}(X) \to [0,1]$ is called a possibility measure when

$$\overline{P}(A \cup B) = \max\{\overline{P}(A), \overline{P}(B)\}$$

para todo $A, B \subseteq X$.

Its conjugate is called a necessity measure, and it satisfies $\underline{P}(A \cap B) = \min\{\underline{P}(A), \underline{P}(B)\}$ for every $A, B \subseteq X$.

A necessity measure is a belief function, and corresponds to the case where the focal elements are nested. They are also called consonant belief functions.

Exercise

Consider $X = \{1, 2, 3, 4\}$.

- Let Π be the possibility distribution associated to the possibility distribution $\pi(1) = 0.3, \pi(2) = 0.5, \pi(3) = 1, \pi(4) = 0.7$. Determine its focal elements and its basic probability assignment.
- Given the basic probability assignment $m(\{1\}) = 0.2, m(\{1, 3\}) = 0.1, m(\{1, 2, 3\}) = 0.4, m(\{1, 2, 3, 4\}) = 0.3$, determine the associated possibility measure and its possibility distribution.

Possibility measures and random sets

Given a random set $\Gamma : \Omega \to \mathcal{P}(X)$ on a finite space X, the following are equivalent:

1. P^* is a possibility measure.
2. There exists a null subset N of Ω such that, for every $\omega_1, \omega_2 \in \Omega \setminus N$, either $\Gamma(\omega_1) \subseteq \Gamma(\omega_2)$ or $\Gamma(\omega_2) \subseteq \Gamma(\omega_1)$.

In other words, the upper probability of a random set is a possibility measure if and only if its images are nested.

Ontic and epistemic interpretations

As discussed by Dubois and Couso, random sets can be given two different interpretations:

- The ontic or conjunctive one: $\Gamma(\omega)$ is a multi-valued random variable. This interpretation is used by Kendall or Mathéron, amongst others.

- The epistemic one: Γ is a model for an ill-known random variable U_0, so that all we know about $U_0(\omega)$ is that it belongs to $\Gamma(\omega)$. This interpretation is used by Dempster and it is closer to imprecise probabilities.

The interpretation we use has implications when modelling conditioning or independence, for instance.

Example

Assume that $X = \{Spanish, French, English\}$ is a set of languages, and that we consider $\Gamma : \Omega \to \mathcal{P}(X)$.

- Under an ontic interpretation, $\Gamma(\omega)$ could be the set of languages a person speaks. Then, the probability that a person speaks English would be

$$\sum_{\text{English} \in A} P(\Gamma^{-1}(A)).$$

- Under an epistemic interpretation, $\Gamma(\omega)$ could be our imprecise knowledge of a person's native language. Then, the probability that a person's native language is English would belong to $[P_*(A), P^*(A)]$, where A='English'.

Conditioning conjunctive random sets

Assume that we know that the value of the random set Γ is included in some set $A \subseteq X$. Then the conditional distribution of the random set can be obtained by applying Bayes' rule on the probability distribution of Γ; this produces

$$P_\Gamma(C|A) := \begin{cases} \frac{P(\Gamma^{-1}(C))}{\sum_{B \subseteq A} P(\Gamma^{-1}(B))} & \text{if } C \subseteq A \\ 0 & \text{otherwise.} \end{cases}$$

Conditioning disjunctive random sets

If instead Γ is understood as a model for the imprecise knowledge of a random variable U_0, then we obtain a set of possible values

$$\{P_U(C|A) : U \in S(\Gamma), P_U(A) > 0\},$$

where

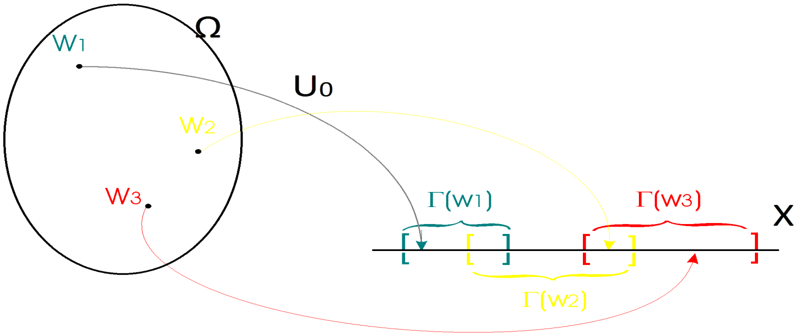
$$S(\Gamma) := \{U : \Omega \to X \text{ r.v.} | U(\omega) \in S(\Gamma)\ \forall \omega\},$$

which, by taking the lower envelope, produces

$$P_*(C|A) = \frac{P_*(C \cap A)}{P_*(C \cap A) + P^*(C^c \cap A)};$$

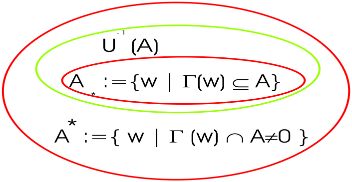
this formula is called the regular extension of the belief function P_*.

Epistemic random sets



Epistemic random sets: upper and lower inverses

Given $A \in \mathcal{A}'$, it is



A^* and A_* are called upper and lower inverses of A by Γ, respectively.

Measurable selections of a random set

Under the epistemic interpretation, the information about the original random variable U_0 is given by the measurable selections of Γ:

$$S(\Gamma) := \{U : \Omega \to X \text{ r.v.} | U(\omega) \in S(\Gamma)\ \forall\omega\},$$

and the set of possible distributions of U_0 is given by

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

In particular, for every $A \subseteq X$ we define

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

Example

Let $\Omega = \{1,2,3,4\} = X, \mathcal{A} = \mathcal{P}(\Omega)$, P the uniform distribution, and $\Gamma : \Omega \to \mathcal{P}(X)$ given by

$$\Gamma(1) = \{1\}, \Gamma(2) = \{1,3\}, \Gamma(3) = \{2,4\}, \Gamma(4) = \{3,4\}.$$

The measurable selections of Γ are:

	1	2	3	4
U_1	1	1	2	3
U_2	1	1	2	4
U_3	1	1	4	3
U_4	1	1	4	4
U_5	1	3	2	3
U_6	1	3	2	4
U_7	1	3	4	3
U_8	1	3	4	4

Upper and lower probabilities

For a given set $A \subseteq X$:

- $P_*(A) = P(\{\omega : \Gamma(\omega) \subseteq A\})$ measures the total degree of support of the set A.
- $P^*(A) = P(\{\omega : \Gamma(\omega) \cap A \neq \emptyset\})$ measures the evidence that is consistent with the set A.

Representation by means of belief function

Denote the core of the belief function P_* by

$$M(P_*) := \{Q \text{ prob. measure} | Q(A) \geq P_*(A)\ \forall A\}.$$

- $P_*(A) = \min\{Q(A) : Q \in P(\Gamma)\}\ \forall A \subseteq X$.
- $M(P_*) = Conv(P(\Gamma))$.
- If $(\Omega, \mathcal{A}, P)$ is non-atomic, then $P(\Gamma) = M(P_*)$.

The key for this result is that the extreme points of $M(P_*)$ belong to $P(\Gamma)$.

Example (cont.)

Given $A = \{1, 3\}$, we deduce that $P(\Gamma)(A) = \{0.5, 0.75\}$. As a consequence, $[P_*(A), P^*(A)] = [0.5, 0.75]$.

We can also see that

$$P_*(A) = P(\{\omega : \Gamma(\omega) \subseteq A\}) = P(\{1, 2\}) = 0.5$$

and

$$P^*(A) = P(\{\omega : \Gamma(\omega) \cap A \neq \emptyset\}) = P(\{1, 2, 4\}) = 0.75$$

Exercise

Consider $\Omega = \{1, 2, 3\}$ with the uniform distribution,
$\Gamma : \Omega \to \mathcal{P}(\{1, 2\})$, $\Gamma(\omega) = \{1, 2\}\ \forall\omega$.

- Show that $P(\Gamma) = \{(0, 1), (1/3, 2/3), (2/3, 1/3), (1, 0)\}$.
- Show that $M(P^*) = \{(\alpha, 1 - \alpha) : \alpha \in [0, 1]\}$.

Let f be the Shannon entropy of a probability measure,
$f(P) = -\sum_i P(x_i) log_2(P(x_i))$.

- $f(P_{U_0}) \in f(P(\Gamma)) = \{0, f(1/3, 2/3)\} = \{0, 0.92\}$.
- $f(M(P_*)) = [0, 1]$.

$M(P_*)$ and mass allocations

Each probability measure $P \in M(P_*)$ corresponds to an allocation of the basic probability assignment of P_*, where the mass $m(A_i)$ of each focal element is distributed between the elements of A_i. The resulting measure has only the singletons as focal elements, and is thus a probability measure.

If for example P_* is associated to the basic probability assignment $m(\{1,2\}) = 0.5, m(\{1,3\}) = 0.3, m(\{2,3\}) = 0.2$, the probability measure $P \equiv (0.6, 0.3, 0.1) \in M(P_*)$ would result for instance:

$$m(\{1,2\}) = 0.5 \rightarrow \text{ mass } 0.4 \text{ for } \{1\} \text{ and } 0.1 \text{ for } 2$$
$$m(\{1,3\}) = 0.3 \rightarrow \text{ mass } 0.2 \text{ for } \{1\} \text{ and } 0.1 \text{ for } 3$$
$$m(\{2,3\}) = 0.2 \rightarrow \text{ mass } 0.2 \text{ for } \{2\} \text{ and } 0 \text{ for } 3$$

The Choquet integral

Given a function $f : X \to \mathbb{R}$, $f = \sum_{i=1}^{n} x_i \mathbb{I}_{A_i}$, for $x_1 > x_2 > \cdots > x_n$ and for some finite partition $\{A_1, \ldots, A_n\}$ of Ω, its Choquet integral with respect to a non-additive measure μ is given by

$$(C) \int f d\mu = \sum_{i=1}^{n} x_i(\mu(S_i) - \mu(S_{i-1}) = \sum_{i=1}^{n} (x_i - x_{i+1})\mu(S_i),$$

where $S_i = \cup_{j=1}^{i} A_j$, $S_0 = \emptyset, x_{n+1} = 0$.

This is a generalization of the expectation of a random variable to non-additive measures

Example

Consider the belief function *Bel* associated with the basic probability assignment

$$m(\{1,2,3\}) = 0.2, m(\{1,3\}) = 0.3, m(\{2,3\}) = 0.4, m(\{1\}) = 0.1$$

and let f be given by $f(1) = 3, f(2) = 5, f(3) = 0$. Then $f = 5I_2 + 3I_1 + 0I_3$, so its Choquet integral is given by

$$(C)\int f d\mu = (5-3)\mu(\{2\}) + (3-0)\mu(\{1,2\}) + 0\mu(\{1,2,3\})$$
$$= 2\cdot 0 + 3\cdot 0.1 + 0\cdot \mu(\{1,2,3\}) = 0.3.$$

Choquet integral: basic properties

- $(C) \int I_A d\mu = \mu(A)$.
- $(C) \int cf d\mu = c(C) \int f d\mu \; \forall c \geq 0$.
- $f \leq g \Rightarrow (C) \int f d\mu \leq (C) \int g d\mu$.
- The Choquet integral is not additive in general, but $\int (f+g) d\mu = \int f d\mu + \int g d\mu$ when f, g are comonotonic.

A more detailed account can be found in the book of Denneberg.

Choquet integral and measurable selections

Given a function $f : X \to \mathbb{R}$, and monotone set functions $\mu_1 \leq \mu_2 \leq \mu_3$, it holds that

$$(C)\int f d\mu_1 \leq (C)\int f d\mu_2 \leq (C)\int f d\mu_3.$$

As a consequence, given a random set $\Gamma : \Omega \to \mathcal{P}(X)$ and $f : X \to \mathbb{R}$ it holds that

$$(C)\int f dP_* \leq \int f dP_U \leq (C)\int f dP^* \ \forall U \in S(\Gamma).$$

The previous results ensures that in fact

$$(C)\int f dP_* = \min\left\{\int f dQ : Q \in P(\Gamma)\right\}.$$

Exercise

Consider $\Omega = \{1, 2, 3, 4\} = X, \mathcal{A} = \mathcal{P}(\Omega), P$ the uniform distribution, and $\Gamma : \Omega \to \mathcal{P}(X)$ given by

$$\Gamma(1) = \{1\}, \Gamma(2) = \{1, 3\}, \Gamma(3) = \{2, 4\}, \Gamma(4) = \{3, 4\}.$$

Determine the Choquet integral of the mapping f given by $f(1) = 10, f(2) = 4, f(3) = 0, f(4) = 2$ with respect to P^* and P_*, and give the measurable selections that attain this lower and upper integral.

Probability boxes

Given a finite set $X = \{x_1, \ldots, x_n\}$ with $x_1 < x_2 < \ldots, < x_n$, a cumulative distribution function is a non-decreasing map $F : X \to [0,1]$ satisfying $F(x_n) = 1$. It represents the cumulative probabilities of a random variable taking values in X.

Given two cumulative distribution functions $F_* \leq F^*$, the set

$$(F_*, F^*) := \{F \text{ cdf} : F_*(x_i) \leq F(x_i) \leq F^*(x_i) \ \forall i\}$$

is called a probability box, or p-box (Ferson et al.). It can be seen as a model for the imprecise knowledge of a distribution function.

Given a p-box (F_*, F^*), the lower envelope of the set $\{P : F_P \in (F_*, F^*)\}$ is a belief function (Troffaes and Destercke).

Distribution functions of a random set

When $X \subseteq \mathbb{R}$, the restriction to cumulative sets of P^* and P_* determine the upper and lower distribution functions $F^*, F_* : X \to [0, 1]$ given by

$$F_*(x) = P_*(\{z \leq x\}) \text{ and } F^*(x) = P^*(\{z \leq x\}).$$

However, the pair (F_*, F^*) does not include all the information of the random set (Couso): we may find probability measures Q such that

$$F_Q(x) \in [F_*(x), F^*(x)] \; \forall x \in X$$

while $Q \notin M(P_*)$.

Example

Let $\Omega = \{1,2,3,4\} = X, \mathcal{A} = \mathcal{P}(\Omega), P$ the uniform distribution, and $\Gamma : \Omega \to \mathcal{P}(X)$ given by

$$\Gamma(1) = \{1\}, \Gamma(2) = \{1,3\}, \Gamma(3) = \{2,4\}, \Gamma(4) = \{3,4\}.$$

The lower and upper distributions are given by:

	1	2	3	4
F_*	0.25	0.25	0.5	1
F^*	0.5	0.75	1	1

The function F given by $F(1) = F(2) = 0.5, F(3) = F(4) = 1$ belongs to (F_*, F^*), but it is not induced by any measurable selection, because $P_U(\{2,4\}) = 0 < 0.25 = P(\{3\}) = P_*(\{2,4\})$.

Exercise

Consider $\Omega = \{1, 2, 3\}$ with $P(\{1\}) = 0.2, P(\{2\}) = 0.3, P(\{3\}) = 0.5$ and the random set Γ given by $\Gamma(1) = \{1, 2\}, \Gamma(2) = \{1, 3\}, \Gamma(3) = \{2, 3\}$.

- Determine the lower and upper distribution functions of Γ.
- Use them to find a probability measure Q such that $F_*(x) \leq F_Q(x) \leq F^*(x)\ \forall x \in \{1, 2, 3\}$ while $Q \notin M(P_*)$.

Random sets and game theory

The lower probability P_* of Γ can be seen as a coalitional game: we regard the elements of X as players and $P_*(A)$ is seen as the gain associated with a coalition among the elements of A.

Then the Shapley value of this fame is the center of gravity of the set $M(P_*)$. It can be determined using the extreme points of $M(P_*)$, which are given by (Dempster; Shapley; Chateauneuf and Jaffray)

$$\{P_\sigma : \sigma \text{ permutation of } X\},$$

where the probability P_σ is determined by

$$P_\sigma(x_{\sigma_1}, \ldots, x_{\sigma_k}) = P_*(x_{\sigma_1}, \ldots, x_{\sigma_k}) \; \forall k = 1, \ldots, |X|.$$

Shapley value of a belief function

The Shapley value is the probability measure Q given by

$$Q(\{x\}) = \sum_{x \in A} \frac{m(A)}{|A|},$$

where m is the Möbius inverse of P_*.

The concept of Shapley value holds (with a different definition) for other types of non-additive measures (not necessarily belief functions or 2-monotone capacities), but games associated with a convex capacity have nicer mathematical properties.

Exercise

Consider the belief function associated with the basic probability assignment:

A	{1}	{2}	{1,2}	{1,3}	{2,3}	{1,2,3}
m(A)	1/9	1/9	1/9	1/6	1/6	1/3

Determine the extreme points of $M(Bel)$ and its Shapley value.

Random sets on infinite spaces

As we shall see, things get more complicated when the random set takes values on an infinite space: many of the results do not translate directly, and we need to impose additional conditions on the images of the random set.

In particular, not all random sets are compatible with the epistemic interpretation, and not all belief functions can be obtained as the lower probability of a random set.

Finitely vs. countably additive probabilities

The source of the problem is the structure of the initial space: if we consider a finitely additive probability P on a field $\mathcal{A}$, then:

- We can assume that $\mathcal{A} = \mathcal{P}(\Omega)$.
- Any belief function is the lower envelope of the set

$$\{P \text{ finitely additive} : P(A) \geq Bel(A)\ \forall A\}.$$

- Any belief function can be obtained as the lower probability of a random set.
- We can characterise the envelopes of sets of finitely additive probability measures.
- P does not satisfy any continuity property.

Finitely vs. countably additive probabilities (II)

However, if we consider a countably additive probability P on a σ-field $\mathcal{A}$, then:

- We cannot assume that $\mathcal{A} = \mathcal{P}(\Omega)$.
- Not every belief function is the lower envelope of the set

$$\{P \text{ countably additive} : P(A) \geq Bel(A)\ \forall A\}.$$

- Not every belief function can be obtained as the lower probability of a random set.
- We cannot characterise the envelopes of sets of countably additive probability measures.
- ... but P satisfies some continuity properties.

Unfortunately, usually the initial probability space $(\Omega, \mathcal{A}, P)$ considers a countably additive P on a σ-field $\mathcal{A}$.

Definition

Let $(\Omega, \mathcal{A}, P)$ be a probability space, $(X, \mathcal{A}')$ a measurable space and $\Gamma : \Omega \to \mathcal{P}(X)$ a non-empty multi-valued mapping. It is called a random set when it is strongly measurable: for every $A \in \mathcal{A}'$,

$$\{\omega : \Gamma(\omega) \cap A \neq \emptyset\} \in \mathcal{A} \ \forall A \in \mathcal{A}',$$

or, equivalently, when

$$\{\omega : \Gamma(\omega) \subseteq A\} \in \mathcal{A} \ \forall A \in \mathcal{A}'.$$

This is not the only measurability condition. Other, non-equivalent, possibilities (Himmelberg) are:

- measurability: $\{\omega : \Gamma(\omega) \cap C \neq \emptyset\} \in \mathcal{A} \ \forall C$ closed.
- weak measurability: $\{\omega : \Gamma(\omega) \cap G \neq \emptyset\} \in \mathcal{A} \ \forall G$ open.

Lower and upper probabilities

As before, the lower and upper probabilities of a random set Γ are defined as:

$$P^*(A) := P(\{\omega : \Gamma(\omega) \cap A \neq \emptyset\})$$

and

$$P_*(A) := P(\{\omega : \Gamma(\omega) \subseteq A\})$$

for every $A \in \mathcal{A}'$, and they are conjugate functions.

Properties of P_*

The lower probability P_* of a random set satisfies the following properties:

- It is upper continuous: given a decreasing sequence $(A_n)_n$, $P_*(\cap_n A_n) = \lim_n P_*(A_n)$.
- It is ∞-monotone (=a belief function when X is finite).

Similarly, P^* is ∞-alternating and lower continuous.

This means in particular that not all ∞-monotone capacities can be obtained as lower probabilities of random sets. Under some conditions, it is possible to characterise those who can.

The hit or miss topology

Denote by $\mathcal{F}, G, K$ the classes of closed, open, and compact subsets of $\mathbb{R}^d$, and let

$$\mathcal{F}^K_{G_1,\ldots,G_n} := \{F \in \mathcal{F} : F \cap K = \emptyset \neq F \cap G_1, \ldots, F \cap G_n\}.$$

This is the basis for a topology on $\mathcal{F}$, called the hit or miss topology. The σ-field it induces is denoted $\mathcal{B}(\mathcal{F})$. Then a random set $\Gamma : \Omega \to \mathbb{R}^d$ with closed values is a $\mathcal{A} - \mathcal{B}(F)$ measurable mapping.

Capacity functionals

A set function $T : \mathcal{K} \to \mathbb{R}$ is called a capacity functional when it satisfies:

- $T(\emptyset) = 0, T(K) \in [0, 1] \ \forall K$.
- T is ∞-alternating.
- If $(K_n)_n \downarrow K$, then $\lim_n T(K_n) = T(K)$.

Choquet's theorem

(Mathéron) If T is a capacity functional, then there is a unique probability measure P on $\mathcal{B}(F)$ such that

$$P(\{F : F \cap K \neq \emptyset\}) = T(K).$$

This determines the distribution of the random closed set (the probability measure on $\mathcal{B}(F)$), because we can use the above equation to determine the probability on the hit or miss topology.

The proof relies heavily on topological properties of $\mathbb{R}^d$, although it can be extended to ∞-dimensional Polish spaces.

Possibility and necessity measures (Dubois and Prade)

Given an infinite space X, a possibility measure is a function $\Pi : \mathcal{P}(X) \to [0,1]$ s.t. for every family of subsets $(A_i)_{i \in I}$ of X,

$$\Pi(\cup_{i \in I} A_i) = \sup_{i \in I} \Pi(A_i).$$

The conjugate function of a possibility measure, given by $Nec(A) = 1 - \Pi(A^c)$, is called a necessity measure, and satisfies

$$Nec(\cap_{i \in I} A_i) = \inf_{i \in I} Nec(A_i)$$

for every family of subsets $(A_i)_{i \in I}$.
Possibility measures are related to fuzzy sets, via their restrictions to singletons.

Connection with maxitive measures

Recall that an upper probability $\overline{P}$ is maxitive when $\overline{P}(A \cup B) = \max\{\overline{P}(A), \overline{P}(B)\}\ \forall A, B \subseteq X$. This is equivalent to being a possibility measure when X is finite.

A related model are the condensable measures (Shafer), which are those satisfying

$$\overline{P}(\cup_{A \in \mathcal{C}} A) = \sup_{A \in \mathcal{C}} \overline{P}(A)$$

for every upward net $\mathcal{C}$, which is class of events such that

$$\forall A_1, A_2 \in \mathcal{C},\ \exists A_3 \in \mathcal{C} \text{ s.t. } A_1 \cup A_2 \subseteq A_3.$$

(Miranda et al., 2004): $\overline{P}$ possibility $\Leftrightarrow$ maxitive and condensable.

Consonant random sets

(Miranda et al., 2004) When Γ is closed valued on a σ-compact metric space, the following are equivalent:

- P^* is a possibility measure.
- P^* is maxitive.
- $\exists N \subseteq \Omega$ null such that $\forall \omega_1, \omega_2 \in \Omega \setminus N$, either $\Gamma(\omega_1) \subseteq \Gamma(\omega_2)$ or $\Gamma(\omega_2) \subseteq \Gamma(\Omega_1)$.

The equivalence does not hold for other types of random sets.

Also, any possibility measure Π can be obtained as the upper probability of a random set.

Measurable selections

As before, if we give Γ an epistemic interpretation as a model for the imprecise knowledge of a random variable U_0, then our information about U_0 is given by

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

from which we determine the probabilistic information:

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

and

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

Again $P(\Gamma) \subseteq M(P_*) := \{Q \text{ prob.} : Q(A) \geq P_*(A)\ \forall A \in \mathcal{A}'\}$, and also $P(\Gamma)(A) \subseteq [P_*(A), P^*(A)]$.

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

In general we have the following inclusion:

$$P(\Gamma)(A) \subseteq [P_*(A), P^*(A)]$$

The study of the equality can be decomposed in two different subproblems:

- Is $P(\Gamma)(A)$ convex?
- $P^*(A) = \max P(\Gamma)(A), P_*(A) = \min P(\Gamma)(A)$?

Properties of $P(\Gamma)(A)$ (Miranda et al., 2010)

- $P(\Gamma)(A)$ closed, whence there are $U_1, U_2 \in S(\Gamma)$ such that $\max P(\Gamma)(A) = P_{U_1}(A), \min P(\Gamma)(A) = P_{U_2}(A), U_1, U_2 \in S(\Gamma)$.
- $P(\Gamma)(A)$ convex $\Leftrightarrow U_1^{-1}(A) \setminus U_2^{-1}(A)$ not an atom.
- If $P^*(A) = P_{U_1}(A)$ and $P_*(A) = P_{U_2}(A)$, then $P(\Gamma)(A) = [P_*(A), P^*(A)] \Leftrightarrow A^* \setminus A_*$ not an atom.

Relationship with the existence of measurable selections

$$P^*(A) = \max P(\Gamma)(A) \Leftrightarrow \exists H \text{ null s.t } S(\Gamma \cap A I_{A^* \setminus H} \oplus \Gamma I_{(A^* \setminus H)^c}) \neq \emptyset$$

- In general, $P^*(A) \neq \max P(\Gamma)(A)$; it may even be $P(\Gamma)(A) = \{0.5\}$ and $[P_*(A), P^*(A)] = [0, 1]$. This is because a random set may not have measurable selections!!

We must study then under which conditions $S(\Gamma) \neq \emptyset$. Early works were summarised by Wagner.

Sufficient conditions for $P(\Gamma)(A) = [P_*(A), P^*(A)]$

By making a study of the existence of measurable selections for different types of random sets, it is possible to show (Miranda et al., 2010) that, under any of the following conditions:

- Ω complete, X Souslin, $Gr(\Gamma) \in \mathcal{A} \otimes \beta_X$.
- Γ closed, (X, d) σ-compact.
- Γ open, (X, d) separable.
- Γ closed, X Polish.
- Γ compact, (X, d) separable.

$$P(\Gamma)(A) = [P_*(A), P^*(A)] \Leftrightarrow A^* \setminus A_* \text{ not an atom}$$

Expectations of random sets

Let $\Gamma : \Omega \to \mathcal{P}(\mathbb{R}^n)$ be a random set. Its Aumann integral is given by

$$(A)\int \Gamma dP := \left\{ \int f dP : f \in L^1(P), f(\omega) \in \Gamma(\omega) \ a.s \right\}.$$

This is the definition of expectation of a random set which is more interesting under the epistemic interpretation. Other definitions are:

- The Debreu integral.
- The Herer integral.

The Choquet integral

Let $(X, \mathcal{A}')$ be a measurable space. Given a measurable function $f : X \to \mathbb{R}$ and a non-additive measure $\mu : \mathcal{A}' \to [0, 1]$, the Choquet integral of f with respect to μ is given by

$$(C) \int f d\mu = \inf f + \int_{\inf f}^{\sup f} \mu(f > t) dt.$$

This definition extends the one of the finite case. Its mathematical properties (monotonicity, homogeneity...) are similar.

Connection between the integrals (Miranda et al., 2010)

Let $\Gamma : \Omega \to \mathcal{P}(X)$ be a random set. If $P^*(A) = \max P(\Gamma)(A)$ for all $A \in \mathcal{A}'$, then for any bounded random variable $f : X \to \mathbb{R}$,

$$(C) \int f dP^* = \sup_{U \in S(\Gamma)} \int f dP_U, \ (C) \int f dP_* = \inf_{U \in S(\Gamma)} \int f dP_U.$$

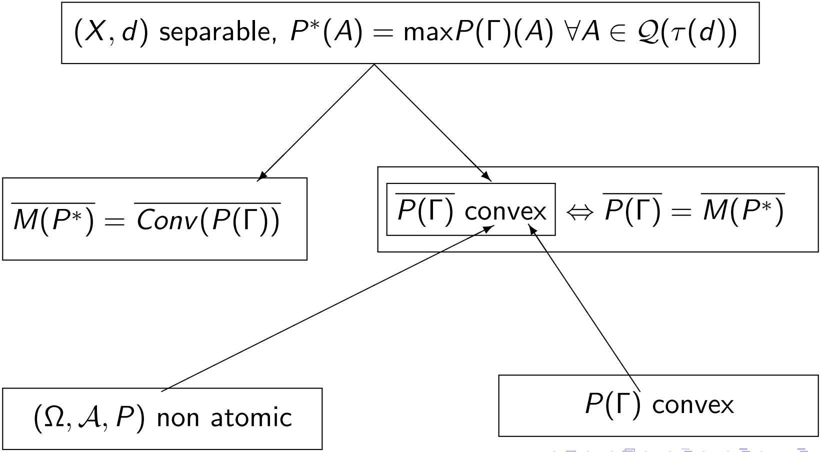
As a consequence,

$$(C) \int f dP^* = \sup(A) \int (f \circ \Gamma) dP, \ (C) \int f dP_* = \inf(A) \int (f \circ \Gamma) dP.$$

Relationships between $P(\Gamma)$ and $M(P^*)$: early results

- (Hart and Köhlberg) $(\Omega, \mathcal{A}, P)$ non atomic complete, $\Gamma_1, \Gamma_2 : \Omega \to \mathcal{P}(\mathbb{R}^n)$ integrably bounded, $P^*_{\Gamma_1} = P^*_{\Gamma_2} \Rightarrow \overline{P(\Gamma_1)} = \overline{P(\Gamma_2)}$.
- (Hess) Γ_1, Γ_2 closed, $(X, || \cdot ||)$ Banach separable, $[P^*_{\Gamma_1} = P^*_{\Gamma_2} \Leftrightarrow \overline{Conv(P(\Gamma_1))} = \overline{Conv(\Gamma_2)}]$.
- (Castaldo y Marinacci) Γ compact, X Polish $\Rightarrow M(P^*) = \overline{Conv(P(\Gamma))}$.

Relationships between $P(\Gamma)$ and $M(P^*)$ (Miranda et al., 2005a)



Interpretation

In particular, this means that, although not in all cases, we can use P^* and P_* to summarise the probabilistic information about the original random variable U_0 for the most important types of random sets, such as for instance:

- Random closed sets on $\mathbb{R}^d$.
- Random open sets on $\mathbb{R}^d$.
- Random sets on finite spaces.

Particular case: random intervals

We next study the case where the images of the random set are subintervals of the real line, determined by two mappings A, B. They have been studied by Demspter and Joslyn, among others.

We focus on random closed intervals $\Gamma = [A, B]$, although some results have been established for random open intervals: $\Gamma = (A, B)$.

Results for random closed intervals

(Miranda et al. (2005b):

- $[A, B]$ strongly measurable $\Leftrightarrow$ A, B measurable.
- $(\Omega, \mathcal{A}, P) = ([0,1], \beta_{[0,1]}, \lambda_{[0,1]}) \Rightarrow P(\Gamma) = M(P^*)$ under any of the following conditions:
  - A, B increasing.
  - A or B constant.
  - A, B strictly comonotonic.

Conclusions

When X is finite:

- Belief functions are equivalent to lower probabilities of random sets.
- They keep all the information about the selections under an epistemic interpretations.
- Possibility measures correspond to the particular case of consonant random sets.
- P-boxes may carry a loss of information.

When X is infinite, the above results do not hold in general, but they do for most random sets of interest (random closed intervals, for instance).

References

- R. Aumann, J. of Math. Anal. and Appl., 12, 1-12, 1965.
- A. Castaldo and M. Marinacci, Sankhya, 66(3), 409-427, 2004.
- G. Choquet, Ann. de l'Ins. Fourier, 5, 131-295, 1953.
- I. Couso, D. Dubois, Int. J. of App. Reasoning, 55(7), 1502-1514, 2014.
- I. Couso, D. Dubois, L. Sánchez, *Random sets and random fuzzy sets as ill-perceived random variables*. Springer, 2014.
- I. Couso, L. Sánchez, P. Gil, Inf. Sciences, 159(1-2), 109-123, 2004.
- A. Dempster, Ann. of Math. Statistics, 38, 325-339, 1967.
- A. Dempster, Ann. of Math. Statistics, 39, 957–966, 1968.

References (II)

- D. Denneberg, *Non-additive measure and integral*. Kluwer, 1994.
- S. Hart and E. Köhlberg, J. of Math. Economics, 1(2), 167-174, 1974.
- C. Hess, J. of Convex Analysis, 6(1), 163-182, 1999.
- C. Himmelberg, Fund. Mathematicae, 87, 53-72, 1975.
- R. Kruse, H. Meyer, *Statistics with vague data*. D. Reidel, 1987.
- G. Mathéron, *Random sets and integral geometry*. Wiley, 1975.
- E. Miranda, I. Couso, P. Gil, Inf. Sciences, 168(1-4), 51–75, 2004.
- E. Miranda, I. Couso, P. Gil, J. of Math. Anal. and Applications, 307(1), 32-47, 2005a.

References (III)

- E. Miranda, I. Couso, P. Gil, Fuz. Sets and Systems, 154(3), 386–412, 2005b.
- E. Miranda, I. Couso, P. Gil. Inf. Sci., 180(8), 1407–1417, 2010.
- I. Molchanov, *Theory of random sets*. Springer, 2005.
- H.T. Nguyen, *An introduction to random sets*. Chapman and Hall, 2006.
- H.T. Nguyen, J. of Math. Anal. and Appl., 65(3), 531-542, 1978.
- G. Shafer, *A mathematical theory of evidence*, Princeton, 1976.
- M. Troffaes, S. Destercke, Int. J. of App. Reasoning, 52(6), 767–791, 2011.