International Journal of Theoretical Physics, Vol. 41, No. 4, April 2002 (© 2002)

On Fuzzy Probability Theory

E. D. Habil[1,2] and T. Z. Nasr[1]

Received March 16, 2001

Our main aim from this work is to see which theorems in classical probability theory are still valid in fuzzy probability theory. Following Gudder's approach [Demonestratio Mathematica **31**(3), 1998, 235–254; Foundations of Physics, **30**, 1663–1678] to fuzzy probability theory, the basic concepts of the theory, that is of fuzzy probability measures and fuzzy random variables (observables), are presented. We show that fuzzy random variables extend the usual ones. Moreover, we prove that for any separable metrizable space, the crisp observables coincide with random variables. Then we prove the existence of a joint observable for any collection of observables, and we prove the weak law of large numbers and the central limit theorem in the fuzzy context. We construct a new definition of almost everywhere convergence. After proving that Gudder's definition implies ours and presenting an example that indicates that the converse is not true, we prove the strong law of large numbers according to this definition.

KEY WORDS: fuzzy sets; fuzzy probability theory; observables; quantum mechanics; σ-morphisms.

1. INTRODUCTION

Fuzzy set theory was born in 1965 in a paper by Zadeh (1965). He introduced the notion of a fuzzy set to describe situations in which certain objects belong to a set "to some extent." In such a way he opened a possibility of studying sets the boundaries of which vanish gradually. Such situations are encountered mainly in "soft" sciences, e.g., psychology, economics, medicine, etc. (Pykacz, 1992).

Beltrametti and Bugajski's approach to fuzzy probability theory is based on the physical foundation of the theory, as they indicated in their papers (Beltrametti and Bugajski, 1995a,b). Their representation of quantum mechanics describes quantum observables by means of fuzzy random variables on a measurable space consisting of quantum-mechanical pure states. Moreover, Bugajski (1996, 1998b) establishes the mathematical foundation of the theory in his papers. Gudder's

[1] Department of Mathematics, Islamic University of Gaza, Gaza, Palestine.
[2] To whom correspondence should be addressed at Department of Mathematics, Islamic University of Gaza, P.O. Box 108, Gaza, Palestine; e-mail: habil@mail.iugaza.edu.

treatment of the theory is slightly different from that of Bugajski and Beltrametti (see Gudder, 1998, 1999, 2000). In this paper, we will consider Gudder's axiomatic quantum mechanics approach to fuzzy probability theory, which is based on the space of all measurable fuzzy sets, and it proved to be very useful not only in quantum mechanics but also in computer science (Dunyak *et al.*, 1999; Gudder, 1998; Ishikawa, 1996). In fact, the most striking feature of standard quantum mechanics is that it rarely provides joint observables for pairs of quantum observables (Bugajski, 1998a), but as we will see, any two fuzzy random variables always have a joint fuzzy random variable.

The most essential difference between standard and fuzzy probability theory lies in the notion of random variables they adopt (Bugajski, 1996). Our main purpose is to clarify the relation between fuzzy and classical probability theory, and to see which of the theorems in classical probability theory are still valid in the new developing field of fuzzy probability theory. Gudder (1998, 2000) addressed this theme in his papers, and we completed some of the work that he mentioned in these papers. We prove Theorem 4.8 which, besides its theoretical interest, is used in proving the weak law of large numbers (Theorem 5.5) and the central limit theorem (Theorem 5.15). The strong law of large numbers (Theorem 5.14) is proved with respect to a weaker version than Gudder's standard definition of almost everywhere convergence for fuzzy observables.

Our proofs of these results are slightly different from those of Gudder (1998, 1999, 2000). The idea is that we use the corresponding theorems in classical probability theory after representing a sequence of observables by a sequence of random variables in a larger probability space. In fact, we can prove the theorems directly but our proofs are much easier. Moreover, in this way the relation between the classical and fuzzy concepts can be detected easily.

2. FUZZY SETS

Let Ω be a nonempty set. In fuzzy set theory, subsets of Ω are replaced by fuzzy sets where the fuzzy sets are defined as follows.

Definition 2.1 (Gudder, 2000). A *fuzzy subset* f of Ω is a function $f : \Omega \to [0, 1]$. We say that a fuzzy set f is *crisp* iff f is an indicator function; that is, $f = I_A$ for some $A \subseteq \Omega$, where

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A. \end{cases}$$

We say $f \subseteq g$ if $f(\omega) \leq g(\omega)$ for any $\omega \in \Omega$. We identify any set $A \subseteq \Omega$ with its indicator function I_A. Hence the system of all fuzzy subsets $[0, 1]^\Omega$ can be treated as a power set, and crisp fuzzy sets correspond to the usual sets (Gudder, 2000). We thus say that a fuzzy set is a generalization of a set. For a measurable space $(\Omega, \mathcal{A})$, a random variable $f : \Omega \to [0, 1]$ is called *an effect or a fuzzy event*. Thus, an effect is just a measurable fuzzy subset of Ω. The set of all effects is denoted by $\mathcal{E}(\Omega, \mathcal{A})$. For $f, g \in \mathcal{E}(\Omega, \mathcal{A})$, we define $f' := 1 - f$, $f \cap g := f \cdot g$ and $f \cup g := f + g - f \cdot g$.

Lemma 2.2 (Gudder, 1998). *Let $f_n \in \mathcal{E}(\Omega, \mathcal{A})$, $n \in \mathbb{N}$. Then $\bigcup_{n=1}^{\infty} f_n$ exists in $\mathcal{E}(\Omega, \mathcal{A})$ and*

$$\bigcup_{n=1}^{\infty} f_n = 1 - \prod_{n=1}^{\infty}(1 - f_n).$$

Definition 2.3. If μ is a probability measure on $(\Omega, \mathcal{A})$ and $f \in \mathcal{E}(\Omega, \mathcal{A})$, we define the *probability* of f to be its expectation $\mu(f) = \int f \, d\mu$.

Definition 2.4 (Gudder, 2000). Let $(\Omega, \mathcal{A})$ and $(\Lambda, \mathcal{B})$ be measurable spaces. A mapping $\phi : \mathcal{E}(\Omega, \mathcal{A}) \to \mathcal{E}(\Lambda, \mathcal{B})$ is called a *σ-morphism* if

(i) $\phi(I_\Omega) = I_\Lambda = 1$ and
(ii) if $f_i \in \mathcal{E}(\Omega, \mathcal{A})$ is a sequence such that $\sum f_i \in \mathcal{E}(\Omega, \mathcal{A})$, then

$$\phi\left(\sum f_i\right) = \sum \phi(f_i).$$

Definition 2.5 (Gudder, 2000). If $(\Lambda, \mathcal{B})$ is a measurable space, a *$\mathcal{B}$-observable* on $(\Omega, \mathcal{A})$ is a map $X : \mathcal{B} \to (\Omega, \mathcal{A})$ such that

(1) $X(\Lambda) = 1$ and
(2) if $B_i \in \mathcal{B}$ are mutually disjoint, then $X(\bigcup B_i) = \sum X(B_i)$ where the convergence of the summation is pointwise.

Let $\mathfrak{B}(\mathbb{R})$ denote the σ-algebra of all Borel subsets of $\mathbb{R}$. A $\mathfrak{B}(\mathbb{R})$-observable on $(\Omega, \mathcal{A})$ is simply called an *observable* on $(\Omega, \mathcal{A})$. For an observable X, if $X(B)$ is crisp for every $B \in \mathcal{B}$, then X is called *crisp*.

Theorem 2.6 (Gudder, 2000). *If $X : \mathcal{B} \to \mathcal{E}(\Omega, \mathcal{A})$ is a $\mathcal{B}$-observable, then X has a unique extension to a σ-morphism $\tilde{X} : \mathcal{E}(\Lambda, \mathcal{B}) \to \mathcal{E}(\Omega, \mathcal{A})$. If $Y : \mathcal{E}(\Lambda, \mathcal{B}) \to \mathcal{E}(\Omega, \mathcal{A})$ is a σ-morphism, then $Y \mid \mathcal{B}$ is a $\mathcal{B}$-observable.*

In fact, if $X : \mathcal{B} \to \mathcal{E}(\Omega, \mathcal{A})$ is an observable, then the corresponding σ-morphism $\tilde{X} : \mathcal{E}(\Lambda, \mathcal{B}) \to \mathcal{E}(\Omega, \mathcal{A})$ has the form $\tilde{X}(f)(\omega) := \int f(\lambda)\mu_\omega(d\lambda)$, where $\mu_\omega : \mathcal{B} \to [0, 1]$ is the probability measure defined by $\mu_\omega(B) = X(B)(\omega)$ for every $B \in \mathcal{B}$.

3. OBSERVABLES

It should be noted that any random variable $\hat{X} : \Omega \to \mathbb{R}$ generates a (crisp) observable $X : \mathfrak{B}(\mathbb{R}) \to \mathcal{E}(\Omega, \mathcal{A})$ given by $X(B) = \hat{X}^{-1}(B)$; i.e., $X(B) = I_{\hat{X}^{-1}(B)}$. But the converse does not hold. However, we give in Theorem 3.2 below sufficient conditions for a crisp observable to represent a random variable. We present here some noncrisp observables.

Example 3.1. (1) For any noncrisp $f \in \mathcal{E}(\Omega, \mathcal{A})$, define $X_f : \mathfrak{B}(\mathbb{R}) \to \mathcal{E}(\Omega, \mathcal{A})$ by

$$X_f(B) := \begin{cases} 0 & \text{if } \{0, 1\} \cap B = \emptyset \\ f & \text{if } \{0, 1\} \cap B = \{1\} \\ 1 - f & \text{if } \{0, 1\} \cap B = \{0\} \\ 1 & \text{if } \{0, 1\} \subseteq B. \end{cases}$$

Then X_f is a noncrisp observable.

(2) Let f, g be random variables on a measurable space $(\Omega, \mathcal{A})$, let $\lambda \in (0, 1)$, and let X_f, X_g be the observables generated by f, g; i.e.,

$$X_f(B) = I_{f^{-1}(B)}, \qquad X_g(B) = I_{g^{-1}(B)} \quad \forall B \in \mathfrak{B}(\mathbb{R}).$$

Define $Y : \mathfrak{B}(\mathbb{R}) \to \mathcal{E}(\Omega, \mathcal{A})$ by

$$Y(B) := \lambda X_f(B) + (1 - \lambda) X_g(B).$$

Then Y is an observable which is not crisp.

Theorem 3.2. *If Λ is a separable metrizable space and $\mathcal{B}$ is the σ-algebra of all Borel subsets of Λ, and X is a crisp $\mathcal{B}$-observable, then there exists a measurable function $f : \Omega \to \Lambda$ such that $X(B) = I_{f^{-1}(B)}$ for all $B \in \mathcal{B}$.*

Proof: Since $\mathcal{B}$ is the σ-algebra generated by a separable metrizable topology τ on Λ, there exists a metric ρ on Λ defining τ and τ is separable. Hence, there exists a countable dense subset D of Λ. For each $n \in \mathbb{N}$, define

$$\mathcal{B}_n := \left\{ B : B = U_\rho\left(d, \frac{1}{4n}\right), d \in D \right\},$$

where $U_\rho(d, \epsilon)$ denotes the open ball centered at d with radius ϵ. Then $\bigcup_{n=k}^{\infty} \mathcal{B}_n$, $k \in \mathbb{N}$ is a base for τ. Fix $\omega \in \Omega$ and define the measure μ_ω on $(\Lambda, \mathcal{B})$ by $\mu_\omega(B) := X(B)(\omega)$. Then construct a sequence of open balls as follows.

(1) Since $\Lambda = \bigcup_{B_i \in \mathcal{B}_1} B_i$, we have

$$1 = X(\Lambda)(\omega) = X\left(\bigcup_{B_i \in \mathcal{B}_1} B_i\right)(\omega) \leq \sum_{B_i \in \mathcal{B}_1} X(B_i)(\omega).$$

Hence, as X is crisp, there exists $B_{i_1} \in \mathcal{B}_1$ such that $X(B_{i_1})(\omega) = 1$. Let $A_1 := B_{i_1}$.

(2) Similarly, B_{i_1} is open and $\bigcup_{n=2}^{\infty} \mathcal{B}_n$ is a base for τ; hence, it covers B_{i_1}. So letting $\mathcal{B}_2' := \{B \in \bigcup_{n=2}^{\infty} \mathcal{B}_n\colon B \subseteq B_{i_1}\}$, there exists $B_{i_2} \in \mathcal{B}_2'$ such that $X(B_{i_2})(\omega) = 1$. Let $A_2 := B_{i_2}$.

(3) Continuing in the same fashion, we get a sequence of open balls $\{A_i\}$ that satisfies $A_{i+1} \subseteq A_i$, $X(A_i)(\omega) = 1\ \forall i \in \mathbb{N}$, and $\rho(a, b) \leq \frac{1}{4i}\ \forall a, b \in A_i$.

Now $\mu_\omega(\bigcap_{i=1}^{\infty} A_i) = \lim \mu_\omega(A_i) = 1$. Hence $(\bigcap_{i=1}^{\infty} A_i) \neq \emptyset$. In fact, $\bigcap_{i=1}^{\infty} A_i$ is a singleton subset of Λ since if $\gamma_1, \gamma_2 \in \bigcap_{i=1}^{\infty} A_i$, $\gamma_1 \neq \gamma_2$, and $l := \rho(\gamma_1, \gamma_2)$, then $l > 0$; so pick $n \in \mathbb{N}$ such that $\frac{l}{2} > \frac{1}{n}$. Since $\bigcap_{i=1}^{\infty} A_i \subseteq A_n$, we get $\rho(\gamma_1, \gamma_2) \leq \frac{1}{n} < \frac{l}{2}$, a contradiction. Thus $\exists \gamma_\omega \in \Lambda \ni \bigcap A_i = \{\gamma_\omega\}$.

We have proved that $\forall\, \omega \in \Omega$, $\exists \gamma_\omega \in \Lambda$ such that $X(\{\gamma_\omega\})(\omega) = 1$. Note that such γ_ω is unique with respect to the property that $X(\{\gamma_\omega\})(\omega) = 1$, since if there exist two such elements $\gamma_1, \gamma_2 \in \Lambda$ corresponding to the same ω, then $1 \geq X(\{\gamma_1, \gamma_2\})(\omega) = X(\{\gamma_1\})(\omega) + X(\{\gamma_2\})(\omega) = 2$, a contradiction. Now define $f : \Omega \to \Lambda$ such that $f(\omega) = \gamma_\omega$. Then f is measurable. To prove this, let $B \in \mathcal{B}$. Then $f^{-1}(B) = \{\omega\colon X(B)(\omega) = 1\}$, since if $\omega \in f^{-1}(B)$ for some $\omega \in \Omega$ then $f(\omega) \in B$ and hence $X(B)(\omega) \geq X(\{f(\omega)\})(\omega) = 1$. Therefore $\omega \in \{\omega\colon X(B)(\omega) = 1\}$. On the other hand, if $\omega \in \{\omega\colon X(B)(\omega) = 1\}$, then $f(\omega) \in B$ since otherwise, $X(B \cup \{f(\omega)\}) = 2$, a contradiction. Hence $f^{-1}(B) = \{\omega\colon X(B)(\omega) = 1\}$, which is measurable since $X(B)$ is measurable.

To prove that $X(B) = I_{f^{-1}(B)}\ \forall B \in \mathcal{B}$, let $\omega \in \Omega$. Then $X(B)(\omega)$ is either 0 or 1 since X is crisp. Now if $X(B)(\omega) = 0$, then $\omega \notin f^{-1}(B)$ and $I_{f^{-1}(B)}(\omega) = 0$, and if $X(B)(\omega) = 1$, then $\omega \in f^{-1}(B)$ and $I_{f^{-1}(B)}(\omega) = 1$. □

A separable topological space is called *Polish* if there is a complete metric defining its topology. Thus, as a consequence of the above theorem, we obtain the following result, which appears in Bugajski (1998b).

Corollary 3.3 (Bugajski, 1998b). *If Λ is a Polish space and $\mathcal{B}$ is the σ-algebra of all Borel subsets of Λ, and X is a crisp $\mathcal{B}$-observable, then there exists a measurable function $f : \Omega \to \Lambda$ such that $X(B) = I_{f^{-1}(B)}$.*

The condition that the space Λ be separable and metrizable in Theorem 3.2 is essential to identify crisp observables with random variables, as the following example shows.

Example 3.4. Let Λ be an uncountable set and let τ be the cocountable topology on Λ. It is well known that τ is neither separable nor metrizable. Let $\mathcal{A}$ be the σ-algebra generated by τ. Then $\mathcal{A}$ is the set of all subsets of Λ that are either countable or whose complements are countable. Define $X : \mathcal{A} \to \mathcal{E}(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ by $X(B) := I_\emptyset$ for a countable $B \in \mathcal{A}$ and $X(B) := I_\mathbb{R}$ for an uncountable $B \in \mathcal{A}$. Although it can be easily checked that X is a crisp observable, there is no random variable that corresponds to it. Indeed, if there exists a measurable function $f : \mathbb{R} \to \Lambda$ such that $X(B) = I_{f^{-1}(B)}\ \forall B \in \mathcal{A}$, then $f^{-1}(B)$ is either $\mathbb{R}$ or $\emptyset\ \forall B \in \mathcal{A}$, which implies that f must be of the form $f(\omega) = c\ \forall \omega \in \mathbb{R}$ for some $c \in \Lambda$. Hence $X(\{c\}) = I_\mathbb{R}$, a contradiction.

Let $(\Omega, \mathcal{A}), (\Lambda_1, \mathcal{B}), (\Lambda_2, \mathcal{B}')$ be measurable spaces and let $Y : \mathcal{B} \to \mathcal{E}(\Lambda_2, \mathcal{B}')$ and $X : \mathcal{B}' \to \mathcal{E}(\Omega, \mathcal{A})$ be observables. In classical probability theory, we can compose any two random variables easily under composition of functions. In fuzzy probability theory, we can compose any two observables X and Y if they are thought of as σ-morphisms (Gudder, 1999). Doing this, we have the σ-morphism $X \circ Y : \mathcal{E}(\Lambda_1, \mathcal{B}) \to \mathcal{E}(\Omega, \mathcal{A})$ which is identified with the observable $X \circ Y : \mathcal{B} \to \mathcal{E}(\Omega, \mathcal{A})$. We then have

$$(X \circ Y)(B)(\omega) = [\tilde{X}(Y(B))](\omega) = \int Y(B)(\lambda)\mu_\omega(d\lambda) \qquad (3.1)$$

where μ_ω is the probability measure defined by the observable X, and $\tilde{X}$ is the unique σ-morphism that extends the observable X to $\mathcal{E}(\Lambda_2, \mathcal{B}')$ as given in Theorem 2.6.

4. JOINT OBSERVABLES

In this section, we continue to present some basic important facts about observables. The major result of this section is Theorem 4.8, which generalizes Theorem 4.1 of Gudder (1998) and will be used in the next section in proving the weak law of large numbers and the central limit theorem.

Definition 4.1 (Gudder, 1998). Let $(\Omega, \mathcal{A})$ and $(\Lambda, \mathcal{B})$ be measurable spaces. If μ is a probability measure on $(\Omega, \mathcal{A})$ and X is a $\mathcal{B}$-observable on $(\Omega, \mathcal{A})$, then *the distribution of* X is the probability measure μ_X on $\mathcal{B}$ given by

$$\mu_X(B) := \mu(X(B)).$$

Note that $\mu_X(B)$ is interpreted as the probability that X has a value in B when the system is in the state μ.

Definition 4.2 (Gudder, 1998). Let X be an observable on $(\Omega, \mathcal{A})$ and let $u : \mathbb{R} \to \mathbb{R}$ be a Borel function. We define

(1) the observable $u(X) : \mathfrak{B}(\mathbb{R}) \to \mathcal{E}(\Omega, \mathcal{A})$ by $u(X)(B) = X(u^{-1}(B))$,
(2) the *expectation* of X by $E(X) = \int \lambda \mu_X(d\lambda)$, and
(3) the *variance* of X by

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \int \lambda^2 \mu_X(d\lambda) - \left[\int \lambda \mu_X(d\lambda)\right]^2.$$

Lemma 4.3 (Chebyshev; Gudder, 1998). *Let X be an observable on a probability space $(\Omega, \mathcal{A}, P)$ and $u : \mathbb{R} \to \mathbb{R}$ be an increasing nonnegative function. If $u(\lambda_\circ) > 0$, then*

$$P(X \geq \lambda_\circ) = P[X([\lambda_\circ, \infty))] \leq \frac{E(u(X))}{u(\lambda_\circ)}.$$

Definition 4.4. Let $X_1, \ldots, X_n$ be observables on $(\Omega, \mathcal{A})$. We say that a $\mathfrak{B}(\mathbb{R}^n)$-observable X on $(\Omega, \mathcal{A})$ is their *joint observable* if

$$\pi_i(X) = X_i, \quad i = 1, \ldots, n,$$

where π_i is the marginal projection map. For finite collections of observables, we have the following theorem.

Theorem 4.5 (Gudder, 1998). *If $X_1, \ldots, X_n$ are observables on $(\Omega, \mathcal{A})$, then there exists a unique n-dimensional observable Z on $(\Omega, \mathcal{A})$ such that for all $B_1, \ldots, B_n \in \mathfrak{B}(\mathbb{R})$,*

$$Z(B_1 \times \cdots \times B_n) = X_1(B_1) \cdots X_n(B_n). \tag{4.1}$$

Note that condition 4.1 is essential for the uniqueness of the joint observable Z, as the following example indicates. This example can be found in Bugajski (1996), and it has a direct connection to the quantum mechanical description of spin-$\frac{1}{2}$ objects.

Example 4.6. Let Ω denote the set of points of the unit sphere in $\mathbb{R}^3$ and let $\omega_1, \omega_2 \in \Omega$. Define $\mathfrak{B}(\mathbb{R})$-observables X_{ω_i}, $i = 1, 2$ on $(\Omega, \mathfrak{B}(\Omega))$ by

$$X_{\omega_i}(B)(\omega) := \begin{cases} 0 & \text{if } \frac{1}{2} \notin B, -\frac{1}{2} \notin B \\ \frac{1}{2}(1 + r_{\omega_i} \cdot r_\omega) & \text{if } \frac{1}{2} \in B, -\frac{1}{2} \notin B \\ \frac{1}{2}(1 - r_{\omega_i} \cdot r_\omega) & \text{if } -\frac{1}{2} \in B, \frac{1}{2} \notin B \\ 1 & \text{if } \frac{1}{2}, -\frac{1}{2} \in B, \end{cases}$$

where r_ω is the unit vector of $\mathbb{R}^3$ pointing to ω. Now define a $\mathfrak{B}(\mathbb{R}^2)$-observable X on $(\Omega, \mathfrak{B}(\Omega))$ generated by

$$X\left(\left\{\left(\frac{1}{2}, \frac{1}{2}\right)\right\}\right)(\omega) = \lambda(\omega),$$

$$X\left(\left\{\left(\frac{1}{2}, -\frac{1}{2}\right)\right\}\right)(\omega) = \frac{1}{2}\left(1 + r_{\omega_1} \cdot r_\omega\right) - \lambda(\omega),$$

$$X\left(\left\{\left(-\frac{1}{2}, \frac{1}{2}\right)\right\}\right)(\omega) = \frac{1}{2}\left(1 + r_{\omega_2} \cdot r_\omega\right) - \lambda(\omega),$$

$$X\left(\left\{\left(-\frac{1}{2}, -\frac{1}{2}\right)\right\}\right)(\omega) = \lambda(\omega) - \frac{1}{2}\left(r_{\omega_1} + r_{\omega_2}\right) \cdot r_\omega,$$

$$X\left(\mathbb{R}^2 \Big\backslash \left\{\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, -\frac{1}{2}\right), \left(-\frac{1}{2}, \frac{1}{2}\right), \left(-\frac{1}{2}, -\frac{1}{2}\right)\right\}\right)(\omega) = 0,$$

where $\lambda(\omega)$ may be one of the following two functions:

(1) $\lambda_1(\omega) = \frac{1}{4}(1 + r_{\omega_1} \cdot r_\omega)(1 + r_{\omega_2} \cdot r_\omega)$.
(2) $\lambda_2(\omega) = \min\{\frac{1}{2}(1 + r_{\omega_1} \cdot r_\omega), \frac{1}{2}(1 + r_{\omega_2} \cdot r_\omega)\}$.

In each case, X is a joint observable of $X_{\omega_1}, X_{\omega_2}$. Hence we have two different joint observables for the same observables $X_{\omega_1}, X_{\omega_2}$.

Definition 4.7 (Gudder, 1998). Let $X_1, \ldots, X_n$ be observables on a probability space $(\Omega, \mathcal{A}, P)$. Then the probability measure $\mu_{X_1,\ldots,X_n}$ on $\mathfrak{B}(\mathbb{R}^n)$ given by

$$\mu_{X_1,\ldots,X_n}(B) := \mu_Z(B) = P(Z(B)),$$

where Z is given by Eq. (4.1), is called the *joint distribution* of $X_1, \ldots, X_n$.

We now extend Theorem 4.5 to any collection of observables. But first we need to establish some notation about infinite product of probability spaces. Let $\{(\Omega_\alpha, \mathcal{A}_\alpha, \mu_\alpha): \alpha \in \Delta\}$ be a family of probability spaces and $\Omega = \prod \Omega_\alpha$. Recall (see Bauer, 1981) that the *product σ-algebra* $\mathcal{U}$ of the σ-algebras $\{\mathcal{A}_\alpha: \alpha \in \Delta\}$ is the smallest σ-algebra on Ω with respect to which each of the projection mappings π_α is $\mathcal{U}$-$\mathcal{A}_\alpha$-measurable. We denote the product σ-algebra $\mathcal{U}$ by

$$\mathcal{U} = \bigotimes_{\alpha \in \Delta} \mathcal{A}_\alpha.$$

By Theorem 5.4.2 of Bauer (1981), there exists a unique probability measure μ on $\mathcal{U}$ such that

$$\mu\left(\prod_{i=1}^{n} B_{\alpha_i} \times \prod_{\alpha \neq \alpha_1,\ldots,\alpha_n} \Omega_\alpha\right) = \mu_{\alpha_1}\left(B_{\alpha_1}\right) \cdots \mu_{\alpha_n}(B\alpha_n)$$

for all $B_{\alpha_i} \in \mathcal{A}_{\alpha_i}$, $i = 1, \ldots, n$, and for all $n \in \mathbb{N}$. The measure μ is called the *product measure* of the probability measures $\{\mu_\alpha : \alpha \in \Delta\}$ and is denoted by

$$\mu = \bigotimes_{\alpha \in \Delta} \mu_\alpha.$$

Theorem 4.8. *Let $\mathcal{U}$ be the product σ-algebra $\bigotimes_{\alpha\in\Delta} \mathfrak{B}(\mathbb{R})$ for some index set Δ. If X_α, $\alpha \in \Delta$ are observables on a measurable space $(\Omega, \mathcal{A})$, then there exists a unique $\mathcal{U}$-observable Z on $(\Omega, \mathcal{A})$ such that*

$$(\star) \qquad Z\left(\prod_{i=1}^{n} B_{\alpha_i} \times \prod_{\alpha \neq \alpha_1,\ldots,\alpha_n} \mathbb{R}_\alpha\right) = X_{\alpha_1}\big(B_{\alpha_1}\big) \cdots X_{\alpha_n}\big(B_{\alpha_n}\big),$$

where $\mathbb{R}_\alpha = \mathbb{R}$ for all $\alpha \in \Delta$.

Proof: Fix $\omega \in \Omega$. For each $\alpha \in \Delta$, define a probability measure $\mu_{\omega,\alpha}$ on $\mathfrak{B}(\mathbb{R})$ by

$$\mu_{\omega,\alpha}(B) := X_\alpha(B)(\omega) \quad \forall B \in \mathfrak{B}(\mathbb{R}).$$

Then, by Theorem 5.4.2 of Bauer (1981), there exists a unique probability measure $\mu_\omega := \bigotimes_{\alpha\in\Delta} \mu_{\omega,\alpha}$, the product measure on $\mathcal{U} = \bigotimes_{\alpha\in\Delta} \mathfrak{B}(\mathbb{R})$, such that for all $B_{\alpha_1}, \ldots, B_{\alpha_n} \in \mathfrak{B}(\mathbb{R})$, we have

$$\mu_\omega\left(\prod_{i=1}^{n} B_{\alpha_i} \times \prod_{\alpha \neq \alpha_1,\ldots,\alpha_n} \mathbb{R}_\alpha\right) = \mu_{\omega,\alpha_1}\big(B_{\alpha_1}\big) \cdots \mu_{\omega,\alpha_n}\big(B_{\alpha_n}\big). \tag{4.2}$$

Now define a mapping $Z : \mathcal{U} \to \mathcal{E}(\Omega, \mathcal{A})$ by $Z(B)(\omega) := \mu_\omega(B)\, \forall B \in \mathcal{U}$. To prove that Z is an observable, we need to prove, first, that $Z(B)$ is an effect $\forall B \in \mathcal{U}$. It is clear that $Z(B)(\omega) \in [0, 1]$ $\forall \omega \in \Omega$. To prove that $Z(B)$ is measurable, let $\mathcal{B} := \{B \in \mathcal{U} : Z(B) \text{ is measurable}\}$, and let

$$\mathcal{S} := \left\{\prod_{i=1}^{n} B_{\alpha_i} \times \prod_{\alpha \neq \alpha_1,\ldots,\alpha_n} \mathbb{R}_\alpha : B_{\alpha_i} \in \mathfrak{B}(\mathbb{R}) \text{ for all } \alpha_1, \ldots, \alpha_n \in \Delta\right\}.$$

It is well known that $\mathcal{S}$ generates $\mathcal{U}$. We also have that $\mathcal{S} \subseteq \mathcal{B}$. In fact, if $B = \prod_{i=1}^{n} B_{\alpha_i} \times \prod_{\alpha \neq \alpha_1,\ldots,\alpha_n} \mathbb{R}_\alpha \in \mathcal{S}$, then we have from Eq. (4.2) that

$$\begin{aligned} Z(B)(\omega) = \mu_\omega(B) &= \mu_{\omega,\alpha_1}\big(B_{\alpha_1}\big) \cdots \mu_{\omega,\alpha_n}\big(B_{\alpha_n}\big) \\ &= X_{\alpha_1}\big(B_{\alpha_1}\big)(\omega) \cdots X_{\alpha_n}\big(B_{\alpha_n}\big)(\omega). \end{aligned}$$

But X_{α_i} is measurable $\forall i = 1, \ldots, n$. Hence $Z(B)$ is measurable. We also have that if

$$B = \prod_{i=1}^{n} B_{\alpha_i} \times \prod_{\alpha \neq \alpha_1,\ldots,\alpha_n} \mathbb{R}_\alpha, \qquad B' = \prod_{i=1}^{m} B_{\alpha_i'} \times \prod_{\alpha \neq \alpha_1',\ldots,\alpha_m'} \mathbb{R}_\alpha \in \mathcal{S},$$

then, by letting $J = \{\alpha_1, \ldots, \alpha_n\} \cup \{\alpha'_1, \ldots, \alpha'_m\}$, we have

$$B \cap B' = \prod_{\alpha \in J} (\pi_\alpha(B) \cap \pi_\alpha(B')) \times \prod_{\alpha \notin J} \mathbb{R}_\alpha \in \mathcal{S}.$$

Hence $\mathcal{S}$ is $\cap$-stable. Therefore, by Theorem 1.2.4 of Bauer (1981),

$$\delta(\mathcal{S}) = \sigma(\mathcal{S}) = \mathcal{U}, \tag{4.3}$$

where $\delta(\mathcal{S})$ denotes the Dynkin system generated by $\mathcal{S}$ and $\sigma(\mathcal{S})$ denotes the σ-algebra generated by $\mathcal{S}$. But $\mathcal{B}$ is a Dynkin system, since it satisfies the following:

(1) $Z(\prod_{\alpha \in \Delta} \mathbb{R}) = 1$.
(2) For $B \in \mathcal{B}$, we have $Z(B)$ is measurable, hence, as

$$Z(B^c)(\omega) = \mu_\omega(B^c) = 1 - \mu_\omega(B) = 1 - Z(B)(\omega) \quad \forall \omega \in \Omega,$$

it follows that $Z(B^c)$ is measurable; therefore $B^c \in \mathcal{B}$.
(3) Let $B_i \in \mathcal{B}$ be a pairwise disjoint sequence. Then for every $\omega \in \Omega$, we have

$$\begin{aligned} Z\left(\bigcup B_i\right)(\omega) &= \mu_\omega\left(\bigcup B_i\right) = \sum \mu_\omega(B_i) \\ &= \sum Z(B_i)(\omega) = \lim \sum_{i=1}^{n} Z(B_i)(\omega), \end{aligned} \tag{4.4}$$

which shows that $Z(\bigcup B_i)$ is measurable; hence $\bigcup B_i \in \mathcal{B}$.

We conclude that $\delta(\mathcal{S}) \subseteq \mathcal{B}$. Then, from Eq. (4.3), we have $\mathcal{U} \subseteq \mathcal{B}$. Hence $Z(B)$ is measurable for each $B \in \mathcal{U}$. We also have from (1) and Eq. (4.4) that Z is a $\mathcal{U}$-observable. The uniqueness of Z follows from the uniqueness of the product measure μ_ω. □

Definition 4.9. We say that

(1) $f, g \in \mathcal{E}(\Omega, \mathcal{A})$ are *independent* if they are independent as random variables;
(2) $f_i \in \mathcal{E}(\Omega, \mathcal{A})$ are *pairwise independent* if f_i, f_j are independent $\forall\, i \neq j$;
(3) $f_i \in \mathcal{E}(\Omega, \mathcal{A})$ are *independent* if they are (*totally*) *independent* as random variables.

If $f, g \in \mathcal{E}(\Omega, \mathcal{A})$ are independent and $\mu(g) \neq 0$, then

$$\mu(fg) = E(fg) = E(f)E(g) = \mu(f)\mu(g).$$

Definition 4.10. Following (Gudder, 1998), a sequence X_i of observables on a probability space $(\Omega, \mathcal{A}, P)$ is said to be (*pairwise*) *independent* if the sequence $X_i(B_i)$ is (pairwise) independent for all possible choices of $\{B_i\}$ *in* $\mathfrak{B}(\mathbb{R})$.

It is clear that if X_i are (pairwise) independent observables and $u_i : \mathbb{R} \to \mathbb{R}$ are Borel functions, then $u_i(X_i)$ are also (pairwise) independent.

Let X, Y be independent observables on $(\Omega, \mathcal{A}, P)$, and let $\mu_{X,Y}$ be their joint distribution. Then for every $B_1, B_2 \in \mathcal{B}(\mathbb{R})$, we have

$$\mu_{X,Y}(B_1 \times B_2) = P(X(B_1)Y(B_2)) = P(X(B_1))P(Y(B_2)) = \mu_X(B_1)\mu_Y(B_2).$$

5. INFINITE SERIES OF OBSERVABLES

In this section we will show that certain theorems on infinite series of observables, namely, the weak and strong laws of large numbers and the central limit theorem, are still valid in the fuzzy probability setting.

Proposition 5.1. *Let $X_1, X_2, \ldots,$ be identically distributed and (pairwise) independent observables on a probability space $(\Omega, \mathcal{A}, P)$ and let Z be their unique joint observable as defined in Theorem* 4.8 *and μ_Z be the distribution of Z. Define the random variables Y_k on $(\prod_{i=1}^{\infty} \mathbb{R}, \bigotimes_{i=1}^{\infty} \mathfrak{B}(\mathbb{R}), \mu_Z)$ by $Y_k((x_i)_{i=1}^{\infty}) = x_k$, $k \in \mathbb{N}$. Then Y_i, $i \in \mathbb{N}$, are identically distributed and (pairwise) independent.*

Proof: Applying Theorem 4.8, we have

$$Z\left(\prod_{i=1}^{n} B_{\alpha_i} \times \prod_{\alpha \neq \alpha_1, \ldots, \alpha_n} \mathbb{R}_\alpha\right) = X_{\alpha_1}\left(B_{\alpha_1}\right) \cdots X_{\alpha_n}\left(B_{\alpha_n}\right).$$

Hence $\forall n \in \mathbb{N}$, if $B \in \mathfrak{B}(\mathbb{R})$ and $\mu_{Y_n(Z)}$ is the distribution of Y_n, then

$$\begin{aligned}\mu_{Y_n(Z)}(B) &= \mu_Z\left(Y_n^{-1}(B)\right) = P\left(Z\left(\prod_{i=1}^{n-1} \mathbb{R}_i \times B \times \prod_{i=n+1}^{\infty} \mathbb{R}_i\right)\right)\\ &= P(X_n(B)) = \mu_{X_n}(B).\end{aligned}$$

Hence $\forall n \in \mathbb{N}$, Y_n and X_n have the same distribution. As the X_n's are identically distributed, so are the Y_n's. We also have that the Y_n's are independent since the X_n's are independent. In fact, since X_i, $i \in \mathbb{N}$, are independent, if $k_i \in \mathbb{N}$, $B_{k_i} \in \mathfrak{B}(\mathbb{R})$, $i = 1, \ldots, n$, then

$$\begin{aligned}\mu_Z\left(\bigcap_{i=1}^{n} Y_{k_i}^{-1}\left(B_{k_i}\right)\right) &= P\left(Z\left(\prod_{i=1}^{n} B_{k_i} \times \prod_{k \neq k_1, \ldots, k_n} \mathbb{R}_k\right)\right)\\ &= P\left(\prod_{i=1}^{n} X_{k_i}\left(B_{k_i}\right)\right) = E\left(\prod_{i=1}^{n} X_{k_i}\left(B_{k_i}\right)\right)\\ &= \prod_{i=1}^{n} E\left(X_{k_i}\left(B_{k_i}\right)\right) = \prod_{i=1}^{n} \mu_{X_{k_i}}\left(B_{k_i}\right) = \prod_{i=1}^{n} \mu_{Y_{k_i}(Z)}\left(B_{k_i}\right).\end{aligned}$$

Similarly, if the X_i's are pairwise independent then so are the Y_n's. □

Corollary 5.2. *Let $X_1, X_2, \ldots,$ be identically distributed and integrable observables on a probability space $(\Omega, \mathcal{A}, P)$ and let Z and $Y_1, Y_2, \ldots,$ be as defined in Proposition* 5.1. *Then we have*

(1) *if $X_1, X_2, \ldots,$ are pairwise independent, then*

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i - E(X_1))$$

converges to 0 *in probability with respect to P, and*

(2) *if $X_1, X_2, \ldots,$ are independent, then*

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i - E(X_1))$$

converges to 0 *almost surely with respect to P.*

Proof: Part (1) follows easily from Proposition 5.1 and the weak law of large numbers in classical probability theory, and part (2) follows from Proposition 5.1 and the strong law of large numbers in classical probability theory. □

In what follows we will denote by $|X_n - X|$ the observable $u(Z)$ generated by the function $u : \mathbb{R}^2 \to \mathbb{R}$ defined by $u((x, y)) = |x - y|$ on the joint observable $Z_{X_n,X}$ of X_n, X, where $Z_{X_n,X}$ is uniquely determined by Eq. (4.1).

Definition 5.3 (Gudder, 1998). Let $X, X_n, n = 1, 2, \ldots,$ be observables on $(\Omega, \mathcal{A}, P)$. We say that X_n *converges to X in probability* if for every $\epsilon > 0$, we have

$$\lim P(|X_n - X| \geq \epsilon) = \lim P(|X_n - X|([\epsilon, \infty))) = 0.$$

Recall that if X is an observable generated by a measurable function f, then $X(B) = I_{f^{-1}(B)}\ \forall B \in \mathfrak{B}(\mathbb{R})$.

Lemma 5.4. *Let X_n, X_c be observables on a probability space $(\Omega, \mathcal{A}, P)$, where X_c is generated by a constant function $f = c$ on $\mathbb{R}$ and let $Z_n, n \in \mathbb{N}$, be their joint observable as defined in Theorem* 4.5. *Then for all $\epsilon > 0$ and all $n \in \mathbb{N}$,*

$$|X_n - X_c|([\epsilon, \infty)) = X_n((-\infty, c - \epsilon] \cup [c + \epsilon, \infty)). \tag{5.1}$$

Moreover, if for every $\epsilon > 0$, $\lim \mu_{X_n}((-\infty, c - \epsilon] \cup [c + \epsilon, \infty)) = 0$, then X_n converges to X_c in probability.

Proof: Let Z_n be the joint observable of X_n, X_c as defined in Theorem 4.5 and let $u : \mathbb{R}^2 \to \mathbb{R}$ be defined by $u((x, y)) := |x - y|$. Then for every $\epsilon > 0$,

$$(\star)\qquad \begin{aligned}\lim P(|X_n - X_c| \geq \epsilon) &= \lim P(u(Z_n))([\epsilon, \infty))\\ &= \lim \mu_{Z_n}(u^{-1}([\epsilon, \infty)))\\ &= \lim \mu_{Z_n}(\{(x, y) : x, y \in \mathbb{R}, |x - y| \geq \epsilon\}).\end{aligned}$$

Now since X_c is the observable generated by a constant function $f = c$, $c \in \mathbb{R}$, then for $B \in \mathfrak{B}(\mathbb{R})$ and $\omega \in \Omega$,

$$X_c(B)(\omega) = I_{f^{-1}(B)}(\omega) = \begin{cases} 1 & \text{if } c \in B, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, for $B_1, B_2 \in \mathfrak{B}(\mathbb{R})$ and $\omega \in \Omega$, we have

$$Z_n(B_1 \times B_2)(\omega) = X_n(B_1)(\omega)X_c(B_2)(\omega) = \begin{cases} X_n(B_1)(\omega) & \text{if } c \in B_2, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, if $B \in \mathfrak{B}(\mathbb{R}^2)$ and $c \notin \pi_2(B)$, then $\mu_{Z_n}(B) = 0$; hence, if $G = \mathbb{R} \times \{c\}$, then $Z_n(B \cap G^c) = 0$ and so $Z_n(B) = Z_n(B \cap G)$. It follows that for every $\epsilon > 0$ and for every $n \in \mathbb{N}$, we have

$$\begin{aligned} |X_n - X_c|([\epsilon, \infty)) &= Z_n((\{(x, y) \colon |x - y| \geq \epsilon\}) \cap G) \\ &= Z_n(\{(x, c) \colon |x - c| \geq \epsilon\}) = Z_n((x \colon |x - c| \geq \epsilon) \times \{c\}) \\ &= X_n((-\infty, c - \epsilon] \cup [c + \epsilon, \infty)). \end{aligned}$$

Hence, this and $(\star)$ yield that for every $\epsilon > 0$,

$$\lim P(|X_n - X_c|([\epsilon, \infty))) = \lim P(X_n((-\infty, c - \epsilon] \cup [c + \epsilon, \infty))),$$

and therefore the second assertion of the lemma now follows. □

In what follows we denote the observable $|X_n - X_c|$, as defined in the previous lemma, by $|X - c|$ for simplicity.

Theorem 5.5 (Weak Law of Large Numbers). *Let $X_1, X_2, \ldots,$ be a sequence of integrable, identically distributed and pairwise independent observables on a probability space $(\Omega, \mathcal{A}, P)$ and let $S_n = s_n(Z_n)$ where $s_n \colon \mathbb{R}^n \to \mathbb{R}$ is defined by $s_n((x_1, \ldots, x_n)) = \frac{1}{n}(x_1 + \cdots + x_n)$ and Z_n is the unique joint observable of $X_1, \ldots, X_n$, as defined in Theorem* 4.5. *Then S_n converges to $E(X_1)$ in probability.*

Proof: Let $X_1, X_2, \ldots,$ be a sequence of integrable, identically distributed and pairwise independent observables on $(\Omega, \mathcal{A}, P)$ and let Z be their joint observable. Let $Y_1, Y_2, \ldots,$ be the random variables defined in Proposition 5.1. Then by Corollary 5.2,

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i - E(X_1))$$

converges to 0 in probability with respect to P. Hence, for every $\epsilon > 0$,

$$0 = \lim \mu_Z \left(\left\{ x \in \prod_{i=1}^{\infty} \mathbb{R} \colon \left| \frac{1}{n} \sum_{i=1}^{n} Y_i(x) - E(X_1) \right| \geq \epsilon \right\} \right)$$
$$= \lim P \left(Z \left(\left\{ x \in \prod_{i=1}^{\infty} \mathbb{R} \colon \frac{1}{n} \sum_{i=1}^{n} Y_i(x) \notin (E(X_1) - \epsilon, E(X_1) + \epsilon) \right\} \right) \right). \tag{5.2}$$

Now let $B_\circ := (-\infty, E(X_1) - \epsilon] \cup [E(X_1) + \epsilon, \infty)$. We claim that $\forall n \in \mathbb{N}$,

$$\begin{aligned} S_n(B_\circ) &:= Z_n\big(S_n^{-1}(B_\circ)\big) \\ &= Z_n \left(\left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \colon \frac{1}{n} \sum_{i=1}^{n} x_i \in B_\circ \right\} \right) \\ &= Z \left(\left\{ x \in \prod_{i=1}^{\infty} \mathbb{R} \colon \frac{1}{n} \sum_{i=1}^{n} Y_i(x) \in B_\circ \right\} \right). \end{aligned} \tag{5.3}$$

To prove the claim, let

$$A_n = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \colon \frac{1}{n} \sum_{i=1}^{n} x_i \in B_\circ \right\},$$
$$D_n = \left\{ x \in \prod_{i=1}^{\infty} \mathbb{R} \colon \frac{1}{n} \sum_{i=1}^{n} Y_i(x) \in B_\circ \right\}.$$

Then it is clear that $D_n = A_n \times \prod_{i=n+1}^{\infty} \mathbb{R}$. Since $\forall \omega \in \Omega$, $Z_n(B)(\omega)$, $B \in \mathfrak{B}(\mathbb{R}^n)$, determines a measure, we have for all $n \in \mathbb{N}$ that

$$\begin{aligned} Z_n(A_n)(\omega) &= \inf \left\{ \sum_{i=1}^{\infty} Z_n(C_i)(\omega) \colon A_n \subseteq \bigcup C_i, C_i \right. \\ &\quad \left. = C_{i1} \times \cdots \times C_{in}, C_{ij} \in \mathfrak{B}(\mathbb{R}) \right\} \\ &= \inf \left\{ \sum_{i=1}^{\infty} \prod_{j=1}^{n} X_j(C_{ij})(\omega) \colon A_n \subseteq \bigcup C_i, C_i \right. \\ &\quad \left. = C_{i1} \times \cdots \times C_{in}, C_{ij} \in \mathfrak{B}(\mathbb{R}) \right\} \end{aligned}$$

$$= \inf \left\{ \sum_{i=1}^{\infty} \prod_{j=1}^{n} X_j(C_{ij})(\omega) \colon D_n \subseteq \bigcup C_i', C_i' \right.$$
$$\left. = \prod_{j=1}^{n} C_{ij} \times \prod_{j=n+1}^{\infty} \mathbb{R}, C_{ij} \in \mathfrak{B}(\mathbb{R}) \right\},$$

which is exactly $Z(D_n)(\omega)$ according to Theorem 4.8. This proves the claim. Hence, by Eq. (5.2) and (5.3), we have

$$\lim \mu_{S_n}(B_\circ) = \lim P(S_n(B_\circ))$$
$$= \lim P\left(Z\left(\left\{x \in \prod_{i=1}^{\infty} \mathbb{R} \colon \frac{1}{n} \sum_{i=1}^{n} Y_i(x) \in B_\circ \right\}\right)\right) = 0.$$

It follows from Lemma 5.4 that S_n converges to $E(X_1)$ in probability, as desired. □

Notice that the observable S_n of the above result is not identically distributed with the observable S_n' on $\mathfrak{B}(\mathbb{R})$ defined by $S_n'(B) := \frac{1}{n}\sum_{i=1}^{n} X_i(B)$. Indeed, $\forall B \in \mathfrak{B}(\mathbb{R})$ and $\forall n \in \mathbb{N}$, we have

$$S_n'(B) = \frac{1}{n}\sum_{i=1}^{n} X_i(B) = \frac{1}{n}\sum_{i=1}^{n} X_1(B) = X_1(B);$$

hence $S_n' \to S_1' = X_1$, while $S_n \to E(X_1)$ in probability.

Definition 5.6. Let $f_1, f_2, \ldots,$ be a sequence of effects on $(\Omega, \mathcal{A})$. We define

$$\lim\ \sup f_n := \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} f_n.$$

Lemma 5.7 (Gudder, 1998). *Let f_n be a sequence of effects on* $(\Omega, \mathcal{A}, P)$. *Then*

(1) $\bigcup f_n$ *exists and is in* $\mathcal{E}(\Omega, \mathcal{A})$, *and* $\bigcup f_n \neq 1$ *iff* $\sum f_n < \infty$,
(2) $\limsup f_n$ *exists and is in* $\mathcal{E}(\Omega, \mathcal{A})$.

Lemma 5.8 (Borel–Cantelli Lemma; Gudder, 1998). *Let f_n be a sequence in* $\mathcal{E}(\Omega, \mathcal{A})$ *and let* $g = \limsup f_n$.

(1) *If* $\sum P(f_n) < \infty$, *then* $P(g) = 0$.
(2) *If* $\sum P(f_n) = \infty$ *and the* f_n *are independent, then* $P(g) = 1$.

Definition 5.9 (Gudder, 1998). Let X, X_n, $n = 1, 2, \ldots$, be observables on $(\Omega, \mathcal{A}, P)$. We say that X_n *converges to* X *almost surely* if for every $\epsilon > 0$, we have

$$P[\lim\ \sup(|X_n - X| \geq \epsilon)] = P[\lim\ \sup(|X_n - X|([\epsilon, \infty)))] = 0.$$

Recall that

$$\lim\ \sup(|X_n - X|([\epsilon, \infty))) = \prod_{k=1}^{\infty}\left(1 - \prod_{n\geq k}[1 - (|X_n - X|([\epsilon, \infty)))]\right).$$

Lemma 5.10 (Gudder, 1998). *If X_n is a sequence of independent observables on $(\Omega, \mathcal{A}, P)$, $c \in \mathbb{R}$, and X_n converges to c almost surely, then X_n converges to c in probability.*

The question about the validity of the strong law of large numbers in the setting of fuzzy probability theory still needs an answer. In fact, we can weaken the definition of almost everywhere convergence for a sequence of observables in fuzzy probability theory which is what we will call almost everywhere* convergence of a sequence of observables. Then we prove our version of the strong law of large numbers according to this definition (see Theorem 5.14 below).

Definition 5.11. Let X, X_n, $n = 1, 2, \ldots$, be observables on $(\Omega, \mathcal{A}, P)$. We say that X_n converges to X almost everywhere* if $\forall \epsilon > 0$,

$$\lim P\left(\sup_{k\geq n} |X_k - X|([\epsilon, \infty))\right) = 0$$

where the supremum is defined as in the usual sense for sequences of functions.

Remark 5.12. The two definitions of almost everywhere convergence and almost everywhere* convergence coincide for crisp observables and hence in the usual case of random variables. But they are not equivalent in general. In fact, the concept of almost everywhere convergence has a very strong condition. To see this, suppose that X_n is a sequence of observables that converges almost everywhere to the observable X. Then $\forall \epsilon > 0$, we have

$$\int \prod_{n}\left(1 - \prod_{k\geq n}(1 - |X_k - X|([\epsilon, \infty)))\right) dP = 0,$$

which implies that $\exists A \in \mathcal{A}$ such that $P(A) = 1$ and $\forall \omega \in A$,

$$\prod_{n}\left(1 - \prod_{k\geq n}(1 - |X_k - X|([\epsilon, \infty)))\right)(\omega) = 0.$$

Hence, by the theory of infinite products, $\forall\omega \in A$, we have

$$\sum_n \prod_{k\geq n}(1 - |X_k - X|([\epsilon, \infty)))(\omega) = \infty \quad \forall\omega \in A.$$

Then $\forall\omega \in \Omega$, $\exists n_\omega \in \mathbb{N}$ such that $\prod_{k\geq n_\omega}(1 - |X_k - X|([\epsilon, \infty)))(\omega) \neq 0$. Again, by the theory of infinite products, $\forall\omega \in A$, we then have

$$\sum_{k\geq n_\omega} |X_k - X|([\epsilon, \infty))(\omega) < \infty.$$

Hence, $\forall\omega \in A$, we have

$$\sum_{n=1}^{\infty} |X_n - X|([\epsilon, \infty))(\omega) < \infty, \tag{5.4}$$

which is a very strong condition. Moreover, $\forall\omega \in A$, it leads to

$$\lim\ \sup|X_n - X|([\epsilon, \infty))(\omega) = 0;$$

and hence X_n converges almost everywhere* to X. Thus the definition of almost everywhere convergence implies the definition of almost everywhere* convergence.

The following example shows that the two definitions of almost everywhere convergence and almost everywhere* convergence are not equivalent.

Example 5.13. Let $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathfrak{B}, (\mathbb{R}))$ and let P be any probability measure on $\mathfrak{B}(\mathbb{R})$. Define a sequence of observables X_n on $(\Omega, \mathcal{A}, P)$ by

$$X_n(B)(\omega) := \begin{cases} 0 & \text{if } 0 \notin B, 1 \notin B, \\ \frac{1}{n} & \text{if } 1 \in B, 0 \notin B, \\ 1 - \frac{1}{n} & \text{if } 1 \notin B, 0 \in B, \\ 1 & \text{otherwise.} \end{cases}$$

Let X be the observable generated by the random variable $f : \Omega \to \mathbb{R}$ such that $f(\omega) = 0\ \forall\omega \in \Omega$. Then X_n converges to X almost everywhere*. To see this, let $\epsilon > 0$ be given. Then, by Lemma 5.4,

$$|X_n - X|([\epsilon, \infty)) = X_n((-\infty, -\epsilon] \cup [\epsilon, \infty)) = \frac{1}{n}.$$

Hence, $P(\inf_n \sup_{k\geq n}|X_k - X|([\epsilon, \infty))) = 0$. But X_n does not converge to X almost everywhere. To see this, suppose on the contrary that X_n converges to X almost everywhere. Then by Inequality 5.4, $\exists A \in \mathcal{A}$ such that $P(A) = 1$ and such that $\forall\omega \in A$, we have

$$\sum_{n=1}^{\infty} |X_n - X|([\epsilon, \infty))(\omega) < \infty.$$

But, $\forall \omega \in A$, we have

$$\sum_{n=1}^{\infty} |X_n - X|([\epsilon, \infty))(\omega) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

a contradiction.

Now we will prove our version of the strong law of large numbers according to the definition of almost everywhere* convergence.

Theorem 5.14 (Strong Law of Large Numbers). *Let $X_1, X_2, \ldots,$ be a sequence of integrable, identically distributed and independent observables on $(\Omega, \mathcal{A}, P)$ and let $S_n = s_n(Z_n)$ where $s_n : \mathbb{R}^n \to \mathbb{R}$ is defined by $s_n((x_1, \ldots, x_n)) = \frac{1}{n}(x_1 + \cdots + x_n)$ and Z_n is the unique joint observable of $X_1, \ldots, X_n$, as defined in Theorem* 4.5. *Then S_n converges to $E(X_1)$ almost everywhere*.*

Proof: Define the joint observable Z of $X_1, X_2, \ldots$ as in the proof of Theorem 5.5 (Weak Law of Large Numbers). And follow the proof to get the sequence $Y_1, Y_2, \ldots,$ of identically distributed random variables on $(\prod_{i=1}^{\infty} \mathbb{R}, \bigotimes_{i=1}^{\infty} \mathfrak{B}(\mathbb{R}),$ $\mu_Z)$ with expectation m. Then, by Proposition 5.1, $Y_1, Y_2, \ldots,$ are also independent and identically distributed. Hence, by Corollary 5.2, we have

$$\frac{1}{n} \sum_{i=1}^{n} (Y_i - E(X_i)),$$

which converges to 0 almost surely with respect to P. Let $m = E(X_1)$. Then for every $\epsilon > 0$ we have that

$$\mu_Z \left(\limsup \left\{ x \in \prod_{j=1}^{\infty} \mathbb{R} \colon \left| \frac{1}{n} \sum_{i=1}^{n} Y_i(x) - m \right| \geq \epsilon \right\} \right) = 0.$$

Hence,

$$\begin{aligned} 0 &= \mu_Z \left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left\{ x \in \prod_{j=1}^{\infty} \mathbb{R} \colon \left| \frac{1}{k} \sum_{i=1}^{k} Y_i(x) - m \right| \geq \epsilon \right\} \right) \\ &= P \left(Z \left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left\{ x \in \prod_{j=1}^{\infty} \mathbb{R} \colon \frac{1}{k} \sum_{i=1}^{k} Y_i(x) \right. \right. \right. \\ &\qquad \left. \left. \left. \in \ (-\infty, m - \epsilon] \cup [m + \epsilon, \infty) \right\} \right) \right). \end{aligned}$$

Now let $\omega \in \Omega$. Then by Eq. (5.3),

$$Z\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\left\{x \in \prod_{j=1}^{\infty}\mathbb{R}\colon \frac{1}{k}\sum_{i=1}^{k}Y_i(x) \in (-\infty, m-\epsilon] \cup [m+\epsilon, \infty)\right\}\right)(\omega)$$

$$= \lim Z\left(\bigcup_{k=n}^{\infty}\left\{x \in \prod_{j=1}^{\infty}\mathbb{R}\colon \frac{1}{k}\sum_{i=1}^{k}Y_i(x) \in (-\infty, m-\epsilon] \cup [m+\epsilon, \infty)\right\}\right)(\omega)$$

$$\geq \lim_n \sup_{k\geq n} Z\left(\left\{x \in \prod_{j=1}^{\infty}\mathbb{R}\colon \frac{1}{n}\sum_{i=1}^{n}Y_i(x) \in (-\infty, m-\epsilon] \cup [m+\epsilon, \infty)\right\}\right)(\omega)$$

$$= \lim_n \sup_{k\geq n} Z_k\left(\left\{x \in \prod_{j=1}^{k}\mathbb{R}\colon s_k(x) \in (-\infty, m-\epsilon] \cup [m+\epsilon, \infty)\right\}\right)(\omega).$$

Hence,

$$P\left(\lim_n \sup_{k\geq n} Z_k\left(\left\{x \in \prod_{j=1}^{k}\mathbb{R}\colon s_k(x) \in (-\infty, m-\epsilon] \cup [m+\epsilon, \infty)\right\}\right)\right)$$

$$\leq \mu_Z\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\left\{x \in \prod_{j=1}^{\infty}\mathbb{R}\colon \left|\frac{1}{k}\sum_{i=1}^{k}Y_i(x) - m\right| \geq \epsilon\right\}\right) = 0,$$

and, therefore, the strong law of large numbers holds. □

Theorem 5.15 (Central Limit Theorem). *Let X_n be a sequence of independent, identically distributed observables with expectation zero and variance $\sigma^2 > 0$. Let $S_n = s_n(Z_n)$ where $s_n : \mathbb{R}^n \to \mathbb{R}$ is defined by $s_n((x_1, \ldots, x_n)) = \frac{1}{\sigma\sqrt{n}}(x_1 + \cdots + x_n)$ and Z_n is the unique joint observable of $X_1, \ldots, X_n$, as defined in Theorem* 4.5. *Let μ_n be the distribution of S_n. Then μ_n converges weakly to μ_σ, where μ_σ denotes the normal distribution with expectation* 0 *and variance* 1.

Proof: Following the proof of Theorem 5.5 (Weak Law of Large Numbers), we construct the sequence of independent, identically distributed random variables Y_i of expectation zero and variance σ^2. For each $n \in \mathbb{N}$, let $T_n = \frac{1}{\sigma\sqrt{n}}(Y_1 + \cdots + Y_n)$ and let μ'_n denote the distribution of T_n. Then by central limit theorem in classical probability theory, we have μ'_n converges weakly to the normal distribution of expectation 0 and variance 1. Now by a similar proof to Eq. (5.3), we can get $\forall B \in \mathfrak{B}(\mathbb{R})$ that

$$S_n(B) := Z_n\left(s_n^{-1}(B)\right) = Z_n\left(\left\{(x_1, \ldots, x_n) \in \mathbb{R}^n\colon \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n}x_i \in B\right\}\right)$$

$$= Z\left(\left\{x \in \prod_{i=1}^{\infty}\mathbb{R}\colon \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n} Y_i(x) \in B\right\}\right) = Z\left(T_n^{-1}(B)\right).$$

Then, $\forall B \in \mathfrak{B}(\mathbb{R})$, we have

$$\mu_n(B) = P(S_n(B)) = P\left(Z\left(T_n^{-1}(B)\right)\right) = \mu_Z\left(T_n^{-1}(B)\right) = \mu'(B).$$

Hence, μ_n and μ'_n are also identically distributed, and therefore μ_n converges weakly to the normal distribution with expectation 0 and variance 1. □

REFERENCES

Bauer, H. (1981). *Probability Theory and Elements of Measure Theory*, 2nd edn., Academic Press, New York.

Beltrametti, E. G. and Bugajski, S. (1995a). A classical extension of quantum mechanics. *Journal of Physics A: Mathematical and General* **28**, 3329–3334.

Beltrametti, E. G. and Bugajski, S. (1995b). Quantum observables in classical framework. *International Journal of Theoretical Physics* **34**(8), 1221–1228.

Bugajski, S. (1996). Fundamentals of fuzzy probability theory. *International Journal of Theoretical Physics* **35**(11), 2229–2244.

Bugajski, S. (1998a). Fuzzy stochastic processes. *Open System and Information Dynamics* **5**, 169–185.

Bugajski, S. (1998b). On fuzzy random variables and statistical maps. *Report on Mathematical Physics* **41**(1), 1–11.

Dunyak, J., Saad, I., and Wunch, D. (1999). A theory of independent fuzzy probability for system reliability. *IEEE Transactions on Fuzzy Systems* **7**(2), 286–295.

Gudder, S. (1998). Fuzzy probability theory. *Demonestratio Mathematica* **31**(3), 235–254.

Gudder, S. (1999). Observables and statistical maps. *Foundations of Physics* **29**, 877–897.

Gudder, S. (2000). What is fuzzy probability theory. *Foundations of Physics* **30**, 1663–1678.

Ishikawa, S. (1996). A quantum mechanics approach to a fuzzy theory. *Fuzzy Sets and Systems* **90**, 276–306.

Pykacz, J. (1992). Fuzzy set ideas in quantum logics. *International Journal of Theoretical Physics* **31**(9), 1767–1783.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control* **8**, 338–353.