Decision-making with belief functions

Thierry Denœux

Université de Technologie de Compiègne
HEUDIASYC (UMR CNRS 7253)
https://www.hds.utc.fr/~tdenoeux

Third School on belief functions and their applications
Stella Plage, France, September 30, 2015

Example of decision problem under uncertainty

Act (Purchase)	Good Economic Conditions	Poor Economic Conditions
Apartment building	50,000	30,000
Office building	100,000	-40,000
Warehouse	30,000	10,000

Formal framework
Acts, outcomes, states of nature

- A decision problem can be seen as a situation in which a decision-maker (DM) has to choose a course of action (an act) in some set $\mathcal{F}$
- An act may have different consequences (outcomes), depending on the state of nature
- Denoting by $\Omega = \{\omega_1, \ldots, \omega_n\}$ the set of states of nature and by $\mathcal{C} = \{c_1, \ldots, c_r\}$ the set of consequences (or outcomes), an act can be formalized as a mapping f from Ω to $\mathcal{C}$
- In this lecture, the three sets Ω, $\mathcal{C}$ and $\mathcal{F}$ will be assumed to be finite

Formal framework
Utilities

- The desirability of the consequences can often be modeled by a utility function $u : \mathcal{C} \to \mathbb{R}$, which assigns a numerical value to each consequence
- The higher this value, the more desirable is the consequence for the DM
- In some problems, the consequences can be evaluated in terms of monetary value. The utilities can then be defined as the payoffs, or a function thereof
- If the actions are indexed by i and the states of nature by j, we will denote by u_{ij} the quantity $u[f_i(\omega_j)]$
- The $n \times r$ matrix $U = (u_{ij})$ will be called a payoff or utility matrix

Payoff matrix

Act (Purchase)	Good Economic Conditions (ω_1)	Poor Economic Conditions (ω_2)
Apartment building (f_1)	50,000	30,000
Office building (f_2)	100,000	-40,000
Warehouse (f_3)	30,000	10,000

Formal framework
Preferences

- If the true state of nature ω is known, the desirability of an act f can be deduced from that of its consequence $f(\omega)$
- Typically, the state of nature is unknown. Based on partial information, is usually assumed that the DM can express preferences among acts, which may be represented mathematically by a preference relation $\succcurlyeq$ on $\mathcal{F}$
- This relation is interpreted as follows: given two acts f and g, $f \succcurlyeq g$ means that f is found by the DM to be at least as desirable as g
- We also define
  - The strict preference relation as $f \succ g$ iff $f \succcurlyeq g$ and not$(g \succcurlyeq f)$ (meaning that f is strictly more desirable than g) and
  - The indifference relation $f \sim g$ iff $f \succcurlyeq g$ and $g \succcurlyeq f$ (meaning that f and g are equally desirable)

Decision problems

- Quite often, the decision problem is to construct a preference relation among acts, from a utility matrix and some description of uncertainty, and to find the maximal elements of this relation
- Depending on the nature of the available information, different decision problems arise:
  1. Decision-making under ignorance
  2. Decision-making with probabilities
  3. Decision-making with belief functions

Outline

1. Classical decision theory
   - Decision-making under complete ignorance
   - Decision-making with probabilities
   - Savage's theorem
2. Decision-making with belief functions
   - Upper and lower expected utility
   - Other approaches
   - Axiomatic justifications

Outline

1. Classical decision theory
   - Decision-making under complete ignorance
   - Decision-making with probabilities
   - Savage's theorem
2. Decision-making with belief functions
   - Upper and lower expected utility
   - Other approaches
   - Axiomatic justifications

Outline

Problem and non-domination principle

- We assume that the DM is totally ignorant of the state of nature: all the information given to the DM is the utility matrix U
- A act f_i is said to be dominated by f_k if the outcomes of f_k are at least as desirable as those of f_i for all states, and strictly more desirable for at least one state

$$\forall j, \; u_{kj} \geq u_{ij} \text{ and } \exists j, \; u_{kj} > u_{ij}$$

- Non-domination principle: an act cannot be chosen if it is dominated by another one

Example of a dominated act

Act (Purchase)	Good Economic Conditions (ω_1)	Poor Economic Conditions (ω_2)
Apartment building (f_1)	50,000	30,000
Office building (f_2)	100,000	-40,000
~~Warehouse (f_3)~~	~~30,000~~	~~10,000~~

Criteria for rational choice

- After all dominated acts have been removed, there remains the problem of ordering them by desirability, and of finding the set of most desirable acts
- Several criteria of "rational choice" that have been proposed to derive a preference relation over acts

1. Laplace criterion

$$f_i \succeq f_k \text{ iff } \frac{1}{n}\sum_j u_{ij} \geq \frac{1}{n}\sum_j u_{kj}.$$

2. Maximax criterion

$$f_i \succeq f_k \text{ iff } \max_j u_{ij} \geq \max_j u_{kj}.$$

3. Maximin (Wald) criterion

$$f_i \succeq f_k \text{ iff } \min_j u_{ij} \geq \min_j u_{kj}.$$

Example

Act	ω_1	ω_2	ave	max	min
Apartment (f_1)	50,000	30,000	**40,000**	50,000	**30,000**
Office (f_2)	100,000	-40,000	30,000	**100,00**	-40,000

Hurwicz criteria

- Hurwicz criterion: $f_i \succeq f_k$ iff

$$\alpha \min_j u_{ij} + (1-\alpha) \max_j u_{ij} \geq \alpha \min_j u_{kj} + (1-\alpha) \max_j u_{kj}$$

  where α is a parameter in $[0,1]$, called the pessimism index
- Boils down to
  - the maximax criterion if $\alpha = 0$
  - the maximin criterion if $\alpha = 1$
- α describes the DM's attitude toward ambiguity

Minimax regret criterion criterion

- (Savage) Minimax regret criterion: an act f_i is at least as desirable as f_k if it has smaller maximal regret, where regret is defined as the utility difference with the best act, for a given state of nature
- The regret r_{ij} for act f_i and state ω_j is

$$r_{ij} = \max_{\ell} u_{\ell j} - u_{ij}$$

- The maximum regret for act f_i is $R_i = \max_j r_{ij}$
- $f_i \succeq f_k$ iff $R_i \leq R_k$

Example

- Pay-off matrix

Act	ω_1	ω_2
Apartment (f_1)	50,000	30,000
Office (f_2)	100,000	-40,000

- Regret matrix

Act	ω_1	ω_2	max regret
Apartment (f_1)	50,000	0	**50,000**
Office (f_2)	0	70,000	70,000

Generalization: OWA criteria

- The Laplace, maximax, maximin and Hurwicz criteria correspond to different ways of aggregating the utilities resulting each act, using, respectively, the average, the maximum, the minimum, and a convex sum of the minimum and the maximum
- These four operators belong to a family of operators called Ordered Weighted Average (OWA) operators (Yager, 1988)

OWA operators

- An OWA operator of dimension n is a function $F : \mathbb{R}^n \rightarrow \mathbb{R}$ of the form

$$F(x_1, \ldots, x_n) = \sum_{i=1}^{n} w_i x_{(i)}$$

  where $x_{(i)}$ is the i-th largest element in the collection $x_1, \ldots, x_n$, and $w_1, \ldots, w_n$ are positive weights verifying $\sum_{i=1}^{n} w_i = 1$
- The four previous operators are obtained for different choices of the weights:
  - Average: $(1/n, 1/n, \ldots, 1/n)$
  - Maximum: $(1, 0, \ldots, 0)$
  - Minimum: $(0, \ldots, 0, 1)$
  - Hurwicz: $(1 - \alpha, 0, \ldots, 0, \alpha)$

Setting the weights of an OWA operator

- In a decision-making context, each weight w_i may be interpreted as a probability that the i-th best outcome will happen
- Yager (1988) defines the degree of optimism of an OWA operator with weight vector $\mathbf{w}$ as

$$OPT(\mathbf{w}) = \sum_{i=1}^{n} \frac{n-i}{n-1} w_i$$

- $OPT(\mathbf{w}) = 1$ for the maximum, $OPT(\mathbf{w}) = 0$ for the minimum, $OPT(\mathbf{w}) = 0.5$ for the mean, $OPT(\mathbf{w}) = 1 - \alpha$ for Hurwicz
- Given a degree of optimism β, we can then choose the OWA operator that maximizes the entropy

$$ENT(\mathbf{w}) = -\sum_{i=1}^{n} w_i \log w_i$$

under the constraint $OPT(\mathbf{w}) = \beta$

Axioms of rational choice

- Let $\mathcal{F}^*$ denote the choice set, defined as a set of optimal acts
- Arrow and Hurwicz (1972) have proposed four axioms a choice operator $\mathcal{F} \to \mathcal{F}^*$ should verify
  1. Axiom A_1: if $\mathcal{F}_1 \subset \mathcal{F}_2$ and $\mathcal{F}_2^* \cap \mathcal{F}_1 \neq \emptyset$, then $\mathcal{F}_1^* = \mathcal{F}_2^* \cap \mathcal{F}_1$
  2. Axiom A_2: Relabeling actions and states does not change the optimal status of actions
  3. Axiom A_3: Deletion of a duplicate state does not change the optimality status of actions
  4. Axiom A_4 (dominance): If $f \in \mathcal{F}^*$ and f' dominates f, then $f' \in \mathcal{F}^*$. If $f \notin \mathcal{F}^*$ and f dominates f', then $f' \notin \mathcal{F}^*$
- Under some regularity assumptions, Axioms $A_1 - A_4$ imply that the choice set depends only on the worst and the best consequences of each act
- In particular, these axioms rule out the Laplace and minimax regret criteria

Violation of Axiom A3 by the Laplace criterion

Act	ω_1	ω_2	ave
Apartment (f_1)	50,000	30,000	**40,000**
Office (f_2)	100,000	-40,000	30,000

Let us split the state of nature ω_1 in two states: “Good economic conditions and there is life on Mars” (ω_1') and “Good economic conditions and there is no life on Mars” (ω_1'')

Act	ω_1'	ω_1''	ω_2	ave
Apartment (f_1)	50,000	50,000	30,000	43,333
Office (f_2)	100,000	100,000	-40,000	**53,333**

Violation of Axiom A1 by minimax regret

- Pay-off matrix

Act	ω_1	ω_2
Apartment (f_1)	50,000	30,000
Office (f_2)	100,000	-40,000
f_4	130,000	-45,000

- Regret matrix

Act	ω_1	ω_2	max regret
Apartment (f_1)	80,000	0	80,000
Office (f_2)	30,000	70,000	**70,000**
f_4	0	75,000	75,000

We had $\mathcal{F}_1 = \{f_1, f_2\}$ and $\mathcal{F}_1^* = \{f_1\}$. Now, $\mathcal{F}_2 = \{f_1, f_2, f_4\}$ and $\mathcal{F}_2^* = \{f_2\}$. So, $\mathcal{F}_1^* \neq \mathcal{F}_2^* \cap \mathcal{F}_1$

Outline

Maximum Expected Utility principle

- Let us now consider the situation where uncertainty about the state of nature is quantified by probabilities $p_1, \ldots, p_n$ on Ω
- These probabilities can be objective (decision under risk) or subjective
- We can then compute, for each act f_i, its expected utility as

$$EU(f_i) = \sum_j u_{ij} p_j$$

- Maximum Expected Utility (MEU) principle: an act f_i is more desirable than an act f_k if i it has a higher expected utility: $f_i \succeq f_k$ iff $EU(f_i) \geq EU(f_k)$

Example

Act	ω_1	ω_2
Apartment (f_1)	50,000	30,000
Office (f_2)	100,000	-40,000

Assume that there is 60% chance that the economic situation will be poor (ω_2). The expected utilities of acts f_1 and f_2 are

$$EU(f_1) = 50,000 \times 0.4 + 30,000 \times 0.6 = 38,000$$
$$EU(f_2) = 100,000 \times 0.4 - 40,000 \times 0.6 = 16,000$$

Act f_1 is thus more desirable according to the maximum expected utility criterion

Axiomatic justification of the MEU principle

- The MEU principle was first axiomatized by von Neumann and Morgenstern (1944)
- Given a probability distribution on Ω, an act $f : \Omega \rightarrow \mathcal{C}$ induces a probability measure P on the set $\mathcal{C}$ of consequences (assumed to be finite), called a lottery
- We denote by $\mathcal{L}$ the set of lotteries on $\mathcal{C}$
- If we agree that two acts providing the same lottery are equivalent, then the problem of comparing the desirability of acts becomes that of comparing the desirability of lotteries
- Let $\succeq$ be a preference relation among lotteries. Von Neumann and Morgentern argued that, to be rational, a preference relation should verify three axioms

Von Neumann and Morgenstern's axioms

1. Complete preorder: the preference relation is a complete and non trivial preorder (i.e., it is a reflexive, transitive and complete relation) on $\mathcal{L}$
2. Continuity: for any lotteries P, Q and R such that $P \succ Q \succ R$, there exists a probabilities α and β in $[0, 1]$ such that

$$\alpha P + (1 - \alpha)R \succ Q \succ \beta P + (1 - \beta)R$$

   where $\alpha P + (1 - \alpha)R$ is a compound lottery, which refers to the situation where you receive P with probability α and Q with probability $1 - \alpha$. This axiom implies, in particular, that there is no lottery R that is so undesirable that it cannot become desirable if mixed with some very desirable lottery P
3. Independence: for any lotteries P, Q and R and for any $\alpha \in (0, 1]$

$$P \succeq Q \Leftrightarrow \alpha P + (1 - \alpha)R \succeq \alpha Q + (1 - \alpha)R$$

Von Neumann and Morgenstern's theorem

The two following propositions are equivalent:

1. The preference relation $\succeq$ verifies the axioms of complete preorder, continuity, and independence
2. There exists a utility function $u : \mathcal{C} \to \mathbb{R}$ such that, for any two lotteries $P = (p_1, \ldots, p_r)$ and $Q = (q_1, \ldots, q_r)$

$$P \succeq Q \Leftrightarrow \sum_{i=1}^{r} p_i u(c_i) \geq \sum_{i=1}^{r} q_i u(c_i)$$

Function u is unique up to a strictly increasing affine transformation

Outline

1. Classical decision theory
   - Decision-making under complete ignorance
   - Decision-making with probabilities
   - Savage's theorem
2. Decision-making with belief functions
   - Upper and lower expected utility
   - Other approaches
   - Axiomatic justifications

Savage's theorem

- We have reviewed some criteria for decision-making under complete ignorance, i.e., when uncertainty cannot be probabilized
- Some researchers have defended the view that a rational DM always maximizes expected utility, for some subjective probability measure and utility function
- Savage's theorem (1954): a preference relation $\succcurlyeq$ among acts verifies some rationality requirements iff there is a finitely additive probability measure P and a utility function $u : \mathcal{C} \rightarrow \mathbb{R}$ such that

$$\forall f, g \in \mathcal{F}, \quad f \succcurlyeq g \Leftrightarrow \int_\Omega u(f(\omega))dP(\omega) \geq \int_\Omega u(g(\omega))dP(\omega)$$

  Furthermore, P is unique and u is unique up to a positive affine transformation
- A strong argument for probabilism, but Savage's axioms can be questioned!

Savage's axioms

- Savage has proposed seven axioms, four of which are considered as meaningful (the other three are technical)
- Axiom 1: $\succcurlyeq$ is a total preorder (complete, reflexive and transitive)
- Axiom 2 [Sure Thing Principle]. Given $f, h \in \mathcal{F}$ and $E \subseteq \Omega$, let fEh denote the act defined by

$$(fEh)(\omega) = \begin{cases} f(\omega) & \text{if } \omega \in E \\ h(\omega) & \text{if } \omega \notin E \end{cases}$$

  Then the Sure Thing Principle states that $\forall E, \forall f, g, h, h'$

$$fEh \succcurlyeq gEh \Rightarrow fEh' \succcurlyeq gEh'$$

- This axiom seems reasonable, but it is not verified empirically!

Ellsberg's paradox

- Suppose you have an urn containing 30 red balls and 60 balls, either black or yellow. Consider the following gambles:
  - f_1: You receive 100 euros if you draw a red ball
  - f_2: You receive 100 euros if you draw a black ball
  - f_3: You receive 100 euros if you draw a red or yellow ball
  - f_4: You receive 100 euros if you draw a black or yellow ball
- Most people strictly prefer f_1 to f_2, but they strictly prefer f_4 to f_3

	R	B	Y
f_1	100	0	0
f_2	0	100	0
f_3	100	0	100
f_4	0	100	100

Now,

$$f_1 = f_1\{R,B\}0, \quad f_2 = f_2\{R,B\}0$$

$$f_3 = f_1\{R,B\}100, \quad f_4 = f_2\{R,B\}100$$

- The Sure Thing Principle is violated!

Summary

- Classically, we distinguish two kinds of decision problems:
  1. Decision under ignorance: we only know, for each act, a set a possible outcomes
  2. Decision under risk: we are given, for each act, a probability distribution over the outcomes
- It has been argued that any decision problem under uncertainty should be handled as a problem of decision under risk. However, the axiomatic arguments are questionable
- In the next part: decision-making when uncertainty is described by a belief functions

Outline

1. Classical decision theory
   - Decision-making under complete ignorance
   - Decision-making with probabilities
   - Savage's theorem

2. Decision-making with belief functions
   - Upper and lower expected utility
   - Other approaches
   - Axiomatic justifications

How belief functions come into the picture

Belief functions become of component of a decision problem in any of the following two situations (or both)

1. The decision maker's subjective beliefs concerning the state of nature may be described by a belief function Bel^Ω on Ω
2. The DM may not be able to precisely describe the outcomes of some acts under each state of nature

Case 1: uncertainty described by a belief function

- Let m^Ω be a mass function on Ω
- Any act $f : \Omega \rightarrow \mathcal{C}$ then carries m^Ω to the set $\mathcal{C}$ of consequences, yielding a mass function $m_f^\mathcal{C}$, which quantifies the DM's beliefs about the outcome of act f
- Each mass $m^\Omega(A)$ is transfered to $f(A)$

$$m_f^\mathcal{C}(B) = \sum_{\{A \subseteq \Omega | f(A) = B\}} m^\Omega(A)$$

  for any $B \subseteq \mathcal{C}$
- $m_f^\mathcal{C}$ is a credibilistic lottery corresponding to act f

Case 2: partial knowledge of outcomes

- In that case, an act may formally be represented by a multi-valued mapping $f : \Omega \rightarrow 2^{\mathcal{C}}$, assigning a set of possible consequences $f(\omega) \subseteq \mathcal{C}$ to each state of nature ω
- Given a probability measure P on Ω, f then induces the following mass function $m_f^{\mathcal{C}}$ on $\mathcal{C}$,

$$m_f^{\mathcal{C}}(B) = \sum_{\{\omega \in \Omega | f(\omega) = B\}} p(\omega)$$

for all $B \subseteq \Omega$

Example

- Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and m^Ω the following mass function

$$
\begin{array}{ll}
m^\Omega(\{\omega_1,\omega_2\}) = 0.3, & m^\Omega(\{\omega_2,\omega_3\}) = 0.2 \\
m^\Omega(\{\omega_3\}) = 0.4, & m^\Omega(\Omega) = 0.1
\end{array}
$$

- Let $\mathcal{C} = \{c_1, c_2, c_3\}$ and f the act

$$
f(\omega_1) = \{c_1\}, \quad f(\omega_2) = \{c_1, c_2\}, \quad f(\omega_3) = \{c_2, c_3\}
$$

- To compute $m_f^{\mathcal{C}}$, we transfer the masses as follows

$$
\begin{aligned}
m^\Omega(\{\omega_1,\omega_2\}) = 0.3 &\rightarrow f(\omega_1) \cup f(\omega_2) = \{c_1, c_2\} \\
m^\Omega(\{\omega_2,\omega_3\}) = 0.2 &\rightarrow f(\omega_2) \cup f(\omega_3) = \{c_1, c_2, c_3\} \\
m^\Omega(\{\omega_3\}) = 0.4 &\rightarrow f(\omega_3) = \{c_2, c_3\} \\
m^\Omega(\Omega) = 0.1 &\rightarrow f(\omega_1) \cup f(\omega_2) \cup f(\omega_3) = \{c_1, c_2, c_3\}
\end{aligned}
$$

- Finally, we obtain the following mass function on $\mathcal{C}$

$$
m^{\mathcal{C}}(\{c_1, c_2\}) = 0.3, \quad m^{\mathcal{C}}(\{c_2, c_3\}) = 0.4, \quad m^{\mathcal{C}}(\mathcal{C}) = 0.3
$$

Decision problem

- In the two situations considered above, we can assign to each act f a credibilistic lottery, defined as a mass function on $\mathcal{C}$
- Given a utility function u on $\mathcal{C}$, we then need to extend the MEU model
- Several such extensions will now be reviewed

Outline

1. Classical decision theory
   - Decision-making under complete ignorance
   - Decision-making with probabilities
   - Savage's theorem

2. Decision-making with belief functions
   - Upper and lower expected utility
   - Other approaches
   - Axiomatic justifications

Upper and lower expectations

- Let m be a mass function on $\mathcal{C}$, and u a utility function $\mathcal{C} \to \mathbb{R}$
- The lower and upper expectations of u are defined, respectively, as the averages of the minima and the maxima of u within each focal set of m

$$\underline{\mathbb{E}}_m(u) = \sum_{A \subseteq \mathcal{C}} m(A) \min_{c \in A} u(c)$$
$$\overline{\mathbb{E}}_m(u) = \sum_{A \subseteq \mathcal{C}} m(A) \max_{c \in A} u(c)$$

- It is clear that $\underline{\mathbb{E}}_m(u) \leq \overline{\mathbb{E}}_m(u)$, with the inequality becoming an equality when m is Bayesian, in which case the lower and upper expectations collapse to the usual expectation
- If $m = m_A$ is logical with focal set A, then $\underline{\mathbb{E}}_m(u)$ and $\overline{\mathbb{E}}_m(u)$ are, respectively, the minimum and the maximum of u in A

Imprecise probability interpretation

- The lower and upper expectations are lower and upper bounds of expectations with respect to probability measures compatible with m

$$\underline{\mathbb{E}}_m(u) = \min_{P \in \mathcal{P}(m)} \mathbb{E}_P(u)$$

$$\overline{\mathbb{E}}_m(u) = \max_{P \in \mathcal{P}(m)} \mathbb{E}_P(u)$$

- The mean of minima (res., maxima) is also the minimum (resp., maximum) of means with respect to all compatible probability measures

Corresponding decision criteria

- Having defined the notions of lower and upper expectations, we can define two preference relations among credibilistic lotteries as

$$m_1 \underline{\succcurlyeq} m_2 \text{ iff } \underline{\mathbb{E}}_{m_1}(u) \geq \underline{\mathbb{E}}_{m_2}(u)$$

  and

$$m_1 \overline{\succcurlyeq} m_2 \text{ iff } \overline{\mathbb{E}}_{m_1}(u) \geq \overline{\mathbb{E}}_{m_2}(u)$$

- Relation $\underline{\succcurlyeq}$ corresponds to a pessimistic (or conservative) attitude of the DM. When m is logical, it corresponds to the maximin criterion
- Symmetrically, $\overline{\succcurlyeq}$ corresponds to an optimistic attitude and extends the maximax criterion
- Both criteria boil down to the MEU criterion when m is Bayesian

Back to Ellsberg's paradox

- Here, $\Omega = \{R, B, Y\}$ and $m^{\Omega}(\{R\}) = 1/3$, $m^{\Omega}(\{B, Y\}) = 2/3$
- The mass functions on $\mathcal{C} = \{0, 100\}$ induced by the four acts are

$$
\begin{aligned}
m_1(\{100\}) &= 1/3, \quad m_1(\{0\}) = 2/3 \\
m_2(\{0\}) &= 1/3, \quad m_2(\{0, 100\}) = 2/3 \\
m_3(\{100\}) &= 1/3, \quad m_3(\{0, 100\}) = 2/3 \\
m_4(\{0\}) &= 1/3, \quad m_4(\{100\}) = 2/3
\end{aligned}
$$

- Corresponding lower and upper expectations

	R	B	Y	$\underline{\mathbb{E}}_m(u)$	$\overline{\mathbb{E}}_m(u)$
f_1	100	0	0	$\mathbf{u(100)/3}$	$u(100)/3$
f_2	0	100	0	$u(0)$	$\mathbf{2u(100)/3}$
f_3	100	0	100	$u(100)/3$	$u(100)$
f_4	0	100	100	$\mathbf{2u(100)/3}$	$\mathbf{2u(100)/3}$

Interval dominance

- If we drop the requirement that the preference relation among acts be complete, then we can consider the interval dominance relation,

$$m_1 \succcurlyeq_{ID} m_2 \text{ iff } \underline{\mathbb{E}}_{m_1}(u) \geq \overline{\mathbb{E}}_{m_2}(u)$$

- Given a collection of credibilistic lotteries, we can then compute the set of maximal (i.e., non dominated) elements of $\succcurlyeq_{ID}$
- Imprecise probability view

$$m_1 \succcurlyeq_{ID} m_2 \Leftrightarrow \forall P_1 \in \mathcal{P}(m_1), \forall P_2 \in \mathcal{P}(m_2), \mathbb{E}_{P_1}(u) \geq \mathbb{E}_{P_2}(u)$$

- The justification for this preference relation is not so clear from the point of view of belief function theory (i.e., if one does not interpret a belief function as a lower probability)

Outline

1. Classical decision theory
   - Decision-making under complete ignorance
   - Decision-making with probabilities
   - Savage's theorem

2. Decision-making with belief functions
   - Upper and lower expected utility
   - Other approaches
   - Axiomatic justifications

Generalized Hurwicz criterion

- The Hurwicz criterion can be generalized as

$$\mathbb{E}_{m,\alpha}(u) = \sum_{A \subseteq \mathcal{C}} m(A) \left(\alpha \min_{c \in A} u(c) + (1-\alpha) \max_{c \in A} u(c) \right)$$
$$= \alpha \underline{\mathbb{E}}_m(u) + (1-\alpha) \overline{\mathbb{E}}(u)$$

  where $\alpha \in [0,1]$ is a pessimism index
- This criterion was introduced and justified axiomatically by Jaffray (1988)
- Strat (1990) who proposed to interpret α as the DM's subjective probability that the ambiguity will be resolved unfavorably

Transferable belief model

- A completely different approach to decision-making with belief function was advocated by Smets, as part of the Transferable Belief Model
- Smets defended a two-level mental model
  1. a credal level, where an agent's belief are represented by belief functions, and
  2. a pignistic level, where decisions are made by maximizing the EU with respect to a probability measure derived from a belief function
- The rationale for introducing probabilities at the decision level is the avoidance of Dutch books
- Smets argued that the belief-probability transformation T should be linear, i.e., it should verify

$$T(\alpha m_1 + (1-\alpha) m_2) = \alpha T(m_1) + (1-\alpha) T(m_2),$$

for any mass functions m_1 and m_2 and for any $\alpha \in [0, 1]$

Pignistic transformation

- The only linear belief-probability transformation T is the pignistic transformation, with $p_m = T(m)$ given by

$$p_m(c) = \sum_{\{A \subseteq \mathcal{C} | c \in A\}} \frac{m(A)}{|A|}, \quad \forall c \in \mathcal{C}$$

- The pignistic probability p_m is mathematically identical to the Shapley value in cooperative game theory
- The expected utility w.r.t. the pignistic probability is

$$\mathbb{E}_p(u) = \sum_{c \in \mathcal{C}} p_m(c) u(c) = \sum_{A \subseteq \mathcal{C}} m(A) \left(\frac{1}{|A|} \sum_{c \in A} u(c) \right)$$

- The maximum pignistic expected utility criterion thus extends the Laplace criterion

Generalized OWA criteria

- A more general family of expected utility criteria can be defined by aggregating the utilities $u(c)$ within each focal set A using OWA operators
- To determine the weights of the OWA operators, Yager (1992) proposes to fix the degree of optimism β and to use the maximum-entropy operators, for each cardinality $|A|$

$$\mathbb{E}_{m,\beta}^{\text{owa}} = \sum_{A \subseteq \mathcal{C}} m(A) F_{|A|,\beta}(\{u(c) | c \in A\})$$

  where $F_{|A|,\beta}$ is the maximum-entropy OWA operator with degree of optimism β and arity $|A|$
- Parameter β has roughly the same interpretation as one minus the pessimism index α in the Hurwicz criterion
- However, each $F_{|A|,\beta}(\{u(c)|c \in A\})$ depends on all the values $u(c)$ for all $c \in A$, and not only on the minimum and the maximum

Generalized minimax regret

- Yager (2004) also extended the minimax regret criterion to belief functions
- We need to consider n acts $f_1, \ldots, f_n$, and we write $u_{ij} = u[f_i(\omega_j)]$
- The regret if act f_i is selected, and state ω_j occurs, is $r_{ij} = \max_k u_{kj} - u_{ij}$
- For a non-empty subset A of Ω, the maximum regret of act f_i is

$$R_i(A) = \max_{\omega_j \in A} r_{ij}$$

- The expected maximal regret for act f_i is

$$\overline{R}_i = \sum_{\emptyset \neq A \subseteq \Omega} m^\Omega(A) R_i(A)$$

- Act f_i is preferred over act f_k if $\overline{R}_i \leq \overline{R}_k$
- The minimax regret criterion is recovered when m^Ω is logical
- The MEU model is recovered when m^Ω is Bayesian

Summary

non-probabilized		belief functions	probabilized
maximin	$\longleftrightarrow$	lower expectation	
maximax	$\longleftrightarrow$	upper expectation	
Laplace	$\longleftrightarrow$	pignistic expectation	expected utility
Hurwicz	$\longleftrightarrow$	generalized Hurwicz	
OWA	$\longleftrightarrow$	generalized OWA	
minimax regret	$\longleftrightarrow$	generalized minimax regret	

Outline

1. Classical decision theory
   - Decision-making under complete ignorance
   - Decision-making with probabilities
   - Savage's theorem

2. Decision-making with belief functions
   - Upper and lower expected utility
   - Other approaches
   - Axiomatic justifications

Linear utility of credibilistic lotteries

- Except for the generalized minimax regret criterion, the previous decision criteria are of the form

$$m_1 \succcurlyeq m_2 \text{ iff } U(m_1) \geq U(m_2)$$

  with

$$U(m) = \sum_{\emptyset \neq A \subseteq \mathcal{C}} m(A) U(m_A)$$

  where m_A is the logical mass function with focal set A
- Writing $U(A)$ in place of $U(m_A)$, and $u(c)$ for $U(\{c\})$
  - $U(A) = \min_{c \in A} u(c)$ for the lower expectation criterion
  - $U(A) = \max_{c \in A} u(c)$ for the upper expectation criterion
  - $U(A) = \alpha \min_{c \in A} u(c) + (1 - \alpha) \max_{c \in A} u(c)$ for the Hurwicz criterion
  - $U(A) = (1/|A|) \sum_{c \in A} u(c)$ for the pignistic criterion
  - $U(A) = F_{|A|,\beta}(\{u(c)|c \in A\})$ for the OWA criterion

Jaffray's theorem

Jaffray (1989) showed that a preference relation among credibilistic lotteries is representable by a linear utility function if and only if it verifies the Von Neumann and Morgenstern axioms extended to credibilistic lotteries, i.e.,

1. Transitivity and Completeness: $\succcurlyeq$ is a transitive and complete relation (i.e., is a weak order)
2. Continuity: for all m_1, m_2 and m_3 such that $m_1 \succ m_2 \succ m_3$, there exists α, β in $(0, 1)$ such that
$$\alpha m_1 + (1-\alpha) m_3 \succ m_2 \succ \beta m_1 + (1-\beta) m_3$$
3. Independence: for all m_1 and m_2 and for all α in (0, 1), $m_1 \succ m_2$ implies
$$\alpha m_1 + (1-\alpha) m_3 \succ \alpha m_2 + (1-\alpha) m_3$$

Consequences of Jaffray's theorem

- Under the previous requirements, we thus have

$$U(m) = \sum_{\emptyset \neq A \subseteq \mathcal{C}} m(A) U(A)$$

- The EU is recovered when m is Bayesian

$$U(m) = \sum_{c \in \mathcal{C}} m(\{c\}) u(c)$$

- Jaffray's theorem does not tell us how to compute $U(A)$. In the general case, we need to elicit the utility values $U(A)$ for each subset $A \subseteq \mathcal{C}$ of consequences, which limits the practical use of the method
- However, Jaffray (1989) showed that a major simplification can be achieved by introducing an additional axiom

Dominance axiom

- Let us write $c_1 \succcurlyeq c_2$ whenever $m_{\{c_1\}} \succcurlyeq m_{\{c_2\}}$
- Furthermore, let $\underline{c}_A$ and $\overline{c}_A$ denote, respectively, the worst and the best consequence in A
- Dominance axiom: for all non-empty subsets A and B of $\mathcal{C}$, if $\underline{c}_A \succcurlyeq \underline{c}_B$ and $\overline{c}_A \succcurlyeq \overline{c}_B$, then $m_A \succcurlyeq m_B$
- Justification:
    - If $\underline{c}_A \succcurlyeq \underline{c}_B$ and $\overline{c}_A \succcurlyeq \overline{c}_B$, it is possible to construct a set Ω of states of nature, and two acts $f : \Omega \rightarrow A$ and $f' : \Omega \rightarrow B$, such that, for any $\omega \in \Omega$, $f(\omega) \succcurlyeq f'(\omega)$
    - As act f dominates f', it should be preferred whatever the information on Ω
    - Hence, f should be preferred to f' when we have a vacuous mass function on Ω, in which case f and f' induce, respectively, the logical mass function m_A and m_B on $\mathcal{C}$
- Consequence: $U(A)$ can be written as $U(A) = u(\underline{c}_A, \overline{c}_A)$

Example

- Assume that $c_1 \succcurlyeq c_2 \succcurlyeq c_3 \succcurlyeq c_4 \succcurlyeq c_5 \succcurlyeq c_6$
- Let $A = \{c_1, c_4, c_5\}$ and $B = \{c_2, c_3, c_6\}$
- Consider the two acts

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
f	c_1	c_4	c_5	c_1	c_1	c_1
g	c_6	c_6	c_6	c_2	c_3	c_6

- f dominates g: it should be preferred whatever the information on Ω
- With m^Ω vacuous, we get $m_f^{\mathcal{C}} = m_A^{\mathcal{C}}$ and $m_g^{\mathcal{C}} = m_B^{\mathcal{C}}$
- Hence, $m_A^{\mathcal{C}} \succcurlyeq m_B^{\mathcal{C}}$

Local Hurwicz criterion

- Adding the dominance axiom to the three previous ones,

$$U(m) = \sum_{\emptyset \neq A \subseteq \mathcal{C}} m(A) u(\underline{c}_A, \overline{c}_A)$$

- We can write

$$u(\underline{c}, \overline{c}) = \alpha(\underline{c}, \overline{c}) u(\underline{c}) + (1 - \alpha(\underline{c}, \overline{c})) u(\overline{c})$$

  where $\alpha(\underline{c}, \overline{c})$ is a local pessism index, defined as the value of α which makes the DM indifferent between:
  1. Receiving at least $\underline{c}$ and at most $\overline{c}$, with no further information, and
  2. Receiving either $\underline{c}$ with probability α or $\overline{c}$ with probability $1 - \alpha$.
- We then have

$$U(m) = \sum_{\emptyset \neq A \subseteq \mathcal{C}} m(A) \left[\alpha(\underline{c}_A, \overline{c}_A) u(\underline{c}_A) + (1 - \alpha(\underline{c}_A, \overline{c}_A)) u(\overline{c}_A) \right]$$

- The generalized Hurwicz criterion corresponds to the case where $\alpha(\underline{c}, \overline{c})$ is equal to a constant α

Other axiomatic justification

- Jaffray's axioms are a counterpart of the axioms of Von Neumann and Morgenstern for decision under risk: they assume that uncertainty on the states of nature is quantified by belief functions
- Jaffray and Wakker (1994) consider a more general situation where probabilities are defined on a finite set S, and there is a multi-valued mapping Γ that maps each element $s \in S$ to a subset $\Gamma(s)$ of Ω
- They justify Jaffray's linear utility for belief functions using a continuity axiom and a weak sure-thing principle (WSTP):
  - A subset $A \subseteq \Omega$ is said to be an ambiguous event if there is a focal set of Bel that intersects both A and $\overline{A}$
  - The WSTP is satisfied if, for any two acts that have common outcomes outside an unambiguous event $A \subset \Omega$, the preference does not depend on the level of those common outcomes

Related work

- Another decision criterion, restricted to partially consonant belief functions, was axiomatized by Giang and Shenoy (2011)
  - A mass function m is said to be partially consonant if its focal sets can be divided into groups such that (a) the focal sets of different groups do not intersect and (b) the focal sets of the same group are nested
  - Appear in some problems of statistical inference (Walley, 1987)
- Decision criteria based on imprecise probabilities are reviewed in (Troffaes, 2007). Some of them may be applicable to belief function, but they remain to be justified in this setting
- The lower and upper expectations are Choquet integrals w.r.t. the belief and plausibility functions, respectively. Savage-like justifications of the Choquet integral (w.r.t. to a non-additive measure) are given in Gilboa (1987), Sarin and Wakker (1992), etc.

Summary

- Several criteria for decision-making with belief functions have been reviewed
- These criteria mix criteria for decision-making under ignorance, and the MEU principle
- A general form of the Hurwicz principle can be justified axiomatically, assuming that uncertainty is quantified by belief functions
- There is no counterpart of Savage's theorem for belief functions

References

J.-Y. Jaffray.
Linear utility theory for belief functions.
Operations Research Letters, 8(2):107 – 112, 1989.

P. Smets.
Decision making in a context where uncertainty is represented by belief functions.
In R. P. Srivastava and T. J. Mock, editors, *Belief functions in business decisions*, pages 17–61. Physica-Verlag, Heidelberg, 2002

J.-Y. Jaffray and P. Wakker.
Decision making with belief functions: compatibility and incompatibility with the sure-thing principle.
Journal of Risk and Uncertainty, 8:255–271, 1994.

I. Gilboa.
Expected utility with purely subjective non-additive probabilities.
Journal of Mathematical Economics, 16:65–88, 1987.