

## Possibility Theory versus Probability Theory in Fuzzy Measure Theory

Parul Agarwal\*, Dr. H.S. Nayal\*\*

\*Research Scholar, P. G. Degree College, Ranikhet, Almora

\*\* Professor, P. G. Degree College, Ranikhet, Almora

### Abstract

The purpose of this paper is to compare probability theory with possibility theory, and to use this comparison in comparing probability theory with fuzzy set theory. The best way of comparing probabilistic and possibilistic conceptualizations of uncertainty is to examine the two theories from a broader perspective. Such a perspective is offered by evidence theory, within which probability theory and possibility theory are recognized as special branches. While the various characteristic of possibility theory within the broader framework of evidence theory are expounded in this paper, we need to introduce their probabilistic counterparts to facilitate our discussion.

**Keywords:** Probability theory, possibility theory, probability measure (Pro), possibility measure (Pos), belief measure (Bel), Plausibility measure (Pl) and necessity measure (Nec).

### I. Introduction

The aim of this paper is to discuss the differences and similarities between probability measures and possibilities measures. This paper thus establishes a new link between probability and possibility theories. These two theories have an important role in uncertainty. Uncertainty-based information was first conceived in terms of classical set theory and in terms of probability theory. The term information theory has almost invariably been used to a theory based upon the well known measure of probabilistic uncertainty established by [2]. Research on a broader conception of uncertainty-based information, liberated from the confines of classical set theory and probability theory, began in the early eighties. The name generalized information theory was coined for a theory based upon this broader conception.

The ultimate goal of generalized information theory is to capture properties of uncertainty-based information formalized within any feasible mathematical framework. Although this goal has not been fully achieved as yet, substantial progress has been made in this direction. In addition to classical set theory and probability theory, uncertainty-based information is now well understood in fuzzy set theory, possibility theory and evidence theory.

Firstly we introduce these three theories i.e. probability theory, evidence theory and possibility theory which is the part of fuzzy measure theory. In this section we discuss some basic terminologies which are used in probability theory and some axioms of probability distribution, which shows that the probability of the union is the sum of the probabilities of these events. In probability theory there are some limitations and disadvantage with probabilistic method. Evidence theory is also called

Dempster-Shafer theory. The theory of evidence is based on two dual non-additive measures: Belief measure (Bel) and Plausibility measure (Pl). In possibility theory, the two measures of evidence theory, belief measure and plausibility measure becomes necessity measure( $\pi$ ) and possibility measure( $\eta$ ) respectively, which is define briefly in section 2.

In section 3, we discuss some mathematical properties of probability theory and possibility theory. These basic properties help us to compare probability theory to possibility theory. Some of properties define the similarity between probability and possibility. In section 4, we define the differences between probability measure and possibility measure in fuzzy measure. At last we discuss some similarity between probability measure and possibility measure in fuzzy measure.

### II. Fuzzy Measure Theory

In mathematics, fuzzy measure considers a number of special classes of measures, each of which is characterized by a special property. Some of the measures used in this theory are plausibility and belief measures, fuzzy set membership function and the classical probability measures. In the fuzzy measure theory, the conditions are precise, but the information about an element alone is insufficient to determine which special classes of measure should be used [8]. The central concept of fuzzy measure theory is the fuzzy measure which was introduced by Choquet in 1953 by [8] and independently defined by Sugeno in 1974 by [15] in the context of fuzzy integrals.

**Definition:**

Given a universal set X and a non empty family  $\rho$  of subsets of X, a “fuzzy measure” on  $\langle X, \rho \rangle$  is a function  $g: \rho \rightarrow [0,1]$  that satisfies the following requirements by [10]:

[g1] Boundary requirements:  $g(\phi) = 0$  and  $g(X) = 1$

[g2] Monotonicity: If  $A \subseteq B$ , then  $g(A) \leq g(B)$  for all  $A, B \in \rho$ .

[g3] Continuity from below: For any increasing sequence  $A_1 \subset A_2 \subset \dots$  in  $\rho$ ,

$$\text{if } \bigcup_{i=1}^{\infty} A_i \in \rho, \text{ then } \lim_{i \rightarrow \infty} g(A_i) = g\left(\bigcup_{i=1}^{\infty} A_i\right)$$

[g4] Continuity from above: For any decreasing sequence  $A_1 \supset A_2 \supset \dots$  in  $\rho$ ,

$$\text{if } \bigcap_{i=1}^{\infty} A_i \in \rho, \text{ then } \lim_{i \rightarrow \infty} g(A_i) = g\left(\bigcap_{i=1}^{\infty} A_i\right)$$

The boundary requirements [g1] state that the element in question definitely does not belong to the empty set and definitely does belong to the universal set. The empty set does not contain any element hence it cannot contain the element of our interest, either; the universal set contains all elements under consideration in each particular context; therefore it must contain our element as well.

Requirement [g2] states that the evidence of the membership of an element in a set must be at least as great as the evidence that the element belongs to any subset of that set. Indeed with some degree of certainty that the element belongs to a set, and then our degree of certainty that it belongs to a larger set containing the former set can be greater or equal, but it cannot be smaller. Requirements [g3] and [g4] are clearly applicable only to an infinite universal set. They can therefore be disregarded when the universal set is finite. Fuzzy measures are usually defined on families  $\rho$  that satisfy appropriate properties (rings, semirings,  $\sigma$  - algebras, etc.). In some cases,  $\rho$  consists of the full power set  $P(X)$  [10].

Fuzzy measure theory is of interest of its three special branches: probability theory, evidence theory and possibility theory. Although our principle interest is in possibility theory and its comparison with probability theory, evidence theory will allow us to examine and compare the two theories from a broader perspective.

**2.1 Probability Theory:**

Probability represents a unique encoding of incomplete information. The essential task of probability theory is to provide methods for translating incomplete information into this code. The code is unique because it provides a method, which satisfies the following set of properties for any system [13].

1. If a problem can be solved in more than one way, all ways must lead to the same answer.
2. The question posed and the way the answer is found must be totally transparent. There must be no “laps of faith” required to understand how the answer followed from the given information’s.
3. The methods of solution must not be “ad-hoc”. They must be general and admit to being used for any problem, not just a limited class of problems. Moreover the applications must be totally honest. For example, it is not proper for someone to present incomplete information, develop an answer and then prove that the answer is incorrect because in the light of additional information a different answer is obtained.
4. The process should not introduce information that is not present in the original statement of the problem.

**2.2 Basic Terminologies Used in Probability Theory:**

**(I) The Axioms of Probability:** The language systems, contains a set of axioms that are used to constrain the probabilities assigned to events [13]. Four axioms of probability are as follows:

1. All values of probabilities are between zero and one i.e.  $0 \leq P(A) \leq 1 \quad \forall A$ .
2. Probabilities of an event that are necessarily true have a value of one, and those that are necessarily false have a value of zero i.e.  $P(\text{True}) = 1$  and  $P(\text{False}) = 0$ .
3. The probability of a disjunction is given by:  
 $P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \forall A, B$
4. A probability measure, “Pro” is required to satisfy the equation  
 $\text{Pro}(A \cup B) = \text{Pro}(A) + \text{Pro}(B)$   
 $\forall A, B \text{ s.t. } A \cap B = \emptyset$ .

This requirement is usually referred to as the “additivity axiom” of probability measures [9]. An important result of these axioms is calculating the negation of a probability of an event. i.e.

$$P(\bar{A}) = 1 - P(A) \quad \forall A.$$

**(II) The Joint Probability Distribution:** The joint probability distribution is a function that specifies a probability of certain state of the domain given the states of each variable in the domain. Suppose that our domain consists of three random variables or atomic events  $\{a_1, a_2, a_3\}$  [13]. Then an example of the joint probability distribution of this domain where  $a_1 = 1, a_2 = 0$ , and  $a_3 = 1$  would be  $P(a_1, a_2, a_3)$  = same value depending upon the values of different probabilities of  $a_1, a_2$ , and  $a_3$ . The joint probability distribution will be one of the many distributions that can be calculated using a probabilistic reasoning system [7].

**(III) Conditional Probability and Bayes Rule:**

Suppose a rational agent begins to perceive data from its world, it stores this data as evidence. This evidence is used to calculate a posterior or conditional probability which will be more accurate than the probability of an atomic event without this evidence, known as a prior or unconditional probability [13].

**Example:** The reliability of a particular skin test for tuberculosis (TB) is as follows:

⇒ If the subject has TB then the sensitivity of the test is 0.98.

⇒ If the subject does not have TB then the specificity of the test is 0.99.

From a large population, in which 2 in every 10,000 people have TB, a person is selected at random and given the test, which comes back positive.

**What is the probability that the person actually has TB?**

Let's define event A as "the person has TB" and event B as "the person tests positive for TB". It is clear that the prior probability  $P(A) = \frac{2}{10,000} = 0.0002$  and  $P(\bar{A}) = 1 - P(A) = 0.9998$ .

The conditional probability  $P(B|A)$ , the probability that the person will test positive for TB given that the person has TB. This was given as 0.98. The other value we need  $P(B|\bar{A})$ , the probability that the person will test positive for TB given that the person does not have TB. Since a person who does not have TB will test negative 99% (given) of the time, he/she will test positive 1% of the time and therefore  $P(B|\bar{A}) = 0.01$ . By Baye's Rule as:

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(\bar{A})P(B|\bar{A})}$$

$$= \frac{0.0002 \times 0.98}{(0.0002 \times 0.98) + (0.9998 \times 0.01)}$$

$$= 0.0192 \text{ (approx)} = 1.92\% \text{ (approx)}$$

We might find this hard to believe, that fewer than 2% of people who test positive for TB using this test actually have the disease. Ever though the sensitivity and specificity of this test are both high, the extremely low incidence of TB in the population has a tremendous effect on the test's positive predictive value, the population of people who test positive that actually have the disease. To see this, we might try answering the same question assuming that the incidence of TB in the population is 2 in 100 instead of 2 in 10,000.

**(IV) Conditional Independence:** When the result of one atomic event, A, does not affect the result of another atomic event, B, those two atomic events are known to be independent of each other and this helps to resolve the uncertainty [1]. This relationship has

the following mathematical property:  $P(A, B) = P(A).P(B)$ . Another mathematical implication is that:  $P(A|B) = P(A)$ . Independence can be extended to explain irrelevant data in conditional relationship.

**2.3 Disadvantages with Probabilistic Method:**

Probabilities must be assigned even if no information is available and assigns an equal amount of probability to all such items. Probabilities require the consideration of all available evidence, not only from the rules currently under consideration [7]. Probabilistic methods always require prior probabilities which are very hard to found out apriority [9]. Probability may be inappropriate where as the future is not always similar to the past. In probabilistic method independence of evidences assumption often not valid and complex statements with conditional dependencies cannot be decomposed into independent parts. In this method relationship between hypothesis and evidence is reduced to a number [1]. Probability theory is an ideal tool for formalizing uncertainty in situations where class frequencies are known or where evidence is based on outcomes of a sufficiently long series of independent random experiments [13].

**2.4 Evidence Theory:**

Dempster-Shafer theory (DST) is a mathematical theory of evidence. The seminal work on the subject is done by Shafer [14], which is an expansion of previous work done by Dempster [3]. In a finite discrete space, DST can be interpreted as a generalization of probability theory where probabilities are assigned to sets as opposed to mutually exclusive singletons. In traditional probability theory, evidence is associated with only one possible event. In DST, evidence can be associated with multiple possible events, i.e. sets of events. DST allows the direct representation of uncertainty. There are three important functions in DST by [3], [14]: (1) The basic probability assignment function (bpa or m), (2) The Belief function (Bel) and (3) The Plausibility function (Pl). The theory of evidence is based on two dual non-additive measures: (2) and (3).

(1) **Basic Probability Assignment:** Basic probability assignment does not refer to probability in the classical sense. The basic probability assignment, represented by m, defines a mapping of the power set to the interval between 0 and 1, s.t.  $p: \rho(X) \rightarrow [0,1]$  satisfying the following properties: [7]

(a)  $p(\emptyset) = 0$ , and (b)  $\sum_{A \in \rho(X)} p(A) = 1$

The value of  $p(A)$  pertains only to the set A and makes no additional claim about any subset of A. Any further evidence on the subset of A would be represented by another basic probability assignment. The summation of the basic probability assignment of

all the subsets of the power set is 1. As such, the basic probability assignment cannot be equated with a classical probability in general [1].

(2) **Belief Measure:** Given a measurable space  $(X, \rho)$ , a belief measure is a function  $\text{Bel}: \rho(X) \rightarrow [0,1]$  satisfying the following properties:

(a)  $\text{Bel}(\emptyset) = 0$ , (b)  $\text{Bel}(X) = 1$ , and

$$\begin{aligned} & \text{(c)} \text{Bel}(X_1 \cup X_2 \cup \dots \cup X_n) \\ & \geq \sum_i \text{Bel}(X_i) - \sum_{i < k} \text{Bel}(X_i \cap X_k) \\ & + \dots \\ & + (-1)^{n+1} \text{Bel}(X_1 \cap X_2 \cap \dots \cap X_n) \end{aligned} \quad (1)$$

Due to the inequality (1), belief measures are called superadditive. When  $X$  is infinite, function  $\text{Bel}$  is also required to be continuous from above. For each  $Y \in \rho(X)$ ,  $\text{Bel}(Y)$  is defined as the degree of belief, which is based on available evidence [3], that a given element of  $X$  belongs to the set  $Y$ . The inequality (1) implies the monotonicity requirement [g2] of fuzzy measure.

Let  $X_1 \subseteq X_2$  where  $X_1, X_2 \in \rho(X)$  and let  $X_3 = X_2 - X_1$ . Then  $X_1 \cup X_3 = X_2$  and  $X_1 \cap X_3 = \emptyset$ . Applying now  $X_1$  and  $X_3$  for  $n = 2$  to (1), we get

$$\begin{aligned} \text{Bel}(X_1 \cup X_3) &= \text{Bel}(X_2) \geq \text{Bel}(X_1) + \text{Bel}(X_3) - \\ &\text{Bel}(X_1 \cap X_3). \end{aligned}$$

Since  $X_1 \cap X_3 = \emptyset$ , and  $\text{Bel}(\emptyset) = 0$ , we have

$$\text{Bel}(X_1 \cup X_3) \geq \text{Bel}(X_1) + \text{Bel}(X_3)$$

Let  $X_1 = A$  and  $X_2 = \bar{A}$  in (1) for  $n=2$ . Then we have,

$$\begin{aligned} \text{Bel}(A) + \text{Bel}(\bar{A}) &\leq \text{Bel}(A \cup \bar{A}) \leq \text{Bel}(X) \leq 1 \\ \Rightarrow \text{Bel}(A) + \text{Bel}(\bar{A}) &\leq 1 \end{aligned} \quad (2)$$

Inequality (2) is called the fundamental property of belief measures.

(3) **Plausibility Measure:** Given a measurable space  $(X, \rho)$ , a plausibility measure is a function  $\text{Pl}: \rho(X) \rightarrow [0,1]$  satisfying the following properties:

(a)  $\text{Pl}(\emptyset) = 0$ , (b)  $\text{Pl}(X) = 1$ , and

$$\begin{aligned} & \text{(c)} \text{Pl}(X_1 \cap X_2 \cap \dots \cap X_n) \leq \sum_i \text{Pl}(X_i) - \\ & \sum_{i < k} \text{Pl}(X_i \cup X_k) + \dots \dots \dots + (-1)^{n+1} \text{Pl}(X_1 \cup X_2 \cup \\ & \dots \cup X_n) \end{aligned} \quad (3)$$

Due to the inequality (2), plausibility measures are called subadditive. When  $X$  is infinite, function  $\text{Pl}$  is also required to be continuous from below [14].

Let  $X_1 = A$  and  $X_2 = \bar{A}$  in (3) for  $n=2$ . Then we have,

$$\begin{aligned} \text{Pl}(A) + \text{Pl}(\bar{A}) - \text{Pl}(A \cup \bar{A}) &\geq \text{Pl}(A \cap \bar{A}) \\ \Rightarrow \text{Pl}(A) + \text{Pl}(\bar{A}) - \text{Pl}(X) &\geq \text{Pl}(\emptyset) \\ \Rightarrow \text{Pl}(A) + \text{Pl}(\bar{A}) &\geq 1 \end{aligned} \quad (4)$$

According to inequality (2) and (4) we say that each belief measure,  $\text{Bel}$ , is a plausibility measure,  $\text{Pl}$ , i.e. the relation between belief measure and plausibility measure is defined by [16] the following equations:

$$\text{Pl}(A) = 1 - \text{Bel}(\bar{A}) \quad \text{and} \quad (5)$$

$$\text{Bel}(A) = 1 - \text{Pl}(\bar{A}) \quad (6)$$

## 2.5 Possibility Theory:

Possibility theory is a mathematical theory for dealing with certain types of uncertainty and is an alternative to probability theory. Possibility theory is an uncertainty theory devoted to the handling of incomplete information. It is comparable to probability theory because it is based on set-functions. Possibility theory has enabled a typology of fuzzy rules to be laid bare, distinguishing rules whose purpose is to propagate uncertainty through reasoning steps, from rules whose main purpose is similarity-based interpolation [4]. The name ‘‘Theory of Possibility’’ was coined by Zadeh [16], who was inspired by a paper by Gaines and Kohout [6]. In Zadeh’s view, possibility distributions were meant to provide a graded semantics to natural language statements. Possibility theory was introduced to allow a reasoning to be carried out on imprecise or vague knowledge, making it possible to deal with uncertainties on this knowledge. Possibility is normally associated with some fuzziness, either in the background knowledge on which possibility is based, or in the set for which possibility is asserted [9].

Let  $S$  be a set of states of affairs (or descriptions thereof), or states for short. A possibility distribution is a mapping  $\pi$  from  $S$  to a totally ordered scale  $L$ , with top 1 and bottom 0, such as the unit interval. The function  $\pi$  represents the state of knowledge of an agent (about the actual state of affairs) distinguishing what is plausible from what is less plausible, what is the normal course of things from what is not, what is surprising from what is expected [12]. It represents a flexible restriction on what is the actual state with the following conventions (similar to probability, but opposite of Shackle’s potential surprise scale):

(1)  $\pi(s) = 0$  means that state  $s$  is rejected as impossible;

(2)  $\pi(s) = 1$  means that state  $s$  is totally possible (= plausible).

For example, imprecise information such as ‘‘X’s height is above 170cm’’ implies that any height  $h$  above 170 is possible any height equal to or below 170 is impossible for him. This can be represented by a ‘‘possibility measure’’ defined on the height domain whose value is 0 if  $h < 170$  and 1 if  $h \geq 170$  (0 = impossible and 1 = possible). when the predicate is vague like in  $X$  is tall, the possibility can be accommodate degrees, the largest the degree, the largest the possibility.

For consonant body of evidence, the belief measure becomes necessity measure ( $\eta$ ) and plausibility measure becomes possibility measure ( $\pi$ ) [16].

Hence

$$\text{Bel}(A \cap B) = \min[\text{Bel}(A), \text{Bel}(B)]$$

$$\text{Becomes} \quad \eta(A \cap B) = \min[\eta(A), \eta(B)]$$

And  $P(A \cup B) = \max[P(A), P(B)]$   
 Becomes  $\pi(A \cup B) = \max[\pi(A), \pi(B)]$   
 And also  $\pi(A) = 1 - \pi(\bar{A})$ ;  $\eta(A) = 1 - \eta(\bar{A})$

### 3. Basic Mathematical Properties of Probability Theory and Possibility Theory:

1. Probability Theory: It is based on measures of one type: Probability measure (P).

Possibility Theory: It is based on measures of two types:

(a) Possibility measure ( $\pi$ ), (b)

Necessity measure ( $\eta$ ).

2. Probability Theory: Here body of evidence consists of singletons.

Possibility Theory: Body of evidence consists of a family of nested subsets.

3. Probability Theory: Probability measures holds additivity i.e.,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Possibility Theory: Possibility measures and Necessity measures follow the max\min rules:

$$\pi(A \cup B) = \max[\pi(A), \pi(B)] \quad \text{and} \quad \eta(A \cup B) = \min[\eta(A), \eta(B)].$$

4. Probability Theory: Unique representation of P by a Probability distribution function  $p: X \rightarrow [0,1]$  via the formula

$$P(A) = \sum_{x \in A} p(x)$$

Possibility Theory: Unique representation of  $\pi$  by a Possibility distribution function  $\pi: X \rightarrow [0,1]$  via the formula

$$\pi(A) = \max_{x \in A} \pi(x)$$

5. Probability Theory: It is normalized by

$$\sum_{x \in X} p(x) = 1$$

Possibility Theory: It is normalized by

$$\max_{x \in X} \pi(x) = 1$$

6. Probability Theory: Total ignorance:

$$p(x) = \frac{1}{|X|}$$

for all  $x \in X$

Possibility Theory: Total ignorance:

$$\pi(x) = 1 \quad \text{for all } x \in X$$

7. Probability Theory:  $P(A) + P(\bar{A}) = 1$

Possibility Theory:  $\pi(A) + \pi(\bar{A}) \geq 1$

$$\begin{aligned} \eta(A) + \eta(\bar{A}) &\leq 1 \\ \max[\pi(A), \pi(\bar{A})] &= 1 \\ \min[\eta(A), \eta(\bar{A})] &= 0 \end{aligned}$$

As obvious from their mathematical properties, possibility, necessity and probability measures do not overlap with one another except for one very special measure, which is characterized by only one focal element, a singleton [10]. The two distribution

functions that represent probabilities and possibilities become equal for this measure: one element of the universal set is assigned the value of 1, with all other element being assigned a value of 0. This is clearly the only measure that represents perfect evidence.

### 4. Difference Between Probability Theory and Possibility Theory:

1. The theory of possibility is analogous to, yet conceptually different from the theory of probability. Probability is fundamentally a measure of the frequency of occurrence of an event, while possibility is used to quantify the meaning of an event.

2. Value of each probability distribution are required to add to 1, while for possibility distributions the largest values are required to be 1.

3. Probability theory is an ideal tool for formalizing uncertainty in situations where class frequencies are known or where evidence is based on outcomes of a sufficiently long series of independent random experiments. On the other hand possibility theory is ideal for formalizing incomplete information expressed in terms of fuzzy propositions.

4. Possibility measures replace the additivity axiom of probability with the weaker subadditivity condition.

5. Probabilistic bodies of evidence consist of singletons, while possibilistic bodies of evidence are families of nested set.

6. A difference between the two theories is in their expressions of total ignorance. In probability theory, total uncertainty is expressed by the uniform probability distribution on the universal set:  $p(x) = \frac{1}{|X|}$  for all  $x \in X$ . In possibility theory, it is expressed in the same way as in evidence theory  $\pi(x) = 1$ , for all  $x \in X$ .

7. Possibility measures degree of ease for a variable to be taken a value where as probability measures the likelihood for a variable to take a value.

8. Possibility theory is still less developed than their probabilistic counterparts, it is already well established that possibility theory provides a link between fuzzy sets and probability theory are connected with probability theory.

### 5. Similarity Between Probability Theory and Possibility Theory:

1. When information regarding some phenomenon is given in both probabilistic and possibilistic terms, the two descriptions should be in some sense consistent. That is, given a probability measure (P) and a possibility measure ( $\pi$ ) both define on  $\rho(X)$ , the two measures should satisfy some consistency condition.

2. Possibility theory and probability theory are suitable for modelling certain type of uncertainty and suitable for modelling other types.
3. Notion of non-interactiveness on possibility theory is analogous to the notion of independence in probability theory. If two random variables  $x$  and  $y$  are independent, their joint probability distribution is the product of their individual distributions. Similarly if two linguistic variables are non-interactive, their joint probabilities are formed by combining their individual possibility distribution through a fuzzy conjunction operator.
4. Possibility theory may be interpreted in terms of interval-valued probabilities, provided that the normalization requirement is applied. Due to the nested structure of evidence, the intervals of estimated probabilities are not totally arbitrary. If  $\pi(A) < 1$ , then the estimated probabilities are in the interval  $[0, \pi(A)]$ ; if  $\text{Bel}(A) > 0$ , then the estimated probabilities are in the interval  $[\text{Bel}(A), 1]$ . Due to these properties, belief measures and plausibility measures may be interpreted as lower and upper probability estimates.

There are multiple interpretations of probability theory and possibility theory. Viewing necessity and possibility measures as lower and upper probabilities opens a bridge between the two theories, which allow us to adjust some of the interpretations of probability theory to the interval-valued probabilities of possibilistic type.

There are two basic approaches to possibility/probability transformations, which both respect a form of probability-possibility consistency. One, due to [7], [11] is based on a principle of information invariance, the other [5] is based on optimizing information content. Klir assumes that possibilistic and probabilistic information measures are commensurate. The choice between possibility and probability is then a mere matter of translation between languages "neither of which is weaker or stronger than the other" [9]. It suggest that entropy and imprecision capture the same facet of uncertainty, albeit in different guises.

### III. Concluding Remark

The theoretical foundations of probability and possibility based methods were examined and compared in design for maximum safety. A major difference between probability and possibility is in the axioms for the union of disjoint events. The probability of the union is the sum of the probabilities of these events, whereas the possibility is equal to the largest possibility.

Since both probabilistic and possibilistic methods provide estimates of uncertainty that depends on highly subjective selection of prior

probability distributions or possibility distributions by designers, it may be desirable to find ways of providing decision- makers with measures of the degree of subjectivity in the final results. To the best of the author's knowledge, there are no well established methods for doing so.

Possibility can be less conservative than probability in risk assessment with many failure modes. In many reliability assessment problems, one can easily determine the most conservative possibilistic model that is consistent with the available information. On the other hand, it is difficult to choose the most conservative probabilistic model if little information is available. This is an advantage of possibility because in design of high consequence systems, for example nuclear power plants, designers prefer to err on the conservative side.

If we have enough information about uncertainties and accurate predictive models, then probability is advantageous. On the other hand, when making design decisions under limited information or using erude predictive models it may be useful to consider both the probability and possibility of failure of a system. When maximizing safety for a given budget, the design with the lowest probability of failure can be more sensitive than its possibilistic counterparts to errors in the models of uncertainty or errors in the deterministic models for predicting the performance of a system.

### References:

- [1] Cayrac, D., Dubois, D., Haziza, M., Prade, H. (1996) Handling uncertainty with possibility theory and fuzzy sets in a satellite fault diagnosis application. IEEE Trans. on Fuzzy Systems, 4, 251-269.
- [2] Claude E. Shannon and Warren Weaver (1949): The Mathematical Theory of Communication, The University of Illinois Press, Urbana, Illinois. ISBN 0-252-72548-4.
- [3] Dempster AP (1967). Upper and lower probabilities induced by a multivalued mapping. Annals Mathematical Statistics 38: 325-339.
- [4] Dubois, D. and Prade, H. (1996) What are fuzzy rules and how to use them. Fuzzy Sets and Systems, 84: 169-185, 1996.
- [5] Dubois, D., Prade, H. and Sandri, S. (1993), On possibility/probability transformations. In R. Lowen, M. Roubens, editors. Fuzzy logic: State of Art, Dordrecht: Kluwer Academic Publ., pp. 103-112.
- [6] Gaines B. R., and Kohout, L. (1975) Possible automata. Proc. Int. Symp. Multiple-Valued logics, Bloomington, IN, pages 183-196.

- [7] Geer, J.F. and Klir, G.J. (1992),“A mathematical analysis of information-preserving transformations between probabilistic and possibilistic formulations of uncertainty.” International Journal of General systems, 20(2): 143-176.
- [8] Gustave Choquet (1953). “*Theory of Capacities*”. Annales de l'Institut Fourier **5**: 131–295.
- [9] Klir G.J. and Parviz B. (1992), Probability–possibility transformations: a comparison, International Journal of General Systems 21 (3),pp. 291–310.
- [10] Klir, G. J. and Yuan BO, (1995), Fuzzy Sets and Fuzzy Logic, Theory and Applications, Prentice-Hall, ISBN-978-81-203-1136-7.
- [11] Klir, G. J. (1990), A principle of uncertainty and information invariance, International Journal of General Systems,17: pp. 249-275.
- [12] Puri M.L. Ralescu D. (1982) A possibility measure is not a fuzzy measure (short communication). Fuzzy Sets and Systems, pp. 311-313.
- [13] Selvin, S. (1975). A Problem in Probability, American Statistician, 29:1, pp. 67
- [14] Shafer, Glenn; A Mathematical Theory of Evidence, Princeton University Press, 1976, ISBN 0-608-02508-9
- [15] Sugeno, M. (1974). “Theory of fuzzy integrals and its applications. Ph.D. thesis”. Tokyo Institute of Technology, Tokyo, Japan.
- [16] Zadeh L.A., (1978), Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1:3-28.