

Some Theoretical Results of Learning Theory Based on Intuitionistic Fuzzy Random Samples

Zhiming Zhang*

College of Mathematics and Computer Science, Hebei University, Baoding 071002, Hebei
Province, PR China

ABSTRACT

In this paper, we discuss the statistical learning theory based on intuitionistic fuzzy random sample. First of all, we introduced the definition of intuitionistic fuzzy number intuitionistic fuzzy random variables. Secondly, we gives some properties of the concept and the mathematical expectation and intuitionistic fuzzy random variable distance. Third, the definitions of intuitionistic fuzzy empirical risk functional, intuitionistic fuzzy expected risk functional, and intuitionistic fuzzy empirical risk minimization principle are presented. Finally, we prove the key theorem of intuitionistic fuzzy random sample of learning theory, and obtain the rate of uniform convergence of learning process based on the intuitionistic fuzzy random sample.

Keywords: Intuitionistic fuzzy numbers; Intuitionistic fuzzy random variables; Intuitionistic fuzzy empirical risk minimization principle; The key theorem; The bounds of the rate of convergence

1. INTRODUCTION

Statistical learning theory (SLT, for short), put forward in 1960s, completely founded by Vapnik et al. in 1990s. [46-48], has become an interesting and good law supporting the development of small samples statistical learning. It provides in small samples, such as encountered in practice learning method of sample is the key problem in the machine, such as over fitting and underfitting, the classification problem of high dimension, multi association are important problems such as local minima obtained in the presence of effective solutions and their architecture, such as for example, nerve network. The SLT has become the fastest growing discipline in machine learning a in the late 1990's. Its essence is to make the limited sample learning machine work effectively, and improve the generalization ability of the learning machine. By doing so, we have established theoretical framework for statistical learning from small data samples. At the same time, the general learning algorithm by SLT, namely the support vector machine (SVM, for short) is a new category. At present, SLT and SVM constitute interesting research avenues, learning in the machine [6, 7, 10, 11, 12, 20, 28, 29, 30, 31, 35, 36, 39, 41, 43, 44, 45, 49, 50, 52, 53, 56].

The SLT covers four main parts [6], namely: (1) the learning process, minimizes the necessary and sufficient conditions for the consistency of the empirical risk - this refers to the key theorem of learning theory form; (2) the scope of generalization; (3) the structural risk minimization principle, according to these limits; (4) algorithm, implement the structure risk minimization principle of support vector machine. The key theorem of boundary convergence of learning theory and the learning process of the STL is the most important foundation. The two basic elements, the pursuit of further research on related topics, such as, the structural risk minimization, SVM and praise.

Although the SLT has reached a level of maturity is high, there are still some problems to be solved, such as the development of such as SLT and SVM realization probability space and

* Tel.: +86 312 5079660; fax: +86 312 5079660.
E-mail address: zhimingzhang@ymail.com.

real valued random sample (real valued random variables) based on. In the real world, there are many non probability space (such as the Sugeno measure space [51], credibility measure space [38]) and non real random samples (such as fuzzy sample [55], fuzzy random sample [32], fuzzy complex random samples [26]). In order to solve these problems, expansion and induction and processing based on non probabilistic space and learning theory line non real random samples based on statistics have become an urgent need to move forward,. Some studies have along this line, as has been achieved. [21-27,42]. For example, Ha, Li, Bai, et al. [21-23] studied the key theorem and the generalized boundary on Sugeno measure probability space, quasi probability measure space and public space; Tian, Ha, Pedrycz, etc. [24,26] investigated the construction of statistical learning theory based on fuzzy random samples and fuzzy complex random samples; Lin et al [36] constructed fuzzy support vector machine based on random sample; Yuki and Zhenxiong [43] presents some comparative study, the error correcting output coding and fuzzy support vector machine between adaptive Liu; Chen [39] the general rate of profit based fuzzy support vector machine on face

Intuitionistic (IF, for short) set originated by Atanassov [1,5], is an important concept used to cope with imperfect and/or imprecise information. It is an intuitively straightforward extension of Zadeh's fuzzy sets [55]: while a fuzzy set gives a degree to which an element belongs to a set, an intuitionistic fuzzy set gives both a membership degree and a non-membership degree. The membership and non-membership values induce an indeterminacy index, which models the hesitancy of deciding the degree to which an object satisfies a particular property. As the basis for the study of IF set theory, many operations and relations over IF sets were introduced [1-5]. Many concepts in fuzzy set theory were also extended to IF set theory, such as IF relations, intuitionistic L-fuzzy sets, IF implication, IF logics, and the degree of similarity between IF sets, etc., [4,9,13,15,16,33,34,57]. For a further study for the structure of IF sets, construction theorems of IF sets, IF topology and the axiomatic characterization of IF sets have been investigated [8,14,18,40]. Recently, IF set theory has been successfully applied in decision analysis and pattern recognition (see, e.g., [19,37,54]).

However, the study for the combination of IF set theory and statistical learning theory is still blank. This paper discusses the statistical learning theory based on intuitionistic fuzzy random samples by combining intuitionistic fuzzy analysis and SLT, revisits the major parts of the SLT and establishes some ground material for further development of classifiers such as support vector machines. This paper is organized as follows: Section 2 introduces some basic definitions, elaborates on a number of properties of intuitionistic fuzzy random variables. In Section 3, we prove the key theorem of learning theory based on intuitionistic fuzzy random samples. In the sequel, in Section 4, the bounds of the rate of uniform convergence of learning process based on intuitionistic fuzzy random samples are discussed. The final section offers the conclusions and brings prospects of potential future developments.

2. PRELIMINARIES

Throughout this paper, we let (Ω, A, P) be a probability measure space and R be the real numbers field.

Definition 2.1 [32]. A fuzzy number is a fuzzy set $X : R \rightarrow [0,1]$ with the following properties:

- (1) X is normal, i.e., there exists $x \in R$ such that $X(x) = 1$;
- (2) X is upper semi-continuous;
- (3) $\text{supp } X = \text{cl}\{x \in R : X(x) > 0\}$ is compact;
- (4) X is a convex fuzzy set, i.e., $X(\lambda x + (1-\lambda)y) \geq \min(X(x), X(y))$, for $\forall x, y \in R$ and $\lambda \in [0,1]$.

94 Let $F(\mathbb{R})$ be the family of all fuzzy numbers on $\mathbb{R}$. For a fuzzy set X , if we define

$$95 \quad (X)_r = \begin{cases} \{x : X(x) \geq r\}, & 0 < r \leq 1 \\ \text{supp } X, & r = 0 \end{cases},$$

96 then it follows that X is a fuzzy number if and only if $(X)_1 \neq \emptyset$ and $(X)_r$ is a closed
97 bounded interval for each $r \in [0,1]$. Therefore, a fuzzy number X is completely determined
98 by the intervals $(X)_r$.

99 Let $X, Y \in F(\mathbb{R})$, $\forall r \in [0,1]$, $(X)_r = [(X)_r^-, (X)_r^+]$ and $(Y)_r = [(Y)_r^-, (Y)_r^+]$. Define

$$100 \quad X \pm Y = \bigcup_{r \in [0,1]} \left\{ [(X)_r^- \pm (Y)_r^-, (X)_r^+ \pm (Y)_r^+] \right\}.$$

101 Let us introduce a partially ordered relation on $F(\mathbb{R})$ as follow: for any $X, Y \in F(\mathbb{R})$, we
102 say $X \leq Y$, iff $\forall r \in [0,1]$, $(X)_r \leq (Y)_r$ i.e., $(X)_r^- \leq (Y)_r^-$ and $(X)_r^+ \leq (Y)_r^+$.

103 **Definition 2.2** [32]. A fuzzy number valued function $\xi : \Omega \rightarrow F(\mathbb{R})$ is called fuzzy random
104 variable if for every closed subset B of $\mathbb{R}$, the fuzzy set $\xi^{-1}(B)$ is measurable when
105 considered as a function from Ω to $[0,1]$, where $\xi^{-1}(B)$ denotes the fuzzy subset of Ω
106 defined by $\xi^{-1}(B)(\omega) = \sup_{x \in B} \xi(\omega)(x)$ for every $\omega \in \Omega$.

107 **Definition 2.3** [32]. A fuzzy random variable $\xi(\omega) = \{[(\xi_r^-(\omega), \xi_r^+(\omega))] | 0 < r \leq 1\}$ is called
108 integrable if for each $r \in [0,1]$, $\xi_r^-(\omega)$ and $\xi_r^+(\omega)$ are integrable, or equivalently $\int \|\xi\| dP < \infty$,
109 where $\|\xi\| = \max\{|\xi_0^-|, |\xi_0^+|\}$. In this case, the expected value of ξ is defined in the following
110 manner

$$111 \quad E\xi = \int \xi dP = \left\{ \left[\int \xi_r^- dP, \int \xi_r^+ dP \right] | 0 < r \leq 1 \right\}.$$

112 **Definition 2.4** [1]. Let U be a nonempty set. An intuitionistic fuzzy set A in U is an object
113 having the form $A = \{ \langle x, \mu_A(x), \nu_A(x) \rangle | x \in U \}$,

114 where $\mu_A : U \rightarrow [0,1]$ and $\nu_A : U \rightarrow [0,1]$ satisfy $0 \leq \mu_A + \nu_A \leq 1$ for all $x \in U$, and $\mu_A(x)$ and
115 $\nu(x)$ are, respectively, called the degree of membership and the degree of non-membership
116 of the element $x \in U$ to A . The complement of an IF set A is denoted
117 by $A^c = \{ \langle x, \nu(x), \mu_A(x) \rangle | x \in U \}$.

118 Obviously, every fuzzy set $A = \{ \langle x, A(x) \rangle | x \in U \} = \{ \langle x, \mu_A(x) \rangle | x \in U \}$ can be identified with
119 the IF set of the form $\{ \langle x, \mu_A(x), 1 - \mu_A(x) \rangle | x \in U \}$ and is thus an IF set.

120 We introduce some basic operations about IF sets as follows [1–5,17,33]:

121 Let A and B be two IF sets, then

122 $A \subseteq B$ iff $\mu_A(x) \leq \mu_B(x)$ and $\nu_A(x) \leq \nu_B(x)$ for all $x \in U$,

123 $A \supseteq B$ iff $B \subseteq A$,

124 $A = B$ iff $A \subseteq B$ and $B \subseteq A$, i.e., $\mu_A(x) = \mu_B(x)$ and $\nu_A(x) = \nu_B(x)$ for any $x \in U$,

125 $A \cap B = \{ \langle x, \mu_A(x) \wedge \mu_B(x), \nu_A(x) \vee \nu_B(x) \rangle | x \in U \},$

126 $A \cup B = \{ \langle x, \mu_A(x) \vee \mu_B(x), \nu_A(x) \wedge \nu_B(x) \rangle | x \in U \},$

$$\forall \alpha, \beta \in [0, 1], \alpha + \beta \leq 1, (\alpha, \beta)A = \{ \langle x, \alpha \wedge \mu_A(x), \beta \vee \nu_A(x) \rangle | x \in U \}.$$

Definition 2.5 [58]. Let A be an IF set and $\alpha, \beta \in [0, 1]$ with $\alpha + \beta \leq 1$, the (α, β) -level cut set of A , denoted by A_{α}^{β} , is defined as follows:

$$A_{\alpha}^{\beta} = \{ x \in U | \mu_A(x) \geq \alpha, \nu_A(x) \leq \beta \}$$

$A_{\alpha} = \{ x \in U | \mu_A(x) \geq \alpha \}$ and $A_{\alpha+} = \{ x \in U | \mu_A(x) > \alpha \}$ are, respectively, called the α -level cut set and the strong α -level cut set of membership generated by A . And $A^{\beta} = \{ x \in U | \nu_A(x) \leq \beta \}$ and $A^{\beta+} = \{ x \in U | \nu_A(x) < \beta \}$ are, respectively, referred to as the β -level cut set and the strong β -level cut set of non-membership generated by A .

Proposition 2.1. Let $A \in IF(U)$, $\alpha, \beta \in [0, 1]$ and $\alpha + \beta \leq 1$. Then

$$A = \bigcup (\alpha, \beta) A_{\alpha}^{\beta} = \bigcup (\alpha, \beta) A_{\alpha+}^{\beta} = \bigcup (\alpha, \beta) A_{\alpha}^{\beta+} = \bigcup (\alpha, \beta) A_{\alpha+}^{\beta+}.$$

Proof. We have only to prove that the equation $A = \bigcup (\alpha, \beta) A_{\alpha}^{\beta}$ holds for any $x \in U$. The other equations can be proved in a similar way. On one hand, we have

$$\begin{aligned} \mu_{\bigcup (\alpha, \beta) A_{\alpha}^{\beta}}(x) &= \vee \mu_{(\alpha, \beta) A_{\alpha}^{\beta}}(x) = \vee (\alpha \wedge A_{\alpha}^{\beta}(x)) \\ &= \left(\bigvee_{0 \leq \alpha \leq \mu_A(x)} (\alpha \wedge A_{\alpha}^{\beta}(x)) \right) \vee \left(\bigvee_{\mu_A(x) < \alpha \leq 1} (\alpha \wedge A_{\alpha}^{\beta}(x)) \right) \\ &= \bigvee_{0 \leq \alpha \leq \mu_A(x)} \alpha = \mu_A(x). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \nu_{\bigcup (\alpha, \beta) A_{\alpha}^{\beta}}(x) &= \wedge \nu_{(\alpha, \beta) A_{\alpha}^{\beta}}(x) = \wedge (\beta \vee (1 - A_{\alpha}^{\beta}(x))) \\ &= \left(\bigwedge_{0 \leq \beta \leq \nu_A(x)} (\beta \vee (1 - A_{\alpha}^{\beta}(x))) \right) \wedge \left(\bigwedge_{\nu_A(x) < \beta \leq 1} (\beta \vee (1 - A_{\alpha}^{\beta}(x))) \right) \\ &= \bigwedge_{\nu_A(x) < \beta \leq 1} \beta = \nu_A(x). \end{aligned}$$

Thus $A = \bigcup (\alpha, \beta) A_{\alpha}^{\beta}$.

Definition 2.6 [59]. The IF set $A = \{ \langle x, \mu_A(x), \nu_A(x) \rangle | x \in \mathbb{R} \}$ is called an IF number if and only if μ_A and ν_A^c are two fuzzy numbers, where $\nu_A^c = 1 - \nu_A$.

The family of all intuitionistic fuzzy numbers is denoted by $IF(\mathbb{R})$.

Proposition 2.2. Let $A \in IF(\mathbb{R})$. Then the membership function $\mu_A(x)$ and non-membership function $\nu_A(x)$ have the following properties:

- (1) $\mu_A : \mathbb{R} \rightarrow [0, 1]$ is an upper semi-continuous function and $\mu_A : \mathbb{R} \rightarrow [0, 1]$ is a lower semi-continuous function;
- (2) $\mu_A(x) = 0, \forall x \in (-\infty, c] \cup [d, +\infty), \nu_A(x) = 1, \forall x \in (-\infty, e] \cup [f, +\infty)$;
- (3) $\mu_A(x) = 1, \nu_A(x) = 0, \forall x \in [a, b]$;
- (4) $\mu_A(x)$ is strictly monotonously increasing on $[c, a]$ and strictly monotonously decreasing on $[b, d]$; $\nu_A(x)$ is strictly monotonously decreasing on $[e, a]$ and strictly monotonously increasing on $[b, f]$.

Definition 2.7. Some Operations on $IF(\mathbb{R})$ are defined as follows:

$$\forall A, B \in IF(\mathbb{R}),$$

$$157 \quad A + B = \left\{ \left\langle z, \sup_{x+y=z} (\mu_A(x) \wedge \mu_B(y)), \inf_{x+y=z} (\nu_A(x) \vee \nu_B(y)) \right\rangle \mid x, y \in U \right\};$$

$$158 \quad \rho A = \left\{ \left\langle z, \mu_A\left(\frac{z}{\rho}\right), \nu_A\left(\frac{z}{\rho}\right) \right\rangle \mid z = \rho x, x \in U \right\}, \quad \rho \neq 0;$$

159 If $\rho = 0$, then

$$160 \quad \mu_{0A}(z) = \begin{cases} \sup \{ \mu_A(x) \mid x \in U \}, & z = 0 \\ 0, & z \neq 0 \end{cases}, \quad \nu_{0A}(z) = \begin{cases} \inf \{ \nu_A(x) \mid x \in U \}, & z = 0 \\ 1, & z \neq 0 \end{cases}.$$

161 **Proposition 2.3.** Let $A, B \in IF(R)$. And (α, β) -level cut set of A and B are denoted by

$$162 \quad A_\alpha^\beta = [\mu_A^L, \mu_A^R, \nu_A^L, \nu_A^R, (\alpha, \beta)], B_\alpha^\beta = [\mu_B^L, \mu_B^R, \nu_B^L, \nu_B^R, (\alpha, \beta)]. \text{ Then we have}$$

$$163 \quad A + B = \left\{ [\mu_A^L + \mu_B^L, \mu_A^R + \mu_B^R, \nu_A^L + \nu_B^L, \nu_A^R + \nu_B^R, (\alpha, \beta)] \mid \alpha, \beta \in [0, 1], \alpha + \beta \leq 1 \right\},$$

$$164 \quad \rho A = \begin{cases} [\rho \mu_A^L, \rho \mu_A^R, \rho \nu_A^L, \rho \nu_A^R, (\alpha, \beta)] \mid \alpha, \beta \in [0, 1], \alpha + \beta \leq 1, & \text{if } \rho \geq 0 \\ [\rho \mu_A^R, \rho \mu_A^L, \rho \nu_A^R, \rho \nu_A^L, (\alpha, \beta)] \mid \alpha, \beta \in [0, 1], \alpha + \beta \leq 1, & \text{if } \rho < 0 \end{cases}.$$

165 **Definition 2.8** [58]. Let $A = \{ \langle x, \mu_A(x), \nu_A(x) \rangle \mid x \in U \}, B = \{ \langle x, \mu_B(x), \nu_B(x) \rangle \mid x \in U \} \in IF(R)$.

166 Define $A \leq B$ if $\mu_A \leq \mu_B$ and $\nu_A^c \leq \nu_B^c$; $A = B$ if $A \leq B$ and $B \leq A$; $A < B$ if $A \leq B$ and $A \neq B$.

167 Obviously, $(IF(R), \leq)$ is a partially ordered set.

168 **Definition 2.9.** Let $S \subseteq IF(R)$ and $M \in IF(R)$. We say that M is the supremum of the set S if
169 the following two conditions are satisfied:

170 (1) for any $A \in S$, $A \leq M$, i.e., M is the upper bound of the set S ;

171 (2) for any the upper bound N of S , $M \leq N$.

172 Similarly, we can define the infimum of the set S .

173 Let $K(R)$ be a family of all the nonempty compact convex subsets of R . If $A, B \in K(R)$,
174 then the Hausdorff metric is defined by

$$175 \quad d_H(A, B) = \max \left\{ \sup_{x \in A} \inf_{y \in B} d(x, y), \sup_{y \in B} \inf_{x \in A} d(x, y) \right\},$$

176 where $d(x, y)$ denotes the distance between two real numbers x and y .

177 Let us define a consistent Hausdorff metric in $IF(R)$ to be in the following form

$$\begin{aligned} d(A, B) &= \sup_{\substack{\alpha, \beta \in [0, 1] \\ \alpha + \beta \leq 1}} \max \{ d_H(A_\alpha, B_\alpha), d_H(A^\beta, B^\beta) \} \\ 178 \quad &= \max \{ d(\mu_A, \mu_B), d(\nu_A, \nu_B) \} \\ &= \max \{ d(\mu_A, \mu_B), d(\nu_A^c, \nu_B^c) \} \end{aligned}$$

179 where $A = \{ \langle x, \mu_A(x), \nu_A(x) \rangle \mid x \in U \}, B = \{ \langle x, \mu_B(x), \nu_B(x) \rangle \mid x \in U \} \in IF(R)$.

180 **Proposition 2.4.** If $A, B, C \in IF(R)$, then $d(A, B) + d(B, C) \geq d(A, C)$.

181 **Proposition 2.5.** If $A, B, C \in IF(R)$ and $A \geq B \geq C$,

182 then $d(A, C) \geq d(A, B)$ and $d(A, C) \geq d(B, C)$.

183 **Definition 2.10.** Let (Ω, A, P) be a probability measure space. An intuitionistic fuzzy number
184 valued mapping

185 $\xi: \Omega \rightarrow IF(R)$,

$$\omega \rightarrow \xi(\omega) = \left\{ \langle x, \mu_{\xi(\omega)}(x), \nu_{\xi(\omega)}(x) \rangle \mid x \in \mathbb{R} \right\}$$

is called an intuitionistic fuzzy random variable if μ_{ξ} and ν_{ξ}^c are two fuzzy random variables defined on $(\Omega, \mathcal{A}, P)$, where $\mu_{\xi} : \omega \rightarrow \mu_{\xi(\omega)}$, $\nu_{\xi} : \omega \rightarrow \nu_{\xi(\omega)}$ and $\nu_{\xi}^c = 1 - \nu_{\xi}$.

Definition 2.11 Let $\xi(\omega) = \left\{ \langle x, \mu_{\xi(\omega)}(x), \nu_{\xi(\omega)}(x) \rangle \mid x \in \mathbb{R} \right\}$ be the intuitionistic fuzzy random variable defined on $(\Omega, \mathcal{A})$. We called

$$F(Z) = P\{\xi(\omega) \leq Z\} = P\{\mu_{\xi(\omega)} \leq \mu_Z, \nu_{\xi(\omega)} \geq \nu_Z\}$$

as the distribution function of ξ , where $Z = \left\{ \langle x, \mu_Z(x), \nu_Z(x) \rangle \mid x \in \mathbb{R} \right\} \in IF(\mathbb{R})$.

Definition 2.12. We call $\xi(\omega) = \left\{ \langle x, \mu_{\xi(\omega)}(x), \nu_{\xi(\omega)}(x) \rangle \mid x \in \mathbb{R} \right\}$ integrable if μ_{ξ} and ν_{ξ}^c are integrable. In this case, we define the mathematical expectation of ξ as the following manner

$$\begin{aligned} E(\xi) &= \int \xi dP = \left\{ \langle x, E(\mu_{\xi})(x), E(\nu_{\xi})(x) \rangle \mid x \in \mathbb{R} \right\} \\ &= \left\{ \left[\int \mu_{\xi}^L dP, \int \mu_{\xi}^R dP, \int \nu_{\xi}^L dP, \int \nu_{\xi}^R dP, (\alpha, \beta) \right] \mid \alpha, \beta \in [0, 1], \alpha + \beta \leq 1 \right\}, \end{aligned}$$

where $E(\nu_{\xi}) = 1 - E(\nu_{\xi}^c)$.

Proposition 2.6. Let ξ be an intuitionistic fuzzy random variable, then the following equalities

$$1. \{E(\xi)\}_{\alpha}^{\beta} = E(\xi_{\alpha}^{\beta}) = E(\xi_{\alpha}) \cap E(\xi^{\beta}) \text{ for } 0 < \alpha < \beta \leq 1;$$

$$2. E(c\xi) = cE\xi, \text{ whenever } c \in \mathbb{R};$$

$$3. E(\xi_1 \pm \xi_2) = E\xi_1 \pm E\xi_2$$

hold.

Definition 2.13. Assume that $\xi_n = \left\{ \langle x, \mu_{\xi_n}(x), \nu_{\xi_n}(x) \rangle \mid x \in U \right\}, n = 1, 2, \dots$ is a collection of intuitionistic fuzzy random variables.

(1) If the fuzzy random variables sequence $\mu_{\xi_n}, n = 1, 2, \dots$ and $\nu_{\xi_n}^c, n = 1, 2, \dots$ are respectively mutually independent, then $\xi_n, n = 1, 2, \dots$ is called a collection of mutually independent intuitionistic fuzzy random variables.

(2) If $\mu_{\xi_n}, n = 1, 2, \dots$ and $\nu_{\xi_n}^c, n = 1, 2, \dots$ are two sequences of identically distributed fuzzy random vectors, then $\xi_n, n = 1, 2, \dots$ is referred to as a sequence of identically distributed intuitionistic fuzzy random variables.

Definition 2.14. Let $\xi_n, n = 1, 2, \dots$ be a sequence of intuitionistic fuzzy random variables and let ξ be an intuitionistic fuzzy random variable (or an intuitionistic fuzzy number). If $\forall \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P\{\omega \mid d(\xi_n(\omega), \xi(\omega)) > \varepsilon\} = 0,$$

then ξ_n converge in probability P to ξ , denoted by

$$\xi_n \xrightarrow{P} \xi, n \rightarrow \infty, \text{ or } \lim_{n \rightarrow \infty} \xi_n = \xi.$$

Theorem 2.1. Suppose that $\xi_n = \left\{ \langle x, \mu_{\xi_n}(x), \nu_{\xi_n}(x) \rangle \mid x \in U \right\}, n = 1, 2, \dots$ is a sequence of intuitionistic fuzzy random variables and $\xi = \left\{ \langle x, \mu_{\xi}(x), \nu_{\xi}(x) \rangle \mid x \in U \right\}$ is an intuitionistic fuzzy random variable. Then, $\xi_n \xrightarrow{P} \xi$ if and only if $\mu_{\xi_n} \xrightarrow{P} \mu_{\xi}$ and $\nu_{\xi_n} \xrightarrow{P} \nu_{\xi}$.

Proof. By the definition of Hausdorff metric in $IF(\mathbb{R})$, we have

$$d(\mu_{\xi_n}, \mu_{\xi}) \leq d(\xi_n, \xi), d(v_{\xi_n}, v_{\xi}) \leq d(\xi_n, \xi), d(\xi_n, \xi) \leq d(\mu_{\xi_n}, \mu_{\xi}) + d(v_{\xi_n}, v_{\xi}).$$

Hence,

$$P\{d(\xi_n, \xi) > \varepsilon\} \leq P\{d(\mu_{\xi_n}, \mu_{\xi}) + d(v_{\xi_n}, v_{\xi}) > \varepsilon\} \leq P\left\{d(\mu_{\xi_n}, \mu_{\xi}) > \frac{\varepsilon}{2}\right\} + P\left\{d(v_{\xi_n}, v_{\xi}) > \frac{\varepsilon}{2}\right\}.$$

The theorem is easily proven.

Theorem 2.2 (The strong law of large numbers of fuzzy random variables) [32].

Let $\{\xi_n, n=1, 2, \dots\}$ be a sequence of independent and identically distributed fuzzy random variables with $E\|\xi_1\| < \infty$. Then we have

$$\frac{1}{n} \sum_{j=1}^n \xi_j \xrightarrow{P} E\xi_1, n \rightarrow \infty.$$

Theorem 2.3 (The strong law of large numbers of intuitionistic fuzzy random variables). Let

$\{\xi_n, n=1, 2, \dots\}$ be a sequence of independent and identically distributed intuitionistic fuzzy random variables with $E\|\mu_{\xi_1}\| < \infty$ and $E\|v_{\xi_1}\| < \infty$. Then

$$\frac{1}{n} \sum_{j=1}^n \xi_j \xrightarrow{P} E\xi_1, n \rightarrow \infty.$$

Proof. It follows from Theorem 2.2 and Proposition 2.3.

233

3. THE KEY THEOREM OF LEARNING THEORY BASED ON INTUITIONISTIC FUZZY RANDOM SAMPLES

236

The key theorem forms an important basis of SLT. Strict consistency requirements that it is consistent with the presence of unilateral convergence means to relax the mathematical expectation of their or an empirical risk minimization method, in other words, there is convergence of some empirical process. In this part, we introduce the theory of intuitionistic fuzzy random variables based on the theory of learning. First, we introduce some basic concepts.

Let $\xi_j, j=1, 2, \dots, l$ be a sequence of independent and identically distributed intuitionistic fuzzy random samples whose distribution is given as $F(\xi)$.

Definition 3.1. We call $R_{if}(\alpha) = E[Q(\xi, \alpha)] = \int Q(\xi, \alpha) dP, \alpha \in \Lambda$, where Λ is an index set and $Q(\xi, \alpha)$ denotes the loss functional, the expected risk functional on the basis of intuitionistic fuzzy random samples. It could be considered as the intuitionistic fuzzy expected risk functional.

Definition 3.2. $R_{ifemp}(\alpha) = \frac{1}{l} \sum_{j=1}^l Q(\xi_j, \alpha), \alpha \in \Lambda$ is called the intuitionistic fuzzy empirical risk functional.

Let the risk functional obtain its minimum at $Q(\xi, \alpha_0)$ and the empirical risk functional obtain its minimum at $Q(\xi, \alpha_l)$.

Definition 3.3. We minimize $R_{ifemp}(\alpha)$ to replace $R_{if}(\alpha)$ and refer to the function $Q(\xi, \alpha_l)$ as an approximation to the function $Q(\xi, \alpha_0)$. We call this principle the intuitionistic fuzzy empirical risk minimization principle.

Definition 3.4. If the following two sequences converge in probability to the same limit:

$$R_{if}(\alpha_l) \xrightarrow{P} \inf_{\alpha \in \Lambda} R_{if}(\alpha),$$

$$R_{ifemp}(\alpha_l) \xrightarrow{P} \inf_{\alpha \in \Lambda} R_{if}(\alpha),$$

then the intuitionistic fuzzy empirical risk minimization principle is consistent for the set of functions $Q(\xi, \alpha)$, $\alpha \in \Lambda$ and for the distribution functional $F(\xi)$

Definition 3.5. If for any nonempty subset $\Lambda(C)$, $C \in IF(\mathbb{R})$ of this set of functions such that

$\Lambda(C) = \left\{ \alpha : \int Q(\xi, \alpha) dP \geq C \right\}$ the convergence

$$\inf_{\alpha \in \Lambda(C)} R_{ifemp}(\alpha) \xrightarrow[l \rightarrow \infty]{P} \inf_{\alpha \in \Lambda(C)} R_{if}(\alpha),$$

then the intuitionistic fuzzy empirical risk minimization principle is strictly consistent for the set of function $Q(\xi, \alpha)$, $\alpha \in \Lambda$ and the distribution function $F(\xi)$.

Remark 3.1. We only consider the sequence of intuitionistic fuzzy random variables that are partially ordered " $\leq$ ". We denote by $d\left(\left(R_{if}(\alpha), R_{ifemp}(\alpha)\right)\right)_{\geq}$ the Hausdorff metric

when $R_{ifemp}(\alpha) \leq R_{if}(\alpha)$.

Theorem 3.1 (The key theorem of learning theory based on intuitionistic fuzzy random samples). Suppose that there exist two intuitionistic fuzzy numbers W_1 and W_2 , such that for all functions in the set $Q(\xi, \alpha)$, $\alpha \in \Lambda$, and for a given distributed function $F(\xi)$, the inequalities $W_1 \leq \int Q(\xi, \alpha) dP \leq W_2$ hold true. Then the sufficient and necessary conditions for the strict consistency of IFERM principle is that the convergence

$$P\left\{\sup_{\alpha \in \Lambda} d\left(\left(R_{if}(\alpha), R_{ifemp}(\alpha)\right)\right)_{\geq} > \varepsilon\right\} \xrightarrow[l \rightarrow \infty]{} 0$$

holds for any $\varepsilon > 0$.

Proof. Necessary: By Definition 3.5 for the set $\Lambda(C) = \left\{ \alpha : R_{if}(\alpha) \geq C \right\}$, we have

$$\inf_{\alpha \in \Lambda(C)} R_{ifemp}(\alpha) \xrightarrow[l \rightarrow \infty]{P} \inf_{\alpha \in \Lambda(C)} R_{if}(\alpha) \quad (3.1)$$

We denote by A the event of the form $\sup_{\alpha \in \Lambda} d\left(\left(R_{if}(\alpha), R_{ifemp}(\alpha)\right)\right)_{\geq} > \varepsilon$.

Suppose that A holds. Then there exists $\alpha^* \in \Lambda(C)$ such that $d\left(\left(R_{if}(\alpha^*), R_{ifemp}(\alpha^*)\right)\right)_{\geq} > \varepsilon$. We

can find $W_1 \leq a_k \leq W_2$, such that $\alpha^* \in \Lambda(a_k)$ and $d\left(\left(R_{if}(\alpha^*), a_k\right)\right)_{\geq} < \frac{\varepsilon}{2}$. Then for these $\Lambda(a_k)$

the inequalities

$$d\left(\left(R_{if}(\alpha^*), \inf_{\alpha \in \Lambda(a_k)} R_{if}(\alpha)\right)\right)_{\geq} < \frac{\varepsilon}{2} \text{ and } R_{ifemp}(\alpha^*) \geq \inf_{\alpha \in \Lambda(a_k)} R_{ifemp}(\alpha)$$

hold true.

Therefore

$$\begin{aligned} & d\left(\left(\inf_{\alpha \in \Lambda(a_k)} R_{if}(\alpha), \inf_{\alpha \in \Lambda(a_k)} R_{ifemp}(\alpha)\right)\right)_{\geq} \\ & \geq d\left(\left(R_{if}(\alpha^*), \inf_{\alpha \in \Lambda(a_k)} R_{ifemp}(\alpha)\right)\right)_{\geq} - d\left(\left(R_{if}(\alpha^*), \inf_{\alpha \in \Lambda(a_k)} R_{if}(\alpha)\right)\right)_{\geq} > \frac{\varepsilon}{2}. \end{aligned}$$

According to (3.1), we have

$$P\left\{d\left(\left(\inf_{\alpha \in \Lambda(a_k)} R_{if}(\alpha), \inf_{\alpha \in \Lambda(a_k)} R_{ifemp}(\alpha)\right)\right)_{\geq} > \frac{\varepsilon}{2}\right\} \xrightarrow[l \rightarrow \infty]{} 0.$$

Denote by T_k the event $d\left(\left(\inf_{\alpha \in \Lambda(a_k)} R_{if}(\alpha), \inf_{\alpha \in \Lambda(a_k)} R_{ifemp}(\alpha)\right)\right)_{\geq} > \frac{\varepsilon}{2}$, then $A \subseteq \bigcup_k T_k$,

$$P(A) \leq P\left(\bigcup_k T_k\right) \xrightarrow[l \rightarrow \infty]{} 0.$$

290 Hence

$$291 \quad P\left\{\sup_{\alpha \in \Lambda} d\left(\left(R_{if}(\alpha), R_{ifemp}(\alpha)\right)_{\geq}\right) > \varepsilon\right\} \xrightarrow{l \rightarrow \infty} 0 \quad (3.2)$$

292 Sufficiency: We denote by A the event $\left\{\omega : d\left(\inf_{\alpha \in \Lambda(C)} R_{if}(\alpha), \inf_{\alpha \in \Lambda(C)} R_{ifemp}(\alpha)\right)_{\geq} > \varepsilon\right\}$

$$293 \quad A_1 = \left\{\omega : d\left(\left(\inf_{\alpha \in \Lambda(C)} R_{ifemp}(\alpha), \inf_{\alpha \in \Lambda(C)} R_{if}(\alpha)\right)_{\geq}\right) > \varepsilon\right\}$$

$$294 \quad A_2 = \left\{\omega : d\left(\left(\inf_{\alpha \in \Lambda(C)} R_{if}(\alpha), \inf_{\alpha \in \Lambda(C)} R_{ifemp}(\alpha)\right)_{\geq}\right) > \varepsilon\right\}.$$

295 Then $A = A_1 \cup A_2$, and $P(A) \leq P(A_1) + P(A_2)$.

296 Suppose that the event A_1 occurs, we can find $\mathcal{Q}(\xi, \alpha^*), \alpha^* \in \Lambda(C)$, such that

$$297 \quad d\left(\left(R_{if}(\alpha^*), \inf_{\alpha \in \Lambda(C)} R_{if}(\alpha)\right)_{\geq}\right) < \frac{\varepsilon}{2} \text{ and } R_{ifemp}(\alpha^*) \geq \inf_{\alpha \in \Lambda(C)} R_{ifemp}(\alpha).$$

298 By $d\left(\left(\inf_{\alpha \in \Lambda(C)} R_{ifemp}(\alpha), \inf_{\alpha \in \Lambda(C)} R_{if}(\alpha)\right)_{\geq}\right) > \varepsilon$, we have

$$299 \quad d\left(\left(R_{ifemp}(\alpha^*), \inf_{\alpha \in \Lambda(C)} R_{if}(\alpha)\right)_{\geq}\right) > \varepsilon.$$

300 Therefore

$$301 \quad d\left(\left(R_{ifemp}(\alpha^*), R_{if}(\alpha^*)\right)_{\geq}\right) > \frac{\varepsilon}{2}.$$

302 In virtue of the monotonicity of probability and the strong law of large numbers of
303 intuitionistic fuzzy random variables, we obtain

$$304 \quad P(A_1) \leq P\left\{d\left(\left(R_{ifemp}(\alpha^*), R_{if}(\alpha^*)\right)_{\geq}\right) > \frac{\varepsilon}{2}\right\} \xrightarrow{l \rightarrow \infty} 0.$$

305 On the other hand, if the event A_2 takes place, we can find $\mathcal{Q}(\xi, \alpha^{**}), \alpha^{**} \in \Lambda(C)$, such that

$$306 \quad d\left(\left(\inf_{\alpha \in \Lambda(C)} R_{if}(\alpha), R_{ifemp}(\alpha^{**})\right)_{\geq}\right) > \frac{\varepsilon}{2} \text{ and } R_{if}(\alpha^{**}) \geq \inf_{\alpha \in \Lambda(C)} R_{if}(\alpha).$$

307 Then

$$308 \quad d\left(\left(R_{if}(\alpha^{**}), R_{ifemp}(\alpha^{**})\right)_{\geq}\right) > \frac{\varepsilon}{2}.$$

309 Therefore

$$310 \quad \begin{aligned} P(A_2) &\leq P\left\{d\left(\left(R_{if}(\alpha^{**}), R_{ifemp}(\alpha^{**})\right)_{\geq}\right) > \frac{\varepsilon}{2}\right\} \\ &\leq P\left\{\sup_{\alpha \in \Lambda} d\left(\left(R_{if}(\alpha), R_{ifemp}(\alpha)\right)_{\geq}\right) > \frac{\varepsilon}{2}\right\} \xrightarrow{l \rightarrow \infty} 0. \end{aligned}$$

311 Hence

$$312 \quad P(A) \leq P(A_1) + P(A_2) \xrightarrow{l \rightarrow \infty} 0.$$

313 The theorem has been proved.

314 The samples become fuzzy random samples when $\nu_{\xi} = 1 - \mu_{\xi}$. Then we have

315 **Corollary 3.1** [26]. Let the set of functions $\mathcal{Q}(\xi, \alpha), \alpha \in \Lambda$, satisfies $X \leq R(\alpha) \leq Y$. Then the
316 sufficient and necessary condition for the strict consistency of FERM principle is the
317 convergence

$$\lim_{l \rightarrow \infty} P \left\{ \sup_{\alpha \in \Lambda} D \left(\left(R_f(\alpha), R_{femp}(\alpha) \right)_\geq \right) > \varepsilon \right\} = 0$$

is valid.

The samples become random samples in the usual sense when the samples are not represented as fuzzy sets, and let d be the subtraction of real numbers. Then we have

Corollary 3.2 [47]. Assume that there exist the constants a and A such that for all functions in the set $Q(\xi, \alpha), \alpha \in \Lambda$, and for a given distribution function $F(\xi)$, the inequalities

$$a \leq R(\alpha) \leq A, \alpha \in \Lambda$$

hold true. Then the following two statements are equivalent:

1. For the given distribution function $F(\xi)$, the empirical risk minimization method is strictly consistent on the set of functions $Q(\xi, \alpha), \alpha \in \Lambda$.

2. For the given distribution function $F(\xi)$, the uniform one-sided convergence of the means to their mathematical expectation takes place over the set of functions $Q(\xi, \alpha), \alpha \in \Lambda$.

330

331 4. BOUNDS OF THE RATE OF UNIFORM CONVERGENCE ON THE BASIS OF 332 INTUITIONISTIC FUZZY RANDOM SAMPLES

333

In statistical learning theory, the important conclusions about the relationship between empirical risk and actual risk form in the promotion of some boundary. They are essential to learning machine capacity analysis and in the development of new learning algorithms. An important part of the rate of convergence of learning processes the boundary of the generalization bounds. In this section, consensus convergence of learning process based on intuitionistic fuzzy random samples of our discussion. We consider the model where a set of intuitionistic fuzzy measurable functions $Q(\xi, \alpha), \alpha \in \Lambda$ contains a finite number N of elements $Q(\xi, \alpha_k), k = 1, 2, \dots, N$. At first, we introduce a basic theorem:

Theorem 4.1[25,26]. Let $\xi_j = \left\{ \left[\left(\xi_j \right)_r^-, \left(\xi_j \right)_r^+ \right] \mid r \in [0, 1] \right\}, j = 1, 2, \dots, n$ be a sequence of fuzzy

random variables, and $\left(\xi_j \right)_r^-, \left(\xi_j \right)_r^+$ have the limited variance and the same upper bounds M .

Then

$$P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \xi_j, E \xi_j \right) \geq \varepsilon \right\} \leq \frac{2M}{n\varepsilon^2}.$$

Theorem 4.2. Let $\xi_j = \left\{ \left\langle x, \mu_{\xi_j}(x), \nu_{\xi_j}(x) \right\rangle \mid x \in U \right\}, j = 1, 2, \dots, n$ be a sequence of intuitionistic

fuzzy random variables, $\left(\mu_{\xi_j} \right)_r^-, \left(\mu_{\xi_j} \right)_r^+$ and $\left(\nu_{\xi_j} \right)_r^-, \left(\nu_{\xi_j} \right)_r^+$ have the limited variance and the same upper bounds M_1 and M_2 , respectively. Then

$$P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \xi_j, E \xi_j \right) \geq \varepsilon \right\} \leq \frac{8(M_1 + M_2)}{n\varepsilon^2}.$$

Proof. According to Theorem 4.1, we have

$$\begin{aligned}
& P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \xi_j, E \xi_j \right) \geq \varepsilon \right\} = P \left\{ \max \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \mu_{\xi_j}, E \mu_{\xi_j} \right), d \left(\frac{1}{n} \sum_{j=1}^n \nu_{\xi_j}, E \nu_{\xi_j} \right) \right\} \geq \varepsilon \right\} \\
& \leq P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \mu_{\xi_j}, E \mu_{\xi_j} \right) + d \left(\frac{1}{n} \sum_{j=1}^n \nu_{\xi_j}, E \nu_{\xi_j} \right) \geq \varepsilon \right\} \\
& \leq P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \mu_{\xi_j}, E \mu_{\xi_j} \right) + d \left(\frac{1}{n} \sum_{j=1}^n \nu_{\xi_j}, E \nu_{\xi_j} \right) \geq \varepsilon \right\} \\
& \leq P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \mu_{\xi_j}, E \mu_{\xi_j} \right) \geq \frac{\varepsilon}{2}, \text{ or }, d \left(\frac{1}{n} \sum_{j=1}^n \nu_{\xi_j}, E \nu_{\xi_j} \right) \geq \frac{\varepsilon}{2} \right\} \\
& = P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \mu_{\xi_j}, E \mu_{\xi_j} \right) \geq \frac{\varepsilon}{2} \right\} + P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \nu_{\xi_j}, E \nu_{\xi_j} \right) \geq \frac{\varepsilon}{2} \right\} - \\
& P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \mu_{\xi_j}, E \mu_{\xi_j} \right) \geq \frac{\varepsilon}{2} \& d \left(\frac{1}{n} \sum_{j=1}^n \nu_{\xi_j}, E \nu_{\xi_j} \right) \geq \frac{\varepsilon}{2} \right\} \\
& \leq P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \mu_{\xi_j}, E \mu_{\xi_j} \right) \geq \frac{\varepsilon}{2} \right\} + P \left\{ d \left(\frac{1}{n} \sum_{j=1}^n \nu_{\xi_j}, E \nu_{\xi_j} \right) \geq \frac{\varepsilon}{2} \right\} \\
& \leq \frac{8M_1}{n\varepsilon^2} + \frac{8M_2}{n\varepsilon^2} = \frac{8(M_1 + M_2)}{n\varepsilon^2}
\end{aligned}$$

351

352 **Theorem 4.3.** Suppose that $Q(\xi, \alpha_k), k=1,2,\dots,N$ is a set of the functions, $E[Q(\xi, \alpha)]$
 353 exists and satisfies the conditions of Theorem 4.2. Then the following inequality

$$354 \quad d \left((R_{if}(\alpha_l), R_{ifemp}(\alpha_l))_{\geq} \right) \leq \sqrt{\frac{8N(M_1 + M_2)}{l\eta}}$$

355 holds with probability of at least $1 - \eta$.

356 **Proof.**

$$\begin{aligned}
357 \quad & P \left\{ \sup_{1 \leq k \leq N} d \left((R_{if}(\alpha_k), R_{ifemp}(\alpha_k))_{\geq} \right) > \varepsilon \right\} \leq \sum_{k=1}^N P \left\{ d \left((R_{if}(\alpha_k), R_{ifemp}(\alpha_k))_{\geq} \right) > \varepsilon \right\} \\
& \leq N \cdot \frac{8(M_1 + M_2)}{l\varepsilon^2}
\end{aligned}$$

$$358 \quad \text{Let } N \cdot \frac{8(M_1 + M_2)}{l\varepsilon^2} = \eta, \text{ we conclude that } \varepsilon = \sqrt{\frac{8N(M_1 + M_2)}{l\eta}}.$$

359 Therefore

$$360 \quad P \left\{ d \left((R_{if}(\alpha_l), R_{ifemp}(\alpha_l))_{\geq} \right) \leq \sqrt{\frac{8N(M_1 + M_2)}{l\eta}} \right\} \geq 1 - \eta.$$

361 **Theorem 4.4.** The inequality

$$362 \quad d \left((R_{if}(\alpha_l), R_{if}(\alpha_0))_{\geq} \right) \leq 2 \sqrt{\frac{2N(M_1 + M_2)}{l\eta}}$$

363 is satisfied with probability of at least $1 - 2\eta$.

364 **Proof.** Considering the properties of d and Theorem 4.3, we have

$$\begin{aligned}
& d\left(\left(R_f(\alpha_l), R_f(\alpha_0)\right)_\geq\right) \leq d\left(\left(R_f(\alpha_l), R_{femp}(\alpha_l)\right)_\geq\right) + d\left(R_{femp}(\alpha_l), R_f(\alpha_0)\right) \\
& \leq d\left(\left(R_f(\alpha_l), R_{femp}(\alpha_l)\right)_\geq\right) + d\left(\left(R_{femp}(\alpha_0), R_f(\alpha_0)\right)_\geq\right) \\
& \leq 2\sqrt{\frac{8N(M_1 + M_2)}{l\eta}}.
\end{aligned}$$

Let the following relationships hold

$$A: d\left(\left(R_f(\alpha_l), R_{femp}(\alpha_l)\right)_\geq\right) \leq \sqrt{\frac{8N(M_1 + M_2)}{l\eta}};$$

$$B: d\left(\left(R_{femp}(\alpha_0), R_f(\alpha_0)\right)_\geq\right) \leq \sqrt{\frac{8N(M_1 + M_2)}{l\eta}};$$

$$C: d\left(\left(R_f(\alpha_l), R_f(\alpha_0)\right)_\geq\right) \leq 2\sqrt{\frac{8N(M_1 + M_2)}{l\eta}},$$

and $P(A) \geq 1 - \eta; P(B) \geq 1 - \eta$.

We can conclude that C must hold true in virtue of A and B true, then $P(C) \geq P(AB)$.

Therefore

$$P(C) \geq P(AB) = 1 - P((AB)^c) \geq 1 - P(A^c) - P(B^c) \geq 1 - 2\eta.$$

The theorem is proved.

When the samples are the fuzzy random samples, let

$$R_f(\alpha_l) = d\left(\int Q(\xi, \alpha) dF(\xi), I_{\{0\}}\right) \text{ and } R_{femp}(\alpha_l) = d\left(\frac{1}{n} \sum_{k=1}^l Q(\xi_k, \alpha), I_{\{0\}}\right). \text{ Then we have}$$

Corollary 4.1[25,26]. Suppose that $Q(\xi, \alpha_k)$, $k = 1, 2, \dots, N$ is a set of functions

and $Q(\xi, \alpha_k)_r = \left[(Q(\xi, \alpha_k))_r^-, (Q(\xi, \alpha_k))_r^+ \right]$. If $(Q(\xi, \alpha_k))_r^-$, $(Q(\xi, \alpha_k))_r^+$ have the limited variance and the same upper bounds M , then

$$1) R_f(\alpha_l) - R_{femp}(\alpha_l) \leq \sqrt{\frac{2M}{l\eta}} \text{ holds true with probability of at least } 1 - \eta.$$

$$2) R_f(\alpha_l) - R_f(\alpha_0) \leq 2\sqrt{\frac{2M}{l\eta}} \text{ is satisfied with probability of at least } 1 - 2\eta.$$

Let d denote the subtraction of real numbers. We use Hoeffding's inequalities when the samples are the random samples in the usual sense. We have

Corollary 4.2[47]. If $\{Q(\xi, \alpha_k), \alpha_k \in \Lambda, k = 1, 2, \dots, N\}$ is bounded, i.e., $A \leq Q(\xi, \alpha_k) \leq B$, then

$$1) R(\alpha_l) - R_{emp}(\alpha_l) \leq (B - A) \sqrt{\frac{\ln N - \ln \eta}{2l}} \text{ is valid with probability of at least } 1 - \eta.$$

$$2) R(\alpha_l) - R(\alpha_0) \leq B \sqrt{\frac{\ln N - \ln \eta}{2l}} + (B - A) \sqrt{\frac{-\ln \eta}{2l}} \text{ holds true with probability of at least } 1 - 2\eta.$$

388

5. CONCLUSIONS

389

This paper discusses the intuitionistic fuzzy numbers and intuitionistic fuzzy random variables. We have proved the strong law of large numbers of intuitionistic fuzzy random variables and showed some related theorems and useful properties. Furthermore, based on

393

intuitionistic fuzzy random samples, we propose the principle of intuitionistic fuzzy empirical risk minimization of learning theory, prove the key theorem of learning theory based on intuitionistic fuzzy random samples, and discuss the bounds of the rate of uniform convergence of learning process. Altogether these findings have laid the foundation for further research in statistical learning theory involving intuitionistic fuzzy random samples. Further investigations might focus on such fundamental issues as intuitionistic fuzzy structural risk minimization and address the applied aspects such as support vector machines based on intuitionistic fuzzy random samples, etc.

ACKNOWLEDGEMENTS

The authors thank the anonymous referees for their valuable suggestions in improving this paper. This work is supported by the National Natural Science Foundation of China (Grant Nos. 61073121, 71271070, and 61375075) and the Natural Science Foundation of Hebei Province of China (Grant Nos. F2012201020 and A2012201033).

COMPETING INTERESTS

AUTHORS' CONTRIBUTIONS

'Zhiming Zhang' designed the study, performed the statistical analysis, wrote the protocol, and wrote the overall draft of the manuscript. The author read and approved the final manuscript.

REFERENCES

1. K. Atanassov, Intuitionistic fuzzy sets, *Fuzzy Sets and Systems* 20 (1986) 87–96.
2. K. Atanassov, More on intuitionistic fuzzy sets, *Fuzzy Sets and Systems* 33 (1989) 37–45.
3. Atanassov, New operations defined over the intuitionistic fuzzy sets, *Fuzzy Sets and Systems* 61 (1994) 137–142.
4. K. Atanassov, Remarks on the intuitionistic fuzzy sets – III, *Fuzzy Sets and Systems* 75 (1995) 401–402.
5. K. Atanassov, *Intuitionistic Fuzzy Sets: Theory and Applications*, Physica-Verlag, Heidelberg, 1999.
6. Zhaoqi Bian and Xuegong Zhang, *Pattern Recognition* (in Chinese), Tsinghua University Press, Beijing, 1999.
7. M. Brunato, R. Battiti, Statistical learning theory for location fingerprinting in wireless LANs. *Computer Networks* 47 (2005) 825–845.
8. P. Burillo, H. Bustince, Construction theorems for intuitionistic fuzzy sets, *Fuzzy Sets and Systems* 84 (1996) 271–281.
9. H. Bustince, P. Burillo, Structures on intuitionistic fuzzy relations, *Fuzzy Sets and Systems* 78 (1996) 293–303.
10. J. L. Castro, L. D. Flores-Hidalgo, C. J. Mantas, J. M. Puche, Extraction of fuzzy rules from support vector machines, *Fuzzy Sets and Systems* 158 (2007) 2057–2077.
11. G. C. Cawley and N. L. C. Talbot, Improved sparse least-squares support vector machines, *Neurocomputing*, 48 (1-4) (2002) 1025–1031.
12. A. Celikyilmaz, I. B. Turksen, Fuzzy functions with support vector machines, *Information Sciences* 177 (2007) 5163–5177.
13. M. D. Cock, C. Cornelis, E.E. Kerre, Intuitionistic fuzzy relational images, *Studies in Computational Intelligence* 2 (2005) 129–145.

- 444 14. D. Coker, An introduction to intuitionistic fuzzy topological spaces, *Fuzzy Sets and*
 445 *Systems* 88 (1997) 81–89.
- 446 15. C. Cornelis, G. Deschrijver, E.E. Kerre, Implication in intuitionistic fuzzy and interval-
 447 valued fuzzy set theory: construction, classification, application, *International Journal of*
 448 *Approximate Reasoning* 35 (2004) 55–95.
- 449 16. S.K. De, R. Biswas, A.R. Roy, Some operations on intuitionistic fuzzy sets, *Fuzzy Sets*
 450 *and Systems* 114 (2000) 477–484.
- 451 17. S.K. De, R. Biswas, A.R. Roy, Some operations on intuitionistic fuzzy sets, *Fuzzy Sets*
 452 *and Systems* 114 (2000) 477–484.
- 453 18. M. Demirci, Axiomatic theory of intuitionistic fuzzy sets, *Fuzzy Sets and Systems* 110
 454 (2000) 253–266.
- 455 19. G. Deschrijver, E.E. Kerre, On the position of intuitionistic fuzzy set theory in the
 456 framework of theories modelling imprecision, *Information Sciences* 177 (2007) 1860–1866.
- 457 20. T. Evgeniou, T. Poggio, M. Pontil, et al, Regularization and statistical learning theory for
 458 data analysis, *Computational Statistics & Data Analysis* 38 (2002) 421–432.
- 459 21. M. H. Ha, Y. C. Bai, P. Wang, et al, The key theorem and the bounds on the rate of
 460 uniform convergence of statistical learning theory on a credibility space, *Advances in Fuzzy*
 461 *Sets and Systems* 1 (2006) 143–172.
- 462 22. M. H. Ha, Z. F. Feng, S. J. Song, L. Q. Gao, The key theorem and the bounds on the
 463 rate of uniform convergence of statistical learning theory on quasi-probability spaces,
 464 *Chinese Journal of Computers* 31 (2008) 476–485.
- 465 23. M. H. Ha, Y. Li, J. Li, D. Z. Tian, The key theorem and the bounds on the rate of uniform
 466 convergence of learning theory on Sugeno measure space, *Science in China: Series F* 49
 467 (2006) 1–14.
- 468 24. M. H. Ha, W. Pedrycz, L. F. Zheng, The theoretical fundamentals of learning theory
 469 based on intuitionistic fuzzy random samples, *Fuzzy Sets and Systems* 178 (2009) 3240–
 470 3246.
- 471 25. Minghu Ha, Wenguang Tang and Yunchao Bai. The key theorem of learning theory
 472 about fuzzy samples, *Proceeding of the Second International Conferences on Machine*
 473 *Learning and Cybernetics*, Xian, 2–5 November 2003, 1213–1216. 3–7.
- 474 26. M. H. Ha, J. Tian, The theoretical foundations of statistical learning theory based on
 475 fuzzy number samples, *Information Sciences* 178 (2008) 3240–3246.
- 476 27. M. H. Ha, J. Tian, C. L. Zhou, The key theorem and bounds on the rate of convergence
 477 of statistical learning theory on the fuzzy samples, *Dynamics of Continuous, Discrete and*
 478 *Impulsive Systems, Series A: Mathematical Analysis* 14(S1) (2007) 72–76.
- 479 28. C. L. Huang, M.C. Chen, C.J. Wang, Credit scoring with a data mining approach based
 480 on support vector machines, *Expert System with Applications* 33 (2007) 847–856.
- 481 29. C. Hwang, D. H. Hong, K. H. Seok, Support vector interval regression machine for crisp
 482 input and output data, *Fuzzy Sets and Systems* 157 (2006) 1114–1125.
- 483 30. B. Jin, Y. C. Tang, Y. Q. Zhang, Support vector machines with genetic fuzzy feature
 484 transformation for biomedical data classification, *Information Sciences* 177 (2007) 476–489.
- 485 31. T. Kikuchi, S. Abe, Comparison between error correcting output codes and fuzzy support
 486 vector machines, *Pattern Recognition Letters* 26 (2005) 1937–1945.
- 487 32. Y. K. Kim, A strong law of large numbers for fuzzy random variables, *Fuzzy Sets and*
 488 *Systems* 111 (2000) 319–323.
- 489 33. D. Li, F. Shan, C. Cheng, On properties of four IFS operators, *Fuzzy Sets and Systems*
 490 154 (2005) 151–155.
- 491 34. Z. Z. Liang, P. F. Shi, Similarity measures on intuitionistic fuzzy sets, *Pattern Recognition*
 492 *Letters* 24 (2003) 2687–2693.

- 493 35. C. C. Lin, S. H. Chen, T. K. Truong and Y. Chang, Auto classification and categorization
494 based on wavelets and support vector machine, *IEEE Transactions on Speech and Audio*
495 *Processing*, 13 (5) (2005), 644-651.
- 496 36. C. F. Lin, S. D. Wang, Fuzzy support vector machines, *IEEE Transactions on Neural*
497 *Networks* 13 (2002) 464-471.
- 498 37. L. Lin, X. H. Yuan, Z. Q. Xia, Multicriteria fuzzy decision-making methods based on
499 intuitionistic fuzzy sets, *Journal of Computer and System Sciences* 73 (2007) 84-88.
- 500 38. B. D. Liu. *Uncertainty Theory: An Introduction to its Axiomatic Foundations*, Springer-
501 *Verlag*, Berlin, 2004.
- 502 39. Y. H. Liu, Y. T. Chen, Face recognition using total margin-based adaptive fuzzy support
503 vector machines, *IEEE Transactions on Neural Networks* 18 (2007) 178-192.
- 504 40. T. K. Mondal, S. K. Samanta, Topology of interval-valued intuitionistic fuzzy sets, *Fuzzy*
505 *Sets and Systems* 119 (2001) 483-494.
- 506 41. C. Su, C. H. Yang, Feature selection for the SVM: an application to hypertension
507 diagnosis, *Expert Systems with Applications* 34 (2008) 754-763.
- 508 42. D. Z. Tian, Z. M. Zhang, M. H. Ha, The key theorem of complex statistical learning theory,
509 *Dynamics of Continuous, Discrete and Impulsive Systems, Series A: Mathematical Analysis*
510 14(S1) (2007) 46-50.
- 511 43. K. Tomonori and A. Shigeo, Comparison between error correcting output codes and
512 fuzzy support vector machines, *Pattern Recognition Letters*, 26 (12) (2005), 1937-1945.
- 513 44. C. F. Tsai, Training support vector machines based on stacked generalization for image
514 classification, *Neurocomputing*, 64 (2005), 497-503.
- 515 45. J. J. Tsong, Hybrid approach of selecting hyper-parameters of support vector machine
516 for regression, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*
517 36 (2006) 699-709.
- 518 46. V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York,
519 1995.
- 520 47. V. N. Vapnik, *Statistical Learning Theory*, A Wiley-Interscience Publication, New York,
521 1998.
- 522 48. V. N. Vapnik, An overview of statistical learning theory, *IEEE Transactions on Neural*
523 *Networks* 10 (1999) 988-999.
- 524 49. M. Vidyasagar, Statistical learning theory and randomized algorithms for control, *IEEE*
525 *Control Systems Magazine* 12 (1998) 69-85.
- 526 50. Y. Q. Wang, S. Y. Wang, K. K. Lai, A new fuzzy support vector machine to evaluate
527 credit risk, *IEEE Transactions on Fuzzy Systems* 13 (2005) 820-931.
- 528 51. Z. Y. Wang, J. K. George, *Fuzzy Measure Theory*, Plenum Press, New York, 1992.
- 529 52. H. Wechsler, Z. Duric, F. Y. Li, V. Cherkassky, Motion estimation using statistical
530 learning theory, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26 (2004)
531 466-478.
- 532 53. W. Y. N. Wing, A. Dorado, D. S. Yeung, W. Pedrycz and E. Izquierdo, Image
533 classification with the use of radial basis function neural networks and the minimization of
534 the localized generalization error, *Pattern Recognition*, 40 (2007), 19 -32.
- 535 54. Z.S. Xu, Intuitionistic preference relations and their application in group decision making,
536 *Information Sciences* 177 (2007) 2363-2379.
- 537 55. L.A. Zadeh, Fuzzy sets, *Information and Control* 8 (1965) 338-353.
- 538 56. Yiqiang Zhan and Dinggang Shen, Design efficient support vector machine for fast
539 classification, *Pattern Recognition*, 38 (1) (2005), 157-161.
- 540 57. C. Y. Zhang, H. Y. Fu, Similarity measures on three kinds of fuzzy sets, *Pattern*
541 *Recognition Letters* 27 (2006) 1307-1317.
- 542 58. Atanassov, K. Intuitionistic fuzzy sets, VII ITKR's Session, Sofia, June 1983 (Deposed
543 in Central Sci. -Techn. Library of Bulg. Acad. of Sci., 1697/84) (in Bulg.)

- 544 59. Atanassov, K. Remark on intuitionistic fuzzy numbers. Notes on Intuitionistic Fuzzy Sets,
545 Vol. 13, 2007, No. 3, 29-32.