ENTROPY AND MUTUAL INFORMATION OF EXPERIMENTS IN THE FUZZY CASE

*Dagmar Markechová**

Abstract: In my previous papers ([18], [19]) the entropy of fuzzy partitions had been defined. The concept of the entropy of a fuzzy partition was used to define the entropy of a fuzzy dynamical system and to propose an ergodic theory for fuzzy dynamical systems ([19], [20]). In this paper, using my previous results related to the entropy of fuzzy partitions, a measure of average mutual information of fuzzy partitions is defined. Some properties concerning this measure are proved. It is shown that the entropy of fuzzy partitions can be considered as a special case of their mutual information. We obtain that subadditivity and additivity of entropy of fuzzy partitions are simple consequences of these properties. The suggested measures can be applied whenever it is need to know the amount of information that we obtain by realization of experiments, the results of which are fuzzy events.

Key words: *Fuzzy probability space, fuzzy partition, entropy, mutual information*

Received: February 20, 2013
Revised and accepted: July 30, 2013

1. Introduction

When planning experiments it is very important to know the amount of information which we obtain from their realization. It is well known that a measure of information is an entropy. A usual mathematical model of a random experiment in the classical information theory is a measurable partition of a probability space. Partitions are standardly defined in the context of classical, crisp sets. It has appeared however, that for a solution of real problems are partitions defined by means of the concept of fuzzy sets more appropriate. That was a motivation for several concepts of generalization of the classical set partition to a fuzzy partition ([1]-[5], [9], [23]-[25], [32]). In our papers [18] and [19], we deal with study of fuzzy partitions defined by K. Piasecki in [26]. In [18], the entropy of fuzzy partitions was defined and its basic properties were proved. In this paper, we define a mutual information $I_m(\boldsymbol{A}, \boldsymbol{B})$ of experiments $\boldsymbol{A}$, $\boldsymbol{B}$, the outcomes of which are fuzzy events.

*Dagmar Markechová
Department of Mathematics, Faculty of Natural Sciences, Constantine the Philosopher University in Nitra, A. Hlinku 1, SK-949 74 Nitra, Slovak Republic, E-mail: dmarkechova@ukf.sk

In the final section, it is proved the inequality $I_m(\boldsymbol{A},\ \boldsymbol{B})\geq 0$, with the equality if and only if the experiments $\boldsymbol{A}$, $\boldsymbol{B}$ are statistically independent. In conclusion, it is shown that the mentioned property implies subadditivity and additivity of the entropy of fuzzy partitions.

2. Fuzzy partitions and their entropy

In the classical probability theory, which is based on the Kolmogorov axiomatic system ([15]), a random event is every element of the $\sigma-$algebra S of subsets of a set X. A probability is a normalized measure defined on the $\sigma-$algebra S. The notion of $\sigma-$algebra S of random events and concept of a probability space $(X,\ S,\ P)$ are a base of the classical concept of probability theory. In doing so, the event in the classical probability theory is understood as an exactly defined phenomenon and from a mathematical point of view (as it is mentioned above) it is a classical set. In real life, however, we often talk about events that carry important information, but they are less exact. For example, "tomorrow will be nice" "a large number falls", "the product operates within a short time" are vaguely defined events, so-called fuzzy events. Their probability can be studied using the apparatus of fuzzy set theory. The first attempts to develop a concept of fuzzy events and their probability come from the founder of the fuzzy set theory L. A. Zadeh (see [34]). Assuming that the probability space $(X,\ S,\ P)$ is given, Zadeh defines fuzzy event as any S-measurable function $f : X \to \langle 0,\ 1\rangle$ and the probability $p(f)$ of fuzzy event f by the formula $p(f) = \int_X f dP$. Axiomatic approaches to the creation of the probability concept of fuzzy events came up in a short time. There are several (see, for example [4], [11], [14]). The object of our studies in the articles [17] and [18] were the fuzzy probability spaces defined by Polish mathematician K. Piasecki.

Definition 2.1 ([26]). A fuzzy probability space is a triplet (X, M, m), where X is a nonempty set; M is a fuzzy $\sigma-$algebra of fuzzy subsets of X (i.e., (i) $1_X \in M$; $(1/2)_X \notin M$; (ii) if $f_n \in M$, $n = 1, 2, ...$, then $\vee_{n=1}^{\infty} f_n \in M$; (iii) if $f \in M$, then $f' := 1_X - f \in M$) and the mapping $m : M \to \langle 0,\ \infty)$ fulfils the following conditions:

(P1) $m(f \vee f') = 1$ for every $f \in M$;

(P2) if $\{f_n\}_{n=1}^{\infty}$ is a sequence of pairwise W-separated fuzzy subsets from M (i.e., $f_i \leq f'_j$ for $i \neq j$), then $m(\vee_{n=1}^{\infty} f_n) = \sum_{n=1}^{\infty} m(f_n)$.

The operations with fuzzy sets are defined here by Zadeh ([33]), i.e., the union of fuzzy subsets f, g of X is a fuzzy set $f \vee g$ defined by $(f \vee g)(x) = \sup(f(x),\ g(x))$ for all $x \in X$ and the intersection of fuzzy subsets f, g of X is a fuzzy set $f \wedge g$ defined by $(f \wedge g)(x) = \inf(f(x),\ g(x))$ for all $x \in X$. The complement of fuzzy subset f of X is a fuzzy set f' defined by $f'(x) = 1 - f(x)$ for all $x \in X$. The empty fuzzy set 0_X is defined by $0_X(x) = 0$ for all $x \in X$. The complement of empty fuzzy set is a fuzzy set 1_X defined by the equality $1_X(x) = 1$ for all $x \in X$. It is called universum. Fuzzy subsets f, g of X such that $f \wedge g = 0_X$ are called separated fuzzy sets. Analogous weak notions (W-notions) were defined by Piasecki in [27] as follows: each fuzzy subset $f \in M$ such that $f \geq f'$ is called W-universum; each fuzzy subset $f \in M$ such that $f \leq f'$ is called W-empty set. Fuzzy subsets $f, g \in M$ such that $f \leq g'$ are called W-separated. It can be proved that a fuzzy set $f \in M$ is a W-universum if and only if there exists a fuzzy set $g \in M$ such that $f = g \vee g'$.

Each mapping $m : M \to \langle 0, \infty)$ having the properties (P1) and (P2) is called in the terminology of Piasecki a fuzzy P-measure. Any fuzzy P-measure has the following properties, which are analogous to the properties of classical probability measure:

(P3) $m(f') = 1 - m(f)$ for every $f \in M$.

(P4) m is a nondecreasing function, i.e., if $f, g \in M$, $f \leq g$, then $m(f) \leq m(g)$.

(P5) $m(f \vee g) = m(f) + m(g) - m(f \wedge g)$ for any $f, g \in M$.

(P6) Let $g \in M$ be given. Then $m(f \wedge g) = m(f)$ for all $f \in M$ if and only if $m(g) = 1$.

(P7) m is a $\sigma-$subadditive function, i.e., $m(\vee_{n=1}^{\infty} f_n) \leq \sum\limits_{n=1}^{\infty} m(f_n)$ for any sequence $\{f_n\}_{n=1}^{\infty} \subset M$.

(P8) $m(\vee_{n=1}^{\infty} f_n) = \lim\limits_{n\to\infty} m(f_n)$ for any sequence $\{f_n\}_{n=1}^{\infty} \subset M$ such that $f_n \leq f_{n+1}$ for $n = 1, 2, \ldots$.

(P9) $m(\wedge_{n=1}^{\infty} f_n) = \lim\limits_{n\to\infty} m(f_n)$ for any sequence $\{f_n\}_{n=1}^{\infty} \subset M$ such that $f_n \geq f_{n+1}$ for $n = 1, 2, \ldots$.

(P10) If $f, g \in M$ are W-separated, then $m(f \wedge g) = 0$.

(P11) If $f, g \in M$ such that $f \leq g$, then $m(g) = m(f) + m(f' \wedge g)$.

The proofs of these properties can be found in [26].

The mapping $m(\,\cdot\,/g) : M \to \langle 0, 1\rangle$ defined for each $g \in M, m(g) > 0$, by the equality

$$m(\,f\,/g) = \frac{m(f \wedge g)}{m(g)},\ f \in M,$$

is a fuzzy P-measure on M ([26]). It is called a conditional probability.

A probability interpretation of the above notions is as follows:

(i) a set X is a set of elementary events;

(ii) a fuzzy set from the system M is a fuzzy event;

(iii) the value $m(f)$ is a probability of fuzzy event f;

(iv) a fuzzy event f' is an opposite event to fuzzy event f;

(v) a fuzzy event with zero probability is an impossible event;

(vi) a fuzzy event with positive probability is a possible event;

(vii) a fuzzy event with probability 1 is a certain event;

($viii$) W-separated fuzzy events are interpreted as mutually exclusive events.

By means of properties of fuzzy P-measure we get that a W-universum is a certain event and a W-empty set is an impossible event. In the following, we give Piasecki‘s definitions of fuzzy partitions of a fuzzy probability space (X, M, m) and our definition of their entropy.

Definition 2.2 ([26]). Let a fuzzy probability space (X, M, m) be given. Each finite or infinite sequence $\{f_i\}_i \subset M$ of pairwise W-separated fuzzy subsets such that $\vee_i f_i$ is a W-universum is called a complete fuzzy partition of the space (X, M, m).

Definition 2.3 ([26]). Let a fuzzy probability space (X, M, m) be given. A finite or infinite sequence $\boldsymbol{A} = \{f_i\}_i$ of pairwise W-separated fuzzy subsets from Msuch that $m(\vee_i f_i) = 1$ is called a stochastic fuzzy partition of the space (X, M, m).

It is evident that each complete fuzzy partition is a stochastic fuzzy partition. In accordance with the classical theory we will in the following consider only finite fuzzy partitions. Let us denote by the symbol $\boldsymbol{P}$ the system of all finite stochastic fuzzy partitions. Each partition $\boldsymbol{A}$ =$\{f_1, f_2, ..., f_n\}$ from the system $\boldsymbol{P}$ represents in the sense of the classical probability theory a random experiment with finite number of outcomes f_i, $\quad i = 1, 2, ..., n$, (which are fuzzy events) with probability distribution $p_i = m(f_i)$, $\quad i = 1, 2, ..., n$, since $p_i \geq 0$ for $i = 1, 2, ..., n$ and $\sum\limits_{i=1}^{n} p_i = \sum\limits_{i=1}^{n} m(f_i) = m(\vee_{i=1}^{n} f_i) = 1$. Therefore the entropy of any experiment $\boldsymbol{A}$ =$\{f_1, f_2, ..., f_n\}$ from the system $\boldsymbol{P}$ had been defined in [19] by Shannon's formula:

$$H_m(\boldsymbol{A}) = H(m(f_1), m(f_2), ..., m(f_n)) = -\sum_{i=1}^{n} F(m(f_i)),$$

where

$$F : \langle 0, \infty) \to R, \quad F(x) = \begin{cases} x \log x, & \text{if} \quad x > 0; \\ 0, & \text{if} \quad x = 0. \end{cases}$$

The notion of entropy of a stochastic fuzzy partition is a generalization of Shannon‘s entropy of a measurable partition (see [19]). In [19], properties of the suggested measure were studied, too. It has been shown that it satisfies all properties analogous to the properties of Shannon's entropy in the crisp case. The entropy of a stochastic fuzzy partition can be considered as a measure of uncertainty of experiment, the outcomes of which are fuzzy events.

Let $\boldsymbol{B}$ =$\{ g_1, g_2, ..., g_m\}$ be a random experiment of a fuzzy probability space (X, M, m). Let us suppose that some circumstances do not allow us to realize the experiment $\boldsymbol{B}$, but we know a result of some other experiment $\boldsymbol{A}$ =$\{ f_1, f_2, ..., f_n\}$ of the same fuzzy probability space (X, M, m). Let us assume that a fuzzy event f_i with nonzero probability is a result of the experiment $\boldsymbol{A}$. Then the fuzzy events $g_1, g_2, ..., g_m$ happen with probabilities $m(g_1/f_i)$, $m(g_2/f_i)$, ..., $m(g_m/f_i)$, where $m(g_j/f_i) := \frac{m(g_j \wedge f_i)}{m(f_i)}$ is a conditional probability of the fuzzy event g_j provided that the fuzzy event f_i occurred. The uncertainty of the experiment $\boldsymbol{B}$ will change from the value $H_m(\boldsymbol{B}) = H(m(g_1), m(g_2), \ldots, m(g_m))$ to the value $H(m(g_1/f_i), m(g_2/f_i), \ldots, m(g_m/f_i))$, which we will denote by the symbol $H_m(\boldsymbol{B} / f_i)$. This is a motivation for the following definition of conditional entropy of experiment $\boldsymbol{B}$ given the fuzzy event f_i and definition of conditional entropy of experiment $\boldsymbol{B}$ assuming a realization of experiment $\boldsymbol{A}$.

Definition 2.4 Let $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$, $\boldsymbol{A} = \{f_1, f_2, ..., f_n\}$, $\boldsymbol{B} = \{g_1, g_2, ..., g_m\}$. A conditional entropy of $\boldsymbol{B}$ given a fuzzy event $f_i \in \boldsymbol{A}$ is defined by

$$H_m(\boldsymbol{B} / f_i) = -\sum_{j=1}^{m} F(\dot{m}(g_j/f_i)),$$

where

$$\dot{m}(g_j/f_i) = \begin{cases} m(g_j/f_i), & \text{if} \quad m(f_i) > 0; \\ 0, & \text{if} \quad m(f_i) = 0. \end{cases}$$

A conditional entropy of the experiment $\boldsymbol{B} \in \boldsymbol{P}$ assuming a realization of the experiment $\boldsymbol{A} \in \boldsymbol{P}$ is defined by the formula

$$H_m(\boldsymbol{B} / \boldsymbol{A}) = -\sum_{i=1}^{n} m(f_i) \cdot H_m(\boldsymbol{B} / f_i) \quad = -\sum_{i=1}^{n} \sum_{j=1}^{m} m(f_i) \cdot F(\dot{m}(g_j/f_i)).$$

Remark 2.5 Put $\alpha = \{i;\ m(f_i) > 0\}$, $\beta = \{(i, j);\ m(f_i \wedge g_j) > 0\}$. Then it holds

$$H_m(\boldsymbol{B} / \boldsymbol{A}) = -\sum_{i=1}^{n} \sum_{j=1}^{m} m(f_i) \cdot F(\dot{m}(g_j/f_i)) = -\sum_{i \in \alpha} \sum_{j=1}^{m} m(f_i) \cdot F(m(g_j/f_i) =$$

$$= -\sum_{(i,j) \in \beta} m(f_i) \cdot \frac{m(f_i \wedge g_j)}{m(f_i)} \cdot \log \frac{m(f_i \wedge g_j)}{m(f_i)} = -\sum_{(i,j) \in \beta} m(f_i \wedge g_j) \cdot \log \frac{m(f_i \wedge g_j)}{m(f_i)}.$$

We define in the set $\boldsymbol{P}$ of all finite stochastic fuzzy partitions of a fuzzy probability space (X, M, m) the relation $\leq$ in the following way: for every $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$, $\boldsymbol{A} \leq \boldsymbol{B}$ iff for every $g \in \boldsymbol{B}$ there exists $f \in \boldsymbol{A}$ such that $g \leq f$. In this case we shall say that the partition $\boldsymbol{B}$ is a refinement of the partition $\boldsymbol{A}$.

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two stochastic fuzzy partitions of a fuzzy probability space (X, M, m). Then the system $\boldsymbol{A} \vee \boldsymbol{B}$: $= \{f \wedge g;\ f \in \boldsymbol{A}, g \in \boldsymbol{B}\}$ is a stochastic fuzzy partition of the space (X, M, m), too. The partition $\boldsymbol{A} \vee \boldsymbol{B}$ represents a joint experiment of experiments $\boldsymbol{A}$, $\boldsymbol{B}$. It easy to see that $\boldsymbol{A} \leq \boldsymbol{A} \vee \boldsymbol{B}$ and $\boldsymbol{B} \leq \boldsymbol{A} \vee \boldsymbol{B}$. The entropy and the conditional entropy of stochastic fuzzy partitions satisfy all properties analogous to properties of Shannon's entropy of measurable partitions in the classical case. We review some basic properties of the considered entropy:

(2.1) $H_m(\boldsymbol{A}) \geq 0$ for each $\boldsymbol{A} \in \boldsymbol{P}$.

(2.2) If $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$, $\boldsymbol{A} \leq \boldsymbol{B}$, then $H_m(\boldsymbol{A}) \leq H_m(\boldsymbol{B})$.

(2.3) Let $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$, $\boldsymbol{A} \leq \boldsymbol{B}$. Then

$$H_m(\boldsymbol{A} / \boldsymbol{C}) \leq H_m(\boldsymbol{B} / \boldsymbol{C}) \text{ for every } \boldsymbol{C} \in \boldsymbol{P}.$$

(2.4) For every $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C} \in \boldsymbol{P}$ it holds

$$H_m(\boldsymbol{B} \vee \boldsymbol{C} / \boldsymbol{A}) \leq H_m(\boldsymbol{B} / \boldsymbol{A}) + H_m(\boldsymbol{C} / \boldsymbol{A}).$$

(2.5) $H_m(\boldsymbol{B}\vee\boldsymbol{C} / \boldsymbol{A}) = H_m(\boldsymbol{C} / \boldsymbol{A}\vee\boldsymbol{B}) + H_m(\boldsymbol{B} / \boldsymbol{A})$ for each $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C} \in \boldsymbol{P}$.

The proofs of the above properties may be found in [19].

The presented concept of entropy of a fuzzy partition was used to define the entropy of a fuzzy dynamical system (X, M, m, U) (the mapping $U : M \to M$ is an m-preserving $\sigma-$homomorphism). In the paper [19], an ergodic theory for fuzzy dynamical systems has been proposed. We have therein defined the generators of a fuzzy dynamical system and using this notion a fuzzy version of Kolmogorov-Sinai theorem on generators was given. In the paper [20], it has proven that the entropy of a fuzzy dynamical system is an isomorphism invariant. Another approaches to a fuzzy generalization of notions of entropy and conditional entropy can be found in articles [1], [4], [5], [8], [10]-[13], [16], [21], [22], [29] and [30].

3. Mutual information of experiments in the fuzzy case

Mutual information measures the amount of information that can be obtained about one experiment by realization of another. At first, we will prove the assertion of the following theorem that will be useful in further considerations.

Theorem 3.1 For every $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$ it holds

$$H_m(\boldsymbol{A} \vee \boldsymbol{B}) = H_m(\boldsymbol{A}) + H_m(\boldsymbol{B} / \boldsymbol{A}). \tag{1}$$

Proof. Put $\boldsymbol{A} = \{f_1, f_2, ..., f_n\}$, $\boldsymbol{E} = \{1_X\}$ ($\boldsymbol{E}$ is an experiment, the result of which is a certain event.). Since

$$H_m(\boldsymbol{A} / \boldsymbol{E}) = -\sum_{i=1}^{n} m(1_X) \cdot F(m(f_i/1_X)) = -\sum_{i=1}^{n} F(m(f_i)) = H_m(\boldsymbol{A}),$$

by means of property (2.5) we get

$$H_m(\boldsymbol{A} \vee \boldsymbol{B}) = H_m(\boldsymbol{A} \vee \boldsymbol{B}/\boldsymbol{E}) = H_m(\boldsymbol{B}/\boldsymbol{E}\vee\boldsymbol{A}) + H_m(\boldsymbol{A}/\boldsymbol{E}) = H_m(\boldsymbol{A}) + H_m(\boldsymbol{B}/\boldsymbol{A}).$$

Corollary 3.2 $H_m(\boldsymbol{A} \vee \boldsymbol{B}) \geq max\,(H_m(\boldsymbol{A}); H_m(\boldsymbol{B}))$ for every $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$.

Proof. This result follows from the previous theorem and from the fact that entropy is nonnegative.

Remark 3.3 From the equality (1) it follows that the value of the entropy $H_m(\boldsymbol{A})$ of experiment $\boldsymbol{A}$ is larger, the conditional entropy $H_m(\boldsymbol{B} / \boldsymbol{A})$ is less for it. Analogously as in the classical theory ([7]), based on the equality (1), we can interpret the value $H_m(\boldsymbol{B} \,/\, \boldsymbol{A})$ as a residual entropy of a joint experiment $\boldsymbol{A} \vee \boldsymbol{B}$ after a realization of experiment $\boldsymbol{A}$ and consider it as an average value of additional information that can be derived from the experiment $\boldsymbol{B}$ after a realization of the experiment $\boldsymbol{A}$. The difference $H_m(\boldsymbol{B})$ - $H_m(\boldsymbol{B} \,/\, \boldsymbol{A})$ can then be regarded as the average amount of information about an experiment $\boldsymbol{B}$ contained in an experiment $\boldsymbol{A}$. This is a motivation for the following definition.

Definition 3.4 Let $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$. The average amount of information $I_m(\boldsymbol{A},\ \boldsymbol{B})$ about an experiment $\boldsymbol{B}$ in an experiment $\boldsymbol{A}$ is given by

$$I_m(\boldsymbol{A},\ \boldsymbol{B}) = H_m(\boldsymbol{B}) - H_m(\boldsymbol{B} \,/\, \boldsymbol{A}). \tag{2}$$

Theorem 3.5 For every $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$ it holds

$$I_m(\boldsymbol{A},\ \boldsymbol{B}) = H_m(\boldsymbol{A}) + H_m(\boldsymbol{B}) - H_m(\boldsymbol{A} \ \vee \boldsymbol{B}). \tag{3}$$

Proof. The equality (3) is a simple consequence of Theorem 3.1.

Corollary 3.6 $I_m(\boldsymbol{A},\ \boldsymbol{B}) \leq\ min\,(H_m(\boldsymbol{A});\, H_m(\boldsymbol{B}))$ for every $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$.

Proof. The result follows immediately from Theorem 3.5 and the property (2.2).

From the preceding theorem it follows that the measure I_m is symmetric, i.e., for every $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$, it holds $I_m(\boldsymbol{A},\ \boldsymbol{B}) = I_m(\boldsymbol{B},\ \boldsymbol{A})$. This means that information about an experiment $\boldsymbol{B}$ in an experiment $\boldsymbol{A}$ is equal to information about an experiment $\boldsymbol{A}$ in an experiment $\boldsymbol{B}$. Hence the value $I_m(\boldsymbol{A},\ \boldsymbol{B})$ is called also mutual information of experiments $\boldsymbol{A}$, $\boldsymbol{B}$. It is easy to see that $H_m(\boldsymbol{A} \,/\, \boldsymbol{A}) = 0$, hence from (2) we obtain $I_m(\boldsymbol{A},\ \boldsymbol{A}) = H_m(\boldsymbol{A}) - H_m(\boldsymbol{A} \,/\, \boldsymbol{A}) = H_m(\boldsymbol{A})$. So that we see that the entropy of fuzzy partitions can be considered as a special case of their mutual information.

Theorem 3.7 Let $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$, $\boldsymbol{A} = \{ f_1, f_2, ..., f_n \}$, $\boldsymbol{B} = \{ g_1, g_2, ..., g_m \}$. Then

$$I_m(\boldsymbol{A},\ \boldsymbol{B}) = \sum_{(i,j)\in\beta} m(f_i \wedge g_j) \cdot \log \frac{m(f_i \wedge g_j)}{m(f_i) \cdot m(g_j)}, \tag{4}$$

where $\beta = \{ (i,\, j);\ m(f_i \wedge g_j) > 0 \}$.

Proof. Let $g_j \in \boldsymbol{B}$. Since for $k \neq l$ we have

$$f_k \wedge g_j \leq f_k \leq f_l' \leq f_l' \vee g_j' = (f_l \wedge g_j)',$$

the system $\{ f_1 \wedge g_j, f_2 \wedge g_j, ..., f_n \wedge g_j \}$ is a system of pairwise W-separated fuzzy subsets from M. Due to the assumption $m(\vee_{i=1}^n f_i) = 1$, the property (P6) and additivity of fuzzy P-measure we get

$$m(g_j) = m((\vee_{i=1}^n f_i) \wedge g_j) \quad = \quad m(\vee_{i=1}^n (f_i \wedge g_j)) = \quad \sum_{i=1}^n m(f_i \wedge g_j).$$

Put next $\alpha = \{ j;\ m(g_j) > 0 \}$. Then, based on the remark from the preceding section, we obtain

$$\begin{aligned} I_m(\boldsymbol{A},\ \boldsymbol{B}) &= H_m(\boldsymbol{B}) - H_m(\boldsymbol{B}\ /\ \boldsymbol{A}) = \\ &= -\sum_{j\in\alpha} m(g_j)\cdot \log m(g_j) + \sum_{(i,j)\in\beta} m(f_i \wedge g_j)\cdot \log \frac{m(f_i\wedge g_j)}{m(f_i)} = \\ &= -\sum_{j\in\alpha}\sum_{i=1}^{n} m(f_i \wedge g_j)\cdot \log m(g_j) + \sum_{(i,j)\in\beta} m(f_i \wedge g_j)\cdot \log \frac{m(f_i\wedge g_j)}{m(f_i)} = \\ &= -\sum_{(i,j)\in\beta} m(f_i \wedge g_j)\cdot \log m(g_j) + \sum_{(i,j)\in\beta} m(f_i \wedge g_j)\cdot \log \frac{m(f_i\wedge g_j)}{m(f_i)} = \\ &= \sum_{(i,j)\in\beta} m(f_i \wedge g_j)\cdot [\ \log \frac{m(f_i\wedge g_j)}{m(f_i)} - \log m(g_j)\,] = \\ &= \sum_{(i,j)\in\beta} m(f_i \wedge g_j)\cdot \log \frac{m(f_i\wedge g_j)}{m(f_i)\cdot m(g_j)}. \end{aligned}$$

Theorem 3.8 Let $\boldsymbol{A}$,$\boldsymbol{B}\in \boldsymbol{P}$. Then $I_m(\boldsymbol{A},\ \boldsymbol{B})\geq 0$ with the equality if and only if the experiments $\boldsymbol{A}$, $\boldsymbol{B}$ are statistically independent.

Proof. Let $\boldsymbol{A}$, $\boldsymbol{B}\in \boldsymbol{P}$, $\boldsymbol{A} =\{ f_1, f_2, ..., f_n\}$, $\boldsymbol{B} =\{ g_1, g_2, ..., g_m\}$. Suppose that $(i,\ j)\ \in\ \beta$, where $\beta\ =\ \{\,(i,\ j);\ m(f_i \wedge g_j) > 0\,\}$. Then using the inequality $\ln x \leq x - 1$, which is valid for all $x > 0$, with the equality if and only if $x = 1$, we get

$$m(f_i \wedge g_j)\cdot \log \frac{m(f_i)\cdot m(g_j)}{m(f_i\wedge g_j)} = m(f_i \wedge g_j)\cdot \ln 2\cdot \ln \frac{m(f_i)\cdot m(g_j)}{m(f_i\wedge g_j)} \leq$$

$$\leq m(f_i \wedge g_j)\cdot \ln 2\cdot \left[\frac{m(f_i)\cdot m(g_j)}{m(f_i\wedge g_j)} - 1\right] = \ln 2\cdot [m(f_i)\cdot m(g_j) - m(f_i\wedge g_j)]\,.$$

The equality holds if and only if $\frac{m(f_i)\cdot m(g_j)}{m(f_i\wedge g_j)} = 1$, i.e., when the fuzzy events $f_i,\ g_j$ are independent.

In view of the preceding theorem and by means of the proven inequality we get

$$\begin{aligned} -I_m(\boldsymbol{A},\ \boldsymbol{B}) &= \sum_{(i,j)\in\beta} m(f_i \wedge g_j)\cdot \log \frac{m(f_i)\cdot m(g_j)}{m(f_i\wedge g_j)} \leq \\ &\leq \ln 2 \sum_{(i,j)\in\beta} (m(f_i)\cdot m(g_j) - m(f_i\wedge g_j)) = \\ &= \ln 2 \left[\sum_{(i,j)\in\beta} m(f_i)\cdot m(g_j) - \sum_{(i,j)\in\beta} m(f_i\wedge g_j)\right] = \end{aligned}$$

$$= \ln 2 \left[\sum_{i=1}^{n} \sum_{j=1}^{m} m(f_i) \cdot m(g_j) - \sum_{i=1}^{n} \sum_{j=1}^{m} m(f_i \wedge g_j) \right] =$$

$$= \ln 2 \left[\sum_{i=1}^{n} m(f_i) \sum_{j=1}^{m} m(g_j) - \sum_{i=1}^{n} m(\vee_{j=1}^{m} (f_i \wedge g_j)) \right] =$$

$$= \ln 2 \left[\sum_{i=1}^{n} m(f_i) \cdot m(\vee_{j=1}^{m} g_j) - \sum_{i=1}^{n} m(f_i) \right] = 0.$$

Evidently the equality holds if and only if $m(f_i \wedge g_j) = m(f_i) \cdot m(g_j)$ for $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$, i.e., when fuzzy events f_i, g_j are independent for $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$.

Theorem 3.9 For every $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$ it holds $H_m(\boldsymbol{A} \ / \ \boldsymbol{B}) \leq H_m(\boldsymbol{A})$ with the equality if and only if the experiments $\boldsymbol{A}$, $\boldsymbol{B}$ are statistically independent.

Proof. The assertion of theorem it follows from Definition 3.4, according to which $I_m(\boldsymbol{A}, \boldsymbol{B}) = H_m(\boldsymbol{A})$ - $H_m(\boldsymbol{A} \ / \ \boldsymbol{B})$ and from the inequality $I_m(\boldsymbol{A}, \boldsymbol{B}) \geq 0$, in which the equality holds if and only if the experiments $\boldsymbol{A}$, $\boldsymbol{B}$ are statistically independent.

Theorem 3.10 For every $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$ it holds $H_m(\boldsymbol{A} \vee \boldsymbol{B}) \leq H_m(\boldsymbol{A}) + H_m(\boldsymbol{B})$ with the equality if and only if the experiments $\boldsymbol{A}$, $\boldsymbol{B}$ are statistically independent.

Proof. The assertion of theorem it follows from Theorem 3.5, according to which for every $\boldsymbol{A}$, $\boldsymbol{B} \in \boldsymbol{P}$ it holds $I_m(\boldsymbol{A}, \boldsymbol{B}) = H_m(\boldsymbol{A}) + H_m(\boldsymbol{B})$ - $H_m(\boldsymbol{A} \vee \boldsymbol{B})$ and from the inequality $I_m(\boldsymbol{A}, \boldsymbol{B}) \geq 0$, in which the equality holds if and only if the experiments $\boldsymbol{A}$, $\boldsymbol{B}$ are statistically independent.

Note that in the previous theorem it is proved subadditivity and additivity of entropy of fuzzy partitions.

4. Conclusion

The object of study in this paper are the fuzzy partitions defined by K. Piasecki. These structures can serve as a mathematical model of random experiments, the results of which are fuzzy events. The average mutual information of fuzzy partitions has been defined using our previous results related to the entropy of fuzzy partitions. We have proved basic properties concerning this measure.

References

[1] Benvenuti P., Vivona D., Divari M.: Fuzzy partitions and entropy. In: Proc. of the III Linz Seminar on Fuzzy Set Theory: Uncertainty Measures, (Ed. P. Klement, S. Weber), 1991, pp. 14-18.

[2] De Baets B., Mesiar R.: T-Partitions. Fuzzy Sets and Systems, **97**, 1998, pp. 211-223.

[3] Dumitrescu D.: Fuzzy partitions with the connectives T_∞, S_∞. Fuzzy Sets and Systems, **47**, 1992, pp. 193-195.

[4] Dumitrescu D.: Fuzzy measures and entropy of fuzzy partitions. J. Math. An. and Appl., **176**, 1993, pp. 359-373.

[5] Dumitrescu D.: Measure-preservig transformation and the entropy of a fuzzy partition. In: Proc. of the 13th Linz seminar on fuzzy set theory (Linz, 1991), 1991, pp. 25-27.

[6] Dvurečenskij A., Tirpáková, A.: Ergodic theory on fuzzy quantum spaces. Bulletin Sous-Ensembl. Flous Appl., **37**, Toulouse : Université Paul Sabatier, 1988, pp. 86-94.

[7] Gray R. M.: Entropy and Information Theory. Springer-Verlag, New York, 2009.

[8] Cheng H. D. et al.: Threshold selection based on fuzzy c-partition - entropy approach, Pattern Recognition, **31**, 1998, pp. 857-870.

[9] Jayaram B., Mesiar R.: I-fuzzy equivalence relations and I-fuzzy partitions. Information Sciences: an International Journal, **179**, 2009, pp. 1278-1297.

[10] Khare M.: Sufficient families and entropy of inverse limit. Mathematica Slovaca, **49**, 1999, pp. 443-452.

[11] Khare M.: Fuzzy σ-algebras and conditional entropy. Fuzzy Sets and Systems, **102**, 1999, pp. 287-292.

[12] Khare M., Roy S.: Entropy of quantum dynamical systems and sufficient families in orthomodular lattices with Bayessian state. Communications in Theoretical Physics, **50**, 2008, pp. 551-556.

[13] Khare M., Roy S.: Conditional entropy and the Rokhlin metric on an orthomodular lattice with Bayessian state. International Journal of Theoretical Physics, **47**, 2008, 1386-1396.

[14] Klement E. P.: Fuzzy $\sigma-$algebras and fuzzy measurable functions. Fuzzy Sets and Systems, **4**, 1980, pp. 83-93.

[15] Kolmogorov A. N.: Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer-Verlag, Berlin, 1933.

[16] Kuriyama K.: Entropy of a finite partition of fuzzy sets. J. Math. Anal. Appl., **94**, 1983, pp. 38-43.

[17] Markechová D.: The conjugation of fuzzy probability spaces to the unit interval. Fuzzy Sets and Systems, **47**, 1992, pp. 87-92.

[18] Markechová D.: Entropy of complete fuzzy partitions. Mathematica Slovaca, **43**, 1993, pp. 1-10.

[19] Markechová D.: The entropy of fuzzy dynamical systems and generators. Fuzzy Sets and Systems, **48**, 1992, pp. 351-363.

[20] Markechová D.: A note to the Kolmogorov-Sinaj entropy of fuzzy dynamical systems. Fuzzy Sets and Systems, **64**, 1994, pp. 87-90.

[21] Mesiar R.: The Bayes formula and the entropy of fuzzy probability spaces. International Journal of General Systems, **20**, 1990, pp. 67-71.

[22] Mesiar R., Rybárik J.: Entropy of Fuzzy Partitions – A General Model. Fuzzy Sets and Systems, **99**, 1998, pp. 73-79.

[23] Mesiar R., Reusch B., Thiele H.: Fuzzy equivalence relations and fuzzy partition. Journal of multiple-valued logic and soft computing, **12**, 2006, pp. 167-181.

[24] Montes S. et al. Fuzzy delta-epsilon partitions. Inform. Sciences, **152**, 2003, pp. 267-285.

[25] Montes S., Couso I., Gil P.: One-to-one correspondence between $\varepsilon-$partitions, $(1-\varepsilon)-$ equivalences and $\varepsilon-$pseudometrics. Fuzzy Sets and Systems, **124**, 2001, pp. 87-95.

[26] Piasecki K.: Probability of fuzzy events defined as denumerable additivity measure. Fuzzy Sets and Systems, **17**, 1985, pp. 271-284.

[27] Piasecki K.: New concept of separated fuzzy subsets. In: Proc. of the Polish Symposium on Interval and Fuzzy Mathematics (Poznaň, 1983).

[28] Piasecki K.: On fuzzy P-measures. In: Proc. of the First Winter Scool on Measure Theory (Liptovský Ján,1988), 1988, pp. 108-112.

[29] Riečan B.: An entropy construction inspired by fuzzy sets. Soft computing, **7**, 2003, pp. 486-488.

[30] Srivastava P., Khare M., Srivastava Y. K.: m-Equivalence, entropy and F-dynamical systems. Fuzzy Sets and Systems, **121**, 2001, pp. 275-283.

[31] Tirpáková A.: Ergodic theorems on fuzzy quantum spaces. Bulletin Sous-Ensembl. Flous Appl., **39**, Toulouse : Université Paul Sabatier, 1989, pp. 34-40.

[32] Yu I., Mu Y. M., Shi H. B.: A short note on fuzzy partition. In: Conference Proc. of IEEE 2002 International Conference on Machine Learning and Cybernetics. Beiing: IEEE Press, **1-4**, 2002, pp. 1451-1454.

[33] Zadeh L. A.: Fuzzy Sets. Inform. and Control, **8**, 1965, pp. 338-358.

[34] Zadeh L. A.: Probability measures of fuzzy events. J. Math. Anal. Appl., **23**, 1968, pp. 421-427.