*International*
*Journal*
*Of Advanced Research in Engineering & Management (IJAREM)*
*ISSN: 2456-2033 || PP. 12-15*

# Uncertain Data Refinement using Evidence and Possibility Theories

Madhavi Kolukuluri[1], and Dr.B. D. C. N. Prasad[2]
[1]*Research scholar, Rayalaseema University, Kurnool, India*
[2]*PSCMR College of Engineering& Technology, Vijayawada, India*

**Abstract:** The uncertainty involved in any problem solving situation is a result of some information deficiency. Information may be incomplete, inprecise, not fully reliable, and contradictory. In general, the lack of information may result in different types of uncertainty. In this paper we discuss the statistical approach for representing uncertain data using evidence theory and possibility theory and applications to the evidence and possibility approximation problems are briefly presented.

**Keywords:** Uncertainty, Fuzzy measures, evidence theory, belief measure, plausibility measure, possibility measure, necessity measure, Dumpster's combination.

## I. INTRODUCTION

In the last decades a number of different uncertainty measures have been proposed. Evidence and possibility theory are popular uncertainty methods [2,4]. Possibility is a subset of evidence theory. Evidence theory is based on 2 dual measures: Belief measures and Plausibility measures. While possibility theory is based on 2 dual measures: Necessity measures and Possibility measures, which are special versions of belief and plausibility measures. 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.

Bel is called belief or lower probability function. Pl is called Plausibility or upper probability function. Belief is what probability will be. Plausibility is what probability could be.
Belief/Plausibility both reduce to probability if evidence is specific.

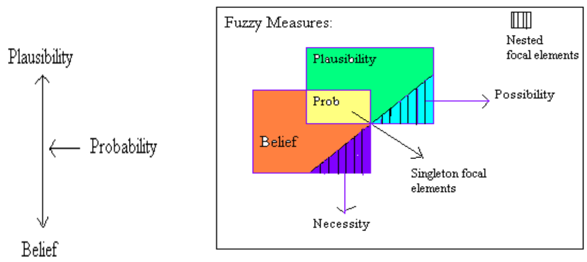


Figure 1.Fuzzy Measures.

## II. EVIDENCE THEORY

The theory of evidence [4] was introduced by Glenn Shafer in the 1970's as a reinterpretation of Dempster's [1] statistical inference. This is one mathematical tool for dealing with probabilities. Let us first review the basic notions of the theory of evidence.

Definition: A Basic Probability Assignment (BPA) over a finite set ρ(X) is a function m: ρ(X)→[0,1] Such that
i) $m(\Phi)=0$
ii) $\sum_{A\epsilon\rho(x)} m(A) = 1$
iii) $m(A)\geq 0$ for all $A\epsilon\ \rho(x)$ , Subsets of ρ(X) associated with non-zero values of m are called focal elements.
Evidence theory is based on 2 dual non additive measures: Belief measures and Plausibility measures.

**A. Belief measures :**

Given a universal set X, assumed here to be finite, a belief measure is a function Bel : ρ(x)⟶[0,1] Such that

i) Bel (Φ)= 0,

ii) Bel(X)=1

iii) $\mathrm{Bel}(A1 \cup A2 \cup \ldots \cup An) \geq \sum_j \mathrm{Bel}(Aj) - \sum_{j<k} \mathrm{Bel}(Aj \cap AK) + \cdots + (-1)^{n+1} \mathrm{Bel}(A1 \cap A2 \cap \ldots \cap An)$

For all possible families of subsets of X.

Belief measures and also called super additive.

The fundamental property of belief measures:

$$\mathrm{Bel}(A) + \mathrm{Bel}(\bar{A}) <= 1$$

Associated with each belief measure is a plausibility measure, Pl, defined by Pl(A)=1-Bel(Ā),for all Aϵρ(x)

$$\mathrm{Bel}(A) = 1 - \mathrm{Pl}(\bar{A})$$

From this we can say that Belief measures and plausibility measures are mutually dual.

**B. Plausibility measures:**

A plausibility measures is a function Pl: ρ(x)⟶[0,1] such that

I) Pl(Φ)=0,

II) ii)Pl(X)=1 and

III) $\mathrm{Pl}(A1 \cap A2 \cap \ldots \cap An) \leq \sum_j Pl(Aj) - \sum_{j<k} Pl(Aj \cup Aj) + \cdots + (-1)^{n+1} Pl(A1 \cup A2 \cup \ldots \cup An)$

For all possible families of subset of X.

Plausibility measures are also called Sub additive.

The fundamental property of plausibility measures:

$$\mathrm{Pl}(A) + \mathrm{Pl}(\bar{A}) \geq 1$$

Belief and plausibility measures have the following properties for all A,Bϵρ(X):

(i) Bel(A⋂B)=min[(Bel(A),Bel(B)]

(ii) Pl(A⋃B)=max[(Pl(A),Pl(B)]

Given a basic assignment m, a belief measures and a plausibility measure are uniquely determined ∀ set Aϵρ(X) by the formula

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

i.e.,Belief is the sum of the masses of all subsets of the hypothesis.

$$Pl(A) = \sum_{B/A \cap B \neq \Phi} m(B)$$

Plausibility is one minus the sum of the masses of all sets whose intersection with the hypothesis is empty. The inverse procedure is also possible. Given a belief measure Bel, the corresponding basic probability assignment m is determined

For all Aϵρ(X) by the formula

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

Evidence obtained in the same context from 2 independent sources (ex. form two experts in the field of enquiry) and expressed by 2 basic assignments $m_1$ & $m_2$ on the power set ρ (X) must be appropriately combined to obtain a joint basic assignment $m_{1,2}$ .

**C. Dempster's rule of combination:**

According to this rule, the degree of evidence $m_1(B)$ from the first source that focuses on set BЄρ(X) & the degree of evidence $m_2$ (C) from the second source the focuses on set CЄρ (X) are combined by taking the product $m_1(B) .m_2(C)$ which focuses on the intersection B∩C.

$$m1,2(A) = \frac{\sum_{B \cap C = A} m1(B).m2(C)}{1-K}$$

For all A≠Φ and $m_{1,2}(\Phi)=0$ where $K = \sum_{B \cap C = \Phi} m1(B).m2(C)$

*Example:* Let 3 sets represents different neurological symptoms about the brain damage of a patient as T= {Temporal lobe damage}, F= {Frontal lobe damage}, P= {Parietal lobe damage}.

There are 2 physicians opinions m1 and m2 regarding the neurological symptoms of the patient. The focal elements are given as m1={0.05,0,0.05,0.15,0.1,0.05,0.6}
and m2={0.15,0,0.05,0.05,0.2,0.05,0.5}.

Applying Dempster's rule to $m_1$ and $m_2$, we obtain the joint basic assignment $m_{1,2}$. To determine the values of $m_{1,2}$ we calculate the normalization factor 1-K first.

K= m1(T).m2(F)+m1(T).m2(P)+m1(T).m2(F∪P)
+m1(D).m2(T)+m1(F).m2(P)+m1(F).m2(T∪P) +m1(P).m2(T)
+m1(P).m2(F)+m1(P).m2(T∪F)
+ m1 (T∪F).m2(P)+m1(T∪P).m2(F)+m1(F∪P).m2(T)
= 0.03.

The normalization factor is then 1-K=0.97. Values of $m_{1,2}$ are calculated as follows:

m1,2 =[m1(T).m2(T)+m1(T).m2(T∪F)+m1(T).m2(T∪P)
+m1(T).m2(T∪F∪P)+m1(T∪F).m2(T)
+ m1(T∪F).m2(T∪P) +m1 (T∪P).m2(T)
+m1 (T∪P).m2(T∪F)+m1(T∪F∪P).m2(T)]/.097
= 0.21.

And similarly the remaining focal elements can be calculated. The joint basic assignment can be used to calculate the joint belief and joint plausibility.

TABLE I
COMBINED EVIDENCE AND COMBINED PLAUSIBILITY

| **Focal elements** | **Expert-1** | | **Expert-2** | | **Combined Evidence** | | $P_{1,2}$ |
|---|---|---|---|---|---|---|---|
| | **m1** | **Bel1** | **m2** | **Bel2** | **m1,2** | **BEL1,2** | |
| T | 0.05 | 0.05 | 0.15 | 0.15 | 0.21 | 0.21 | 0.84 |
| F | 0 | 0 | 0 | 0 | 0.01 | 0.01 | 0.5 |
| P | 0.05 | 0.05 | 0.05 | 0.05 | 0.09 | 0.09 | 0.66 |
| TUF | 0.15 | 0.2 | 0.05 | 0.2 | 0.12 | 0.34 | 0.91 |
| TUP | 0.1 | 0.2 | 0.2 | 0.4 | 0.2 | 0.5 | 0.99 |
| FUP | 0.05 | 0.1 | 0.05 | 0.1 | 0.06 | 0.16 | 0.79 |
| TUFUP | 0.6 | 1 | 0.5 | 1 | 0.31 | 1 | 1 |

## III. POSSIBILITY THEORY

Possibility theory represents a non classical uncertainty theory, first introduced by Zadeh [7] and then developed by several authors ( eg. Dubois and Prade [2]).Possibility theory deals with bodies of evidence whose focal elements are nested (A1⊆A2⊆A3⊆…….⊆An ) i.e., focal elements can be ordered so that each one is contained in the following one. Nested focal elements are also called consonants. Similarly a belief function whose focal elements are nested is said to be consonant. Possibility theory is closely connected with fuzzy set theory and plays an important role in some of its application.

Special counter parts of belief measures and plausibility measures in possibility theory are called Necessity measures and Possibility measures.

The basic equations of possibility theory :

$$Nec(A\cap B)=\min[(Nec(A),Nec(B)]$$
$$Pos(A\cup B)=\max[(Pos(A),Pos(B)]$$

### A. Necessity measures:

*Definition1:* Let Nec denote a fuzzy measures on (X, ρ) then Nec is called a necessity measures if

$$\boldsymbol{Nec}(\cap_{k\in K}(\boldsymbol{A_k})) = \boldsymbol{inf}_{k\in K}\boldsymbol{Nec}(\boldsymbol{A_k})$$

for any family $\{A_k/k\in K\}$ in ρ such that $\cap_{k\in K} \boldsymbol{Ak\in\rho}$ where k is an arbitrary index set.

*Definition2*: Let pos denote a fuzzy measures a (X, ρ) then pos is called a possibility measures iff

$$\boldsymbol{Pos}\Big(\bigcup_{k\in K}(\boldsymbol{A_k})\Big) = \boldsymbol{Sup_{k\in K}Pos}(\boldsymbol{A_k})$$

for any family $\{A_k/k \in K\}$ in ρ such that $\bigcup_{k\in K} \boldsymbol{A_k} \in \boldsymbol{\rho}$ where k is an arbitrary index set.

Since necessity measures are special belief measures and possibility measures are special plausibility measures we have

$Nec(A)+Nec(\bar{A})\leq 1$

$Pos(A)+Pos(\bar{A})\geq 1$

$Nec(A)=1-Pos(\bar{A})$

$Min[Nec(A),Nec(\bar{A})]=0$

$Max[Pos(A),Pos(\bar{A})]=1$

Given a possibility measure Pos on ρ(X), let function r: X→ [0, 1]. Such that r(x)=Pos({x}) for all x Є X be called a possibility distribution function associated with Pos. An important property of possibility theory is that every possibility measure is uniquely represented by the associated possibility distributed function $\boldsymbol{r_i \sum_i^n m_k}$ .

*Example*: Let the temperature measured in centigrade's and assume that only its integer values are recognized. Let information about the actual value of the temperature be given in terms of the proposition "Temperature is around 310C"

TABLE III

POSSIBILITY AND NECESSITY MEASURES

| **Focal Elements** | **m(A)** | **Possibility disribution** | **Pos(A)** | **Nec(A)** | **C(A)** |
|---|---|---|---|---|---|
| A1={31} | 0.1 | 1 | 1 | 0.1 | 0.1 |
| A2={30, 31,32} | 0.4 | 0.9 | 1 | 0.4 | 0.4 |
| A3={29, 30,31,32, 33} | 0.5 | 0.5 | 1 | 0.5 | 0.5 |

As explained that possibility theory is based on 2 dual functions: necessity measures, Nec, and possibility measures, Pos. the 2 functions whose range is [0,1], can be converted to a single combined function C, whose range is [-1,1].for each A Є ρ(X),

C (A) = Nec (A) + Pos (A) - 1.

## IV. CONCLUSION

In this paper we discussed representation of uncertain data using Evidence theory and Possibility theory. Here a statistical approach is used to representing uncertain data. Currently we are working to prove that we can also represent uncertain data using Utility theory and Value Based theory.

## REFERENCES

[1] Dempster, A. P. [1967] "Upper and lower probabilities induced by a multivalued mapping". Annals Mathematical Statistics.

[2] Dubois, D., Prade, H.: Possibility Theory: An approach to Computerized Processing of Uncertainty. Plenium Press, New York(1988).

[3] George J.Klir/Bo Yuan "Fuzzy Sets and Fuzzy Lozic Theory and Applications.

[4] Glenn Shafer, A Mathematical Theory Of Evidence, Princeton University Press,1976.

[5] Klir, G. J.[1993]," Developments in uncertainty-based information."In:Yovits, M.C., ed., Advances in computers. Academic press, San Diego.

[6] Wang, P. P., ed.[1983], Advances in Fuzzy Theory and Technology, Vol.I. Bookwrights Press, Durham,NC.

[7] Zadeh, L.A.," Fuzzy sets as a basis for a theory of possibility." Fuzzy Sets and Systems.

[8] A. P. Dempster, Upper and lower probability inferences based on a sample from a finite univariate population, Biometrika 54 (1967), 515–528.