Probability- Possibility Transformations:A Brief Revisit

Mamoni Dhar

Abstract. In this article, an annotated survey of the approaches to the transformation from probability to possibility or conversely is provided and noticeable properties of the transformation are discussed. This article proposes an approach which is different from the one described by some authors in the literature of fuzzy set theory. The alternative proposal which advocates reconstructions is presented hereby. The principle suggested has a considerable potential for practical as well as for mathematical applications.

1. Introduction

Possibility theory is a mathematical theory dealing with certain types of uncertainties and is considered as an alternative to probability theory. Possibility theory is devoted to the handling of incomplete information. The process of transformation from probability to possibility had received attention in the past. This question is philosophically interesting as a part of debate between probability and fuzzy sets.

The conversion problem between probability and possibility has its roots in possibility-probability consistency principle of Zadeh [11], that he introduced in the paper founding possibility theory. The transformation between probability and possibility has been studied by many researchers. Most of these studies examined principles that must be satisfied for transformations and devised an equation satisfying them. Dubios and Prade further contributed to its development. In Zadeh's view, possibility distributions were meant to provide a graded semantics to natural statements. The transformation between probabilities to possibility is useful in some practical problems as: constructing a fuzzy membership function from a statistical data Krishnapuram [13], combining probabilities and possibilities in expert systems Klir [3]

and reducing complicated complexity Dubois [2]. In other words, the transformation from possibility to probability or conversely is useful in case of decision making when the experts need precise informations to take any decision. In literature, we can see three most common and well known principles relating probability with possibility which are Zadeh consistency principle, Klir consistency principle and Dubois and Prade consistency principles. Following these three principles we can see various other principles too. All these principles were applied to different fields but they were found not too appropriate for every circumstances. As a consequence of which we can see the existence of many principles relating probability with possibility. Most of these studies examined the principles that must be satisfied for transformations and devised an equation satisfying them in a heuristic way. The problem is that the new comers in the domain are overwhelmed by the multitude of models. Their reaction may to accept one of them and use it in every context .Another reaction may be to accept all of them and to apply them more or less at random. Both the attitudes as can be found in the newcomers are not at all appreciable because these can be seriously misleading. There is certainly a relationship between possibility and probability and this relationship is seen differently by different researchers at different point of time. The theories differ from one another in their meaningful interpretations, generality, computational complexity and other aspects. In this article, we shall try to specify this relationship in a way that is expected to replace the ones which exist in the literature of fuzzy set theory. In literature, the reader will find numerous different approaches which from a mathematical point of view are quite interesting.We shall however mention a few of such transformation principles which are most commonly used.

2. TRANSFORMATION CONSISTENCY PRINCIPLES

Zadeh consistency principle:

Zadeh defined the probability-possibility consistency principle such as "a high degree of possibility does not imply a high degree of probability, nor does a low degree of probability imply a low degree of possibility" ,Zadeh [11]. He defined the degree of consistency between a probability distribution $p = (p_1, p_2, \cdots, p_n)$ and a possibility distribution $\pi = (\pi_1, \pi_2, \cdots, \pi_n)$as:

$$(2.1) \qquad C_z = \sum \pi_i.p_i, i = 1, 2, \cdots, n$$

From the above, it can be seen that Zadeh in the process of finding consistency between probability and possibility simply found an index by summing up the scalar multiplication of p_i and π_i. Later on Zadeh pointed out that the probability-possibility consistency, defined in (2.1), is not a precise law or a relationship between possibility and probability distributions. It is an approximate formalization of the heuristic connection that a lessening of the possibility of an event tends to lessen its probability but not vice-versa. As a result many more probability- possibility consistency principles were forwarded thereafter by others. It is important to mention here that there are some researchers who found some problems with the Zadehian principle and this should be the case because Zadeh himself was not satisfied with the one developed by him.Consequently, many other principles developed in due course.

Among various principles, we shall mention here the two well known principles as can be found in the literature.

Klir consistency principle:

Klir defined the consistency principle in the following manner: Let $X = (w_1, w_2, \cdots, w_n)$ be a finite universe of singletons, let $p_i = p_i(w_i)$ and $\pi_i = \pi(w_i)$. Assume that the elements of X are ordered in such a way that: $\forall i = 1, 2, \cdots, n$, $p_i > 0$ and $p_i \geq p_{i+1}$. According to Klir, the transformation from p_i to π_i must preserves some appropriate scale and the amount of information contained in each distribution Klir [4]. The information contained in p or π_i can be expressed by the equality of their uncertainties. Klir has considered the principle of uncertainty preservation under two scales: - The ratio scale: This is a normalization of the probability distribution. The transformations $p \to \pi$ and $\pi \to p$ were named as the normalized transformation and they are defined by

$$\pi_i = \frac{p_i}{p_1}$$

and

$$p_i = \frac{\pi_i}{n \sum_1^n \pi_i}$$

But later on it was found that the transformations found on ratio scales do not have enough flexibility to preserve uncertainty and consequently are not applicable. -The log-interval scales: the corresponding transformations $p \to \pi$ and $\pi \to p$ are defined by:

$$\pi_i = (\frac{p_i}{p_1})^\alpha$$

and

$$p_i = \frac{\pi_i^{\frac{1}{\alpha}}}{\sum_{i=1}^n \pi_i^{\frac{1}{\alpha}}}$$

These transformations, which are named Klir transformations, satisfy the uncertainty preservation principle defined by Klir [4]. a is a parameter that belongs to the open interval $]0, 1[$. Klir's assumptions are debatable from various aspects. The uncertainty invariance equation $E(\pi) = H(P)$, along with a scaling transformation assumption $(\pi(x) = \alpha p(x) + \beta, \forall x)$, reduces the problem of computing π from p to that of solving an algebraic equation with one or two unknowns. Then, the scaling assumption leads to assume that $\pi(x)$ is a function of $p(x)$ only. This point-wise assumption may conflict with the probability-possibility consistency principle that requires for all events. (See Dubois and Prade [2], pp. 258-259) for an example of such a violation. Then, the nice link between possibility and probability, casting possibility measures in the setting of upper and lower probabilities cannot be maintained. The second most questionable prerequisite assumes that possibilistic and probabilistic information measures are commensurate. The basic idea is that the choice between possibility and probability is a mere matter of translation between languages "neither of which is weaker or stronger than the other" (quoting Klir and Parviz [3]). It means that entropy and imprecision capture the same facet of uncertainty,albeit in different guises.

The last point of divergence is that Klir did not try to respect the probability-possibility consistency principles which enable a nice link between possibility and

probability to be maintained. Defining p_i from π_i in the way to satisfy uncertainty preservation principles defined by Klir himself is nothing but trying to define a probability space in the measure theoretic sense from the knowledge of possibilities concerned.It seems that it was done to normalize the values of π_i so that the total probability is equal to 1.

Dubois and Prade consistency principle:

The possibilistic representation is weaker than the probabilistic one because it explicitly handles imprecision (e.g. incomplete data) and because possibility measures are based on ordering structure than an additive one in the probability measures Dubois [2]. Thus in going from a probabilistic representation to a possiblistic one, some information is lost because we go from point-valued probabilities to interval valued ones; the converse transformation adds information to some possibilistic incomplete knowledge. The transformation $p \to \pi$ is guided by the principle of maximum specificity, which aims at finding the most informative possibility distribution. While the transformation $\pi \to p$ is guided by the principle of insufficient reason which aims at finding the probability distribution that contains as much uncertainty as possible but that retains the features of possibility distribution Dubois [2]. This leads to write the consistency principle of Dubois and Prade such as:

$$(2.2) \qquad A \subset X : \Pi(A) \geq P(A)$$

The transformations $p \to \pi$ and $\pi \to p$ are defined by

$$(2.3) \qquad \pi_i = \sum_{j=1}^{n} p_j$$

$$(2.4) \qquad p_i = \frac{\sum_{j=1}^{n} \pi_j - \pi_{j+1}}{j}$$

The transformations defined by (2.2) , (2.3)and (2.4) are not converse of each other because they are not based on the same informational principle. For this reason, the transformation defined by equations (2.3) and (2.4) can be named as asymmetric. Dubois and Prade suggested a symmetric $p \to \pi$ transformation which is defined by:

$$(2.5) \qquad \pi_i = \sum_{j=1}^{n} min(p_i, p_j)$$

Dubois and Prade proved that the symmetric transformation $p \to \pi$, defined by (2.5), is the most specific transformation which satisfies the condition of consistency of Dubois and Prade [2] defined by (2.2) . From our standpoints this principle too is not free from defects, which can be seen from the fact that the consistency principle was not derived in accordance to his definition of a normal fuzzy number.If this be the case then what is the use of defining a normal fuzzy number with the help of two functions? Another noticeable thing in the principle is that it was defined for discrete cases only and nothing was mentioned about continuous cases as already been discussed in Dhar.et.al[12]. On the other hand, in Dubious- Prade's consistency principle there was the use of possibility measure which is not a measure

in the classical sense.A possible justification of this is as follows: The measure of a point is zero, Possibility of occurrence of a point is defined by membership function and therefore in this case the possibility of occurrence of the point is not zero.Hence there should not be any formalism with reference to the membership function. It is for these reasons we do not prefer this principle too.

Some other papers dealing with membership function:

Du, Choi and Young [10] were of the opinion that unlike the probability based methods in which the probability density function and the cumulative distribution function of the random variable is well known the selection of the membership function of the fuzzy variable in possibility based methods are not clear. They introduced a probability-possibility consistent principle to generate the membership function of a fuzzy variable from temporary probability density function. Moreover, the kernal smoothing method was recommended to generate the temporary probability density function of the fuzzy variable from the insufficient data.

Here, our basic aim is to inform that there is no need of introducing temporary probability density function with the help of kernal smoothing methods because possibility distribution can be expressed as two distribution functions which are associated with some densities.Hence the process of finding probability density function seems to have no logical meaning. For showing these, we are to take the help of Superimposition of sets as defined by Baruah[5] Moreover, it is to be worth mentioning here the Dubois and Prade's (see for example Kaufmann and Gupta[1] definition of a fuzzy number because it is this definition which plays a very important role in our work. Dubious-Prade defined a fuzzy number $X = [a, b, c]$ with membership function $\mu_X(x) = L(x)$ if $a \leq x \leq b$,$R(x)$ if $b \leq x \leq c$ and 0 otherwise.$L(x)$ being continuous and non decreasing in the interval $[a, b]$and $R(x)$ being continuous and non increasing in the interval $[b, c]$. In this article we are going to demonstrate a possible link between probability and possibility on the basis of the fact that Dubious-Prade left reference function can be expressed as a distribution function and Dubious- Prade left reference function as a complementary distribution function. As we proceed with this, we would be able to see everything very clearly and hence let us have a brief view of the operation of set superimposition because it is this concept which leads to formulate a new consistency principle.In other words, the concept of operation of superimposition of sets plays the principal role in the relationship between probability and possibility. This can be viewed as a bridge by which probability and possibility can be connected.

3. The operation of set superimposition:

The operation of set superimposition is defined by Baruah [5] and is expressed as follows: If the set A is superimposed over the set B, we get

$$A(S)B = (A - B) \cup (A \cap B)^2 \cup (B - A)$$

where S represents the operation of superimposition, and $(A \cap B)^{(2)}$ represents the elements of $(A \cap B)$ is not void.With the application of superimposition of sets on uniformly fuzzy intervals, we can define a normal fuzzy number of the type $N = [\alpha, \beta, \gamma]$ as

$$\mu_N(x) = \psi_1(x) if \alpha \leq x \leq \ \gamma, = \psi_2(x) if \beta \leq x \leq \gamma, and = 0,$$

5

otherwise.

where $\psi_1(x)$ and $1 - \psi_2(x)$ are probability distribution functions which would be associated with densities $\frac{d\psi_1(x)}{dx}$ and $\frac{1-\psi_2(x)}{dx}$ and this would in turn lead to a very important principles which is named as the Randomness- Fuzziness Consistency Principle. Thus proceeding with the operation of set superimposition we have derived a result which is different from those existing in the literature of fuzzy sets theory. From the above discussion, we can say that two distributions with reference to two probability measures defined on two disjoint spaces can construct a fuzzy membership function.Hence, we would like to say that integrating a probability distribution function is totally meaningless because physically a probability distribution function defines an area and integrating the probability distribution again means we are trying to obtain the area under a function which already defines an area.Hence we can establish our claim that the triangular probability distribution function obtained by the said authors in the manner proposed by them seems to have some defects if it is seen from the standpoints of superimposition of sets.In the next section, we would like to discuss about the suggested principle in short.

4. Randomness-Fuzziness Consistency Principles:

Baruah ([5],[6],[7],[8]) introduced a framework for reasoning with the link between probability and possibility. The development of this principle focused mainly on the existence of two laws of randomness which are required to define a law of fuzziness. In other words , not one but two laws of fuzziness is required to define a law of randomness on two disjoint spaces which in turn can construct a fuzzy membership function. Fundamental to this approach is the idea that possibility distribution can be viewed as a combination of distributions of which one is a probability distribution and the other is a complementary probability distribution. The consistency principle introduced in the manner can be explained mathematically in the following form: For a normal fuzzy number of the type $N = [\alpha, \beta, \gamma]$ with membership functions

$$\mu_N(x) = \psi_1(x) if \alpha \leq x \leq \ \gamma, = \psi_2(x) if \beta \leq x \leq \gamma, and = 0, otherwise$$

with

$$\psi_1(\alpha) = \psi_2(\gamma) = 0, \psi_1(\beta) = \psi_2(\beta) = 1$$

the partial presence of a value x of the variable X in the interval $[\alpha, \gamma]$ is expressible as

$$\mu_N(x) = \theta Prob[\alpha \leq X \leq x] + (1 - \theta)(1 - Prob[\beta \leq X \leq x])$$

where $\theta = 1$ if $\alpha \leq x \leq \beta$ and $\theta = 0$ if $\beta \leq x \leq \gamma$.

The above relationship between probability and possibility is named as "The Randomness- Fuzziness Consistency Principle" which is more mathematical or formal in character. Another thing to be mentioned here that the triangular probability distribution function satisfy the consistency principles of Zadeh and Dubious-Prade which are also not free from defects from our standpoints because it is seen that Zadeh tried to define a probability law over the same space over which possibility law has been defined. Zadeh himself had some weaker constraint in mind for his own principles. Moreover, with the help of two probability measures we can study possibility independently. From those results results it is clear that possibility can be

expressible either as a probability or as a complementary probability. The need for a fundamentally different approach to the study of a possible link between possibility and probability can be realized from the above discussions.

5. Conclusions

In this article, efforts have been made to show that since possibility is expressible either as probability or as complementary probability and hence integrating the membership function of a normal fuzzy set means integrating the probability distribution function which itself represents an area and so it can never yield any expected result. Moreover, this article also does not agree with the consistency principles provided to us by different researchers at different point of time, who tried to modify the principles put forward by their predecessors but still lagged behind to reach logical results due to some facts or others. It seems that none of them were satisfied with the principles and as a result we can see more than one such principle. Since the existence of many principles would lead to a chaotic state, we would like to stress on the fact that there should be only one principle which is established within mathematical frameworks. Through this article, we intend to inform that the consistency obtained with the help of superimposition of sets can give us a logical result which the researchers are trying to get for years.

References

[1] A. Kaufmann , M.M. Gupta,Introduction to Fuzzy Arithmetic, Theory and Applications,, Berkshire, 1784

[2] D. Dubois and H. Prade, On possibility/probability transformations,,Fuzzy Logic,(1993),103-112.

[3] G.J. Klir, B. Parvez, Probability - Possibility Transformation : A Comparision, Int. J. of General Systems , 21(1992) 291-310.

[4] G.J .Klir ,Information preserving probability- possibility transformations: recent developments, Fuzzy Logic, (1993) 417- 428.

[5] H. K. Baruah, Set Superimposition and Its Application to the Theory of Fuzzy Sets, Journal of the Assam Science Society, 40(1)& (2), (1999) 25-31.

[6] H.K.Baruah, The Randomness- Fuzziness Consistency Principles, IJEIC, 1(1),(2010) 37-48.

[7] H.K.Baruah , Theory of Fuzzy sets:Beliefs and Realities, Int. Journal of Energy, Information and Communications ,1(2011)1-22.

[8] H.K.Baruah, In Search of Root of Fuzziness: The Measure Theoritic Meaning of Partial Presence,Annals of Fuzzy Mathematics and Informatics, 2(1), (2011) 57-68.

[9] H.K .Baruah,An Introduction to the Theory of Imprecise sets: the Mathematics of Partial Presence,Journal of Mathematical and Computational Sciences, 2(2), (2012)110-124.

[10] L. Du, K.K.Choi and B.D.Young,A New Fuzzy Analysis for Possibility-Bases Design Optimization,AIAA Journal,(2005).

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

[12] M. Dhar, A note on the Coherence between Probability and Possibility Measures,IJCA, 43(7), (2012)28-31

[13] R. Krishnapuram, J.M. Keller, A possibilistic approach to clustering, IEEE transcation on fuzzy systems , 1(2),(1993) 98-109.