The Intersections Based on Joint Observables In Fuzzy Probability

Murat Ozcek

Gaziantep University, Department of Mathematics, Gaziantep, Turkey

Email: muratozcek.12@gmail.com

Abstract

"Fuzzy probability theory" appeared as a smooth extension of classical probability theory in 1995. It was expected that it will be of great importance in quantum mechanics, but the theory doesn't keep its development as it was expected. This necessitates revising some of its fundamental basic concepts. We argue that if quantum probability theory should have less constrained than classical probability theory as can be seen in the case of joint random variables, we surely need to weaken the definition of the intersection operation. In this paper, discuss the definition validity in quantum probability theory and to discuss the consistency of the given definitions with the whole theory and the possibility to have a more suitable definition.

Keywords: probability theory; quantum mechanics; fuzzy set; fuzzy probability

I. ntroduction

The connection between the pure concepts of classical probability theory and its applications is strongly maintained as they go hand in hand along the various steps of the theory. Moreover, it can be easily seen that the definitions in the theory accurately translate the concepts that they represent.

The main definitions of the theory such as probability measure, events, random variables, independence, and joint random variables are very consistent with common sense and general understanding of the word "probability". To indicate this, notice that the widespread definition of the "probability of an event" is the limit of the ratio of the frequency in which the event falls when the experiment is repeated n times as n goes to infinity. Although the axioms of the probability space are not based on the previous definition, they succeeded to prove the "laws of large numbers", which emphasizes the existence of such limit. The strength of the theory was then evident since the mathematical structure of probability theory translates precisely its applied nature.

Till the appearance of the field of quantum mechanics, there was no need to construct a new probability theory or to extend its scope. But as quan- tum mechanics was developing, the demand to extend probability theory was urgent [BB95]. There were several trials to build other probability theories that fulfill the quantum mechanics needs [Gra89, Rie00]. "Operational prob- ability theory" which appeared as a smooth extension of classical probability theory is one such [BLM96, Gud00, BS08]. Moreover, it seems to possess the ability to extend all the other new theories, so it was expected that it will be of great importance [Gud98, Gud00]. However, the theory did not develop as expected.

A review of the literature on operational probability theory has been conducted and shown very little progress in its development. The liter-ature was limited by the work appeared in [FP10, CF10, Fri07] over the last fifteen years. The weak progress of the theory especially after the death of S. Bugajski (March 2003) [Bug98a] poses questions regarding the consistency of the theory structure. On the other hand, since any quantum theory has a statistical nature, we need to know to what degree the basic definitions of the theory translate the quantum mechanics nature. This paper aims to answer this question in its mathematical aspects. In the first section, the main bases of the theory are presented. In the second section, we focus on the importance of the intersection operation and demonstrate the need to extend its definition in operational probability theory. In the third section we exploit the concept of joint observable by explaining its relationship with the intersection operation, and hence use that relation to explain and establish the need to extend it. The last section suggests the approach that can be used to define the intersection operation.

Doi: https://doi.org/10.54216/GJMSA.050203
Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023

2. Operational probability theory

Pykacz (1992) introduced the notion of a fuzzy quantum logic using the notion of fuzzy set theory only. Pykacz's approach was based on a theo- rem proving that any quantum logic L with an ordering set of states can be isomorphically represented by a special family of fuzzy subsets [Mes95]. Operational probability theory appears as a natural result of the development of quantum structures and the corresponding probability theories. So the definitions of the operations on fuzzy sets are derived from the corresponding definitions on the several quantum structures by extending its range to the set of all fuzzy sets.

Before introducing fuzzy sets, note that if Ω is a nonempty set and 2^{Ω} is its power set, we can identify any set $A \in 2^{\Omega}$ with its indicator function I_A since $A = B$ if and only if $I_A = I_B$. In fuzzy set theory, subsets of Ω are replaced by fuzzy sets, where the fuzzy sets are defined as follows.

Definition 1.1. A *fuzzy subset* f of Ω is a function $f : \Omega \to [0, 1]$. Hence the system of all fuzzy subsets $[0,1]^{\Omega}$ can be treated as a power set. We say that a fuzzy set f is *a crisp set* if the values of f are contained in $\{0, 1\}$. Thus, f is crisp if and only if f is an indicator function.

It is clear that crisp fuzzy sets correspond to the usual sets. We thus say that a fuzzy set is a generalization of a set. We say $f \subseteq g$ if $f(x)\square g(x)$for any$x \in \Omega$.

Let (Ω, A) be a measurable space. A random variable $f : \Omega \to [0, 1]$ is called *an effect* or a *fuzzy event*. In operational probability theory, [Bug96, Bug98b, Gud98, Gud00], we replace the σ-algebra A with $E(\Omega, A)$, the set of all effects $f : \Omega \to [0, 1]$.

Definition 1.2. [Gud98] Let $f, g \in E(\Omega, A)$. Define $f' := 1 - f, f \cap g \coloneqq fg$ and $f \cup g = f + g - fg$.

These definitions correspond to the usual properties of indicator functions.

Lemma 1.3. [HN02] *Let $fn \in E(\Omega, A), n \in N$. Then $\cup \infty n = 1\, fn$ exists and*

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

Lemma 1.4. [HN02] *For $f \in [0,1]^{\Omega}$ the following are equivalent:*

(1) f is crisp,

(2) $f \cap f = f$,

(3) $f \cup f' = 1$.

The following examples indicate the need to generalize classical probability theory.

Example 1.5. [Bug98b]

(1) *In real situations, even if the measurable quantity represented by a random variable $f : \Omega \to$ R on a measurable space (Ω, A) assumes in fact a single value $f(\omega)$, the single readings of a measuring apparatus are subject to some errors which can be described by an error map ε that to every value $x \in$ Rattaches a probability measure $v_x$such that*

$\mathcal{E}^{\circ} f(\omega) = v_{f(\omega)}$

Now $\forall B \in$ B(R), define $X_f(B) : \Omega \to [0,1]$such that $X_f(B)(\omega) = v_{f(\omega)}(B)$. Then $X_f(B)$ determines a fuzzy set. Now if $X_f(B)$ is measurable for every $B \in$ B(R) it is clear that X_f corresponds to no random variable (see Example 1.14,1).

(2) *Let (Ω, A, μ) be a probability measurable space. The probability distribution*

$\mu P^X(B) = \mu\left(X^{-1}(B)\right) \quad \forall\, B \in$ B.

Let $f, g : \Omega \to$R be two random variables. The mixed measurement of f and g in the ratio $\lambda : 1 - \lambda,\ 0 < \lambda < 1$ yields the probability distribution.

$\lambda P_{\mu}^{g} + (1 - \lambda) P_{\mu}^{f}$

which in general does not correspond to any standard random variable as we will see later (see Example 1.14,2).

For more details we refer the reader to [Bug98b]. Thus, standard random variables fail to represent properties of events in the previous two examples, while, as we will see in Example 1.14, we can represent them by the corresponding concept of random variable, that is; fuzzy random variable. The standard random variables represent special properties of elementary events.

So, various properties lie outside the scope of standard probability theory or, at best, can be modeled there in an indirect way [Bug96, Gud98]. Operational probability theory introduces new probability space generated from the classical one in a natural way. As before, the basic structure is a measurable space (Ω, A). We identify any set $A \in$ A with its indicator function I_A.

Definition 1.6. If μ is a probability measure on (Ω, A) and $f \in E$, we define *the probability* of f to be its expectation $\mu(f) = \int f d\mu$..

Corresponding to the concept of probability measures in classical probability theory, we have states in quantum mechanics.

Doi: https://doi.org/10.54216/GJMSA.050203
Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023

Definition 1.7. *A state* on $E(\Omega, A)$ is a map $s : E(\Omega, A) \to [0, 1]$ that satisfies

(i) $s(I_\Omega) = 1$, and

(ii) if $f_i \in E(\Omega, A)$ is a sequence such that $\sum f_i \in E(\Omega, A)$, then $\varphi(\sum f_i) = \sum \varphi(f_i)$.

A state corresponds to a condition of a system and $s(f)$ is interpreted as the probability that the effect f occurs when the system is in the condition corresponding to s.

Using the important concept of *σ-morphism*, it can be proved that μ is a probability measure on (Ω, A), if and only if μ is a state on $E(\Omega, A)$.

Definition 1.8. [Gud99, Gud00] A mapping $\varphi: E(\Omega, A) \to E(\Lambda, B)$is called *a morphism* if

(1) $\varphi(1) = 1$, and

(2) $f \perp g$ in $E(\Omega, A)$implies that $\varphi(f) \perp \varphi(g)$ in $E(\Lambda, B)$and

$$\varphi(f + g) = \varphi(f) + \varphi(g)$$

A morphism is *σ-morphism* if for every sequence $f_i \in E(\Omega, A)$such that $\sum f_i \in E(\Omega, A)$, e have $\sum \varphi(f_i) \in E(\Lambda, B)$ and $\varphi(\sum f_i) = \sum \varphi(f_i)$.

Hence when we restrict the codomain of a σ-morphism φ to [0, 1], then $\varphi : E(\Omega, A) \to [0, 1]$ is simply a state.

Theorem 1.9. [Gud98]

(1) *If $\varphi: E(\Omega, A) \to E(\Lambda, B)$, is a σ- morphism, then $\varphi(\lambda f) = \lambda\varphi(f)$ for every $\lambda \in [0,1], f \in E(\Omega, A)$.*

(2) *If s is a state on $E(\Omega, A)$, then there exists a unique probability measure μ on (Ω, A) such that $s(f) = \mu(f)$ for every $f \in E(\Omega, A)$.*

Thus we have that states coincide with probability measures in classical probability theory. In fact, the main difference between operational probability theory and classical probability theory lies in the definition of fuzzy random variables, which are called observables.

Definition 1.10. [Gud98] If (Γ, B) is a measurable space, *a B-observable on* (Ω, A) is a map $X : B \to E(\Omega, A)$ such that

(1) $X(\Gamma)=1$, and

(2) If $A_i \in B$ are mutually disjoint, then $X(\bigcup A_i) = \sum X(A_i)$where the convergence of the summation is pointwise.

A B(R)-observable is simply called *an observable.*

For an observable X, if $X(B)$ is crisp for every $B \in$ B, then X is *crisp*. It was proved that there exists a natural one-to-one correspondence be- tween observables and σ-morphisms (see [HN02]).

Theorem 1.11. [HN02] *If $X : B \to E(\Omega, A)$ is a B-observable, then X has a unique extension to a σ-morphism $\tilde{X}: E(\Gamma, B) \to E(\Omega, A)$. If $Y E(\Gamma, B) \to E(\Omega, A)$ is a σ-morphism, then $Y|B$ is a B-observable.*

If $f : \Omega \to \Gamma$ is measurable function, then the corresponding crisp observable $X_f$is given by; $X_f(B) = I_{f^{-1}}(B)$.

It was proved that most of the results in classical probability theory are still valid in operational probability theory [Gud98, Gud00, HN02]. Let (Ω, A) and (Γ, B) be measurable spaces. If μ is a probability measure on (Ω, A) and X is a B-observable on (Ω, A), *the distribution of* X is the prob- ability measure μX on B given by, $\mu X(B) = \mu(X(B))$.

We interpret $\mu X(B)$ as the probability that X has a value in B when the system is in the state μ.

Definition 1.12. If X is an observable on (Ω, A), μ is a probability measure on (Ω, A) and $u : R \to R$ is a Borel function, we define the observable $u(X) : B(R) \to E(\Omega, A)$ by$u(X)(B) = X(u^{-1}(B))$. The distribution of $u(X)$ becomes;

$$\mu_{u(X)}(B) = \mu\left(X\left(u^{-1}(B)\right)\right) \quad \forall B \in B(R).$$

Definition 1.13. [Gud98] If X is an observable and u: R →R is a Borel function, then $E\left(u(X)\right) = \int \lambda\mu_{u(X)}(d\lambda)$.

Here are some examples of noncrisp observables.

Example 1.14.

(1) *Refer to Example 1.5,1, and define the observable $X : B(R) \to E(\Omega, A)$ such that $X(B) = X_f(B)\ \forall B \in B(R)$. Then X is an observable.*

(2) *If f, g are random variables on (Ω, A), $\lambda \in (0, 1)$, let X_f, X_g be the observables generated by f, g; i. e. ,* $X_f(B) = I_{f^{-1}(B)}, X_g(B) = I_{g^{-1}(B)}\quad \forall B \in B(R)$.
Define Y: $B(R) \to E(\Omega, A)$ *such that*

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

Doi: https://doi.org/10.54216/GJMSA.050203
Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023

Then Y is an observable which is not crisp. Now if μ *is a probability measure and hence a state on* (Ω, A), *then the distribution of Y will be* $P^Y\mu$
Where;
$P^Y\mu(B) = P\left(Y^{P^Y}(B)\right) = P(\lambda X_f(B) + (1-\lambda)X_g(B))$
Applying Theorem 1.9, we have
$P^Y\mu(B) = \lambda P\left(X_f^{P^Y}(B)\right) + (1-\lambda)P(X_g(B))$
Hence;
$P_\mu^Y = \lambda P_\mu^f + (1-\lambda)P_\mu^g,$
which is exactly the probability distribution defined in Example 1.5,2.

(3) *For every* $f \in E(\Omega, A)$, *we can define the observable* X_f *on* (Ω, A)*by*

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

In general, X_f *has no corresponding random variable since if so, then according to the above remark the observable* X_f *will be crisp but it's clear that*
if f is not crisp, then so is the observable X_f.

Theorem 1.15. [Nas00, HN02] *For an observable X on* (Ω, A), *X is crisp if and only if there exists a random variable* $f: \Omega \to \mathrm{R}$ *such that* $X(B) = I_{f^{-1}(B)}$ $\forall B \in \mathrm{B}(R)$

Definition 1.16. Let $X_1, \ldots, X_n$ be observables on (Ω, A). We say that a $\mathrm{B}(R^n)$-observable X on (Ω, A) is their *joint observable* if
$\pi_i(X) = X_i, i = 1, \ldots, n,$
where π_i is the marginal projection map. For finite collections of observables, we have the following theorem.

Theorem 1.17. [Gud98] *If* $X_1, \ldots, X_n$ *are observables on* (Ω, A), *then there exists a unique n-dimensional observable Z on* (Ω, A) *such that* (1. 1)
$Z(B_1 \times \ldots \times B_n) = X_1(B_1) \ldots X_n(B_n)$ *for all* $B_1, \ldots, B_n \in \mathrm{B}(\mathrm{R})$.

Note that condition 1. 1 is essential for the uniqueness of the joint observable *Z*, as the following example indicates. This example can be found in [Bug96] and it has a direct connection to the quantum mechanical description of spin- 1 objects.

Example 1.18. *Let* Ω *denote the set of points of the unit sphere in* R^3 *and let* $\omega_1, \omega_2 \in \Omega$. *Define* $\mathrm{B}(\mathrm{R})$-*observables* $X_{\omega_i}, i = 1,2$ *on* $(\Omega, \mathrm{B}(\Omega))$ *by;*

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

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

$$X\left(\left\{\left(\frac{1}{2}, \frac{1}{2}\right)\right\}\right) = \lambda(\omega),$$
$$X\left(\left\{\left(\frac{1}{2}, -\frac{1}{2}\right)\right\}\right) = \frac{1}{2}(1 + r_{\omega_1}.r_\omega) - \lambda(\omega),$$
$$X\left(\left\{\left(-\frac{1}{2}, \frac{1}{2}\right)\right\}\right) = \frac{1}{2}(1 + r_{\omega_2}.r_\omega) - \lambda(\omega),$$
$$X\left(\left\{\left(-\frac{1}{2}, -\frac{1}{2}\right)\right\}\right) = \lambda(\omega) - \frac{1}{2}(r_{\omega_1} + r_{\omega_2}).r_\omega, and$$
$$X(\mathbb{R}^2 \backslash \{\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, -\frac{1}{2}\right), \left(-\frac{1}{2}, \frac{1}{2}\right), \left(-\frac{1}{2}, -\frac{1}{2}\right)\}) = 0,$$

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

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

Doi: https://doi.org/10.54216/GJMSA.050203
Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023

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 1.19. [Gud98] Let $X_1, \ldots, X_n$ be observables on a probability space (Ω, *A*, *P*), then the probability measure $\mu X_1, \ldots, X_n$ on B(R) given by

$\mu X_1, \ldots, X_n(B) \coloneqq \mu z(B) = P(Z(B))$

where *Z* is given by Equation 1. 1, is called the *joint distribution of* $X_1, \ldots, X_n$.

Definition 1.20. We say that;

(1) $f, g \in E(\Omega, A)$are *independent* if they are independent as random variables;

(2) a sequence of events (f_i) from *E*(Ω, *A*) are *pairwise independent* if f_i, f_j are independent $\forall i = j$;

(3) a sequence of events $(f_i)$from *E*(Ω, *A*) are *independent* if they are *(totally) independent* as random variables.

If $f, g \in E(\Omega, A)$are independent and $\mu(g) = 0$, then

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

Definition 1.21. Following [Gud98], a sequence $X_i$of observables on a probability space (Ω, *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* B(R).

It is clear that if the $X_i's$ are (pairwise) independent observables and u_i: R → Rare Borel functions, then $u_i(X_i)$ are also (pairwise) independent.

Let *X*, *Y* be independent observables on (Ω, *A*, *P*), let *μX,Y* be their joint distribution, then for every $B_1, B_2 \in B$(R), we have;

$\mu X, Y(B_1, B_2) = P\big(X(B_1)Y(B_2)\big) = P\big(X(B_1)\big)P\big(Y(B_2)\big) = \mu X(B_1)\mu Y(B_2)$.

3. A closer look at the operations on fuzzy events

As we saw in the previous section, operational probability theory appeared as an extension of the standard probability theory by its natural extended concepts to obtain a probability theory that can be applied to classical statistics as well as to quantum physics. On the other hand, we lost some features which characterize the classical theory as we can see in the following remark.

***Remark* 2.1.** By generalizing the definitions of operations on fuzzy sets, we obtain some properties, which appeared very strange. For example,

(1) we may have a fuzzy set which has a nonempty intersection with its complement.

(2) the intersection of two noncrisp fuzzy sets *u*, *v* is strictly less than any of them. Moreover,

(3) if *u* is a noncrisp fuzzy set then $u \cap u < u$.

Some will argue that "these properties are not strange from the nature of quantum physics. In fact, these properties prove the capability of the theory to describe the events in quantum physics." But is it true that defining the intersection operation of two effects (fuzzy set) *u*, *v* as $u \cap v = uv$ translates exactly the needed properties of quantum events? We will try to answer this question mathematically and physically. We concentrate on the intersection operation since the definitions of other operations can be determined according to it.

Let us begin to discuss the advantages of the given definition of the intersection operation in operational probability theory. Besides the fact that this definition was a natural result of the development in quantum structures and the fact that this definition generalizes the corresponding operation for crisp sets, it was rational to define the independence condition of two fuzzy events *u*,*v* to be "*u*, *v* are independent as random variables". The aim can be explained if we notice that in this way we can assure that if *u*, *v* are independent then *P*(*uv*) = *P*(*u*)*P*(*v*) which means that we really extend the classical definitions.

Now since the probability space defined on the *σ*- algebra (Ω, *A*) is arbitrary, and we used to define the independence of events by using the intersection operation as $P(u \cap v) = P(u)\,P(v)$, it was natural to define $u \cap v = uv$. But it is worth to mention here that if all we need to be satisfied is $P(uv) = P(u)\,P(v)$ if *u*, *v* are independent, the condition that *u*, *v* are independent as random variables is a very strong condition.

On the other hand, let *X* be an observable. If $\omega \in \Omega$, then $X(B)(\omega), B \in B$(R) determines a probability measure μ_ω where $\mu_\omega(A) = X(A)(\omega)$. According to the given definition of the intersection operation, for any two sets $A, B \in B$(R), we have

$X(A \cap B)(\omega) = X(A)(\omega)X(B)(\omega)$.

Hence *A*, *B* are independent classical events with respect to the measure μ_ω. Although the measure μ_ω determined only by the observable *X* and has no relation with the original probability space on Ω, the

Doi: https://doi.org/10.54216/GJMSA.050203

Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023

independence is an ideal case for the relation between events in arbitrary probability measures especially when we deal with fuzzy sets. Do we really need this strong condition? To be more obvious, let us consider first the classical case. If f is a classical random variable, then $X_f(B)(\omega) = I_{f^{-1}(B)}(\omega)$. So
$X_f(A \cap B)(\omega) = I_{f^{-1}(A)}(\omega) I_{f^{-1}(B)}(\omega) = I_{f^{-1}(A \cap B)}(\omega)$.

Thus, the independence is satisfied in the classical case. But note that the probabilities here related to A, B have only the values 0 or 1 and we know that the independence is trivially satisfied in such situation. That is, for any two events if one of them has probability 0 or 1, then they are automatically independent. But if we lose this condition as the case when we deal with fuzzy sets, the independence needn't be satisfied unless we have some specified conditions. So we think that the used intersection definition is very strong since it is not compatible with the need to weaken the conditions in classical probability theory to in order to extend it.

Moreover, we think that it will be more appropriate that the independence condition needn't be satisfied in fuzzy random variables since the possibility of the belonging degree of some element ω is affected mainly by our partial knowledge and with the nature of the system and the events we deal with, which affect the whole experiment in some direction besides the observable itself. So we can not translate these effects by supposing that the events are independent, since the independence reflects only the relations and the homogeneity between the subsets themselves, while fuzzyness is a problem that concerns the system itself besides the events and it varies from event to event. To solve the problem, we suggest that we make use of the simple fact that for any two events u, v we have $P(u \cap v) = P(u)P(v|u)$. We think that it is more reasonable to define $X(A \cap B)(\omega) = X(A)(\omega)X(B|A)(\omega)$, where the event $B|A$ is explained according to the problem at hand which may be affected by the whole system and by the element ω itself.

4. A closer look at joint observables

Here, we will try to make use of the concept of joint observable to confirm the need to review the definition of the intersection operation. To this end, we think that we need to revise the original concepts in classical probability theory and how it was translated mathematically to see how we need to extend the intersection operation to be consistent with the whole theory. As we will see, the intersection operation and the concept of joint random variables are very correlated. First, we will consider the classical case. Clas- sically, in any experimental situation, it would be unusual to observe only the value of one random variable. That is, it would be an unusual experiment in which the total data collected consist of one numeric value. In almost all applications, random variables don't occur singly. We will have a need for the tools necessary to describe or model the behavior of n random variables simultaneously [CB90, Lar82]. For example, consider an electronic system containing two components, one for backup, but both underloads. Suppose that the only way the system will fail if both components fail. The distribution, then, of Z, the system life, depends "jointly" on the distribution of X and Y, the components lives. Knowing only the probability distribution of X and Y, though, will not necessarily provide us with enough information to determine the probability distribution of Z. What we need is a probability function (the joint probability distribution) describing the "simultaneous" behavior of X and Y [LM86]. To be more precise this is the fact which makes the definition of dependence or independent the greatest feature that characterizes probability theory from measure theory.

Mathematically, let f, g be any two random variables with probability distributions μ_f, μ_g respectively and let $\mu_{f,g}$ be their joint distribution. Note firstly that we use the joint distribution $\mu_{f,g}$ of the two random variables f, g, these random variables in general may or may not be independent, to get the probabilities of the events in the product space $B(\mathrm{R}^2)$; that is, for any two measurable sets $A, B \in B(R)$, $\mu_{f,g}(A \times B)$ is the probability that the two events $f^{-1}(A), g^{-1}(B)$ fall on the same time; that is (see [Bau81]),

(3.1) $\mu_{f,g}(A \times B) = P(f^{-1}(A) \cap g^{-1}(B))$.

So we can see that the joint distribution $\mu_{f,g}$ of the two random variables f,g can be defined to be the distribution generated by the joint mapping $h: \Omega \to \mathrm{R}^2$ defined by $h(\omega) = (f(\omega), g(\omega))$ [Bau81, Bil86], which satisfies

$$\mu_h(A \times B) = P\big(h^{-1}(A \times B)\big) = P\big(f^{-1}(A) \cap g^{-1}(B)\big) = \mu_{f,g}(A \times B)$$

Hence the joint random variable of f, g is uniquely determined with the joint mapping h. Moreover, $\mu_{f,g}(R \times B) = \mu_g(B)$ and $\mu_{f,g}(A \times R) = \mu_f(A)$. If the random variables are independent, we will have
$\mu_{f,g}(A \times B) = P\big(f^{-1}(A) \cap g^{-1}(B)\big) = \mu_f(A)\mu_g(B)$.

We have similar situation in operational probability theory. But this situation has two stages here. It is not represented only for the joint probability distribution, but it is also represented in the definition of the joint observable (fuzzy random variable).

Doi: https://doi.org/10.54216/GJMSA.050203

Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023

To explain, let X, Y be two observables (fuzzy random variables) and let Z be the joint observable of X, Y . Note that while in classical probability theory we can deal with the joint distribution of two random variables without any need to determine their joint random variable since it is uniquely determined by the relation 3. 1, in operational probability theory we must determine explicitly their joint observable (joint fuzzy random variable) since it is not uniquely determined. In fact, the fuzzy set Z ($A \times B$) is the fuzzy set which represents that a simultaneous measure of X and Y has value in A, B, respectively. That is,
$Z(A \times B) = X(A) \cap Y(B)$
But we must be careful when dealing with the intersection operation here since there are many constrains in quantum mechanics which affect it other than the fuzzy sets themselves such that the uncertainty relation [BHL07], unsharpness of the state [Bus10] and the effect of the measuring process and the apparatus [BLM96, BL10]. Therefore, the joint observable or the simultaneous measure of X and Y doesn't represent the values of the observable X and the observable Y as if we measure each alone, the matter which is im- possible in quantum mechanics [Bus85]. So, we surely will not deal with the product observable, and this is the basic difference from classical probability theory. E. G. Beltrametti and S. Bugajski separate the two concepts of correlation between quantum observables in their papers [BB03, BB04, BB05], the classical correlation and probabilistic entanglement. The probabilistic entanglement can emerge only when the joint observable referred to differs from the product observable [BB03, BB05]. In fact, a consistent separation of a total correlation into a classical and quantum term requires the knowledge of the statistical content of the mixed state we are dealing with [BB04]. So, by connecting the intersection operation with the concept of joint observable, we can deal with the concept of entanglement between fuzzy sets. Moreover, the concept of independence, when it is only defined on the range of observables, can be used to classify those cases in which the intersection of two given fuzzy sets u, v is their product and other cases in which their intersection could not be their product. Moreover, we can define the concept of correlation between fuzzy sets to determine the strength of the relation between them. We think that this relation should be specified when we describe the space we deal with. We leave it to the people interested in operational quantum physics to come up with a suitable definition of correlation in measurements. In all cases we needn't to have Z ($A \times B$) = $X(A).Y(B)$ except in the case we deal with crisp observable since they are another shape of classical random variables. We can notice from the previous brief notes that.

Remark **3.1.** (1) Any joint random variable or joint observable represents some intersection operation, and the converse is also true since if C, D are two events in a probability space (Ω, A, P), define the two random variables $f = I_C, g = I_D$. Hence
$P(C \cap D) = P(f^{-1}(\{1\}) \cap g^{-1}(\{1\}) = \mu_{f,g}(\{1\} \times \{1\})$.
So, the joint distribution for any set of the form $A \times B$ is another expression of the intersection operation of two events. This proves the great connection between the intersection operation and the joint probability distributions.

(2) Classically, the existence and uniqueness of the joint random variable correspond to the fact that the intersection operation of events is a closed operation. Although in classical probability theory the joint random variable is unique, the joint measure of two random variables is not. This fact refers to the relation between the random variables; that is, whether they are independent or not. If the random variables are not independent then knowing all the information about the distribution of any two random variables f and g doesn't provide enough information about their joint distribution. Example

1.18 is an applied example of the need to dependent joint measures and here are some examples of simple joint measures of dependent random variables.

Example 3.2. *Let* (Ω, A, P) *be a probability space and let f, g be two random variables each of them has a range* $\{1, 2, 3\}$ *and their probability distributions* μ_f, μ_g *are given by* $\mu_f(\{x\}) = \mu_g(\{x\}) = \frac{1}{3}, x = 1,2,3$. *Define the joint probability distributions* μ_1, μ_2, μ_3 *on* $(\mathrm{R}^2, B(\mathrm{R}^2))$ *such that;*

(a) $\mu_1(\{(1,1)\}) = \mu_1(\{(2,2)\}) = \mu_1(\{(3,3)\}) = \frac{1}{3}$

(b) $\mu_2(\{(1,3)\}) = \mu_2(\{(2,2)\}) = \mu_2(\{(3,1)\}) = \frac{1}{3}$

(c) $finally$

$$\mu_3(\{(1,1)\}) = mu_3(\{(1,3)\}) = \mu_3(\{(2,1)\}) = \mu_3(\{(2,2)\}) = \mu_3(\{(3,2)\}) = \mu_3(\{(3,3)\}) = \frac{1}{6}$$

Note that the joint probability distributions in the previous example are certainly joint measures, each of them differs from the unique product measure which is the probability measure μ' generated by $\mu'(\{(i,j)\} = \frac{1}{9}\ \forall i,j = 1,2,3$. Also note that μ_3 in part (c) is not commutative. That is, although the two random variables are identically distributed, we have $\mu'(\{(i,j)\}) = \mu'(\{(j,i)\}$ in general. I hope that these facts can be considered as starting point to model the noncommutative nature of quantum mechanics. Other complicated examples can also be easily found in probability theory textbooks. This situation discriminates classical probability theory from its mathematical background measure theory.

Doi: https://doi.org/10.54216/GJMSA.050203

Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023

4.1. **From measure theory to probability theories.**

The mathematical background of probability theory is measuring theory. But measure theory is concerned with the product measure as the unique well-defined joint measure; that is, the product measure $\mu_{1,2}$ of μ_1 and μ_2 is the measure on R^2which satisfies.

$$\mu_{1,2}(B_1 \times B_2) = \mu_1(B_1)\mu_2(B_2) \quad \forall B_1, B_2 \in \mathrm{B}(\mathrm{R}).$$

The situation in probability theory is different. In fact, in probability theory we may deal with various distributions of dependent or independent random variables which are generated in a common probability space. This feature as Kolmogorov and others have remarked, "*it is the concept of independence more than anything else which has given probability theory a life of its own, distinct from other branches of analysis"* [Lam66]. So joining the random variables always has a meaning in probability theory, while in measure theory the product measure is the unique meaningful joint measure since there is no information that can connect the measures to each other. This leads us to the following remarks.

***Remark* 3.3.**

(1) Comparing the condition for the product measure and the joint measure for independent random variables we conclude that the product measure in probability theory treats only the case of random variables that are independent.

(2) If each of the probability spaces (Ω_i, A_i, P_i) describes an experiment E_i with a random outcome, then the product of these probability spaces should describe that experiment E which consist of performing $E_1, \ldots, E_n$one after the other or simultaneously "without mutual influence". The terminology "without mutual influence" can and could henceforth be replaced by "independently". The random variable X_i just defined describes the outcome of the experiment E_i in the joint experiment E. Such experiments are usually possible in classical probability theory [Bau81]. So classical probability theory gives a great importance to the product measure, and we found that nearly all advanced topics in probability theory deal with the product measure and hence consider the independence condition to be satisfied.

(3) The question that arises is the following: how does probability. theory deal with the joint random variables which are dependent? We can separate the treatment for the problem in probability theory into two parts. The first deals directly with the probability of the events in the sample space which correspond to the measurable sets in the product space on $\mathrm{R}^2$by using the conditional probability. That is, $\mu_{f,g}(A \times B) = \mu_f(A|B) = \mu_g(B)$where $\mu_f(A|B)$is called the conditional probability of A given B. To do so we need a complete knowledge of the nature of the probability space we deal with. The second treats the problem partially by dealing only with the moments and correlation coefficient between two random variables which carry only partial information about the strength of the relation between random variables.

According to the previous remarks, and by recognizing the relation between joint observables and the intersection operation, we realize that the nonuniqueness of the joint observable means that there is some intersection operation with various results. This contradicts the uniqueness of the intersection operation as shown in Definition 1.2. In fact, as we have in Theorem. 1.17 the product observable is unique. But as indicated in Example 1.18 we do need in quantum mechanics to deal with joint observables rather than the product one. Moreover, such observables always exist.

Example 3.4. *According to Example 1.18 we have two possible fuzzy sets for the effect $X(\{(\frac{1}{2},\frac{1}{2})\})$, namely*

$$X\left(\left\{\left(\frac{1}{2},\frac{1}{2}\right)\right\}\right)(\omega) = \frac{1}{4}(1 + r_{\omega_1}.r_\omega)(1 + r_{\omega_2}.r_\omega)$$

Or

$$X\left(\left\{\left(\frac{1}{2},\frac{1}{2}\right)\right\}\right)(\omega) = \min\{\frac{1}{2}(1 + r_{\omega_1}.r_\omega), \frac{1}{2}(1 + r_{\omega_2}.r_\omega)\}$$

But if we insist to define the intersection operation as in Definition 1.2, we will have a unique result of $X\left(\left\{\left(\frac{1}{2},\frac{1}{2}\right)\right\}\right)(\omega)$ since

$$\begin{aligned} X\left(\left\{\left(\frac{1}{2},\frac{1}{2}\right)\right\}\right)(\omega) &= X\left(\left(\left\{\frac{1}{2}\right\} \times \mathbb{R}\right) \cap \left(\mathbb{R} \times \left\{\frac{1}{2}\right\}\right)\right)(\omega) \\ &= (X(\left\{\frac{1}{2}\right\} \times \mathbb{R}) \cap X(\left(\mathbb{R} \times \left\{\frac{1}{2}\right\}\right))(\omega) \\ &= (X_{\omega_1}(\left\{\frac{1}{2}\right\}) \cap X_{\omega_2}(\left\{\frac{1}{2}\right\}))(\omega) \\ &= ((X_{\omega_1}\left(\left\{\frac{1}{2}\right\}\right).X_{\omega_2}(\left\{\frac{1}{2}\right\}))(\omega) \\ &= \frac{1}{4}(1 + r_{\omega_1}.r_\omega)(1 + r_{\omega_2}.r_\omega) \end{aligned}$$

In fact, the nonuniqueness of the joint observable is basic in operational probability theory since it has a great importance in illustrating the Bell phenomenon in quantum mechanics which can't be explained by classical

Doi: https://doi.org/10.54216/GJMSA.050203

Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023

probability theory [BB02]. To indicate, let us refer to the reference [BB02] where the authors said: *"The occurrence of the Bell phenomenon in operational probability theory and its nonoccurrence in standard probability theory mirrors the fact that in operational probability theory we have the non-uniqueness of the joint fuzzy random variable while in standard prob- ability theory there is uniqueness of the joint random variable."* Thus, we conclude that we really need to redefine the intersection operation to be consistent with the nonuniqueness of the joint observable.

5. Another approach to define fuzzy operations

Far away from quantum mechanics, there is another totally different well-constructed fuzzy probability theory. The concept of fuzzy random variables was firstly introduced by Puri and Ralesscu(1986) as a generalization of compact random sets to combine both randomness and imprecision. The stochastic variability is represented by use of probability theory while the vagueness by use of fuzzy sets introduced by Zadeh [Zad65]. The operations on fuzzy sets are defined there as follows.

Definition 4.1. Let A, B be two fuzzy sets.

(1) The complement of A, denoted by A^c, is defined by $\mu_{A^c}(x) = 1 - \mu_A(x)\ \ \forall x \in X$

(2) The intersection of A, B is a fuzzy set defined by$\mu_A \cap B(x) = T(\mu_A(x), \mu_B(x))$

where T is a triangular norm (i. e., commutative, associative, non-decreasing in each argument, and $T(a, 1) = a\ \ \forall a \in [0,1]$

(3) The union of A, B is a fuzzy set defined by $\mu_A \cup B(x) = S\big(\mu_A(x), \mu_B(x)\big)$ where S is a triangular conorm (i. e., commutative, associative, nondecreasing in each argument, and $S(a, 0) = a\ \ \forall a \in [0,1]$.

We will exemplify with two widely used t-norms and t-conorms [Pop04]:

(1) (*Standard*) $T_S(a, b) = \min\{a, b\}$and $S_S(a, b) = \max\{a, b\}$

(2) (*Lukasiewicz*) $T_L(a, b) = \max\{a + b - 1, 0\}$ and $S_L(a, b) = \min\{a + b - 1\}$.

Some other examples of t-norms and the corresponding t-conorm are given in the following table.

t-norm	t-conorm
T(x,y)=xy	*S(x,y)=x+y-xy*
$T(x,y) = \begin{cases} \min(x,y), & if \max(x,y) \\ 0, & otherwise \end{cases}$	$S(x,y) = \begin{cases} \max(x,y), & if \min(x,y) = \\ 0, & otherwise \end{cases}$

Note that the definition of intersection in operational probability theory is a special case of t-norm. Moreover, each of these definitions generate the corresponding definitions on crisp sets. So it can easily see that the standard t-norm is also natural generalization of the intersection operation on crisp sets. In fact, the standard t-norm is the largest t-norm we may have. Moreover, it satisfies the familiar property, if u is any fuzzy set then $u \cap u = u$, which is different to the case in remark 2.1.

6. Conclusion and Further Work

The need to extend the definition of the intersection operation has been established through this paper. Now to get a suitable clear definition of the intersection operation, we firstly need to determine the conditions that the intersection operation should have. Further research can use those conditions to examine whether the concept of t-norms can be used to define the intersection operation or not. Even in the case where the conditions of t-norms are not satisfied in quantum mechanics, it is worth to know if we can define the intersection operation by a flexible concept such as the concept of t-norms. Finally, the author concludes that, defining the intersection of two fuzzy events to be their product may give us a very nice theory, However, it will not be the appropriate one for operational quantum physics. In fact, if we manage to extend the definition of the intersection operation, a lot can be done to develop operational probability theory, take for example, the field of conditional probability theory [Mye06] and quantum Markov chains [Gud08, Gud09]. It is thus recommended that specialists of quantum physics should take the lead in studying the validity of the previous suggestions in order to derive the possibilities and the restrictions that should be applied to any intersection operation in quantum physics.

References

[1] [Bau81] H. Bauer, (1981). Probability Theory and Element of Measure Theory. Academic Press. second edition.

[2] [BB95] E. G. Beltrametti, & S. Bugajski, (1995). "Quantum observables in classical framework". Int. J. Theor. Phys. 34(8). 1221–1228.

[3] [BB02] E. G. Beltrametti, & S. Bugajski, (2002). "Quantum mechanics and operational probability theory". Found. Scien. 7. 197–212.

[4] [BB03] E. G. Beltrametti, & S. Bugajski, (2003). "Entanglement and clas- sical correlations in quantum frame". Int. J. Theor. Phys. 42(5). 969–981.

[5] [BB04] E. G. Beltrametti, & S. Bugajski, (2004). "Separating classical and quantum correlations". Int. J. Theor. Phys. 43. 1793– 1801.

Doi: https://doi.org/10.54216/GJMSA.050203

Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023

[6] [BB05] E. G. Beltrametti, & S. Bugajski, (2005). "Correlations and entan-glement in probability theory". Int. J. Theor. Phys. 44(7). 827–837. - [BHL07] P. Busch, T. Heinonen, & P. Lahti, (2007). "Heisenberg'suncertainty principle". Phys. Rep. 452. 155–176.

[7] [Bil86] P. Billingsley, (1986). Probability And Measure. John Wiley & Sons. second edition.

[8] [BL10] Kiukas, J. Busch, P. & P. Lahti, (2010). "On the notion of coexistence in quantum mechanics". Mathematica Slovaca. 60. 665–680 .

[9] [BLM96] P. Busch, P. Lahti, & P. Mittelstaedt, (1996). The Quantum Theory of Measurement. Springer. second edition.

[10] [BS08] P. Busch, & W. Stulpe, (2008). The structure of classical extensions of quantum probability theory. J. Math. Phys. 49. 1–22.

[11] [Bug96] S. Bugajski, (1996). "Fundamentals of fuzzy probability theory". Int. J. Theor. Phys. 35(11). 2229–2244.

[12] [Bug98a] S. Bugajski, (1998)."Fuzzy stochasticprocesses". Open Sys. Info. Dyn. 5. 169–185.

[13] [Bug98b] S. Bugajski, (1998). "On fuzzy random variables and statistical maps". Rep. Math. Phys. 41(1). 1–11.

[14] [Bus85] P. Busch, (1985). "Indeterminacy relations and simultaneous measurements in quantum theory". Int. J. Theor. Phys. 24(1). 63– 91.

[15] [Bus10] G. Busch, P. & Jaeger, (2010). "Unsharp quantum reality". Found. Phys. 40. 1341–1367.

[16] [CB90] G. Casella, & R. L. Berger, (1990). Statistical Inference. Duxbury Press.

[17] [CF10] F. Chovanec, & Roman, Frič. (2010). "States as morphisms". Int. J. Theor. Phys. 49. 3050–3060.

[18] [FP10] R. Frič, & M. Papčo, (2010). "A categorical approach to probability theory". Studia Logica. 94. 215–230.

[19] [Fri07] R. Frič, (2007). "Statistical maps. A categorical approach". Mathematica Slovaca. 57. 41–58.

[20] [Gra89] M. Grabowski, (1989)."What is an observable". Found. Phys. 9(7). 923– 930.

[21] [Gud98] S. Gudder, (1998). "Fuzzy probability theory". Demonestratio Math. 31(3). 235–254.

[22] [Gud99] S. Gudder, (1999). "Observablesand statistical maps". Found. Phys. 29. 877– 897.

[23] [Gud00] S. Gudder, (2000). "What is fuzzy probability theory". Found. Phys. 30. 1663–1678.

[24] [Gud08] S. Gudder, (2008). "Quantum Markov chains". Jour. Math. Phys. 49(7). 72105 – 72300.

[25] [Gud09] S. Gudder, (2009). "Transition effect matrices and quantum Markov chains". Found. Phys. 39. 573–592.

[26] [HN02] E. D. Habil, & T. Z. Nasr, (2002). "On fuzzy probability theory". Int. J. Theor. Phy. 41(4). 791–810.

[27] [Lam66] J. Lamperti, (1966). Probability. a survey of the mathematical theory. W. A. Benjamin.

[28] [Lar82] H. J. Larson, (1982). Introduction to Probability Theory and Statistical Inference. John Wiley and sons. third edition.

[29] [LM86] R. J. Larsen, & M. Marks, (1986). An Introduction to Mathematical Statistics and its Applications. Library of congress cataloging-in- publication data. second edition.

[30] [Mes95] R. Mesiar, (1995). "Do fuzzy quantum structures exists". Int. J. Theor. Phys. 34(8). 1609–1613.

[31] [Mye06] J. M. Myers, (2006). "Conditional probabilities and density operators in quantum modeling". Foundations of Physics. 36. 1012–1035.

[32] [Nas00] T. Z. Nasr, (2000). "Fuzzy probability theory". Master's thesis. Islamic university of Gaza.

[33] [Pop04] H. F. Pop, (2004). "Data analysis with fuzzy sets. A short survey". Stdia univ. Babes-Bolyai. info. XLIX(2). 111–122.

[34] [Rie00] B. Riečan, (2000). "On the probability theory on mv algebras". Soft Computing. 4. 49–57.

[35] [Zad65] L. Zadeh, (1965). "Fuzzy sets". Info. Cont. 8. 338– 353.

Doi: https://doi.org/10.54216/GJMSA.050203

Received: December 15, 2022 Revised: April 11, 2023 Accepted: May 02, 2023