JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 23, 421–427 (1968)

Probability Measures of Fuzzy Events

L. A. ZADEH*

Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, California

I. INTRODUCTION

In probability theory [1], an *event*, A, is a member of a σ-field, $\mathcal{A}$, of subsets of a sample space Ω. A *probability measure*, P, is a normed measure over a measurable space $(\Omega, \mathcal{A})$; that is, P is a real-valued function which assigns to every A in $\mathcal{A}$ a *probability*, $P(A)$, such that (a) $P(A) \geqslant 0$ for all $A \in \mathcal{A}$; (b) $P(\Omega) = 1$; and (c) P is countably additive, i.e., if $\{A_i\}$ is any collection of disjoint events, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i). \tag{1}$$

The notions of an event and its probability constitute the most basic concepts of probability theory. As defined above, an event is a precisely specified collection of points in the sample space. By contrast, in everyday experience one frequently encounters situations in which an "event" is a fuzzy rather than a sharply defined collection of points. For example, the ill-defined events: "It is a *warm* day," "x is *approximately* equal to 5;" "In twenty tosses of a coin there are *several* more heads than tails," are fuzzy because of the imprecision of the meaning of the underlined words.

By using the concept of a fuzzy set [2], the notions of an event and its probability can be extended in a natural fashion to fuzzy events of the type exemplified above. It is possible that such an extension may eventually significantly enlarge the domain of applicability of probability theory, especially in those fields in which fuzziness is a pervasive phenomenon.

The present note has the limited objective of showing how the notion of a fuzzy event can be given a precise meaning in the context of fuzzy sets. Thus, it consists mostly of definitions and is largely preliminary in nature. We make

* Currently on leave as Visiting Professor of Electrical Engineering and Guggenheim Fellow with the Department of Electrical Engineering and Project MAC, M.I.T., Cambridge, Massachusetts. This work was supported in part by Project MAC, sponsored by the Advanced Research Projects Agency, Department of Defense, under Office of Naval Research Contract Nonr-4102(01).

no attempt to formulate our definitions in the most general setting, nor do we attempt to explore in detail any of the paths along which classical probability theory may be generalized through the use of concepts derived from the notion of a fuzzy event.

II. Fuzzy Events

We shall assume for simplicity that Ω is an Euclidean n-space R^n. Thus our probability space will be assumed to be a triplet $(R^n, \mathcal{Ol}, P)$, where $\mathcal{Ol}$ is the σ-field of Borel sets in R^n and P is a probability measure over R^n. A point in R^n will be denoted by x.

Let $A \in \mathcal{Ol}$. Then, the probability of A can be expressed as

$$P(A) = \int_A dP \tag{2}$$

or equivalently

$$\begin{aligned} P(A) &= \int_{R^n} \mu_A(x)\, dP \\ &= E(\mu_A), \end{aligned} \tag{3}$$

where μ_A denotes the characteristic function of A ($\mu_A(x) = 0$ or 1) and $E(\mu_A)$ is the expectation of μ_A.

Equation (3) equates the probability of an event A with the expectation of the characteristic function of A. It is this equation that can readily be generalized to fuzzy events through the use of the concept of a fuzzy set.

Specifically, a fuzzy set A in R^n is defined by a characteristic function $\mu_A : R^n \to [0, 1]$ which associates with each x in R^n its "grade of membership," $\mu_A(x)$, in A. To distinguish between the characteristic function of a nonfuzzy set and the characteristic function of a fuzzy set, the latter will be referred to as a *membership* function. A simple example of a fuzzy set in R^1 is $A = \{x \mid x \gg 0\}$. A membership function for such a set might be subjectively defined by, say,

$$\begin{aligned} \mu_A(x) &= (1 + x^{-2})^{-1}, \qquad && x \geqslant 0 \\ &= 0, && x < 0. \end{aligned} \tag{4}$$

We are now ready to define a fuzzy event in R^n.

Definition. Let $(R^n, \mathcal{Ol}, P)$ be a probability space in which $\mathcal{Ol}$ is the σ-field of Borel sets in R^n and P is a probability measure over R^n. Then, a *fuzzy event* in R^n is a fuzzy set A in R^n whose membership function, $\mu_A(\mu_A : R^n \to [0, 1])$, is Borel measurable.

The *probability* of a fuzzy event A is defined by the Lebesgue-Stieltjes integral:

$$P(A) = \int_{R^n} \mu_A(x)\, dP = E(\mu_A). \tag{5}$$

Thus, as in (3), the probability of a fuzzy event is the expectation of its membership function. The existence of the Lebesgue-Stieltjes integral (5) is insured by the assumption that μ_A is Borel measurable.

The above definitions of a fuzzy event and its probability form a basis for generalizing within the framework of the theory of fuzzy sets a number of the concepts and results of probability theory, information theory and related fields. In many cases, the manner in which such generalization can be accomplished is quite obvious. We shall illustrate this in the sequel by a few simple examples.

There are several basic notions relating to fuzzy sets which we shall need in our discussion. These are summarized below. A more detailed discussion of these and other notions may be found in [2].

Containment $A \subset B \Leftrightarrow \mu_A(x) \leqslant \mu_B(x)\ \forall x$ (6)

Equality $A = B \Leftrightarrow \mu_A(x) = \mu_B(x)\ \forall x$ (7)

Complement A' = complement of $A \Leftrightarrow \mu_{A'}(x) = 1 - \mu_A(x)\ \forall x$ (8)

Union $A \cup B$ = union of A and $B \Leftrightarrow \mu_{A\cup B}(x) = \mathrm{Max}[\mu_A(x), \mu_B(x)]\ \forall x$ (9)

Intersection $A \cap B$ = intersection of A and B

$$\Leftrightarrow \mu_{A\cap B}(x) = \mathrm{Min}[\mu_A(x), \mu_B(x)]\ \forall x \tag{10}$$

Product AB = product of A and $B \Leftrightarrow \mu_{AB}(x) = \mu_A(x)\, \mu_B(x)\ \forall x$ (11)

Sum $A \oplus B$ = sum of A and B

$$\Leftrightarrow \mu_{A\oplus B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\, \mu_B(x)\ \forall x \tag{12}$$

We are now ready to draw some elementary conclusions from (5)–(12). First, as an immediate consequence of (6), we have

$$A \subset B \Rightarrow P(A) \leqslant P(B) \tag{13}$$

Similarly, as immediate consequences of (9)-(12), we have the identities

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{14}$$

$$P(A \oplus B) = P(A) + P(B) - P(AB). \tag{15}$$

A level set, $A(\alpha)$, of a fuzzy set A is a non-fuzzy set defined by

$$A(\alpha) = \{x \mid \mu_A(x) \leqslant \alpha\} \tag{16}$$

A fuzzy set A will be said to be a *Borel fuzzy set* if all of its level sets (for $0 \leqslant \alpha \leqslant 1$) are Borel sets. Since the membership function of a fuzzy event is measurable, it follows that all of the level sets associated with a fuzzy event are Borel sets and hence that a fuzzy event is a Borel fuzzy set.

It is well-known that if μ_A and μ_B are Borel measurable, so are Max $[\mu_A, \mu_B]$, Min $[\mu_A, \mu_B]$, $\mu_A + \mu_B$ and $\mu_A\mu_B$ [3]. The same holds, more generally, for any infinite collection of Borel measurable functions. Consequently, we can assert that, like the Borel sets, Borel fuzzy sets form a σ-field with respect to the operations (8), (9) and (10). In this connection, it should be noted that fuzzy sets obey the distributive law.

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \tag{17}$$

but not

$$(A \oplus B)\, C = AC \oplus BC. \tag{18}$$

Employing induction and making use of (14) and (15), we obtain for fuzzy sets the familiar identities for nonfuzzy sets:

$$P\left(\bigcup_{i=1}^{m} A_i\right) = \sum_i P(A_i) - {\sum_{i,j}}' P(A_i \cap A_j) + \cdots + (-1)^m P\left(\bigcap_i A_i\right) \tag{19}$$

$$P(A_1 \oplus \cdots \oplus A_m) = \sum_i P(A_i) - {\sum_{i,j}}' P(A_i A_j) + \cdots + (-1)^m P(A_1 \cdots A_m). \tag{20}$$

In a similar fashion, (14) and (15) yield the generalized Boole inequalities for fuzzy sets

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} P(A_i) \tag{21}$$

$$P(A_1 \oplus A_2 \oplus \cdots) \leqslant \sum_{i=1}^{\infty} P(A_i). \tag{22}$$

We turn next to the notion of independence of fuzzy events. Specifically, let A and B be two fuzzy events in a probability space $(R^n, \mathcal{Cl}, P)$. Then A and B will be said to be *independent* if

$$P(AB) = P(A)\, P(B). \tag{23}$$

Note that in defining independence we employ the product AB rather than the intersection $A \cap B$.

An immediate consequence of the above definition is the following: Let $\Omega_1 = R^n$, $\Omega_2 = R^m$ and let P be the product measure $P_1 \times P_2$, where P_1 and P_2 are probability measures on Ω_1 and Ω_2, respectively. Let A_1 and A_2 be events in Ω_1 and Ω_2 characterized by the membership functions $\mu_{A_1}(x_1, x_2) = \mu_{A_1}(x_1)$ and $\mu_{A_2}(x_1, x_2) = \mu_{A_2}(x_2)$, respectively. Then A_1 and A_2 are independent events in the sense of (23). Note that this would not be true if independence were defined in terms of $P(A \cap B)$ rather than $P(AB)$.

To be consistent with (23), the *conditional probability* of A given B is defined by

$$P(A \mid B) = \frac{P(AB)}{P(B)}, \tag{24}$$

provided $P(B) > 0$. Note that if A and B are independent, then $P(A \mid B) = P(A)$, as in the case of nonfuzzy independent events.

Many of the basic notions in probability theory, such as those of the mean, variance, entropy, etc., are defined as functionals of probability distributions. The concept of a fuzzy event suggests that it may be of interest to define these notions in a more general way which relates them to both a fuzzy event and a probability measure. For example, the mean of a fuzzy event A relative to a probability measure P may be defined as follows:

$$m_P(A) = \frac{1}{P(A)} \int_{R^u} x\mu_A(x)\, dP \tag{25}$$

where μ_A is the membership function of A and $P(A)$ serves as a normalizing factor. Similarly, the variance of a fuzzy event in R^1 relative to a probability measure P may be defined as

$$G_P^2(A) = \frac{1}{P(A)} \int_{R^1} (x - m_P(A))^2\, \mu_A(x)\, dP \tag{26}$$

The subscript P in (25) and (26) may be omitted when the dependence on P of the quantities in question is implied by the context.

Turning to the notion of entropy, we note that its usual definition in information theory is as follows: Let x be a random variable which takes the values $x_1, \ldots, x_n$ with respective probabilities $p_1, \ldots, p_n$. Then, the entropy of x—or, more properly, the entropy of the distribution $P = \{p_1, \ldots, p_n\}$—is given by

$$H(x) = -\sum_{i=1}^{n} p_i \log p_i . \tag{27}$$

This definition suggests that the entropy of a fuzzy subset, A, of the finite set $\{x_1, ..., x_n\}$ with respect to a probability distribution $P = \{p_1, ..., p_n\}$ be defined as follows

$$H^P(A) = - \sum_{i=1}^{n} \mu_A(x_i)\, p_i \log p_i\,, \tag{28}$$

where μ_A is the membership function of A. Note that whereas (27) expresses the entropy of a distribution P, (28) represents the entropy of a fuzzy event A with respect to the distribution P. Thus, (28) does not reduce to (27) when A is nonfuzzy, unless A is taken to be the whole space $\{x_1, ..., x_n\}$. Intuitively, $H^P(A)$ may be interpreted as the uncertainty associated with a fuzzy event.

Let x and y be independent random variables with probability distributions $P = \{p_1, ..., p_n\}$ and $Q = \{q_1, ..., q_m\}$, respectively. One of the basic properties of the joint entropy of x and y is that when x and y are independent, we can write

$$H(x, y) = H(x) + H(y). \tag{29}$$

It is easy to verify that for fuzzy events this identity generalizes to

$$H^{PQ}(AB) = P(A)\, H^P(A) + P(B)\, H^Q(B), \tag{30}$$

where

$$PQ = \{p_i q_j\}, \qquad i = 1, ..., n, \qquad j = 1, ..., m$$

$$P(A) = \sum_{i=1}^{n} \mu_A(x_i)\, p_i$$

$$P(B) = \sum_{j=1}^{m} \mu_B(y_j)\, q_j$$

$$H^P(A) = - \sum_{i=1}^{n} \mu_A(x_i)\, p_i \log p_i$$

and

$$H^Q(B) = - \sum_{j=1}^{m} \mu_B(y_j)\, q_j \log q_j\,.$$

Note that (30) reduces to (29) when $A = \{x_1, ..., x_n\}$ and $B = \{y_1, ..., y_m\}$.

The foregoing examples are intended merely to demonstrate possible ways of defining some of the elementary concepts of probability theory in a more general setting in which fuzzy events are allowed. It appears that there are many concepts and results in probability theory, information theory and related fields which admit of such generalization.

References

1. H. G. Tucker. "A Graduate Course in Probability." Academic Press, New York, 1967.
2. L. A. Zadeh. Fuzzy Sets. *Inform. Control* 8 (1965).
3. W. Rudin. "Principles of Mathematical Analysis." McGraw-Hill, New York, 1967.