Journal of Mahani Mathematical Research Center
Print ISSN: 2251-7952 Online ISSN: 2645-4505

FUZZY, POSSIBILITY, PROBABILITY, AND GENERALIZED UNCERTAINTY THEORY IN MATHEMATICAL ANALYSIS

Weldon A. Lodwick*

Dedicated to sincere professor Mehdi Radjabalipour on turning 75

Article type: Research Article

(Received: 12 August 2021, Revised: 04 September 2021, Accepted: 13 September 2021)

(Communicated by M. Mashinchi)

Abstract. This presentation outlines from a quantitative point of view, the relationships between probability theory, possibility theory, and generalized uncertainty theory, and the role that fuzzy set theory plays in the context of these theories. Fuzzy sets, possibility, and probability entities are defined in terms of a function. In the case of fuzzy sets, it is called a **membership function**, in the case of possibility it is called a **possibility measure**, in the case of probability, it is called a **probability distribution function.** In each case, these three functions map the domain to the interval [0,1]. However, each of these functions are defined with different properties. There are generalizations associated with these three theories that lead to intervals (sets of connected real numbers bounded by two points) and interval functions (sets of functions that are bounded by known upper and lower functions). An interval or interval function encodes the fact that it is unknown which of the points or functions is the point or function in questions, that is, the numerical value or real-valued function is unknown, it is uncertain. For generalizations given by pairs of numbers or functions, a case is made for a particular type of generalized uncertainty theory, **interval-valued probability measures**, as a way to unify the generalizations of probability, possibility theory, as well as other generalized probability theories via fuzzy intervals and fuzzy interval functions. This presentation brings a new understanding of quantitative fuzzy set theory, possibility theory, probability theory, and generalized uncertainty and gleans from existing research with the intent to organize and further clarify existing approaches.

Keywords: Possibility Theory, Probability Theory, Interval-Valued Probability Measures, Fuzzy Set Theory
2020 MSC: 60Yxx

1. Introduction

Real-valued mathematical analysis of processes and systems that incorporate uncertainty in its theory, has a history that dates back to at least Archimedes (see [1]) where the perimeter of a circle was approximated via outer and inner

*Corresponding author, ORCID: 0000-0003-0606-5916
E-mail: Weldon.Lodwick@ucdenver.edu
DOI: 10.22103/jmmrc.2021.18055.1160

How to cite: Weldon A. Lodwick, *Fuzzy, Possibility, Probability, and Generalized Uncertainty Theory in Mathematical Analysis*, J. Mahani Math. Res. Cent. 2021; 10(2): 73-101.

polygonal perimeters. These notes will outline newer methods for mathematical analyses inherently characterized by uncertainty of various types where this presentation restricts itself to real-valued uncertainties. Consequently quantitative real analysis is implied in what follows and "real" is dropped when the context is clear.

Many of these notes draw heavily on the research done by Dubois and Prade as well as Klir (see the bibliography) as well as on the author's book [22]. Thus, these notes are a synthesis of many ideas that have already been published. This presentation assumes that the reader is familiar with the basics of set theory and fuzzy set theory at the level of [25] and interval analysis at the level of [26].

The types of uncertainties of interest to this presentation are those that are represented by bounds, either as pairs of numbers (intervals) or pairs of functions enclosing what is unknown, a real number or a real function respectively. Loosely speaking, then, uncertainty will mean, for these sets of notes, the fact that in the analysis it is not known which number or function is the number or function that applies to the problem. However, it is always assumed that at least one number or function is the one that is in fact the number or function in question. Moreover, it is assumed that the bounds are a-priori known either given or readily computable and available. In the case of Archimedes' outer/inner approximation of the perimeter, the circumference, of a circle radius 1, it is known with mathematical certainty that the circumference resides in the interval $[6.2832, 6.2833]$ since $6.2832 \leq 2\pi \leq 6.2833$.

The definitions, properties, representations as mathematical entities, and relationships of uncertainty types are presented. After an introduction to fuzzy set theory where we will define fuzzy interval numbers as the entity of our quantitative analysis, we continue with a discussion of the key theme of these notes, which centers on three types of uncertainties that can be represented by bounds: (1) *probability and its generalizations*, (2) *possibility*, and (3) *interval-valued probability measures* (IVPMs). It will be shown that IVPMs are a type of generalized uncertainty that unifies many types of uncertainties. Let us start with an example.

Example 1.1. *Consider five mutually disjoint states of an event, such that* $X = S_1 \cup S_2 \cup S_3 \cup S_4 \cup S_5$. *However, the probability of each event is not known with certainty although the probabilities are know to be contained, with certainty, in the following intervals* $\Pr S_1 \in \left[\frac{5}{16}, \frac{7}{16}\right]$ *with the "best guess"* $\Pr S_1 = \frac{3}{8}$, $\Pr S_2 \in \left[\frac{1}{10}, \frac{3}{10}\right]$ *with the "best guess"* $\Pr S_2 = \frac{1}{5}$, $\Pr S_3 \in \left[\frac{1}{10}, \frac{3}{10}\right]$ *with the "best guess"* $\Pr S_3 = \frac{1}{5}$, $\Pr S_4 \in \left[\frac{1}{16}, \frac{3}{16}\right]$ *with the "best guess"* $\Pr S_4 = \frac{1}{8}$, *and* $S_5 \in \left[\frac{1}{20}, \frac{3}{20}\right]$ *with the "best guess"* $\Pr S_5 = \frac{1}{10}$. *Note that for this example, the sum of the best guess probabilities is* $\frac{3}{8} + \frac{1}{5} + \frac{1}{5} + \frac{1}{8} + \frac{1}{10} = 1$. *Suppose an event,* E_1, *is composed of two states,* S_1, S_5, *that is,* $E_1 = \{S_1 \cup S_5\}$. *And suppose a second event,* E_2, *is composed of three states,* S_2, S_3, S_4, *that is,*

$E_2 = \{S_2 \cup S_3 \cup S_4\}$. *Given the interval probability of set A, the upper probability needs to be the sum of the upper probabilities. However, it should be no more than* $1 - lower(\Pr(A^c)$, *where* A^C *is the complement of set A. Moreover, the lower probability needs to be the sum of the lower probabilities but greater than or equal to* $1 - upper\Pr(A^c)$. *That is,*

$$upper\,\Pr(A) = (1 - \Pr(A^c)) \leq 1 - \min lower\,\Pr(A^c)$$
$$lower\,\Pr(A) = (1 - \Pr(A^c)) \geq 1 - \max upper\,\Pr(A^c),$$

Thus,

$$\begin{aligned} &upper(\Pr(E_1)) = \\ &= \min\{\{\max\Pr(S_1) + \max\Pr(S_5)\}, 1 - \{\min\Pr(S_2) + \min\Pr(S_4)\}\} \\ &= \min\left\{\left\{\frac{7}{16} + \frac{3}{20}\right\}, \left\{1 - (\frac{1}{10} + \frac{1}{10} + \frac{1}{16}\right\}\right\} \\ &= \min\{0.5875, 7375\} = 0.5875 \end{aligned}$$

and

$$\begin{aligned} &upper(\Pr(E_2)) = \\ &= \min\{\{\max\Pr(S_2) + \max\Pr(S_3) + \max\Pr(S_4)\}, 1 - \{\min\Pr(S_1) + \min\Pr(S_5)\}\} \\ &= \min\left\{\left\{\frac{3}{10} + \frac{3}{10} + \frac{3}{16}\right\}, 1 - \left\{\frac{5}{16} + \frac{1}{20}\right\}\right\} \\ &= \min\{0.7875, 6375\} = 0.6375. \end{aligned}$$

On the other hand,

$$\begin{aligned} &lower(\Pr(E_1)) = \\ &= \max\{\{\min\Pr(S_1) + \min\Pr(S_5)\}, 1 - \max\{\Pr(S_2) + \Pr(S_3) + \Pr(S_4)\}\} \\ &= \max\left\{\left\{\frac{5}{16} + \frac{1}{20}\right\}, 1 - \left\{\frac{3}{10} + \frac{3}{10} + \frac{3}{16}\right\}\right\} \\ &= \max\{0.3625, 2125\} = 0.3625 \end{aligned}$$

and

$$\begin{aligned} &lower(\Pr(E_2)) = \\ &= \max\{\{\min\Pr(S_2) + \min\Pr(S_3) + \Pr(S_4)\}, 1 - \{\max\Pr(S_1) + \max\Pr(S_5)\}\} \\ &= \max\left\{\left\{\frac{1}{10} + \frac{1}{10} + \frac{1}{16}\right\}, 1 - \left\{\frac{7}{16} + \frac{3}{20}\right\}\right\} \\ &= \max\{0.2625, 0.4125\} = 0.4125. \end{aligned}$$

Thus,

$$\Pr E_1 = [0.3625, 0.5875],$$
$$\Pr(E_2) = [0.4125, 0.6375]$$

and since $X = E_1 \cup E_2$

$$\Pr(X) = 1 \in [\mathit{lower}\Pr(X), \mathit{upper}\Pr(X)] = [0.775, 1.225]\,,$$

that is,

$$0.775 \leq \Pr(X) = 1 \leq 1.225.$$

Without using the property that

$$\Pr(A) + \Pr(A^C) = 1,$$

$$\mathit{upper}(E_2) = 0.7875, \mathit{low}(E_1) = 0.2625,$$

where the upper/lower probabilities for E_1 *remain the same. That is, there are potential reductions to upper/lower bounds given that the entities obey the laws of probabilities.*

Remark 1.2. Example 1.1 highlights that when there is only partial information (of the interval in which the probability lies) available, a different approach to classical probabilistic analysis is important to consider. One such approach is possibility theory. Another more general and unifying theory is IVPM, where this latter theory encompasses many generalized uncertainty theories including probability and possibility theory.

Remark 1.3. To obtain more precise results, more information would need to be obtained. For example, given the axiom (definition) of probability that the sum of disjoint probabilities whose union is the entire set is captured by,

$$x_1 + x_2 + x_3 + x_4 + x_5 = 1, \tag{1}$$

such that,

$$x_i \in [\mathit{lower}\Pr(S_i), \mathit{upper}\Pr(S_i)]\,, i = 1, ..., 5.$$

Even with this one constraint (1), there are an uncountably infinite number of solutions. With more information about the nature of the problem, more constraints would result in more precise results. It is noted that Example 1.1 used the fact that $\Pr(A) + \Pr(A^C) = 1$.

The word “uncertainty” and phrase “incomplete information” have been used above. These words for these notes will have a restricted meaning that is articulated next.

Definition 1.4. (Moderately modified from [11]) **Uncertainty** in the context of a quantitative entity is the state of not knowing the exact (crisp, deterministic) real-value of the entity. That is, a piece of information or data of or about a quantitative entity is said to be uncertain when its value is an unknown real-number or real-valued function, contained in a set of two or more numbers or functions prescribing its possible values.

Definition 1.5. (Moderately modified from [11]) A piece of information or data is said to be **incomplete** (or imprecise, not completely specified, lacks information) in the context of quantitative entities, if it is not sufficient to allow for the quantitative entity to be determined as a single real number or a single real-valued function (as a unique fuzzy membership function or as a unique probability distribution).

Functions over sets of real numbers are needed for our exposition. Recall that set-valued functions are functions whose domains are sets, often not singleton sets. As will be seen, generalized uncertainty types of interest begin with set-valued functions. Regardless, the domain of set-valued functions are sets of real numbers, which require an underlying set structure that will allow the construction of these entities of interest and its measure, that is, sets of numbers (intervals) or upper/lower bounding distribution functions. Mathematical analyses (such as integration, optimization, differential equations) often require measures of such functions. These are defined as needed. Moreover, for applications in the context of mathematical analyses one needs a way to account for entities that are pairs of numbers and/or functions and a way to operate on these pairs.

It is clear from Example 1.1, that a more general uncertainty theory than classical probability is necessary. The mathematics of uncertainty has, arguably, been the domain of probability until 1978 when Zadeh's (see [31]) article on possibility theory appeared. Zadeh's motivation was to quantify systems described by linguistic descriptions. We note that toward this aim of quantified systems described by linguistic variables, an application, radiation therapy of tumors, [19,21], incorporates linguistic entities in an "industrial strength" mathematical analysis (possibilistic optimization). That is, what is presented herein has applications beyond illustrative examples.

Fuzzy sets, possibility entities, and probability are defined as functions. That is, fuzzy theory is developed from the properties of its membership function. Then, possibility is defined as a function with properties different from fuzzy membership functions. It is shown that fuzzy intervals can be used to construct possibility entities. Next probability theory, as a function with properties different from fuzzy membership functions, and possibility measures, is defined. Generalizations to probability leading to generalized uncertainty theory ensues. This is followed by a discussion of the relationships between possibility theory and probability theory. The penultimate section proposes a unifying approach to uncertainty theory. The last section contains some conclusions.

2. **Fuzzy Sets**

Fuzzy set theory is a generalization of classical set theory. Mathematically, therefore, fuzzy set theory sits "above" set theory, that is, it is in the realm of set theory. The "fuzzy" generalization is that classical set theory is characterized by entities we call sets whose elements are deterministically known to exclusively belong or not belong. That is, given an entity, say A, called set, in a universe, say X, elements of the universal set $x \in X$ either belong to A or do not belong to A. Fuzzy sets can be approached from a set function point of view, which is presented next.

One of the most general set-valued functions, also a measure since the range is a subset of $\mathbb{R}^{+}$, is

$$g : A \subseteq X \to [0,1], A \subseteq X \subseteq \text{Power set of } \mathbb{R}, \tag{2}$$

together with a structure of the subsets of X. Denote the power set of $\mathbb{R}$ by $P_S(\mathbb{R})$. Some requirements on g to be associated with fuzzy sets are the following:

$$g(\emptyset) = 0, \text{ and } g(X) = 1 \tag{3}$$

$$A \subseteq B \Rightarrow g(A) \leq g(B). \tag{4}$$

The assumptions (2), (3) and (4) define a *fuzzy measure*, a set function. The use of the word "measure" is because the range of the function is a subset of the non-negative real numbers. A consequence of (3) and (4) is:

$$g(A \cup B) \geq \max\{g(A), g(B)\}. \tag{5}$$

$$g(A \cap B) \leq \min\{g(A), g(B)\} \tag{6}$$

For X consisting of a finite set of elements, (2), (3), and (4) suffice. For infinite universal sets X endowed with a sigma algebra (defined below), σ_X, M. Sugeno (see [29]) adds two more axioms. If, $A_1 \subset A_2 \subset ...$,then

$$g\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} g(A_n). \tag{7}$$

If $A_1 \supset A_2 \supset ...$,

$$g\left(\bigcap_{n}^{\infty} A_n\right) = \lim_{n\to\infty} g(A_n). \tag{8}$$

Formally, we have the following definition.

Definition 2.1. A function that satisfies (2), (3), (4), (7), and (8), is a **fuzzy measure**.

The case where the sets of the domain are points, singletons (individual real numbers in the quantitative case), the set function is called a *distribution function* or in the context of fuzzy set theory, it is called a **fuzzy membership function** denoted $\mu(x)$. A classical set A of real numbers from a distribution function point of view can be defined as follows.

Definition 2.2. A **classical set** A, which resides in a universal set X, is characterized by a (unique) membership function $\mu_A(x)$ such that for any x belonging to the universal set X,

$$\mu_A(x) = \begin{cases} 1, x \in A \\ 0, x \notin A \end{cases}, x \in X,$$

that is, $\mu_A(x) \in \{0,1\}$.

Note that a classical set defined via the membership function (the characteristic function) satisfies the conditions of being a fuzzy measure.

Definition 2.3. A **fuzzy set** A, which resides in a universal set X, is characterized by a (unique) membership function $\mu_A(x)$ such that for any x belonging to the universal set X,

$$\mu_A(x) \in [0,1], \tag{9}$$

$$\mu(x) = 0, x \notin A, \mu(x) = 1, x \in A, \tag{10}$$

$$0 < \mu(x) < 1 \text{ otherwise.} \tag{11}$$

Our main interest are in the case of real-valued domains.

Proposition 2.4. *A fuzzy set defined by a membership function* $g : X \subset \mathbb{R} \to [0,1]$ *that satisfies (9), (10), and (11), also satisfies (2), (3), (4), (7), and (8).*

Proof. Let a membership function g be defined by (9), (10), and (11).
1) $g : A \subseteq X \to [0,1], A \subseteq X \subseteq P_S(\mathbb{R})$, by (9) and thus (2) is satisfied.
2) $g(\emptyset) = 0$ since no $x \in X$ belongs to $\{\emptyset\}$ and $g(X) = 1$ since all $x \in X$ and therefore (3) is satisfied.
3) Let $A \subseteq B$. From (10)

$$g_A(x) = 0, x \in B - A.$$

This means that

$$g_A(x) = g_B(x), x \in A \cap B$$
$$g_A(x) = 0 \leq g_B(x) \in [0,1], x \in B - A.$$

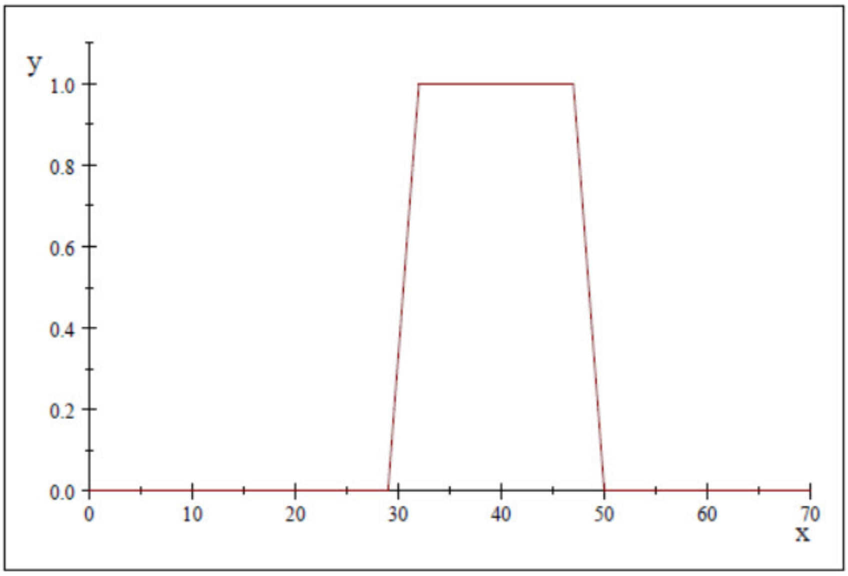
Thus

$$g(A) \leq g(B)$$

and (4) is satisfied.
4) Property (7) follows from the fact that $A = \bigcup_{n=1}^{\infty} A_n = \lim_{n\to\infty} A_n$.
5) Property (8) follows from the fact that $A = \bigcap_{n=1}^{\infty} A_n = \lim_{n\to\infty} A_n$. □

Remark 2.5. Recall that the semantic of a fuzzy set is that it is a set whose elements **transitionally** belong to the set of interest, such that, if the membership function value of an element is zero, this means that the element definitely does not belong to the set. If the membership function value of an element is one, this means that the element definitely does belong to the set. Values between zero and one indicate the grade to which an element belongs. Moreover, if one thinks of set "belonging", ∈, as an operation, then fuzzy set theory "relaxes" or generalizes the set belonging operation from a binary value to a value lying in the interval $[0,1]$.

Example 2.6. *Suppose a fuzzy set is "middle age person", where if the person is between zero years old and 29 years old, the person is definitely not "middle aged" and if the person is older than 50 years, the person is definitely not "middle aged". But between 29 and 50 years, the belonging has the following distribution.*



Middle Age

Measure is one of the key concepts in mathematical analysis. Measure is the mapping from a subset of an entity to a subset of the non-negative real numbers and evaluates the "extent" of the entity as a number or in the case of uncertainty, an interval of numbers. Let us restate fuzzy measure using notation that Puri/Ralescu employ since it will be used to show that a fuzzy measure is not a possibility measure except in the trivial case when the function is identically zero. The Puri/Ralescu characterization of fuzzy measure is simply the function defined by (2), (3), (4), (7), and (8.

Definition 2.7. (See [27]) Let X be a classical set and let $\Im$ be a $\sigma-$algebra (sigma algebra - see definition 4.1 below) of the subsets of X. A **fuzzy measure**, using Puri/Ralescu's notation, is the set function $\mu : \Im \to [0,1]$ with the following properties.
(FM1) $\mu\{\emptyset\} = 0$, and $0 < \mu(A) \leq 1$, for $A \neq \{\emptyset\}$,
(FM2) $A \subset B \Rightarrow \mu(A) \leq \mu(B)$,
(FM3) $A_1 \subset A_2 \subset ... \Rightarrow \mu(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$,
(FM4) $A_1 \supset A_2 \supset ... \Rightarrow \mu(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$.

The operations associated with quantitative fuzzy set theory are generalizations of the operations on sets and can be found in [25]. These operations can be thought of as set "aggregation" operations. When the set is what is called a *fuzzy number*, the operations are those of traditional interval analysis (see [26]) applied to fuzzy sets and are called *fuzzy interval arithmetic* (see [18]).

Operations generate algebraic spaces. In the case of traditional fuzzy interval arithmetic, the space is the fuzzy interval space. Newer representations of intervals than traditional interval representation have been used to endow fuzzy interval spaces with a more ample set of algebraic properties. The traditional approach to fuzzy interval arithmetic does not have additive or multiplicative inverses whereas changing the interval representation to what is called constraint intervals does (see [20, 22]).

A real fuzzy interval is uniquely defined by its alpha levels and therefore its arithmetic is defined via these alpha levels where the alpha levels are closed and bounded intervals. What is crucial to understand is that given the representation of a real fuzzy interval, the associated mathematical analyses are intimately related to those of interval analysis.

3. Possibility

We begin the discussion of possibility theory with a concrete example, which is an application of possibility theory (see [19]).

Example 3.1. *An example of uncertainty that is not captured by conventional probability theory is the uncertainty associated with "the minimum radiation dosage that will kill a cancer cell", as a unique real number, located at a particular voxel of a particular type of cancer of a particular person's computed tomography (CT) image. It is clear that there exists a radiation that will not kill a cancer tumor cell, zero units. And there exists a radiation level that will kill the cancer tumor cell (and kill the patient). So, one posits that there exists a minimal radiation, but what this value is as a unique real number is unknown. Now, suppose a radiation oncologist represents his/her knowledge about the minimum radiation dosage as a distribution of preferred values (see Figure 1, the trapezoid 58/59/61/62) where less than 58 units of radiation definitely does not kill a cancer tumor cell while 62 units of radiation is definitely more radiation than required to kill a tumor cell, with the range 59 to 61 being the most preferred and is the interval that is certain to contain the minimum radiation dosage to kill a cancerous cell. Here 58/59/61/62 (Figure 1) encodes what is known about the minimal radiation level that will kill a cancer cell. Note that this distribution is not a probability distribution since the area under the curve of the trapezoid, 58/59/61/62, is not 1 but 3. Clearly, a probability distribution could be used to describe "minimal tumorcidal dose for a cancer cell", though it is clear that to obtain a value for "minimum dose that will kill a cancer cell" is hard if not impossible to obtain as a single unique real number.*

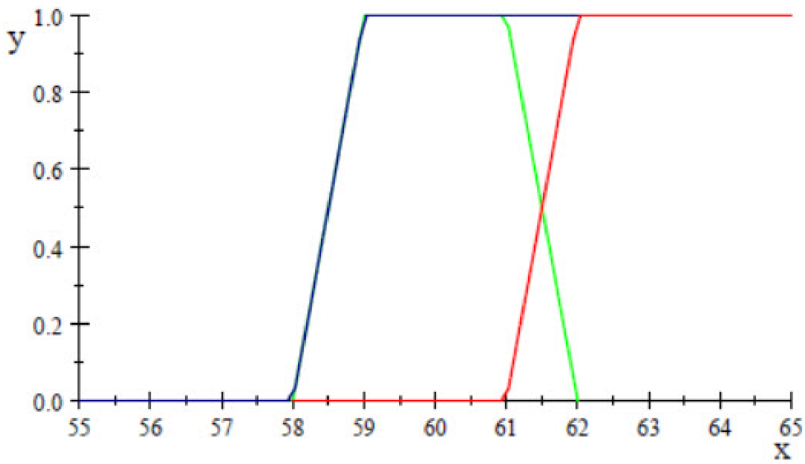


FIGURE 1. Minimal Tumorcidal Dose

This fuzzy interval, the trapezoid 58/59/61/62, associated with "minimal radiation dosage that will kill the cancer cell", is a possibility distribution from which upper and lower bounding possibility measures can be constructed and used in mathematical analysis, for example, in optimization (kill the tumor cells while minimizing the negative effects of radiation to the body). The trapezoidal distribution function is the encoding of what the current state of the partial and incomplete knowledge is. It is noted that the trapezoid itself models the fact that it is not known which distribution is the one which will be the one to be "the" minimal to kill the tumor cell and so we have an uncertainty. In this context, various distributions could be used to represent the uncertainty information encoded in the trapezoid 58/59/61/62:

$$\mu_{trap}(x) = \begin{cases} 0 \text{ for } x < 58, x > 62 \\ x - 58 \text{ for } 58 \leq x \leq 59 \\ 1 \text{ for } 59 < x < 61 \\ -x + 62 \text{ for } 61 \leq x \leq 62 \end{cases} ;$$

$$pos(x) = \begin{cases} 0 \text{ for } x < 58 \\ x - 58 \text{ for } 58 \leq x \leq 59 \\ 1 \text{ for } x > 59 \end{cases};$$

$$nec(x) = \begin{cases} 0 \text{ for } x < 61 \\ x - 61 \text{ for } 61 \leq x \leq 62 \\ 1 \text{ for } x > 62 \end{cases};$$

$$uni(x) = \begin{cases} 0 \text{ for } x < 58 \\ \frac{1}{4}x - 14.5 \text{ for } 58 \leq x \leq 62 \\ 1 \text{ for } x > 62 \end{cases}.$$

Remark 3.2. Note that $pos(x)$ and $nec(x)$ (also $uni(x)$) are cumulative distributions.

The distributions $pos(x)$ and $nec(x)$, where $nec(x)$ is the dual of $pos(x)$ called the necessity, are both possibilities. Moreover, the functions $pos(x)$, and $nec(x)$ are cumulative distributions that enclose ***all*** cumulative distributions whose uncertainty is described by the uncertainty as modeled by the trapezoid $\mu_{trap}(x)$. Of course, it is assumed that the trapezoid $\mu_{trap}(x)$ is the correct and complete uncertainty information associated with the minimal tumorcidal dose, that is, no dose below 58 units has the ability to kill a tumor cell and doses above 62 units are always more than the minimal. Between these two values there is a radiation level that will kill the tumor cell. The function can vary depending on the tumor (type, aggressiveness, stage, location, etc.), research results, oncologists's experience, and the particular patient characteristics.

Remark 3.3. Any distribution like $\mu_{trap}(x)$ generates three principle distributions - (1) itself, (2) its associated possibility, and (3) its associated necessity. From a mathematical analysis point of view, generalized uncertainties are characterized by: (1) A distribution that represents the state of partial knowledge, in our example, $\mu_{trap}(x)$; and (2) A pair of distributions that enclose the possible distributions that can describe the actual situation, $pos(x)$ and $nec(x)$ in our example. This is the salient characteristic of what we call **generalized uncertainty**, which is formally defined subsequently. There are a variety of theories in which the uncertainty is defined by extremal functions such as Dempster-Shafer theory [28], P-Boxes [15] to name two such theories that are discussed below.

Possibility theory was first proposed by L. Zadeh [31] and also articulated by D. Dubois and H. Prade [8]. Necessity was first developed by D. Dubois and H. Prade [8]. Possibility is easier to use than probability as is, perhaps, apparent from Example 3.1. The trapezoid, in fact, was developed from a 30 minute conversation with a radiation oncologist. Additionally, Dubois [3] states:

> "Limited (minimal) specificity can be modeled in a natural way by possibility theory. The mathematical structure of possibility

> theory equips fuzzy sets with set functions, conditioning tools, notions of independence and dependence, decision-making capabilities [lattices]. Lack [deficiency] of information or lack of specificity means we do not have 'the negation of a proposition is improbable if and only if the proposition is probable.' In the setting of lack of specificity, 'the negation of a proposition is impossible if and only if the proposition is necessarily true.' Hence, in possibility theory pairs of possibility and necessity are used to capture the notions of plausibility [possibility] and certainty [necessity]. When pairs of functions are used we may be able to capture or model lack of information. *A membership function is a possibility only when the domain of a fuzzy set is decomposable into mutually exclusive elements.* A second difference [between probability and possibility besides possessing a dual necessity] lies in the underlying assumption regarding a probability distribution; namely, all values of positive probability are mutually exclusive. A fuzzy set is a conjunction of elements. For instance, in image processing, imprecise regions are often modeled by fuzzy sets. However, the pixels in the region are not mutually exclusive although they do not overlap. Namely the region contains several pixels, not a single unknown one. *When the assumption of mutual exclusion of elements of a fuzzy set is explicitly made, then, and only then, the membership function is interpreted as a possibility distribution; this is the case of fuzzy intervals describing the ill-located unique value of a parameter*." (here, the braces, [ ], indicate the authors' comments not in the original and the *italics* are the authors' emphases)

Moreover, possibility is normalized since the semantics of possibility is tied to existential entities. That is, models that use possibility are of existential entities. It is crucial to distinguish gradualness, fuzzy sets, from lack of information, possibility.

3.1. **Possibility Theory From a Function Point of View.** There are several ways that quantitative possibility measures/distribution are derived, defined. This presentation highlights two: (1) Possibility generated via fuzzy sets, which is the way that L. Zadeh [31] introduced the theory, (2) Possibility stated as a function/measure via defining properties that we call *definitional* or *axiomatic* possibility theory. However, we develop possibility and necessity theory from the function/measure point of view. A detailed development can be found in [8] and [25].

A quantitative *possibility measure* is a set-valued map $Pos(A)$,

$$Pos(A) : A \subseteq X \subseteq P_S(\mathbb{R}) \to [0,1], \tag{12}$$

which is precisely (2) for fuzzy measures and also how the general initial part of the definition of a probability distribution. When singletons are in the algebraic structure of sets of the domain, X, as is true of the power set, there is a possibility *distribution,* $pos(x)$, that can be defined from the possibility measure, $Pos(A)$. That is,

$$pos(x) = Pos(\{x\}).$$

On the other hand, if a possibility distribution $pos(x), x \in X$ is known, the possibility measure $Pos(A), A \subseteq X$, can be derived as follows:

$$Pos(A) = \sup_{x \in A} pos(x).$$

We next add properties to (12). It is these properties that distinguish fuzzy set membership functions from possibility measures from probability distribution functions. Then we will show how to construct possibility functions from fuzzy interval (membership) functions.

Figure 1 illustrated how a possibility could be constructed from a fuzzy interval, but this will be more formally done after developing the definitional or axiomatic approach to possibility. The definitions of possibility measures (functions) are the following.

Definition 3.4. A **possibility measure/function** obeys (2), (3), (4) and (5), where (5) is satisfied with equality. In particular, given a universal set X, and its power set $P(X)$, following [31], a **possibility measure** $Pos(A)$, satisfies the following:

$$Pos(A) : A \subseteq X \to [0,1], A \subseteq X, \tag{13}$$

$$Pos(\emptyset) = 0, \text{ and } Pos(X) = 1, \tag{14}$$

$$A \subseteq B \Rightarrow Pos(A) \leq Pos(B), \tag{15}$$

$$Pos(A \cup B) = \max\{Pos(A), Pos(B)\}, \tag{16}$$

The infinite case,

$$Pos\left(\bigcup_{i=1}^{\infty} A_i\right) = \sup_{i=1,2,\ldots} Pos(A_i). \tag{17}$$

And when the universal set X is uncountable, and indexing set I is used so that

$$Pos\left(\bigcup_{i \in I} A_i\right) = \sup_{i \in I} Pos(A_i). \tag{18}$$

However, we restrict ourselves to the infinite countable case (16).

Note that (16) is (5) where $\geq$ is replace with equality. Next, we define a *dual* measure to Pos, the necessity measure, via functions.

Definition 3.5. A necessity measure obeys (2), (3), (4) and (6), where (6) is satisfied with equality. That is, a **necessity measure** satisfies the following:

$$Nec(A) : A \subseteq X \to [0,1], A \subseteq X, \tag{19}$$

$$Nec(\emptyset) = 0, \text{ and } Nec(X) = 1 \tag{20}$$

$$A \subseteq B \Rightarrow Nec(A) \leq Nec(B) \tag{21}$$

$$Nec(A \cap B) = \min\{Nec(A), Nec(B)\}. \tag{22}$$

The infinite case for necessity is

$$Nec\left(\bigcap_{i \in I} A_i\right) = \inf_{i \in I} Nec(A_i). \tag{23}$$

Remark 3.6. Given a possibility measure $Pos(A)$, a **dual** necessity measure can also be defined (see Dubois/Prade [8]) as

$$Nec(A) = 1 - Pos(A^C). \tag{24}$$

Proposition 3.7. *Necessity defined by (24) satisfies (19), (20), (21), and (22) for the finite case.*

Proof. (1) To see (19)

$$0 \leq 1 - Pos(A^C) = Nec(A) \leq 1$$

since $0 \leq Pos(A^C) \leq 1$.
(2) To see (20),

$$Nec(\emptyset) = 1 - Pos(\emptyset^C) = 1 - Pos(X) = 1 - 1 = 0$$

(3) To see (21), let $A \subseteq B$. This means that $B^C \subseteq A^C$ which means $Pos(A^C) \geq Pos(B^C)$. Thus

$$Nec(A) = 1 - Pos(A^C) \leq 1 - Pos(B^C) = Nec(B).$$

(4) To see (22), $(A \cap B)^C = A^C \cup B^C$. Therefore,

$$\begin{aligned} Nec(A \cap B) &= 1 - Pos((A \cap B)^C) \\ &= 1 - Pos(A^C \cup B^C) \\ &= 1 - \sup\left\{Pos(A^C), Pos(B^C)\right\} \\ &= \inf\left\{\left(1 - Pos(A^C)\right), \left(1 - Pos(B^C)\right)\right\} \\ &= \inf\left\{Nec(A), Nec(B)\right\}. \end{aligned}$$

The case of arbitrary number of intersection is done using finite induction. For the infinite countable case the definition of infimum is used. □

3.2. **Fuzzy and Possibility Measures are Different.** We end this section by showing that a possibility measure is ***not*** a fuzzy measure except in the trivial case. This result, to the best of our knowledge, was first observed by Puri/Ralescu (see [27]). Thus, there not only is the semantic difference between possibility and fuzzy, there is also an underlying measure theoretical difference. It should be noted that the domain space $X \subseteq \mathbb{R}$ to which these examples and theorems refer are infinite.

Note that a possibility *measure* Π can be uniquely defined by any function

$$f : X \to [0, 1]$$

via

$$\Pi(A) = \sup_{x \in A} f(x), A \subseteq X. \tag{25}$$

On the other hand, given a possibility *measure* $\Pi(A), A \subseteq X$, where X contains singleton elements (as in the power set of X) a function, its possibility *distribution*, can be generated as follows.

$$\pi(x) = f(x) = \Pi(\{x\}), x \in X. \tag{26}$$

That is, given a function, f, we can generate a possibility measure, as Zadeh did [31]. In Zadeh's case, the function f that maps $X \subseteq \mathbb{R}$ into $[0, 1]$, was the membership function. Given a possibility measure, Π, we can generate a possibility distribution function, $\pi(x)$. Bearing this in mind, Puri/Ralescu [27] use the following example to show that fuzzy and possibility measures are distinct except when the associated with the trivial distribution $\pi(x) = f(x) = 0, \forall x \in X$. The example is the following, where Puri/Ralescu use f rather than π. However, we will keep both for clarity.

Example 3.8. *([27]) Let $X = \mathbb{R}, \pi(x) = f(x) = 1$. Define a possibility measure, Π, as*

$$\Pi(A) = \sup_{x \in A} f(x).$$

Let $A_n = (n, \infty)$. We have $\Pi\left(\bigcap_{n=1}^{\infty} A_n\right) = \Pi(\emptyset) = 0$. However

$$\lim_{n \to \infty} \Pi(A_n) = 1.$$

Thus, Π is not a fuzzy measure since it violates property FM4.

Example 3.9. *([27]) Let $X = [0, 1]$,*

$$\pi(x) = f(x) = 1, x \in [0, 1)$$
$$\pi(1) = f(1) = 0.$$

Let $A_n = [1 - \frac{1}{n}, 1]$. Therefore

$$\Pi\left(\bigcap_{n=1}^{\infty} A_n\right) = \Pi(1) = \pi(1) = 0.$$

However

$$\lim_{n\to\infty} \Pi(A_n) = \sup_{x\in A_n} \pi(x) = 1.$$

Puri/Ralescu go on to prove the following theorem.

Theorem 3.10. *([27], page 312). Given*

$$\pi(x) = f(x) : \mathbb{R}^k \to [0,1]$$

with

$$\Pi(A) = \sup_{x\in A} \pi(x) = \sup_{x\in A} f(x), A \subseteq \mathbb{R}^k,$$

its associated measure. If Π *is a fuzzy measure, then* $\pi(x) = f(x) = 0$ *at every point of the continuity of* π, f.

Corollary 3.11. *([27], page 312). Let* Π *be a possibility measure with continuous "density" (distribution) function* π. *If* Π *is a fuzzy measure, then* $\pi = 0$.

Remark 3.12. What the examples, theorem, and corollary mean is that when fuzzy measures are possibility measures, then the corresponding possibility distribution is the zero function. Thus, fuzzy measures and possibility measures are the same only for the "trivial" case.

3.3. **Possibility Theory From Fuzzy Sets.** The original work by L. Zadeh [31] on possibility theory started with a possibility distribution derived from a fuzzy set. Let $\mu_A(x)$ be a membership function of a given fuzzy set A.

Definition 3.13. The possibility distribution associated with the fuzzy set A, as defined by Zadeh [31], is

$$pos_A(x) = \mu_A(x). \tag{27}$$

Remark 3.14. Zadeh's definition of possibility (27) can be, and indeed is, confusing since it appears that there is no difference between fuzzy and possibility.

Klir and Yuan [25] state the following:

> "A *fuzzy set* F defines the degree to which x belongs to F, not the degree to which evidence supports the fact that x is F. A *possibility measure* is one that assigns a degree of certainty that an element is F. It is the degree to which the evidence supports that x is F."

Remark 3.15. For the example "Middle Age" (see Figure 2.6) as a fuzzy set, x is precise age and $\mu_{MidAge}(x)$ is the degree to which the age x belongs to the set "Middle Age". On the other hand, "Middle Age" as a possibility, $pos_{MidAge}(x)$ is the degree of certainty, our ranking, that the age x is "Middle Age".

A possibility measure defined via a fuzzy membership function satisfies the properties of possibility measure (13), (14), (15), and (16). This is proved next.

Proposition 3.16. *Given $Pos(A)$ defined via fuzzy sets according to [31], it satisfies (13), (14), (15), and (16) for the finite union case. Recall that, for possibility, the domain X is always assumed to contain a point that definitely exists, that is, has possibility value* 1.

Proof. Given $pos(x) = \mu_A(x)$, then

$$\begin{aligned} Pos(A) &= \sup_{x \in A} pos(x) = \sup_{x \in A} \mu_A(x) : A \subseteq X \to [0,1], A \subseteq X \\ Pos(\emptyset) &= \sup_{x \in \emptyset} \mu_\emptyset(x) = 0 \text{ by convention} \\ Pos(X) &= \sup_{x \in X} \mu_X(x) = 1. \end{aligned}$$

Given two sets $A \subseteq B$

$$Pos(A) = \sup_{x \in A} \mu_A(x) \leq \sup_{x \in B} \mu_B(x) = Pos(B).$$

$$\begin{aligned} Pos(A \cup B) &= \sup_{x \in A \cup B} \mu_{A \cup B}(x) = \sup_{x \in A} \mu_A(x) \text{ or } \sup_{x \in B} \mu_B(x) \\ &= \max_{x \in A \cup B} \left\{ \sup_{x \in A} \mu_A(x), \sup_{x \in B} \mu_B(x) \right\}. \end{aligned}$$

The finite case is done by finite induction. For the countable case the definition of supremum is used. □

The Zadeh definition of possibility satisfies the conditions of a possibility measure. That is, an arbitrary fuzzy membership function $\mu(x)$ generates a possibility measure $pos(x)$. A fuzzy interval is a particular case of a membership function. Thus, it generates a possibility and therefore a necessity distribution. This is formalized next.

Definition 3.17. A fuzzy interval A is a fuzzy set over the real numbers, $\mathbb{R}$, such that the membership function μ_A has the following properties:

$$\begin{aligned} &\mu_A(x) \in [0,1], \forall x \in \mathbb{R}, \\ \{x \mid \mu_A(x) > 0\} &= (a,b), -\infty < a \leq b < \infty, \\ &\exists x \in [c,d] \subset [a,b], \mu_A(x) = 1, c \leq d, \\ &\text{if } x \in [a,c],\ \mu_A(x) \text{ is continuous nondecreasing}, \\ &\text{if } x \in [d.b],\ \mu_A(x) \text{ is continuous nonincreasing.}, \\ &x \in (-\infty, a] \text{ or } x \in [b, \infty),\ \mu_A(x) = 0 \end{aligned}$$

There are more general ways to define a fuzzy interval, but the above is sufficient for our purposes. Note that since a fuzzy interval is a membership function, its possibility is directly defined as the membership functions. The possibility and necessity associated with a fuzzy interval as presented next was introduced by Dubois (for a more recent exposition, see [5]).

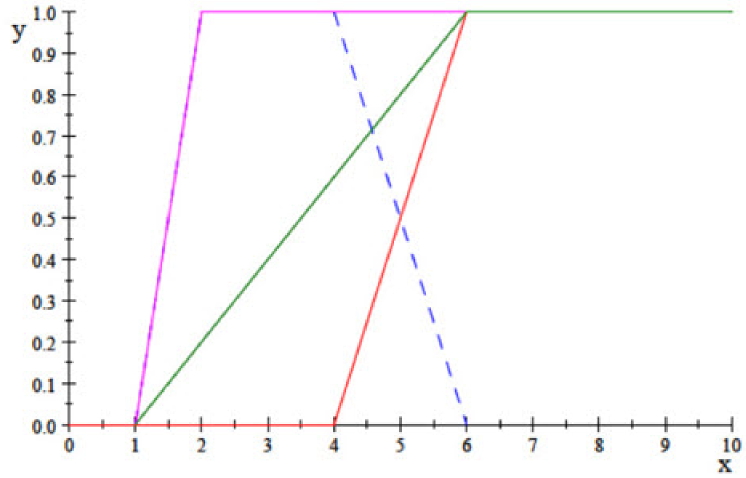


FIGURE 2. Fuzzy Interval: Trapezoid 1/2/4/6 - Possibility, Necessity

A fuzzy interval 1/2/4/6 28, like the trapezoid depicted in Figure 2, is given by

$$Trap_{1/2/4/6}(x) = \begin{cases} 0 \quad x \leq 1 \\ x-1 \quad 1 < x \leq 2 \\ 1 \quad 2 < x \leq 4 \\ -\frac{1}{2}x + 3, 4 \leq x \leq 6 \\ 0, x > 6 \end{cases} . \tag{28}$$

As previously mentioned, a fuzzy interval like (28) can be viewed as encoding a family of cumulative probability distributions. Both upper possibility and lower necessity, respectively, the magenta and light red distributions of Figure 2, are cumulative probability distribution functions. Therefore, given a fuzzy interval as a piece of incomplete information, it generates a family of CDFs bounded by a upper possibility and lower necessity pair as depicted in Figure 2. A possibility description of the value of a parameter (or entity) which encapsulates what is known about the possible values of that parameter or entity, generally uses a (single) fuzzy interval as its distribution function, for example, the trapezoid in Figure 2.

Let us next show that the upper function depicted in Figure 2 is indeed a possibility whereas the lower function, depicted in the same figure, is indeed its corresponding dual necessity. Without loss of generality, we assume that our fuzzy interval is a trapezoidal fuzzy interval.

Proposition 3.18. *Given a trapezoidal fuzzy interval $a/b/c/d$, with$a < b < c < d \in \mathbb{R}$ and*

$$pos_{trap}(x) = \begin{cases} 0 \text{ if } x < a, x > d \\ \frac{1}{b-a}x - \frac{a}{b-a} \text{ if } a \leq x \leq b \\ 1 \text{ if } b \leq x \leq c \\ -\frac{1}{d-c}x + \frac{d}{d-c} \text{ if } c \leq x \leq d \end{cases}. \tag{29}$$

Define the upper function

$$\overline{pos}_{trap}(x) = \begin{cases} 0 \text{ if } x < a \\ \frac{1}{b-a}x - \frac{a}{b-a} \text{ if } a \leq x \leq b \\ 1 \text{ if } x > b \end{cases}. \tag{30}$$

Both 29 and 30 are possibility distributions.

Proof. Let

$$Pos(A) = \sup_{x \in A \subseteq \mathbb{R}} \overline{pos}_{trap}(x).$$

(1) $Pos(A) \in [0, 1]$ by construction.
(2) $Pos(\emptyset) = \sup_{x\in\emptyset} \overline{pos}_{trap}(\emptyset) = 0$, and $Pos(X) = \sup_{x\in X} \overline{pos}_{trap}(x) = 1$.
(3) $Pos(A \cup B) = \sup_{x\in A\cup B} \overline{pos}_{trap}(x) = \sup \left\{ \sup_{x\in A} \overline{pos}_{trap}(x), \sup_{x\in B} \overline{pos}_{trap}(x) \right\}$ □

Proposition 3.19. *Given a trapezoidal fuzzy interval $a/b/c/d$, with $a < b < c < d \in \mathbb{R}$ and $pos_{trap}(x)$ given by distribution (29), the lower function*

$$\underline{pos}_{trap}(x) = nec_{trap}(x) = \begin{cases} 0 \text{ if } x < c \\ \frac{1}{d-c}x - \frac{c}{d-c} \text{ if } c \leq x \leq d \\ 1 \text{ if } x > d \end{cases} \tag{31}$$

is the dual necessity distribution with respect to the possibility Equation (29). Moreover, Equation (31) is also a possibility.

Proof. Let fuzzy interval be $a/b/c/d$ be given by (29) and (30).

$$\begin{aligned} Nec(A) &= 1 - Pos(A^C) \\ &= 1 - \sup_{x\in A^C} \underline{pos}_{trap}(x). \end{aligned}$$

Let $A = (-\infty, c) \Rightarrow A^C = [c, \infty)$. Therefore

$$\begin{aligned} Nec(A) &= 1 - \sup_{x\in[c,\infty)} \underline{pos}_{trap}(x) \\ &= 1 - 1 = 0. \end{aligned}$$

This means that

$$nec(x) = 0 \ \forall x \in (-\infty, c).$$

Now let $A = (-\infty, x), x \in (c, d) \Rightarrow A^C = [x, \infty), c < x < d.$

This means that

$$
\begin{aligned}
Nec(A) &= 1 - \sup_{[x,\infty)} \underline{pos}_{trap}(x), c < x < d \\
&= 1 - \left(-\frac{1}{d-c}x + \frac{d}{d-c} \right) \\
&= 1 + \frac{1}{d-c}x - \frac{d}{d-c} \\
&= \frac{1}{d-c}x + \frac{d-c}{d-c} - \frac{d}{d-c} \\
&= \frac{1}{d-c}x - \frac{c}{d-c}.
\end{aligned}
$$

Lastly, let $A = (-\infty, x), x \in [d, \infty) \Rightarrow A^C = [x, \infty), x \geq d$.

This means that

$$
\begin{aligned}
Nec(A) &= 1 - \sup_{[x,\infty)} \underline{pos}_{trap}(x), d \leq x \\
&= 1 - 0 \\
&= 1.
\end{aligned}
$$

Thus, $nec(x) = 1, \forall x \in [d, \infty)$ and the first part of the theorem is proved. Since $Nec(A)$ is similar to $Pos(A)$ except it is shifted, it too is a possibility. □

Remark 3.20. We emphasize that the dual necessity is with respect to the **fuzzy interval** $pos_{trap}(x)$, Equation (29), and it is also a possibility.

Remark 3.21. It can be shown that nested sets can be generated by possibility measures and possibility measures can be generated from nested sets. Moreover, if we are to construct possibility/necessity pairs, we either need to begin with a set of nested subsets, or if we construct possibility/necessity pairs, we have nestedness underlying the domain of the sets over which the pair of distribution measures operate. Given a fuzzy interval as defined by Definition 3.17, since the $\alpha-$levels of a fuzzy interval are nested, the underlying Pos/Nec measures associated with fuzzy intervals exists. That is, nestedness is a characteristic of alpha levels of a fuzzy interval. When considering generalizations of probability, such as Dempster/Shafer theory of evidence or random sets, nestedness is used to related these generalized probability theories to possibility. That is, theory of evidence and random sets will be related to possibility as long as the underlying set structure over which these two generalized theories operate are nested. The third generalized probability theory we present called P-Boxes, have nestedness so are directly related to possibility.

Some simple standard properties associated with possibility and necessity measures (see [25]) are the following.

(1) $Nec(A) + Nec(A^C) \leq 1$
(2) $Pos(A) + Pos(A^C) \geq 1$

(3) $Nec(A) + Pos(A^C) = 1$
(4) $\min\{Nec(A), Nec(A^C) = 0$
(5) $\max\{Pos(A), Pos(A^C) = 1$
(6) $Nec(A) > 0 \Rightarrow Pos(A) = 1$
(7) $Pos(A) < 1 \Rightarrow Nec(A) = 0$.

4. Probability

We begin with a very brief introduction to probability measures highlighting only those aspects that are relevant to our exposition. The domain, a subset of sets, needs a structure and for our purposes, the set structure is a sigma algebra, which is discussed next. Note that in these notes, by a **partition** of a set X, is an at most countable collection of pair-wise disjoint subsets whose union is X. For possibility, we used the power set as a structured set of sets. A sigma algebra set structure is used for probability, the power set being one example of a sigma algebra.

Definition 4.1. Given a (universal) set $X \neq \emptyset$, a **sigma algebra** defined on X, denoted σ_X, is a family of subsets of X such that:
1) $\emptyset \in \sigma_X$;
2) $X \in \sigma_X$;
3) $A \in \sigma_X \Rightarrow A^C \in \sigma_X,\ A^C$;
4) $A_i \in \sigma_X$, for any countable set (could be finite) $\Rightarrow \cup_i A_i \in \sigma_X$.

The power set of X, $P_S(X)$ is a common sigma algebra structure of the domain. However, the power set need not be the structure of the subsets of the domain. The pair (X, σ_X) is called a ***measurable space***. Let (X, σ_X) be a measurable space. By a ***measure*** μ on this space we mean a set-valued function

$$\mu : \sigma_X \to \mathbb{R}^+$$

such that $\mu(\emptyset) = 0$ and for any partition of $A \subseteq X$, $A_i \in \sigma_X, \mu(A) = \mu(\cup_i A_i) = \sum_i \mu(A_i)$. The triple (X, σ_X, μ) is called a ***measure space***.

Definition 4.2. (Probability) If the mapping $\mu : \sigma_X \to [0, 1]$ has the property that $\mu(X) = 1$, then the measure is called a ***probability measure*** with μ now denoted $\Pr_X$ and the measure space is called a ***probability measure space*** denoted $(X, \sigma_X, \Pr_X)$.

Remark 4.3. By definition of a probability measure, we have the fundamental additive property of probabilities, that is,

$$\Pr(A \cup B) = \Pr(A) + \Pr(B), A \cap B = \emptyset.$$

5. Generalization of Probability Theory

There are many generalizations of probability theory that have been developed and have relationships to possibility theory. We present three such generalizations, each of which has a relationship with possibility theory. There are at least 8 other generalizations (see, for example, [22]) that can be found in the literature.

5.1. **Evidence Theory of Dempster/Shafer.** Given X and $P_S(X)$, the power set of X, the Dempster/Shafer theory is based on the function

$$m : P_S(X) \to [0,1]$$

such that

$$m(\emptyset) = 0 \text{ and } \sum_{A \in P_S(X)} m(A) = 1.$$

The function m is called the *basic assignment function*. Every set $A \in P_S(X)$ for which $m(A) \neq 0$ is called a *focal element*. The pair (F, m) where F denotes the set of all focal elements of m, is called a body of evidence. There are two measures they define,

$$Bel(A) = \sum_{B \subseteq A} m(B) \text{ Belief measure,}$$

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

(1) When the focal elements are singletons, then $Bel(A) = Pl(A)$ and they become a probability.
(2) When all focal elements are nested (ordered by inclusion), the body of evidence is called consonant. In this case, the plausibility measures become possibility measures and the belief measures become necessity measure.

Recall that a possibility measure Π is determined by a function

$$pos : X \to [0,1]$$

via the formula

$$\Pi(A) = \max_{x \in A} pos(x)$$

for all $A \in P_S(X)$. The corresponding necessity measure Nec is determined by

$$Nec(A) = 1 - \Pi(A^c)$$

where A^c is the complement of the set A. This results in

$$\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$$
$$Nec(A \cap B) = \min\{Nec(A), Nec(B)\}.$$

On the other hand, a consonant body of evidence

$$F = \{A_1, A_2, ..., A_n\}, A_1 \subset A_2 \subset ... \subset A_n$$

and a possibility distribution pos can be used to generated a basic assignment function via

$$m(A_i) = pos(x_i) - pos(x_{i+1})$$

for some $x_i \in A_i, x_{i+1} \in A_{i+1}$, where $pos(x_{n+1}) = 0$ by convention, $i = 1, 2, ..., n$. Similarly, a possibility distribution is given by

$$r(x_i) = \sum_{k=1}^{n} m(A_k)$$

for each $x_i \in A_i$. That is, given a consonant body of evidence, we can generate a belief and plausibility pair and also the pair possibility and necessity.

5.2. **Random Sets.** A random set can be considered as a convenient name for a convex combination of a weighted family of sets, with positive weights summing to 1. Let $U \subseteq \mathbb{R}$ be a non-empty set.

Definition 5.1. A random set on U is a pair (F, m) where F is a family of distinct non-empty subsets of U and m is a mapping $F \to [0, 1]$, such that $\sum_{A \in F} m(A) = 1$.

The family of sets F is called the *support* of the random set and m is called the *basic probability assignment*. Each $A \in F$ contains the possible values of a variable x, and $m(A)$ can be viewed as the probability that A is the actual range of x. Such a random set (F, m) is equivalent to a belief function in the sense of Shafer (see [28]) Given a random set (F, m), a belief function Bel and plausibility function Pl, can be defined as

$$Bel(A) = \sum \{m(B), B \subseteq A, \forall A\}. \tag{32}$$

$$Pl(A) = 1 - Bel(A^C), \forall A \subseteq U. \tag{33}$$

A fuzzy interval generates a random set as follows.

Example 5.2. *Given a fuzzy interval, say the fuzzy number 2 given by a triangular fuzzy interval whose representation is 1/2/3. Discretize the fuzzy interval 2 into alpha levels of, for example, 1/4 units apart. Thus,* $\alpha_0, \alpha_1, \alpha_2, \alpha_3.\alpha_4 = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$ *respectively. Now, define* m*, the basic probability assignment as follows.*

$$\begin{aligned}
m([1,3]) &= 0\\
m([\frac{5}{4}, \frac{11}{4}]) &= \alpha_1 - \alpha_0 = \frac{1}{4} - 0 = \frac{1}{4}\\
m([\frac{3}{2}, \frac{5}{2}]) &= \alpha_2 - \alpha_1 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}\\
m([\frac{7}{4}, \frac{9}{4}]) &= \alpha_3 - \alpha_2 = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}\\
m([2,2]) &= \alpha_4 - \alpha_3 = 1 - \frac{3}{4} = \frac{1}{4}.
\end{aligned}$$

In general, for a fuzzy interval $A \subset \mathbb{R}$ whose alpha levels are partitioned in N levels of increasing values from 0 to 1, define

$$\mu_A([a_k, b_k]) = \alpha_k,$$
$$\mu_A([a_0, b_0]) = \alpha_0 = 0,$$
$$\mu_A([a_N, b_N]) = \alpha_N = 1,$$

and

$$[a_{k-1}, b_{k-1}] \subseteq [a_k, b_k].$$

Then, the basic probability assignment function is

$$m([a_k, b_k]) = \alpha_k - \alpha_{k-1}.$$

Note that we have nested sets and normalization. In this case we also have a possibility distribution and a necessity distribution.

5.3. **P-Boxes.** Another generalization associated with probability is what is called a probability box (see [15]).

Definition 5.3. A set of cumulative distribution functions P, $\mathcal{P} = \{P \mid F_* \leq P \leq F^*\}$, induced by two bounding cumulative distributions, F_* and F^*, is called a **probability box** (P-Box). A P-Box is a special random interval with focal sets E_α whose upper and lower bounds induce the same ordering. That is, given $F^*(a) = \alpha, F_*(b) = \alpha, E_\alpha = [a, b]$,

$$P^*([a,b]) = F^*(b) - F_*(a), P_*([a,b]) = \max\{0, F_*(b) - F^*(a)\}.$$

Remark 5.4. Consider fuzzy interval, A, whose membership function is $\mu_A(x)$ induces a P-Box as follows. Let

$$F^*(a) = \Pi_m((-\infty, a]) = \left\{ \begin{array}{c} \mu_A(a) \text{ if } a \leq \inf Core_A \\ 1 \text{ otherwise} \end{array} \right\},$$

and

$$F_*(a) = Nec_m((-\infty, a]) = \left\{ \begin{array}{c} 0 \text{ if } a \leq \sup Core_A \\ 1 - \lim_{x \downarrow a} \mu_A(x) \text{ otherwise} \end{array} \right\}.$$

The P-Box is the set $\{F \mid F_* \leq F \leq F^*\}$.

6. An Unified Theory of Uncertainty - Generalized Uncertainty: Interval-Valued Probability

Definition 6.1. Generalized uncertainty theory, for this presentation, is a mathematical theory of incompleteness or lack of information, or lack of specificity, or imprecision, whose representation is given by a set of functions that are between two a-priori given bounding functions.

Note that generalized uncertainty includes possibility. A different approach to generalized uncertainty than probability, possibility, Dempster/Shafer theory of evidence, random sets, and P-Boxes, is Interval-Valued Probability Measures. IVPMs are general in that it includes these five types of generalized uncertainties as well as other generalizations (see [22]). Many situations possess insufficient information to accurately construct an underlying unique probability measure as is the case for possibility, Dempster/Shafer, random sets, or P-Boxes. The IVPM approach utilizes subsets of the interval [0,1], which is denoted $Int_{[0,1]} = \{[a,b] \mid 0 \leq a \leq b \leq 1\}$. There is a similar yet distinct theory to IVPMs, interval probabilities (see [30]), which we note but do not discuss except to say that it is a type of IVPM.

Definition 6.2. [16]Given measurable space (X, σ_X), then $i_m : \sigma_X \to Int_{[0,1]}$ is called an **interval-valued probability measure (IVPM)** on (X, σ_X) if it satisfies the following:
1) $i_m(\phi) = [0,0]$
2) $i_m(X) = [1,1]$
3) $\forall A \in \sigma_X,\ i_m(A) = \left[A^l, A^u\right] \subseteq [0,1]$
4) for every partition of X, $\{A_{k\in K}\}$, $\{B_{j\in J}\} \subseteq \sigma_X$ such that $A = \cup_{k\in K} A_k$ and $A^c = \cup_{j\in J} B_j$ then

$$i_m(A) \subseteq \left[\begin{array}{c} \max\left\{1 - \Sigma_{j\in J} B_j^u, \Sigma_{k\in K} A_k^l\right\}, \\ \min\left\{1 - \Sigma_{j\in J} B_j^l, \Sigma_{k\in K} A_k^u\right\} \end{array} \right]$$

We call (X, σ_X, i_m) an **interval-valued probability measure space**.

Remark 6.3. To see the motivation behind the above definition of an IVPM suppose we have an interval valued measure space $(\mathcal{R}, \mathcal{B}, i_m)$, which provides a model for an unknown random variable X such that $\forall A \in \mathcal{B}$ we know that $\Pr(X \in A) \in i_m(A)$. Let A,B,C,D be mutually disjoint with union all of $\mathcal{R}$ (i.e. they are a partition of $\mathcal{R}$). Consider $\Pr(X \in A \cup B)$. Since these sets are disjoint, the maximum this probability can be is $A^u + B^u$, the maximum probabilities of A and B separately. Similarly, the minimum is $A^l + B^l$. Combined we have

$$\Pr(X \in A \cup B) \in \left[A^l + B^l, A^u + B^u\right]$$

Similarly we have

$$\Pr(X \in C \cup D) \in \left[C^l + D^l, C^u + D^u\right]$$

But since $C \cup D = (A \cup B)^c$ and $\Pr(X \in A \cup B) = 1 - \Pr(X \in (A \cup B)^c)$ we know that

$$\Pr(X \in A \cup B) \in \left[1 - (C^u + D^u), 1 - \left(C^l + D^l\right)\right]$$

Combining these two bounds on $\Pr(X \in A \cup B)$ gives

$$\Pr(X \in A \cup B) \in \left[\begin{array}{c} \max\left\{1 - (C^u + D^u), A^l + B^l\right\}, \\ \min\left\{1 - \left(C^l + D^l\right), A^u + B^u\right\} \end{array} \right], \tag{34}$$

which is consistent with our definition for an IVPM. Thus, for any $A \in \mathcal{B}$ the interval $i_m(A)$ has the smallest width possible for the given information (data) at hand.

Example 6.4. *Let $X \subset \mathbb{R}$ with partition A, B, C, D (mutually disjoint whose union is X) and $i_m(A) = [0.2, 0.5], i_m(B) = [0.1, 0.15].i_m(C) = [0.15, 0.2]$, $i_m(D) = [0.3, 0.4]$.*

$$\begin{aligned} i_m(A) &\subseteq [\max\{(1-(B^u + C^u + D^u)), A_l\}, \min\{(1-(B^l + C^l + D^l), A^u\} \\ &= [\max\{(1-(0.15+0.2+0.4)), 0.2\}, \min\{(1-(0.1+0.15+0.3)), 0.5\}] \\ &= [\max\{0.25, 0.2\}, \min\{0.45, 0.5\}] = [0.25, 0.45]. \end{aligned}$$

$$\begin{aligned} i_m(A \cup B) &\in \left[A^l + B^l, A^u + B^u\right] = [0.2 + 0.1, 0.5 + 0.15] \\ &= [0.3, 6.5] \end{aligned}$$

However, using (34)

$$\begin{aligned} i_m(A \cup B) &\in \left[\begin{array}{c} \max\left\{1 - (C^u + D^u) A^l + B^l\right\}, \\ \min\left\{1 - \left(C^l + D^l\right), A^u + B^u\right\} \end{array} \right] \\ &= \left[\begin{array}{c} \max\{1 - (0.2 + 0.4), 0.2 + 0.1\}, \\ \min\{1 - (0.15 + 0.4), 0.5 + 0.15\} \end{array} \right] \\ &= [0.4, 0.45] \end{aligned}$$

That is, the bounds on $i_m(A)$ and $i_m(A \cup B)$ were reduced. The uncertainty in A and $A \cup B$ are related to the uncertainty of their respective compliments, which means that the uncertainty on collected interval data can be used to reduce the uncertainties in the data, especially when the sets form a partition. Therefore, preprocessing to obtain the smallest width on the interval probability bounds possible associated with the given uncertainty data should be implemented prior to any mathematical analysis.

IVPMs include at least 8 uncertainty types (see [22]). This means that IVPMs are a useful way to look at uncertainty.

7. Probability and Possibility - Relationships and Differences

Possibility theory (as well as generalized uncertainty, and IVPMs) is an uncertainty theory devoted to handling of incomplete information. It is similar to probability theory because it is based on set-functions. It differs by the use of a pair of dual set-functions (possibility and necessity measures) instead of only one. Possibility is not additive and makes sense on ordinal structures.

7.1. **Upper Probabilities as Possibilities.** It has been established that a possibility measure is a special case of upper probability (see [9]), or in Shafer terminology, a plausibility function. In particular, let π be a possibility distribution where $\pi(s) \in [0, 1]$ and Π the corresponding possibility measure. Next, let $\mathbf{P}_\pi$ be the set of probability measures P such that

$$P \leq \Pi,$$

that is,

$$\forall A \subseteq X, P(A) \leq \Pi(A).$$

Then the possibility measure Π coincides with the upper probability function P^* such that

$$P^*(A) = \sup\{P(A), P \in \boldsymbol{P}_\pi\} \tag{35}$$

while the necessity measure Nec is the lower probability function P_* such that

$$P_*(A) = \inf\{P(A), P \in \boldsymbol{P}_\pi\}.$$

Thus, we can go from possibility/necessity to a bounded set of probabilities. We note that this is similar to interval probabilities (see [30]).

On the other hand, we can go from a family of probabilities where we assume that this family is bounded above and below by $\overline{P}(A), \underline{P}(A)$ respectively to possibility and necessity measures. That is, given a set of probabilities, $\{P(A) \mid \forall A \subseteq X, \underline{P}(A) \leq P(A) \leq \overline{P}(A)\}$, a possibility and necessity is obtained by setting

$$\Pi(A) = \overline{P}(A),$$
$$Nec(A) = \underline{P}(A).$$

7.2. **Belief/Plausibility and Necessity/Possibility.** Recall that when the set of focal elements, within the theory of evidence [28], F, is a nested family

$$A_1 \subset A_2 \subset ... \subset A,$$

then Belief and Plausibility satisfy the decomposability properties (see Shafer [28]),

$$Bel(A \cap B) = \min\{Bel(A), Bel(B)\}$$
$$Pl(A \cup B) = \max\{Pl(A), Pl(B)\}.$$

Plausibility in this case is a possibility measure in the sense of Zadeh ([31]) and belief is its dual necessity in the sense of Dubois/Prade ([8]). That is, from the generalization of probability theory, we can obtain possibility and necessity measures as long as the underlying set of focal elements are nested. And from possibility and necessity, we can obtain belief and plausibility functions over these nested sets.

7.3. **Random Sets and Possibility.** Suppose a random set, (F, m) is given, where F is a family of nested sets. In this case, the random sets can be translated into belief/plausibility functions using Equations (32), and (33). In turn, the belief and plausibility functions are necessity and possibility functions given that we have nested sets.

8. Conclusion

A way to understand possibility theory, probability theory, and their relationship was presented. It is noted that probability and possibility coincide when for the case of the Dirac delta function. A unifying theory of generalized uncertainty, IVPM, was proposed as a way to represent possibility theory and probability theory. The advantage of unification is that the associated representation and its resulting properties are given by one approach. Moreover, for problems in which the various types of uncertainties occur in the input data (parameters), the overarching theory (IVPM) can be used as a consistent representation and thus mathematical analysis becomes possible.

References

[1] Archimedes of Siracusa, "On the measurement of the circle," In: Thomas L. Heath (editor), *The Works of Archimedes*, Cambridge University Press, Cambridge, 1897; Dover edition, 1953.

[2] Dubois, D. (2005) "On the links between probability and possibility theories", Plenary talk at *IFSA 2005*, Beijing, China.

[3] Dubois, D. (2010), "The role of fuzzy sets in decision sciences: old techniques and new directions," *Fuzzy Sets and Systems*, **184:**1, pp. 3–28.

[4] Dubois, D. (2014). "Uncertainty Theories: A Unified View", Plenary talk at *FLINS and CBSF III 2014*, João Pessoa, Brazil.

[5] Dubois, D., Kerre, E., Mesiar,R., and Prade, H. (2000). Chapter 10, "Fuzzy Interval Analysis", in Didier Dubois and Henri Prade, (editors) *Fundamentals of Fuzzy Sets*, Kluwer Academic Publishers, Boston.

[6] Dubois, D., Nguyen, H. T., and Prade, H. (2000). Chapter 7, "Possibility Theory, Probability and Fuzzy Sets: Misunderstandings, Bridges and Gaps", in Didier Dubois and Henri Prade, (editors) *Fundamentals of Fuzzy Sets*, Kluwer Academic Publishers, Boston.

[7] Dubois, D. and Prade, H. (1982). "On several representations of an uncertainty body of evidence", in M. M. Gupta and E. Sanchez, editors. *Fuzzy Information and Decision Processes*, North-Holland, Amsterdam, pp. 167-181.

[8] Dubois, D. and Prade, H. (1988). *Possibility Theory,* Plenum Press, New York.

[9] Dubois, D. and Prade, H. (1992). "When upper probabilities are possibility measures", *Fuzzy Sets and Systems,***49**, pp. 65-74.

[10] Dubois, D. and Prade H., editors (2000). *Fundamentals of Fuzzy Sets.* Kluwer Academic Press.

[11] Dubois, D. and Prade, H. (2009), "Formal representations of uncertainty," Chapter 3 in. D. Bouyssou, D. Dubois, H. Prade, Editors, *Decision-Making Process,* ISTE, London, UK & Wiley, Hoboken, N.J., USA.

[12] Dubois, D. and Prade, H. (2011). *Mathware&Soft Computing Magazine* **18**:1, December 2011, pp. 18-71.

[13] Dubois, D., and Prade, H. (2012). "Gradualness, uncertainty and bipolarity: Making sense of fuzzy sets." *Fuzzy Sets and Systems*, **192**, pp. 3-24.

[14] Dubois, D. and Prade, H. (2015). "Possibility theory and its applications: Where do we stand?" in Kacprzyk, J. and Pedrycz, W., (editors) *Handbook of Computational Intelligence*, Springer, pp. 31-60.

[15] Ferson, S., Kreinovich, V., Ginzburg, R., Sentz, K., and Myers, D. S., (2003). *Constructing Probability Boxes and Dempster-Shafer Structures.* Sandia National Laboratories, Technical Report SAND2002-4015, Albuquerque, New Mexico.

[16] Jamison, K. D., Lodwick, W. A. (2004), "Interval-Valued Probability Measures," *UCD/CCM Report No. 213*, March 2004.

[17] Jamison, K. D. and Lodwick, W. A. (2020) "A new approach to interval-valued probability measures, A formal method for consolidating the languages of information deficiency: Foundations", *Information Sciences*, **507**, pp. 86-107.

[18] Kaufmann, A. and Gupta, M. M. (1985) . *Introduction to Fuzzy Arithmetic -Theory and Applications*, Van Nostrand Reinhold.

[19] Lodwick, W. A. and Bachman, K., (2005), "Solving large scale fuzzy possibilistic optimization problems", *Fuzzy Optimization and Decision Making*, **4**:4, pp. 257-278.

[20] Lodwick, W. A., and Jenkins, O. (2013), "Constrained intervals and interval spaces," *Soft Computing*, **17**: 8, pp. 1393-1402.

[21] Lodwick, W. A., Neumaier, A. and Newman, F. D., "Optimization under uncertainty: Methods and applications in radiation therapy," *Proceedings 10th IEEE International Conference on Fuzzy Systems 2001*, **3**, pp. 1219-1222.

[22] Lodwick, W. A. and Sales Neto, L. (2021). *Flexible and Generalized Uncertainty Optimization.* Springer-Verlag.

[23] Klir, G. J., (1989). "Is there more to uncertainty than some probability theorists might have us believe?" *International Journal of General Systems*, **15**, pp. 347-378.

[24] Klir, G. J., (1995). "Principles of uncertainty: What are they? Why do we need them?" *Fuzzy Sets and Systems*, **74**, pp.15-31.

[25] Klir, G. J., and Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications.* Prentice Hall, Upper Saddle River, New Jersey.

[26] Moore, R. E., Kearfott R. B. , and Cloud, M. J. (2009). *Introduction to Interval Analysis.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

[27] Puri, M. L., and Ralescu, D. (1982). "A possibility measure is not a fuzzy measure", *Fuzzy Sets and Systems,* **7,** pp. 311-313.

[28] Shafer G., (1976). *A Mathematical Theory of Evidence.* Princeton University Press. Princeton N.J.

[29] Sugeno M., (1974). "Theory of fuzzy integrals and its applications," Ph.D. Thesis, Tokyo Institute of Technology.

[30] Weichselberger, K. (2000), "The theory of interval-probability as a unifying concept for uncertainty," *International Journal of Approximate Reasoning* **24,** pp. 149-170.

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

Weldon A. Lodwick
Orcid number: 0000-0003-0606-5916
Department of Mathematical and Statistical Sciences
University of Colorado Denver
Denver, Colorado, USA
E-mail address: `Weldon.Lodwick@ucdenver.edu`