ISSN 2007-9737

Fuzzy Distribution Sets

Ildar Z. Batyrshin

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Mexico

batyr1@gmail.com

Abstract. We introduce a new type of fuzzy set called Fuzzy Distribution Set (FDS). Fuzzy distribution sets are fuzzy sets defined on a finite domain subject to a sum of membership values equal to 1. Such fuzzy sets can serve as models of subjective probability distributions and subjective weight distributions. Considering these distributions as fuzzy sets gives a possibility to extend on such distributions the operations of fuzzy sets and, more generally, the calculus of fuzzy restrictions developed during the last decades. Recently Yager introduced the concept of negation of probability distributions. In our works, we studied several classes of such negations. Here we consider an involutive negation of probability distributions as a complement of FDS. We introduce the operations of union and intersection of fuzzy distribution sets. These basic operations on FDS can serve as a basis for the application of fuzzy logic methods to subjective probability and weight distributions. These operations can be used for the development of reasoning models with subjective probability distributions and subjective weighting functions. Weight distributions can be used in multi-criteria, multi-person, and multi-attribute decision-making models.

Keywords. Fuzzy set, probability distribution, weight distribution, complement of distributions, union of distributions, intersection of distributions.

1 Introduction

Recently R. Yager introduced an operation of the negation of probability distributions [1]. Such negation operation can be used for formalizing sentences like "NOT High Price," where "High Price" denotes a probability distribution defined on a set of prices. Dozen of papers used Yager's negation for modeling uncertain information related to probability distributions (pd) [4].

Also, some new negations of pd have been proposed [2-4]. Since probability distributions can be considered as some kind of probabilistic predicates, the natural question appears: How to define the operations of disjunction and conjunction of probability distributions? The disjunction operation could give a possibility to formalize sentences like "High OR Very High Price." On the other hand, one can use the conjunction operations for formalizing sentences like "Which of my friends called my father AND promised to come today?". Here, subjective probability distributions "called" and " promised to come today " are defined on the set of my friends.

We construct new logical operations on the set of probability distributions as extensions of operations used in fuzzy set theory [5-12]. A probability distribution is considered a special type of fuzzy set called Fuzzy Distribution Set (FDS). We define operations AND, OR, and NOT on the set of probability distributions as operations of intersection, union, and complement of these fuzzy sets. The involutive negation of probability distributions proposed in [3] is used as a complement of fuzzy distribution sets. The union and intersection operations of FDS are constructed as extensions of similar operations defined on the set of fuzzy sets. Formally, as a fuzzy distribution set, one can consider a weight distribution used in multi-criteria, multi-person, and multi-attribute decision-making. For this reason, fuzzy distribution sets can also serve as models of subjective weight distributions.

In such an interpretation, the fuzzy set-theoretic operations considered in this paper can also be used as operations on the set of subjective weight distributions.

The paper has the following structure. Section 2 gives a definition of a fuzzy distribution set and defines a complement of such fuzzy sets as an involutive negation of probability distributions. In Section 3, we introduce the union operations of FDS. Sections 4 and 5 propose the methods of construction of the intersection of FDS. Section 6 contains an example of the intersection and union of subjective probability distributions. Section 7 contains the conclusion.

2 Complement of Fuzzy Distribution Sets

Let $X = \{x_1, \dots, x_n\}$ be a finite non-empty set, $(n > 1)$.

Definition 1. A *fuzzy distribution set* (FDS) defined on X is a function $d: X \to [0,1]$ subject to:

$$\sum_{i=1}^{n} d(x_i) = 1. \quad (1)$$

Denote $d_i = d(x_i)$, $i = 1, \dots, n$. From (1), we have:

$$\sum_{i=1}^{n} d_i = 1. \quad (2)$$

The set of *membership values* $D = \{d_1, \dots, d_n\}$ will be called a *distribution*. When several distributions considered in some problem are defined on the same domain $X = \{x_1, \dots, x_n\}$ it is convenient to represent a distribution $D = \{d_1, \dots, d_n\}$ as n-tuple $D = (d_1, \dots, d_n)$. If it is not confusing, we will use both of these representations of distributions defined on a domain X. In such notations, the distribution $D = (0.8, 0.2, 0, \dots, 0)$ will denote the distribution D with: $d_1 = 0.8$, $d_2 = 0.2$, and $d_i = 0$ for $i = 3, \dots, n$.

The distribution $D_{(i)} = (d_1, \dots, d_n)$ satisfying the property: $d_i = 1$ for some $i = 1, \dots, n$, and $d_j = 0$ for all $j \neq i$, will be referred to as a *degenerate* or *point distribution*. For example, for $i = 1$ and $i = n$ we have the following point distributions: $D_{(1)} = (1,0, \dots,0)$, and $D_{(n)} = (0, \dots,0,1)$.

The simplest example of distribution is the *uniform distribution*:

$$D_U = \left(\frac{1}{n}, \dots, \frac{1}{n}\right).$$

Let $\mathcal{D}_n$ be the set of all fuzzy distribution sets defined on the set $X = \{x_1, \dots, x_n\}$.

Definition 2. A *complement* of a fuzzy distribution set is a function $com: \mathcal{D}_n \to \mathcal{D}_n$ such that for any fuzzy distribution set $D = (d_1, \dots, d_n)$ in $\mathcal{D}_n$ the distribution $C = com(D) = (c_1, \dots, c_n)$ satisfies for all $i, j = 1, \dots, n$, the following properties:

$$\text{if } d_i \leq d_j, \text{then } c_i \geq c_j.$$

From the definition of the complement, it follows for all $i, j = 1, \dots, n$:

$$0 \leq c_i \leq 1, \quad \sum_{i=1}^{n} c_i = 1,$$

$$\text{if } d_i = d_j, \text{then } c_i = c_j.$$

A *negator* N is a function of distribution values d_i taking values in [0,1] point-by-point transforming distribution $D = (d_1, \dots, d_n)$ into its complement:

$$com(D) = \big(N(d_1), \dots, N(d_n)\big).$$

Hence, for all $i, j = 1, \dots, n$, the following properties are satisfied [2]:

$$0 \leq N(d_i) \leq 1, \quad \sum_{i=1}^{n} N(d_i) = 1,$$

$$\text{if } d_i \leq d_j, \text{then } N(d_i) \geq N(d_j),$$

$$\text{if } d_i = d_j, \text{then } N(d_i) = N(d_j).$$

We will say that a negator N *generates* a complement $com(D) = \big(N(d_1), \dots, N(d_n)\big)$ of FDS D and distribution $D = (d_1, \dots, d_n)$.

A negator N is called a *distribution-independent* [2] if for any distribution $D = (d_1, \dots, d_n)$ in $\mathcal{D}_n$ the negator $N(d_i)$ depends only on the value d_i but not on other values d_j from D. A negator that is not distribution-independent will be referred to as *distribution-dependent*.

Consider examples of negators [2-4].

Yager negator is defined for all d in [0,1] as follows [1]:

$$N_Y(d) = \frac{1-d}{n-1}.$$

It is a distribution-independent negator. For any distribution $D = (d_1, \dots, d_n)$ in $\mathcal{D}_n$ it defines a complement of D as follows:

$$com_Y(D) = \left(\frac{1-d_1}{n-1}, \dots, \frac{1-d_n}{n-1}\right).$$

The *uniform negator* is defined for all d in [0,1] as follows [2]:

$$N_U(d) = \frac{1}{n}.$$

It is another example of a distribution-independent negator. For any distribution $D = (d_1, \dots, d_n)$ in $\mathcal{D}_n$ negator N_U defines its complement as follows:

$$com_U(D) = \left(\frac{1}{n}, \dots, \frac{1}{n}\right) = D_U.$$

The following negator N_B, introduced in [3], is defined for any distribution $D = (d_1, \dots, d_n)$ in $\mathcal{D}_n$ and all d_i, $i = 1, \dots, n$ as follows:

$$N_B(d_i) = \frac{\max(D)+\min(D)-d_i}{n(\max(D)+\min(D))-1} = \frac{MD-d_i}{nMD-1}. \quad (3)$$

where $\max(D) = \max\{d_1, \dots, d_n\}$, $\min(D) = \min\{d_1, \dots, d_n\}$ and $MD = \max(D) + \min(D)$. This negator is an example of a distribution-dependent negator. It depends not only on membership value d_i, but also on maximal and minimal values of the distribution D.

Generally, this distribution-dependent negator can be denoted, for example, as $N_B(D, d_i)$, but, for simplicity of notations, we denote it here as $N_B(d_i)$.

The complement com of fuzzy distribution sets is called *involutive*, if for all distributions $D = (d_1, \dots, d_n)$ in $\mathcal{D}_n$ it satisfies the following *involutivity property*:

$$com(com(D)) = D.$$

This property is fulfilled for the complement $\overline{A}$ of crisp, non-fuzzy sets: $\overline{\overline{A}} = A$, so it is also convenient to have involutive complements for fuzzy distribution sets. The complements of FDS based on Yager negator, uniform negator, and many other negators are non-involutive [2-4]. The complement of FDS based on the negator N_B (3):

$$com_B(D) = (N_B(d_1), \dots, N_B(d_n)).$$

is involutive, satisfying the property:

$$com_B(com_B(D)) = D.$$

3 Union of Fuzzy Distribution Sets

As in fuzzy sets theory, we can define the union of fuzzy distribution sets element-by-element using a disjunction operation (disjunctor) applied to elements of FDS. Let us use t-conorms [9] as such disjunctors for constructing the union of FDS.

T-conorm is a function $S: [0,1] \times [0,1] \to [0,1]$, such that for all a, b, c in [0,1] the following properties are satisfied [9]:

$S(a, b) = S(b, a)$ (commutativity),

$S(a, S(b, c)) = S(S(a, b), c)$ (associativity),

$S(a, b) \le S(a, c)$, whenever $b \le c$ (monotonicity),

$S(a, 0) = S(0, a) = a$ (boundary condition).

The following boundary condition follows from the properties of t-conorms:

$S(1, a) = S(a, 1) = 1.$

Here are the examples of the basic t-conorms:

$S_M(a, b) = \max(a, b)$ (maximum),

$S_P(a, b) = a + b - ab$ (probabilistic sum).

For all t-conorms S and all a, b in [0,1] it is fulfilled:

$$S_M(a, b) \le S(a, b).$$

Definition 3. A *union* of fuzzy distribution sets based on T-conorm S is a function $un: \mathcal{D}_n \times \mathcal{D}_n \to \mathcal{D}_n$, defined for all FDS $A = (a_1, \dots, a_n)$ and $B = (b_1, \dots, b_n)$ in $\mathcal{D}_n$ as follows:

$$un(A, B) = C = (c_1, \dots, c_n), \quad (4)$$

where

$$c_i = \frac{S(a_i, b_i)}{\sum_{i=1}^{n} S(a_i, b_i)}, i = 1, \dots, n. \quad (5)$$

It is easy to see that (4) is a distribution. From the boundary conditions and (2), it follows that the denominator of (5) is positive. From the definition of t-conorms for all $i = 1, \dots, n$, it follows:

$$0 \le c_i \le 1,$$

and from (5) we have:

$$\sum_{i=1}^{n} c_i = \sum_{i=1}^{n} \frac{S(a_i, b_i)}{\sum_{i=1}^{n} S(a_i, b_i)} = 1.$$

From the properties of t-conorms, it follows that for all A, B in $\mathcal{D}_n$ the property of commutativity is satisfied:

$$un(A, B) = un(B, A).$$

The function $DIS: [0,1] \times [0,1] \to [0,1]$ defined for any distributions $A = (a_1, \dots, a_n)$ and $B = (b_1, \dots, b_n)$ in $\mathcal{D}_n$ and all $i = 1, \dots, n$ by

ISSN 2007-9737

$$DIS(a_i, b_i) = \frac{S(a_i, b_i)}{\sum_{i=1}^{n} S(a_i, b_i)} \quad (6)$$

will be called a *disjunctor*. The following properties of disjunctor (6) follow from the properties of t-conorms for all $i, j = 1, \dots, n$:

Commutativity:

$$DIS(a_i, b_i) = DIS(b_i, a_i).$$

Monotonicity:

if $a_i \le a_j$ and $b_i \le b_j$, then $DIS(a_i, b_i) \le DIS(a_j, b_j)$.

For t-conorm $S_M(a, b) = \max(a, b)$, we obtain the following disjunctor:

$$DIS_M(a_i, b_i) = \frac{\max(a_i, b_i)}{\sum_{i=1}^{n} \max(a_i, b_i)}.$$

For t-conorm $S_P(a, b) = a + b - ab$, we obtain the following disjunctor:

$$DIS_P(a_i, b_i) = \frac{a_i + b_i - a_i b_i}{\sum_{i=1}^{n}(a_i + b_i - a_i b_i)} = \frac{a_i + b_i - a_i b_i}{2 - \sum_{i=1}^{n} a_i b_i}.$$

4 Intersection of Fuzzy Distribution Sets Using t-Norms

T-norm is a function $T: [0,1] \times [0,1] \to [0,1]$, such that for all a, b, c in [0,1] the following properties are satisfied [9]:

$T(a, b) = T(b, a)$ (commutativity),

$T(a, T(b, c)) = T(T(a, b), c)$ (associativity),

$T(a, b) \le T(a, c)$, whenever $b \le c$ (monotonicity),

$T(a, 1) = T(1, a) = a$ (boundary condition).

The following boundary condition follows from the properties of t-norms:

$T(0, a) = T(a, 0) = 0.$

Here are the examples of the basic t-norms:

$T_M(a, b) = \min(a, b)$ (minimum),

$T_P(a, b) = ab$ (product).

For all t-norms T and all a, b in [0,1], it is fulfilled:

$$T(a, b) \le T_M(a, b).$$

Definition 4. An intersection of fuzzy distribution sets based on t-norm T is a function $int: \mathcal{D}_n \times \mathcal{D}_n \to \mathcal{D}_n$, defined for all FDS $A = (a_1, \dots, a_n)$ and $B = (b_1, \dots, b_n)$ in $\mathcal{D}_n$, as follows:

$$int(A, B) = C = (c_1, \dots, c_n),$$

where, for all $i = 1, \dots, n$, c_i is defined by:

$$c_i = \frac{1}{n}, \text{ if } T(a_i, b_i) = 0 \text{ for all } i = 1, \dots, n,$$

otherwise:

$$c_i = \frac{T(a_i, b_i)}{\sum_{i=1}^{n} T(a_i, b_i)}. \quad (7)$$

The function $CON: [0,1] \times [0,1] \to [0,1]$ defined for any distributions $A = (a_1, \dots, a_n)$ and $B = (b_1, \dots, b_n)$ in $\mathcal{D}_n$ by

$$CON(a_i, b_i) = \begin{cases} \frac{1}{n}, if\ T(a_i, b_i) = 0\ for\ all\ i = 1, \dots, n, \\ \frac{T(a_i, b_i)}{\sum_{i=1}^{n} T(a_i, b_i)}\ otherwise, \end{cases}$$

will be called a *conjunctor*. The following properties of conjunctor follow from the properties of t-norms for all $i, j = 1, \dots, n$:

Commutativity:

$$CON(a_i, b_i) = CON(b_i, a_i).$$

Monotonicity:

if $a_i \le a_j$ and $b_i \le b_j$, then $CON(a_i, b_i) \le CON(a_j, b_j)$.

5 Intersection of Fuzzy Distribution Sets Using Union and Complement

From De Morgan law of crisp sets:

$$\overline{A \cap B} = \overline{A} \cup \overline{B},$$

and the involutivity of complement $\overline{\overline{A}} = A$ we have:

$$A \cap B = \overline{\overline{A} \cup \overline{B}}.$$

We will use this formula to define a new intersection of fuzzy distribution sets.

Definition 5. An intersection of fuzzy distribution sets is a function $int: \mathcal{D}_n \times \mathcal{D}_n \to \mathcal{D}_n$, defined for all FDS $A = (a_1, \dots, a_n)$ and $B = (b_1, \dots, b_n)$ in $\mathcal{D}_n$, as follows:

$$int(A, B) = com_B(un(com_B(A), com_B(B))),$$

such that in the distribution:

ISSN 2007-9737

$$int(A,B) = C = (c_1, \dots, c_n),$$

c_i is defined for all $i = 1, \dots, n$, by:

$$c_i = N_B\left(DIS\big(N_B(a_i), N_B(b_i)\big)\right).$$

Step-by-step, we obtain:

$$N_B(a_i) = \frac{\max(A)+\min(A)-a_i}{n(\max(A)+\min(A))-1},$$

$$N_B(b_i) = \frac{\max(B)+\min(B)-b_i}{n(\max(B)+\min(B))-1},$$

$$DIS\big(N_B(a_i), N_B(b_i)\big) = \frac{S(N_B(a_i), N_B(b_i))}{\sum_{i=1}^{n} S(N_B(a_i), N_B(b_i))}.$$

Denoting $e_i = DIS\big(N_B(a_i), N_B(b_i)\big)$ and $E = (e_1, \dots, e_n)$, finally we obtain:

$$c_i = N_B(e_i) = \frac{\max(E)+\min(E)-e_i}{n(\max(E)+\min(E))-1}.$$

6 Example

Suppose you with your wife want to evaluate how many neighbors will come to your open party. She supposes that you will have 5-6 guests, and you suppose to have 8-9 guests. Suppose you can represent these evaluations as subjective probability distributions defined on the set:

$$X = \{4, \mathbf{5}, \mathbf{6}, 7, \mathbf{8}, \mathbf{9}, 10\},$$

for your wife as:

$$W = (0.1, \mathbf{0.3}, \mathbf{0,3}, 0.2, 0.1, 0, 0)$$

and for you, as:

$$Y = (0, 0, 0, 0.1, \mathbf{0.4}, \mathbf{0}, \mathbf{4}, 0.1).$$

Using the formula (7) with minimum and product conjunctions, you obtain intersections of these distributions as possible solutions to your problem.

Using minimum t-norm T_M you obtain

$$int_M(W,Y) = (0, 0, 0, \mathbf{0.5}, \mathbf{0.5}, 0, 0),$$

i.e., $p(7) = p(8) = 0.5$, and $p(k) = 0$, for $k = 4{,}5{,}6{,}9{,}10$.

Using product t-norm T_P you obtain:

$$int_P(W,Y) = (0, 0, 0, \mathbf{0.33}, \mathbf{0.67}, 0, 0),$$

i.e., $p(7) = 0.33, p(8) = 0.67$, and $p(k) = 0$, for $k = 4{,}5{,}6{,}9{,}10$.

Compared with initial subjective evaluations of the number of guests: 5-6, and 8-9, the applied models give 7-8 as the more probable number of guests for both t-norms used in the model, but the product t-norm evaluates the number of 8 guests as the most probable.

Evaluation of the probable number of guests using union operation gives the following results for maximum S_M and probabilistic sum S_P t-conorms:

$$un_M(W,Y) = (0.056, \mathbf{0.167}, \mathbf{0.167}, 0.111, \mathbf{0.222}, \mathbf{0.222}, 0.056)$$

$$un_P(W,Y) = (0.052, \mathbf{0.155}, \mathbf{0.155}, 0.144, \mathbf{0.237}, \mathbf{0.206}, 0.052)$$

These distributions are defined on the domain: $X = \{4, \mathbf{5}, \mathbf{6}, 7, \mathbf{8}, \mathbf{9}, 10\}$.

All elements of this domain have non-zero probabilities with local maximums for 5-6 and 8-9 guests corresponding to the initial subjective distributions but considering 8-9 as more probable.

The product t-conorm refines the probability values giving the most probable value for 8 guests.

Similar methods can be used for aggregating subjectively defined weight distributions widely used in multi-criteria, multi-objective, multi-person, and multi-alternative decision-making.

7 Conclusion

The paper proposed a new type of fuzzy set that can be used to model subjective probability distributions and subjective weight distributions.

Fuzzy logic allows you to build models of subjective reasoning and human behavior It has numerous applications in control, pattern recognition, decision-making, machine learning, and other research areas. The paper paves a new way to apply the power of fuzzy logic for modeling and processing subjective probability distributions and subjective weight distributions.

In future works, we plan to extend the methods of processing fuzzy distribution sets by applying generalized fuzzy conjunction and disjunction operations and other fuzzy logic techniques (in the wide sense) [7-12].

Acknowledgments

This work was partially supported by the project SIP 20220857 of the Mexican National Polytechnic Institute. The author thanks Dr. Grigori Sidorov for his valuable comments.

References

1. **Yager, R. R. (2014).** On the maximum entropy negation of a probability distribution. IEEE Transactions on Fuzzy Systems, Vol. 23, No. 5, pp. 1899–1902. DOI: 10.1109/TFUZZ.2014. 2374211.
2. **Batyrshin, I., Villa-Vargas, L. A., Ramirez-Salinas, M. A., Salinas-Rosales, M., Kubysheva, N. (2021).** Generating negations of probability distributions. Soft Computing, Vol. 25, No. 12, pp. 7929–7935. DOI: 10.1007/s00500-021-05802-5.
3. **Batyrshin, I. Z. (2021).** Contracting and involutive negations of probability distributions. Mathematics, Vol. 9, No. 19, pp. 1–11, DOI: 10.3390/math 9192389.
4. **Batyrshin, I. Z., Kubysheva, N. I., Bayrasheva, V. R., Kosheleva, O., Kreinovich, V. (2021).** Negations of Probability Distributions: A Survey. Computación y Sistemas, Vol. 25, No. 4, pp. 775–781. DOI: 10.13053/cys-25-4-4094.
5. **Zadeh, L. A. (1965).** Fuzzy Sets. Information and Control, Vol. 8, pp. 338–353.
6. **Zadeh, L. A. (1975).** Calculus of fuzzy restrictions. In **Zadeh, L. A., Fu, K. S., Tanaka, K., Shimura, M., editors,** Fuzzy sets and their applications to cognitive and decision processes. Academic Press, pp. 1–39. DOI: 10.1016/B978-0-12-775260-0.50006-2.
7. **Klir, G. J., Yuan, B. (Eds.). (1996).** Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh. Vol. 6. World Scientific.
8. **Jang, J. S. R., Sun, C. T., Mizutani, E. (1997).** Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Prentice Hall.
9. **Klement, E. P., Mesiar, R., Pap, E. (2000).** Triangular norms. In Trends in Logic. Springer Dordrecht. DOI: 10.1007/978-94-015-9540-7.
10. **Nguyen, H. T., Walker, C., Walker, E. A. (2018).** A first course in fuzzy logic. Chapman and Hall/CRC, DOI: 10.1201/9780429505546.
11. **Batyrshin, I., Kaynak, O. (1999).** Parametric classes of generalized conjunction and disjunction operations for fuzzy modeling. IEEE Transactions on Fuzzy Systems, Vol. 7, No. 5, pp. 586–596. DOI: 10.1109/91.797981.
12. **Batyrshin, I., Kaynak, O., Rudas, I. (2002).** Fuzzy modeling based on generalized conjunction operations. IEEE Transactions on Fuzzy Systems, Vol. 10, No. 5, pp. 678–683. DOI: 10.1109/TFUZZ. 2002.803500.

Article received on 20/04/2022; accepted on 08/07/2022. Corresponding author is Ildar Z. Batyrshin.