Decision Making with Imprecise and Fuzzy Probabilities - a Comparison

Sven-Hendrik Lossin
University of Hannover
Department for financial economics
losso@losso.de

Abstract

The standard framework of decision theory has no answer to the question how to deal with partial or fuzzy information. In this article two frameworks are presented and compared. The first one uses fuzzy probabilities as in [Buckley 2003] and has been developed by Dubois/Prade [Dubois 1979]. The data-based case is added here. The second framework deals with imprecise probabilities as in [Walley 1991] and proposes a model similar to that by [Kofler 1976]. Furthermore the two frameworks are compared with classical statistical decision theory. It is shown that both of them are similar concerning the mathematical techniques they require but are different regarding the knowledge the decision maker has about the probabilities.

Keywords. decision making, fuzzy probabilities, imprecise probabilities

1 Introduction

Decision theoretic models are widely used especially in economics but also in other disciplines. Recent research is still focused on decisions under risk. This means that most parts of decision theory assume that decision makers have complete knowledge about the distribution from which states of nature are drawn. This assumption is often unrealistic and leads possibly to wrong conclusions if the decision maker is averse to ambiguity. That economic agents are mostly ambiguity averse has been shown in the experiments of Ellsberg[1]. One well known proposal to overcome this problem is the Hodges-Lehmann-Rule[2] that proposes a mixture of the maximin rule and the classical Bayesian rule. But also the Hodges-Lehmann-Rule is not always satisfactory as a unique prior distribution is still needed. The approaches that are presented in this paper cope with the problem of ambiguity by us-

ing soft probabilities such as imprecise probabilities [3] or fuzzy probabilities [4]. In chapter 2 decision making with fuzzy probabilities is presented. In section 2.1 the foundations of fuzzy probabilities and the corresponding probability theory is characterized, followed by decision making with fuzzy probabilities in section 2.2. The third chapter is likely structured as there will be the basics of imprecise probabilities in section 3.1 and the decision problems in section 3.2. In chapter 4 the frameworks are compared with the classical statistical decision theory.

Please note that all probability measures in this article are discrete. The continuous case will bring no new insight but is mathematically more demanding. The models in this paper are kept simple, further readings are given in footnotes.

2 Decision Theory with Fuzzy Probabilities

2.1 The Theory of Fuzzy Probabilities

Since the invention of fuzzy sets by Zadeh[5] it often has been proposed to use fuzzy set theory to express probabilities. Zadeh himself worked upon this[6], the work we refer to is [Buckley 2003]. In Buckley's framework every singleton has a fuzzy probability which is, in contrast to the fuzzy measure theory[7], a fuzzy number. The reason to use fuzzy probabilities instead of classical is that in most applications you do not exactly know the objective probabilities and that you furthermore are not able to compute unique prior subjective probabilities. Instead you only know approximate quantities or linguistic estimations for your probabilities like: this is improbable or that is most probable. Rommelfanger states that "the

---

[1] see [Ellsberg 1961]
[2] see [Hodges 1952]

[3] see [Walley 1991]
[4] see [Buckley 2003]
[5] see [Zadeh 1965]
[6] see [Zadeh 1984]
[7] cf. [Zadeh 1978]

case that extensive information about the entry of the states of nature may not be available has [...] to be considered"[8] and for that he proposes to use fuzzy probabilities. In these cases fuzzy set theory is suitable to express the ambiguity concerning the probabilities. Certainly, an introduction in fuzzy theory would go beyond the scope of this paper so that only a few definitions are given[9].

Let X be a set with elements x and let $\mu_{\overline{A}}$ be a membership function with

$$\mu_{\overline{A}}(x) : X \longrightarrow [0,1] \ .$$

Then a set of pairs

$$\overline{A} = \{(x; \mu_{\overline{A}}(x)) | x \in X\}$$

is called fuzzy set.
A fuzzy set is called a fuzzy number if its membership function increases monotonously to a single culmination in which the membership function takes on the value 1 and then decreases monotonously.

Let $\overline{A}$ be a fuzzy set. Then the (classical) set

$$\overline{A}_\alpha := \{x \in X | \mu_{\overline{A}}(x) \geq \alpha\}$$

is called the α-cut of $\overline{A}$.
Clearly, the α-cuts of a fuzzy number are intervals. Fuzzy numbers are overlined so that they can be distinguished from quantities of the classical set theory.

To overcome the problem of assigning a distribution to a random variable with incomplete knowledge the probabilities in this chapter are given as fuzzy numbers so that the partial ignorance can be modeled. For instance, if you know that the probability for an event is about 0.2, then you will certainly assign the culmination of the fuzzy probability to 0.2 and you will chose the limits of the α-cuts according to the degree of your knowledge.

You may ask what more has to be said about fuzzy probabilities than that you use fuzzy set theory for probabilities. The point is that the extended arithmetic operations proposed by Zadeh[10] cannot be used with probabilities because the sum of discrete probabilities can never exceed 1. Hence, fuzzy probabilities have to be calculated in another way.

For the framework of decision making with fuzzy

probabilities we need not much more then the concept of fuzzy numbers and α-cuts as determined. The fuzzy probability function that is the basis of the fuzzy probability theory is defined as follows.

Definition 1: Let $X = \{x_1, \cdots, x_n\}$ be a discrete random variable. Then the function

$$\overline{P}(X = x_i) = \overline{A}_i \text{ with}$$
$$\mu_{\overline{A}_i}(p) = 0 \text{ for all } p \not\in [0,1] \text{ and } \sum_{i=1}^{n} \overline{A}_{i,1} = 1$$

that assigns to every realization a fuzzy number that stands for its probability is called *fuzzy probability function*.

The last condition - $\sum_{i=1}^{n} \overline{A}_{i,1} = 1$ - will guarantee that the culminations of the fuzzy probabilities - these are the single points at which the membership function takes on the value 1 - will sum to 1 so that the probability function can be analysed for every α. The fuzzy probability function avoids sure loss in the sense of [Walley 1991] and is going to enable us to calculate fuzzy expected utilities which is presented next.

2.2 Decision Making with Fuzzy Probabilities

In the following A stands for the set of possible acts $A = \{a_1, a_2, \cdots, a_m\}$ and $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ for the possible states of nature. The set of natural numbers from 1 to n is denoted by $\mathbb{N}_n$. The utility function u is given by $u : A \times \Theta \longrightarrow \mathbb{R}$. The decision maker prefers high u-values so that it is possible to bring all the pairs of $\{a, \theta\}$ in an order of preference. This is just like in the classical decision theory as in von Neumann/Morgenstern[11]. To adopt the techniques used in classical decision theory we have to calculate an expected utility for every act. After this we have to bring the measured expected utilities in an order so that the highest ranked act can be chosen. Both can certainly easier be done with non-fuzzy (crisp) probabilities. Then you have to weigh the utilities with the probabilities and the act with the highest expected utility has to be chosen. With fuzzy probabilities you have to consider first how a fuzzy expected utility can be computed and second how fuzzy expected utilities can be brought in an order. Both aspects will be dealt with in this chapter.

As has been mentioned above, it is not correct to use extended arithmetic operations on fuzzy probabilities. So we have to chose an indirect way by calculating intervals for every α-cut of the fuzzy expected utility.

---

[8]see [Rommelfanger 1999]
[9]Further readings to fuzzy set theory are [Kruse 1993], [Zadeh 1965] and [Klir 1988].
[10][Zadeh 1965]

[11]see [Neumann 1944]

Definition 2: Let A be a set of acts, Θ a set of states of nature, u a utility function and $\overline{P}$ a fuzzy probability function for Θ. Then for every α with $0 \leq \alpha \leq 1$:

$$\overline{E}_\alpha(a_i) = [\,{}^L\overline{E}_\alpha(a_i);\; {}^U\overline{E}_\alpha(a_i)] :=$$

$$\left[\inf_{p \in \overline{P}_\alpha} \sum_{j=1}^n p_j \cdot u(a_i, \theta_j);\; \sup_{p \in \overline{P}_\alpha} \sum_{j=1}^n p_j \cdot u(a_i, \theta_j) \right]$$

with $p \in \overline{P}_\alpha$ means that $p_j \in \overline{P}_\alpha(\theta_j)$ for all $j \in \mathbb{N}_n$ is called the α-*cut of the expected fuzzy utility*. ${}^L\overline{E}_\alpha(a_i)$ is termed as the *lower α expected utility* and ${}^U\overline{E}_\alpha(a_i)$ as the *upper α expected utility*. The fuzzy number that consists of the α-cuts $\overline{E}_\alpha(a_i)$ $(\alpha \in [0,1])$ is called *fuzzy expected utility* and written $\overline{E}(a_i)$.

The fuzzy expected utility is well defined since there are no mathematical subtleties like open sets of action. The algorithm to calculate these α-cuts is provided by Dubois/Prade[12] and has been simplified by Rommelfanger in [Rommelfanger 1999]. It is presented in the following theorem.

Theorem 3: For every $\alpha \in [0,1]$ the α-cut of the fuzzy expected utility of an act a_i $\overline{E}_\alpha(a_i)$ can be computed by the following algorithm - the normal text points out what has to be done to calculate the lower bound ${}^L\overline{E}_\alpha(a_i)$ and in brackets what for the upper bound ${}^U\overline{E}_\alpha(a_i)$.
1. Reindex the states of nature so that for the given a_i $u(a_i, \theta_j)$ is ordered: $u(a_i, \theta_1) \leq u(a_i, \theta_2) \leq \cdots \leq u(a_i, \theta_m)$[13].
2. Specify for all probabilities the smallest value: $p_j = \inf \overline{P}(\theta_j)$ $\forall j \in \mathbb{N}_n$.
3. Increase p_1 (p_n) as high as possible so that the condition $\sum_{j=1}^n p_j \leq 1$ is still satisfied.
4. If the inequality is fulfilled in the strong sense then continue with the increasement of p_2 (p_{n-1}) in the same way.
5. Repeat this procedure with the next (previous) index as long as the inequality is not fulfilled as equation.
6. Use the calculated probability to compute the lower (upper) α expected utility in the classical way.

After the fuzzy expected utilities have been calculated for every act you have to bring the computed fuzzy numbers in an order. Certainly, this is not as easy as ordering crisp numbers. Several approaches to do this have been proposed. The first one to be presented here is according to [Buckley 2003]. The

---
[12]see [Dubois 1979]
[13]For reasons of simplicity we dispense with a probably mathematically more correct new identifier for the new index.

ordering works with a threshold function.

Definition 4: Let $\overline{E}(a_{i_1})$ and $\overline{E}(a_{i_2})$ be fuzzy expected utilities of two acts (a_{i_1}) and (a_{i_2}). Then the *threshold function* $\nu : A \times A \longrightarrow [0,1]$ which measures how much $\overline{E}(a_{i_1})$ is higher than $\overline{E}(a_{i_2})$ is defined as follows:

$$\nu(\overline{E}(a_{i_1}), \overline{E}(a_{i_2})) =$$
$$\max\{\min\{\mu_{\overline{E}(a_{i_1})}(p_1), \mu_{\overline{E}(a_{i_2})}(p_2)|p_1 \leq p_2\}\}.$$

For given threshold η we define $\overline{E}(a_{i_1}) :\prec \overline{E}(a_{i_2})$ if $\nu(\overline{E}(a_{i_1}), \overline{E}(a_{i_2})) = 1$ and $\nu(\overline{E}(a_{i_2}), \overline{E}(a_{i_1})) < \eta$ and $\overline{E}(a_{i_1}) :\approx \overline{E}(a_{i_2})$ if neither $\overline{E}(a_{i_1}) :\prec \overline{E}(a_{i_2})$ nor $\overline{E}(a_{i_2}) :\prec \overline{E}(a_{i_1})$.

Suppose i.e. that $\eta = 1$. Then the ordering of the fuzzy numbers only depends on the culminations and the two fuzzy expected utilities are only in that case equal iff the culminations are at the same position. For $\eta < 1$ the points of intersection of the membership functions have significance[14].

A second option to order the fuzzy expectations is to shrink the fuzzy number to one representing crisp number. Two alternative ways to do that are to be shown here.
The first way to represent fuzzy numbers with crisp numbers is to calculate the centroid. The centroid is sort of a barycentre of a fuzzy number.

Definition 5: The *centroid of a fuzzy expected utility* is given by

$$c(\overline{E}(a_{i_1})) = \frac{\int_x x \cdot \mu_{\overline{E}(a_{i_1})}(x)}{\int_x \mu_{\overline{E}(a_{i_1})}(x)}.$$

A fuzzy expected utility $\overline{E}(a_{i_1})$ is then *less than* another fuzzy number $\overline{E}(a_{i_2})$ *regarding the centroidal order*, written $\overline{E}(a_{i_1}) <_c \overline{E}(a_{i_2})$, if $c(\overline{E}(a_{i_1})) < c(\overline{E}(a_{i_2}))$.

The second way is more discrete and works with the lower and upper bound of an α-cut for given α.

Definition 6: Let $\overline{E}(a_{i_1})$ and $\overline{E}(a_{i_2})$ be fuzzy expected utilities of two acts (a_{i_1}) and (a_{i_2}). A fuzzy expected utility $\overline{E}(a_{i_1})$ is then *less than* another fuzzy number $\overline{E}(a_{i_2})$ *regarding the Minα* respectively *Maxα order* if $\min\{\overline{E}(a_{i_1})_\alpha\} < \min\{\overline{E}(a_{i_2})_\alpha\}$ respectively $\max\{\overline{E}(a_{i_1})_\alpha\} < \max\{\overline{E}(a_{i_2})_\alpha\}$.

Although definition 4 looks a bit unpleasant its

---
[14]see [Buckley 2003], p.22

concept is quite easy and intuitive. It is computationally less demanding than the calculating of centroids and has the additional merit that a higher degree of fuzziness coincides with a higher degree of indecision[15]. This characteristic maintains the fuzziness of the information about the probabilities in the decisional process which gives it sort of a philosophical justification. On the other hand there are only few situations where indecision is desirable so that definitions 5 and 6 are prefered then. Definition 5 has despite its complexity the advantage that the decision is not only made at one single α-cut but the hole membership function is used for calculation. Definition 6 is a simple rule that can be used if definition 4 leads to indecision. This method is also suggested in [Rommelfanger 1999]. It has to be mentioned that you can mix the orders, in which the fuzzy numbers are represented by crisp numbers, by weighing its representatives.

We have shown an algorithm to compute fuzzy expected utilities from fuzzy probabilities and some ways to bring the calculated fuzzy numbers in an order so that a preference can be scaled. We are going to cope with the data-based decision now.

As in clasical statistical decision theory the prior fuzzy probability will be converted into a posterior fuzzy probability. After that the posterior fuzzy expected utility will be computed by using this posterior knowledge and an order will be used accordingly. So the only thing that has to be shown is how the conversion works. In classical decision theory the Bayes-theorem is applied. This theorem has to be adopted for the fuzzy case. Again the Bayes theorem cannot directly brought into the fuzzy world by using extended arithmetic operations but the α-cuts have to be computed. We assume that there is given an additional random variable Z that contains information about the true state of nature. We define the set of strategies S as the product of the set of acts and $Z : S := A \otimes Z$ like in statistical decision theory. This extension to the set of acts has to be done because the decision depends now on the realization of Z.

Definition 7: Given a set of strategies S, a set of states of nature Θ, $\alpha \in [0, 1]$, a fuzzy probability function $\overline{P}$ for Θ and a random variable Z with the conditional distribution $f : Z \times \Theta \longrightarrow \mathbb{R}$ with $f = f(z|\theta)$. Then the fuzzy probability function

$$\overline{q}(\theta_i|z)_\alpha :=$$

$$\left\{ \frac{f(z|\theta_i) \cdot P_i}{\sum_{j=1}^n f(z|\theta_j) \cdot P_j} \, \middle| \, P_i \in \overline{P}(\theta_i)_\alpha, \sum_{j=1}^n P_j = 1 \right\}$$

is called *posterior fuzzy probability*.

The definition 7 is the Bayes rule for fuzzy probabilities and works precisely like the generalized Bayes rule by Walley[16] applied on the α-cuts of the natural extension[17] of the prior fuzzy probability.
Definition 7 has been extended for the case that also $f(z|\theta)$ is a fuzzy probability function. For further readings see [Buckley 2003] or [Lossin 2004].
With the posterior fuzzy probability we are able to calculate a fuzzy expected utility by using the Dubois/Prade-algorithm (theorem 3) and afterwards these fuzzy utilities can be brought in an order of preference with one of the proposed methods.

3 Decision Theory with Imprecise Probabilities

3.1 The Theory of Imprecise Probabilities

The idea of imprecise probabilities was mentioned first by Boole[18] and has been developed especially by Walley[19]. The basis of this chapter are [Kofler 1976] and [Ehemann 1981]. Their theory of linear partial information has only been discussed in the German-speaking part and bears analogy to [Walley 1991]. Recent research is among others from Augustin[20].

The classical statistical decision theory distinguishes between complete ignorance and risk. Complete ignorance in this case means that the decision maker has no idea about the probabilities that determine the state of nature to arise. The word risk in this context stands for the situation when a decision maker knows exactly the probability density function for Θ. The assumption that in every decision situation one of these two cases - complete ignorance or risk - takes place is highly critical[21]. In decision situations you usually have an idea about the probability density function for Θ, but you don't know it exactly. The assumption with imprecise probabilities is that you can restrict the pdf[22]-space that contains every possible probability function to

---

[15]Indecision is generated in this context whenever two acts are $\approx$ in definition 4 or have the same value c respective $Min\alpha$ or $Max\alpha$ in definition 5 and 6.

[16]see [Walley 1991]
[17]see [Walley 1991]
[18]see [Boole 1854], chapter 16-21
[19]see [Walley 1991]
[20]see [Augustin 2002] and [Augustin 2004]
[21]Many economic researchers even go a step further and argue that you can always construct a subjective probability density function so that you always decide under risk. But this can not only lead to wrong decisions but also has the disadvantage that you lose probably important information about the degree of ignorance.
[22]pdf=probability density function

a subset with the shape of a convex polyhedron. Considering this, the following definitions are easy to understand.

Definition 8: Let p with $p := (p_1, p_2, \cdots, p_n)$ and $\sum_{j=1}^{n} p_j = 1$ be a probability vector. The set V_n, with

$$V_n := \{(p_1, p_2, \cdots, p_n) | p_j \geq 0 \; \forall j \in \mathbb{N}_n, \sum_{j=1}^{n} p_j = 1\}$$

that contains every possible probability density function of Θ with $|\Theta| = n$ is called *probability simplex*.

Definition 9: A set M_n with $M_n \subset V_n$, $M_n \neq \emptyset$ and $|M_n| > 1$ is called *partial information*.

$|M_n| > 1$ ensures that the set M_n contains more than one probability density function as it is the case with risk. M_n is a real subset of V_n so that the decision maker does not decide under complete ignorance. Partial information as defined is not necessarily a convex polyhedron. But it is sensible to focus on convex polyhedrons as partial information because in practice it is realistic that a finite set of information is given. From the algebraic point of view convex polyhedrons are expressed by linear relations. This leads us to the following definition.

Definition 10: Let M_n^* be a partial information. If there is a matrix A and a vector b with $M_n^* = \{p | A \cdot p \leq b\}$, then M_n^* is called a *linear partial information (LPI)*.

The inequality $A \cdot p \leq b$ in definition 10 leads to a convex structure of the LPI. This means that for every $\theta \in \Theta$ you can calculate a lower and an upper probability so that the real probability is between the two of them.

Definition 11: For given M_n we define: $l(\theta_j) = \min\{p_j' | \exists p : p_j = p_j' \wedge p \in M_n\}$ as the *lower probability of event j* and $u(\theta_j) = \max\{p_j' | \exists p : p_j = p_j' \wedge p \in M_n\}$ as the *upper probability of event j*.

Definition 12: Lets consider a special case of LPI in which the upper and the lower probability of every state of nature does not depend directly on the probability of the other states of natures so that the relations of $A \cdot p \leq b$ can be divided to $0 \leq l(\theta_j) \leq p(\theta_j) \leq u(\theta_j) \leq 1 \; \forall j \in \mathbb{N}_n$. This kind of LPI is called *autonomous linear partial information* or *probability intervals*.

Probability intervals have been analysed in [Campos

1994]. For further calculation the extreme-points-matrix is needed. This is a matrix in which vertex distributions of a linear partial information are given.

Definition 13: Let M be an LPI. Then the matrix $EX(M)$ that contains in its columns the vertex distributions of M is called extreme-points-matrix.

Example 14: Let M_3 be an autonomous LPI with $0.4 \leq p(\theta_1) \leq 0.65$, $0.2 \leq p(\theta_2) \leq 0.4$, and $0.1 \leq p(\theta_3) \leq 0.2$. Then:

$$EX(M) = \begin{pmatrix} 0.65 & 0.65 & 0.6 & 0.5 & 0.4 \\ 0.25 & 0.2 & 0.2 & 0.4 & 0.4 \\ 0.1 & 0.15 & 0.2 & 0.1 & 0.2 \end{pmatrix}.$$

You can extend the theory of imprecise probabilities to events which are not singletons and realize that there is no additivity for imprecise probabilities as with usual probabilities. This aspect is interesting but goes beyond the scope of this paper[23].

3.2 Decision Making with Imprecise Probabilities

In this section, criteria for data-free and data-based decisions on the basis of a linear partial information about the state of nature are derived. Considering the data-free decision, there is given the set of acts A, the set of states of nature Θ, the utility function u and the linear partial information M. In classical decision theory you now calculate the expected utility for every act $a_i \in A$ using the given probability to weigh the resulting utilities. With imprecise probabilities this is not possible because you do not have a unique prior distribution but a set of vertex prior distributions. Clearly, with interval probabilities the resulting expected utility is not a number but an interval. It has to be calculated a lower and an upper bound for this interval.

Definition 15: Let A be a set of acts, Θ a set of states of nature, u a utility function and M a linear partial information for Θ. Then

$$E_M(a_i) := [\, {}^L E_M(a_i),\; {}^U E_M(a_i)] :=$$
$$\left[\min_{p \in M} \sum_{j=1}^{n} p_j \cdot u(a_i, \theta_j),\; \max_{p \in M} \sum_{j=1}^{n} p_j \cdot u(a_i, \theta_j) \right]$$

is the *(interval-valued) expected utility of* a_i. ${}^L E_M(a_i)$ and ${}^U E_M(a_i)$ are called *lower and upper expected utility* respectively.

[23]For further information regarding this aspect see [Walley 1991].

$E_M(a_i)$ is well defined and the calculation of this interval can easily be done with the following theorem that works for linear partial informations. If the linear partial information is autonomous then it is even possible to use the Dubois-Prade-algorithm (theorem 3). You simply interpret the interval probabilities as α-cuts by setting $\inf \overline{P}(\theta_j) = l(\theta_j)$ and $\sup \overline{P}(\theta_j) = u(\theta_j)$ and get the interval-valued expectation of the utility. This is methodically the bridge between decision making with fuzzy and with imprecise probabilities. The α-cuts of the fuzzy probabilities of a state of nature are nothing else than a monotone decreasing set-function of probability intervals. Now consider the decision making with non-autonomous linear partial information.

Theorem 16: Let $[\,{}^L E_M(a_i),\ {}^U E_M(a_i)]$ be an expected utility for an act a_i and M the linear partial information for Θ. We define u_{a_i} as a row-array with the utility of the act i for every state of nature as its components. Let 1_j be an array with the dimension n that consists of zeroes but its j-th component is one. Then for all $i \in \mathbb{N}_m$:

$$[\,{}^L E_M(a_i),\ {}^U E_M(a_i)]$$
$$= [\min_{j \in \mathbb{N}_n}\{1_j \cdot u_{a_i} \cdot EX(M)\},\ \max_{j \in \mathbb{N}_n}\{1_j \cdot u_{a_i} \cdot EX(M)\}].$$

In practice you calculate $u_{a_i} \cdot EX(M)$ and the minimum and the maximum of this array define the interval-valued expectation.

After we calculated interval-valued expectation with imprecise probabilities we now have to derive criterions for chosing an optimal act. Analogous to the fuzzy decision theory we have to bring the interval-valued expectations in an order so that a preference order for A can be developed. A criterion that is similar to the minα-criterion (definition 6) is the max$E$min according to [Kofler 1976].

Definition 17: An act a_i is called *optimal with respect to the maxEmin-criterion* if it maximizes $\varphi(a_i) = {}^L E_M(a_i)$ for all $i \in \mathbb{N}_m$.

This criterion is pessimistic because it estimates an act under the assertion that the worst distribution is the true one. Nevertheless it is widely used in scientific literature as it corresponds to the Gamma-Minimax criterion[24], the Maxmin expected utility[25] and the Choquet expected utility[26]. Of course there are other options to bring intervals in an order to derive optimal acts. An important class of criterions bases upon weighted averages.

Definition 18: An act a_i is called *optimal with respect to the η-criterion* if it maximizes for given $\eta \in [0,1]$ $\varphi(a_i) = \eta \cdot {}^L E_M(a_i) + (1 - \eta) \cdot {}^U E_M(a_i)$ for all $i \in \mathbb{N}_m$[27].

It is easy to see that the η-criterion with $\eta = 1$ brings the max$E$min-criterion. For ambiguity-averse agents the parameter η, that has been called caution by [Weichselberger 2001], should be greater than 0.5 so that ${}^L E_M(a_i)$ is higher weighted then ${}^U E_M(a_i)$. An η greater than 0.5 and less or equal than 1 also solves the Ellsberg-paradox[28]. The η-criterion is obviously related to the Hurwicz-criterion in classical decision theory. In imprecise probability theory there are some other interesting concepts that have the disadvantage that they can lead to indecision. Nevertheless they should be mentioned here because they also can be used to preselect the acts that should be taken into account when using the max$E$min or the η-criterion. E-admissibility[29] is one of those concepts. An act is E-admissible if there is a $p \in M$ so that the act has the highest expected utility. Another concept is the maximality by Walley[30] that bases on pairwise comparings of acts. A sophisticated review to decision rules is [Troffaes 2004]. This should be enough to understand how no-data problems with linear partial information can be solved.

Now the data-based decision has to be considered. We have to cope with the aspect how to transform prior knowledge into posterior knowledge with imprecise probabilities. As you will see, this works analogously to fuzzy probabilities.

Definition 19: Given a set of strategies S, a set of states of nature Θ, a linear partial information M for Θ and a random variable Z whose realization contains information about the true state of nature and a conditional distribution $f : Z \times \Theta \longrightarrow \mathbb{R}$ with $f = f(z|\theta)$. Then the interval-valued probability function

$$P(\theta_i|z) :=$$
$$\left[\min_{p \in M}\left\{\frac{f(z|\theta_i) \cdot p_i}{\sum_{j=1}^{n} f(z|\theta_j) \cdot p_j}\right\}; \max_{p \in M}\left\{\frac{f(z|\theta_i) \cdot p_i}{\sum_{j=1}^{n} f(z|\theta_j) \cdot p_j}\right\}\right]$$

---

[24]see [Berger 1984]
[25]see [Gilboa 1989]
[26]see [Schmeidler 1989]

[27]cf. [Weichselberger 2001], ch. 2.6
[28]For further information about the Ellsberg-paradox see [Ellsberg 1961], that the Ellsberg-paradox can be solved by the η-criterion with $0.5 < \eta \le 1$ was mentioned by [Lossin 2004], p. 63.
[29]see [Good 1952] and [Levi 1983]
[30][Walley 1991]

is called *(interval-valued) posterior probability function*[31].

After the interval-valued posterior probability function is calculated, the expected utility can be computed by using theorem 16 and the max$E$min- or the η-criterion can be used accordingly. The calculation is precisely done by using the generalized Bayes rule by [Walley 1991]. This updating mechanism has been criticized by Augustin[32], recently Seidenfeld[33] argued against it when he compared E-admissibility and max$E$min.

4 Comparison of the Frameworks to Classical Decision Theory and Conclusions

Two frameworks have been given to overcome the informational problems with probabilities. It has been shown that there are several similarities between them regarding the mathematical techniques and that autonomous linear partial informations are related to fuzzy probabilities. This has been already mentioned by [Rommelfanger 1999] but he was wrong when he asserted that "this model with linear partial information (LPI)[34] can be interpreted as the special case where all [fuzzy probabilities] have constant membership functions". This is not true because with fuzzy probabilities it is not possible - in the given framework - to model relations between probabilities as with LPI even for singletons. Hence, there is still some space for a model that merges these two concepts. For example, every α-cut could be a coherent imprecise probability measure[35]. [Cooman 2002] contains such a model in terms of possibility measures.

To compare these frameworks with the classical decision theory means to explain what the additional performance of these frameworks is. As has been shown above, the imprecise probabilities allow for relations between probabilities so that the given information for these relations can be used precisely. With fuzzy probabilities these relations can not be modeled but the given framework has the advantage relating to LPI that the culmination as the most probable value of the probability can be taken into account. This can be of great relevance if an information about the probabilities like "the probability for

θ_j is about 0.1" is given. With imprecise probabilities this information can not be used in an obvius way. These are the differences between the two frameworks but what about classical theory? We think that using subjective probabilities is highly controversial because an important information, namely the degree of knowledge, gets lost by computing them. Somebody who is not as good informed as another one has a higher risk to make wrong decisions. This has to be taken into account by making decisions because - not only due to Ellsberg - you have to consider that economic agents are most probably risk averse concerning making right or wrong decisions. But this underlying risk can only be measured in one of the given frameworks and not in classical decision theory.

Acknowledgements

References

[Augustin 2002] T. Augustin (2002); *Expected utility within a generalized concept of probability - a comprehensive framework for decision making under ambiguity*; Statistical Papers 43, 5-22

[Augustin 2003] T. Augustin (2003); *On the Suboptimality of the Generalized Bayes Rule and Robust Bayesian Procedures from the Decision Theoretic Point of View: A Cautionary Note on Updating Imprecise Priors*; ISIPTA '03: Proceedings of the First International Symposium on Imprecise Probabilities and their Applications; Lugano; http://www.sipta.org/isipta03

[Augustin 2004] T. Augustin (2004); *Optimal decisions under complex uncertainty - basic notions and a general algorithm for data-based decision making with partial prior knowledge described by interval probability*; ZAMM - Zeitschrift für angewandte Mathematik und Mechanik 10-11, 1-10

---

[31]A generalization of this definition for the case that also the conditional distribution for Z is imprecise and given by M_f is provided in [Lossin 2004].

[32][Augustin 2003]

[33][Seidenfeld 2004]

[34]Meant by this is the model of [Kofler 1976].

[35]coherence: see [Walley 1991]

[Berger 1984] J.O. Berger (1984); *Statistical Decision Theory and Bayesian Analysis (2nd edition)*; Springer; New York

[Boole 1854] G. Boole (1854); *An Investigation of the Laws of Thought on which are founded the Mathematical Theories of Logic and Probabilities*; Macmillan; London

[Buckley 2003] J.J. Buckley (2003); *Fuzzy Probabilities*; Physica-Verlag; Heidelberg

[Campos 1994] L.M. Campos, J.F. Huete, S. Moral (1994); *Probability intervals: a tool for uncertain reasoning*; International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 2, 167-196

[Cooman 2002] G. de Cooman, P. Walley (2002); *A possibilistic hierarchical model for behaviour under uncertainty*; Theory and Decision 52, 327-374

[Dubois 1979] D. Dubois, H. Prade (1979); *Decision-making under fuzziness*; In : Advances in Fuzzy Set Theory and Applications; Amsterdam

[Ehemann 1981] K. Ehemann (1981); *Entscheidungen bei partieller Information*; Verlagsgruppe Athenäum/Hain/Scriptor/Hanstein; Königstein

[Ellsberg 1961] D. Ellsberg (1961); *Risk, ambiguity, and the Savage axioms*; Quarterly Journal of Economics 75, 643-669

[Gilboa 1989] I. Gilboa, D. Schmeidler (1989); *Maxmin expected utility with nonunique prior*; Journal of Mathematical Economics 18, 141-153

[Good 1952] I.J. Good (1952); *Rational decisions*; Journal of the Royal Statistical Society B 14, 107-114

[Hodges 1952] J.L. Hodges Jr., E.L. Lehmann (1952); *The use of previous experience in reaching statistical decisions*; Ann. M. Stat. 23, 396-407

[Klir 1988] G.J. Klir, T.A. Folger (1988); *Fuzzy Sets, Uncertainty, and Information*; Prentice-Hall; London

[Kofler 1976] E. Kofler, G. Menges (1976); *Entscheidungen bei unvollständiger Information*; Springer; Berlin

[Kruse 1993] R. Kruse, J. Gebhardt, F. Klawonn (1993); *Fuzzy-Systeme*; Teubner; Stuttgart

[Levi 1983] I. Levi (1980); *The Enterprise of Knowledge*; MIT Press; Cambridge

[Lossin 2004] S.H. Lossin (2004); *Erweiterte Konzepte der Wahrscheinlichkeit in der Entscheidungstheorie*; Diploma Thesis; University Hannover

[Neumann 1944] J.v. Neumann, O. Morgenstern (1944); *Theory of Games and Economic Behaviour*; Princeton University Press; Princeton

[Rommelfanger 1999] H. Rommelfanger (1999); *Decision making in fuzzy environment - Ways for getting practical decision models*; In: Decision Theory and Optimization in Theory and Practice; University Chemnitz

[Schmeidler 1989] D. Schmeidler (1989); *Subjective probability and expected utility without additivity*; Econometrica 57, 571-587

[Seidenfeld 2004] T. Seidenfeld (2004); *A contrast between two decision rules for use with (convex) sets of probabilities: Gamma-maximin versus E-admissibility*; Synthese 140 (12), 69-88

[Trofaes 2004] M.C.M. Trofaes (2004); *Decision Making with Imprecise Probabilities: A Short Review*; SIPTA Newsletter Vol. 2 No. 1, 4-7

[Walley 1991] P. Walley (1991); *Statistical Reasoning with Imprecise Probabilities*; Chapman & Hall; London

[Weichselberger 2001] K. Weichselberger (2001); *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I, Intervallwahrscheinlichkeit als umfassendes Konzept*; Physika; Heidelberg

[Zadeh 1965] L.A. Zadeh (1965); *Fuzzy Sets*; Information and Control 8, 338-353

[Zadeh 1978] L.A. Zadeh (1978); *Fuzzy Sets as a basis for a theory of possibility*; Fuzzy Sets and Systems 1, 3-28

[Zadeh 1984] L.A. Zadeh (1984); *Fuzzy Probabilities*; Information Processing and Management 20, 363-372