DE GRUYTER

Mathematica Slovaca

DOI: 10.1515/ms-2017-0441
Math. Slovaca **70** (2020), No. 6, 1445–1456

DIVISIBLE EXTENSION OF PROBABILITY

Roman Frič*, c — Peter Eliaš* — Martin Papčo**

(*Communicated by Mirko Navara*)

ABSTRACT. We outline the transition from classical probability space $(\Omega, \mathbf{A}, p)$ to its "divisible" extension, where (as proposed by L. A. Zadeh) the σ-field $\mathbf{A}$ of Boolean random events is extended to the class $\mathcal{M}(\mathbf{A})$ of all measurable functions into [0,1] and the σ-additive probability measure p on $\mathbf{A}$ is extended to the probability integral $\int(\cdot)\,\mathrm{d}p$ on $\mathcal{M}(\mathbf{A})$. The resulting extension of $(\Omega, \mathbf{A}, p)$ can be described as an epireflection reflecting $\mathbf{A}$ to $\mathcal{M}(\mathbf{A})$ and p to $\int(\cdot)\,\mathrm{d}p$.

The transition from $\mathbf{A}$ to $\mathcal{M}(\mathbf{A})$, resembling the transition from whole numbers to real numbers, is characterized by the extension of two-valued Boolean logic on $\mathbf{A}$ to multivalued Łukasiewicz logic on $\mathcal{M}(\mathbf{A})$ and the divisibility of random events: for each random event $u \in \mathcal{M}(\mathbf{A})$ and each positive natural number n we have $u/n \in \mathcal{M}(\mathbf{A})$ and $\int(u/n)\,\mathrm{d}p = (1/n)\int u\,\mathrm{d}p$.

From the viewpoint of category theory, objects are of the form $\mathcal{M}(\mathbf{A})$, morphisms are observables from one object into another one and serve as channels through which stochastic information is conveyed.

We study joint random experiments and asymmetrical stochastic dependence/independence of one constituent experiment on the other one. We present a canonical construction of conditional probability so that observables can be viewed as conditional probabilities.

In the present paper we utilize various published results related to "quantum and fuzzy" generalizations of the classical theory, but our ultimate goal is to stress mathematical (categorical) aspects of the transition from classical to what we call divisible probability.

1. Introduction

There are numerous attempts to upgrade the classical probability space $(\Omega, \mathbf{A}, p)$ introduced by A. N. Kolmogorov in [16]. For example, in order to capture fuzzy phenomena, L. A. Zadeh in [23] proposed to replace it by $(\Omega, \mathcal{M}(\mathbf{A}), \int(\cdot)\,\mathrm{d}p)$, where $\mathcal{M}(\mathbf{A})$ is the class of all $\mathbf{A}$-measurable fuzzy subsets of Ω and $\int(\cdot)\,\mathrm{d}p$ is the probability integral with respect to p. In [20], M. Navara observed that no justification to define the probability of a fuzzy event $f \in \mathcal{M}(\mathbf{A})$ by the formula $\int(f)\mathrm{d}p$ was given by Zadeh and he discussed two distinct approaches to generalized probability, probability on tribes and probability on MV-algebras with products ([22]). Another supportive argument for

2010 Mathematics Subject Classification: Primary 60A02, 60A05, 60A86; Secondary 28A35.
Keywords: Divisible probability, full Łukasiewicz tribe, observable, probability integral, state, statistical map, stochastic channel, fuzzy outcome, degenerated observable, g-joint experiment, stochastic independence, conditional probability, conditional expectation.
The authors acknowledge the support by the grant of the Slovak Research and Development Agency under contract [APVV-16-0073], by the grant of the Slovak Scientific Grant Agency under contract [VEGA project 2/0097/20], and by the Faculty of Education, Catholic University in Ružomberok, grant [GAPF 2/27/2017].
c Corresponding author.

the transition from $(\Omega, \mathbf{A}, p)$ to $(\Omega, \mathcal{M}(\mathbf{A}), \int(\cdot)\mathrm{d}p)$ can be found in [1] (cf. [12, 14]): categorical approach.

In a series of papers ([1, 2, 8, 12–14], see also papers cited therein) we have investigated various generalizations of the classical probability space, where the fuzzy and quantum aspects are dual faces of the same generalized probability space. In the present paper we describe a canonical extension of $(\Omega, \mathbf{A}, p)$ and its basic stochastic properties which can be viewed as a "mathematical optimization" respecting the following principles:

1. Systems of generalized random events constitute objects of a category and classical Boolean random events are treated as a special case.
2. Relevant maps are morphisms of the category in question.
3. Relevant stochastic notions can be defined in terms of morphisms and relevant stochastic constructions can be described via diagrams.
4. The resulting extension of $(\Omega, \mathbf{A}, p)$ is determined by canonical mathematical properties (divisibility, logic, and closedness with respect to sequential limits) and can be described as an epireflection reflecting $\mathbf{A}$ to $\mathcal{M}(\mathbf{A})$ and p to $\int(\cdot)\,\mathrm{d}p$.

Basic category theory makes the transition more transparent. Relevant maps, in particular probability measures, become morphisms, relevant constructions can be schematized via simple diagrams, and the transition can be characterized via epireflection.

Morphisms, called observables, represent channels through which stochastic information flows from one object $\mathcal{M}(\mathbf{A})$ to another object $\mathcal{M}(\mathbf{B})$. Observables can be interpreted via Markov kernels or via suitable conditional probabilities. The added value is a construction of conditional probability on full Łukasiewicz tribes and that "the flow of no relevant stochastic information" yields asymmetrical stochastic independence of $\mathcal{M}(\mathbf{B})$ on $\mathcal{M}(\mathbf{A})$; it amounts to the independence of the outcomes of $\mathcal{M}(\mathbf{B})$ on the outcomes of $\mathcal{M}(\mathbf{A})$.

Answering a problem put forward by B. Riečan and D. Mundici ([22]), conditional probabilities for σ-complete MV-algebras were studied by A. Dvurečenskij and S. Pulmannová ([6]) and for σ-complete MV-algebras with products by T. Kroupa ([17]). This yields, as a special case, conditional probabilities on Łukasiewicz tribes, where the product is the usual product of functions. Our construction of conditional probability sheds light on this special case (interpretation and the role of product).

In the extended theory, some nonclassical, namely, "quantum and fuzzy" phenomena can be modeled. In particular, an observable can map a classical random event $B \in \mathbf{B} \subset \mathcal{M}(\mathbf{B})$ to a nonclassical event $u \in \mathcal{M}(\mathbf{A}) \setminus \mathbf{A}$ and a classical outcome $\omega \in \Omega$, viewed as the Dirac probability measure δ_ω, can "correspond" to a genuine probability measure q on $\mathbf{B}$.

2. Bottom-up

In this section, we recall some properties of probability measures and probability integrals, resp. random variables and measurable maps which lead to the fixing of objects and morphisms in the next section.

2.1. Probability measures and probability integrals

Let $(\Omega, \mathbf{A}, p)$ be a *classical probability space*, where Ω is the set of outcomes, $\mathbf{A}$ is a σ-field of subsets of Ω, and p is a probability measure on $\mathbf{A}$. Let $\mathcal{M}(\mathbf{A})$ be the set of all $\mathbf{A}$-measurable functions on Ω into $[0,1]$, equipped with pointwise operations, pointwise partial order, pointwise sequential convergence, and let $\int(\cdot)\,\mathrm{d}p$ be the *probability integral* on $\mathcal{M}(\mathbf{A})$. For $A \subset \Omega$, denote

$\chi_A \in [0,1]^\Omega$ the indicator function of A ($\chi_A(\omega) = 1$ for $\omega \in A$ and $\chi_A(\omega) = 0$ otherwise). Observe that if $|\Omega| = 1$, $\{a\} = \Omega$, then $\mathbf{A}$ reduces to the trivial σ-field $\mathbf{T} = \{\emptyset, \{a\}\}$ and $[0,1]$ can be viewed as $\mathcal{M}(\mathbf{T})$ (we identify $r \in [0,1]$ and the function $r\chi_{\{a\}} \in \mathcal{M}(\mathbf{T})$). Clearly, the following hold:

(i) $\int \chi_\emptyset \, \mathrm{d}p = 0$, $\int \chi_\Omega \, \mathrm{d}p = 1$;

(ii) If $u, v \in \mathcal{M}(\mathbf{A})$, $v \le u$, then $\int v \, \mathrm{d}p \le \int u \, \mathrm{d}p$ and $\int (u - v) \, \mathrm{d}p = \int u \, \mathrm{d}p - \int v \, \mathrm{d}p$ (subtractivity);

(iii) $\int (\cdot) \, \mathrm{d}p$ is sequentially continuous (with respect to the pointwise sequential convergence of functions in $\mathcal{M}(\mathbf{A})$).

Hence, the probability integral as a map of $\mathcal{M}(\mathbf{A})$ into $\mathcal{M}(\mathbf{T}) \equiv [0,1]$ has the following properties: "*preserves constants, partial order, subtraction and it is sequentially continuous*" (sequential continuity follows from the Dominated Lebesgue Convergence Theorem). The properties characterize probability integrals on $\mathcal{M}(\mathbf{A})$ and the corresponding theorem (cf. [9: Corollary 4.2]) is a cornerstone of the categorical approach to generalized probability. Its proof can be divided into two steps. First, from [21: Lemma 3.3] follows

THEOREM 2.1. *Probability measures on* $\mathbf{A}$ *are exactly subtractive and sequentially continuous maps of* $\mathbf{A}$ *into* $\mathcal{M}(\mathbf{T})$ *which preserve constants and partial order.*

Second, using the Dominated Lebesgue Convergence Theorem (see [9: Lemma 4.1 and Corollary 4.2]) we get

THEOREM 2.2. *Let* h *be a map of* $\mathcal{M}(\mathbf{A})$ *into* $\mathcal{M}(\mathbf{T})$. *If* h *preserves constants, partial order, it is subtractive, and sequentially continuous, then there exists a unique probability measure* p *on* $\mathbf{A}$ *such that* $h = \int (\cdot) \, \mathrm{d}p$.

2.2. Random variables and measurable maps

A *random variable* f is a measurable map sending Ω into the real line R. It "pushes forward" p into the distribution P_f, a probability measure on the real Borel sets $\mathbf{B}_R$ defined by $P_f\big((-\infty, r)\big) = p\big(f^{\leftarrow}((-\infty, r))\big) = p\big(\{\omega \in \Omega; f(\omega) < r\}\big)$. The corresponding preimage $f^{\leftarrow}$ yields a map of $\mathbf{B}_R$ into $\mathbf{A}$.

In what follows, we identify each subset $A \subseteq X$ and its indicator function $\chi_A \colon X \to \{0,1\}$; 0_X and 1_X will denote the indicator functions of the empty set and the universe, respectively.

Let $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ be classical probability spaces and let f be a measurable map of Ω into Ξ. Then f, in a natural way, "pushes forward" p into a probability measure $P_f = p \circ f^{\leftarrow}$ on $\mathbf{B}$, and if $q = P_f$, then f might be called a *generalized random variable*. In fact, f defines a distribution map D_f, $D_f(p) = P_f$, of the set $\mathcal{P}(\mathbf{A})$ of all probability measures on $\mathbf{A}$ into the set $\mathcal{P}(\mathbf{B})$ of all probability measures on $\mathbf{B}$. Note, that if we identify each $\omega \in \Omega$ and the corresponding Dirac probability measure δ_ω, $\delta_\omega(A) = 1$ for $\omega \in A$ and $\delta_\omega(A) = 0$ otherwise, $A \in \mathbf{A}$ (in general, points and Dirac probability measures), then $D_f \restriction \Omega = f$. The preimage map $f^{\leftarrow}$ sending $B \in \mathbf{B}$ to its preimage $f^{\leftarrow}(B) = \{\omega \in \Omega; f(\omega) \in B\} \in \mathbf{A}$ is a Boolean homomorphism (the preimage map $f^{\leftarrow}$ preserves set operations) and it can be uniquely extended to a map $f^{\triangleleft}$ on $\mathcal{M}(\mathbf{B})$ into $\mathcal{M}(\mathbf{A})$ satisfying suitable conditions. Unlike in the "divisible" probability theory, in the classical probability theory $f^{\leftarrow}$ plays only an auxiliary role. Using $\chi_B \circ f = \chi_{f^{\leftarrow}(B)}$, the measurability of f can be reformulated as follows: f is measurable iff for each $B \in \mathbf{B}$ there exists $A \in \mathbf{A}$ such that $\chi_B \circ f = \chi_A$. Consequently, the preimage map $f^{\leftarrow}$ preserves constants, partial order, it is subtractive, and sequentially continuous.

The pair $(f, f^{\leftarrow})$ can be seen as a stochastic channel through which stochastic information flows from $(\Omega, \mathbf{A}, p)$ to $(\Xi, \mathbf{B}, q)$.

The divisible extension of classical probability ([16,18]) can be viewed as a quantum modification of f and a fuzzy modification of $f^{\leftarrow}$. Surprisingly, the two modifications "*are dual*"! The duality has two meanings. Categorical: the two generalizations are morphisms in two dually equivalent categories related by a pair of contravariant functors. Real: the two generalizations represent two dual and equivalent faces of the resulting stochastic channel, fuzzy and quantum. They can be studied in two dual languages.

The generalization of f, called *statistical map*, or operational random variable, can map a Dirac probability measure (a point of Ω) to a probability measure q (remember, a classical f maps points to points). The generalization of $f^{\leftarrow}$, called observable, can map a crisp ({0,1}-valued) random event to a genuine fuzzy random event (remember, $f^{\leftarrow}$ maps crisp sets to crisp sets). Strange, but $f^{\leftarrow}$ does not have a name in the classical probability.

As indicated, maps "preserving constants, partial order, subtractive, and sequentially continuous" will play an important role and will be called *observables.* The composition of observables is defined as the composition of maps and this makes the calculations with observables (diagrams) transparent.

The next theorem guarantees that the generalized probability extends the classical one.

Theorem 2.3. *Let* $(\Omega, \mathbf{A})$, $(\Xi, \mathbf{B})$ *be measurable spaces. Let* $h_{\mathbf{B}}$ *be an observable which maps* $\mathbf{B}$ *into* $\mathcal{M}(\mathbf{A})$.

(i) *There exists a unique observable* h *mapping* $\mathcal{M}(\mathbf{B})$ *into* $\mathcal{M}(\mathbf{A})$ *such that* $h_{\mathbf{B}}$ *is the restriction* $h \restriction \mathbf{B}$ *of* h *to* $\mathbf{B}$*;*

(ii) *If* $h_{\mathbf{B}}(B) \in \mathbf{A}$ *for all* $B \in \mathbf{B}$*, then* $h_{\mathbf{B}}$ *preserves set operations: complement, union, intersection.*

P r o o f. (i) follows directly from [10: Theorem 4.1] and (ii) can be verified by straightforward calculations. □

Observe that a nondegenerated probability q on $\mathbf{B}$ is an observable mapping $\mathbf{B}$ into $\mathcal{M}(\mathbf{T})$ and there exists $B \in \mathbf{B}$ such that $0 < q(B) < 1$, hence $q(B) \notin \{0, 1\}$.

Next, we recall the notion of Łukasiewicz tribe − a generalization of Boolean algebra and a mathematical must to understand the transition from Boolean to multivalued logic and from classical to divisible probability theory ([5,19,20]). A Boolean random event $A \in \mathbf{A}$ can be viewed as a classical propositional function χ_A (having a stochastic interpretation). A generalized random event can be viewed as a manyvalued propositional function $u \in \mathcal{M}(\mathbf{A})$. Since the Łukasiewicz logic extends the Boolean one and probability integrals and probability measures are in one-to-one correspondence, the divisible probability theory can be viewed as an extension of the classical one.

Recall that a bold algebra is a system $\mathcal{X} \subseteq [0,1]^{\Omega}$ containing the constant functions $0_{\mathcal{X}}$, $1_{\mathcal{X}}$ and closed with respect to the usual Łukasiewicz operations: for $u, v \in \mathcal{X}$ put $(u \oplus v)(\omega) = u(\omega) \oplus v(\omega) = \min\{1, u(\omega) + v(\omega)\}$, $u^*(\omega) = 1 - u(\omega)$, $\omega \in \Omega$. Each bold algebra $\mathcal{X} \subseteq [0,1]^{\Omega}$ is a lattice, where for $u, v \in \mathcal{X}$ we have $(u \vee v)(\omega) = u(\omega) \vee v(\omega)$ and $(u \wedge v)(\omega) = u(\omega) \wedge v(\omega)$, $\omega \in \Omega$. If for each $u \in \mathcal{X}$ and for each (positive) natural number n we have $u/n \in \mathcal{X}$, then $\mathcal{X}$ is said to be *divisible.* If a bold algebra $\mathcal{X} \subseteq [0,1]^{\Omega}$ is sequentially closed in $[0,1]^{\Omega}$ (with respect to the pointwise sequential convergence), then $\mathcal{X}$ is said to be a *Łukasiewicz tribe* ($\mathcal{X}$ is closed not only with respect to finite, but also with respect to countable Łukasiewicz sums). It is known that if $\mathcal{X} \subseteq [0,1]^{\Omega}$ is a Łukasiewicz tribe then there is a unique σ-field $\mathbf{A}_{\mathcal{X}}$ of subsets of Ω such that $\mathbf{A}_{\mathcal{X}} \subseteq \mathcal{X} \subseteq \mathcal{M}(\mathbf{A}_{\mathcal{X}})$. Hence, with respect to inclusion, σ-fields of sets and measurable functions into $[0,1]$ are extremal objects. Extremal Łukasiewicz tribes form a distinguished category and play an important role in the next section. Further, $\mathcal{X} = \mathcal{M}(\mathbf{A}_{\mathcal{X}})$ iff $\mathcal{X}$ contains all constant $[0,1]$-valued functions; such Łukasiewicz tribes are called *full* (generated). Clearly, a Łukasiewicz tribe is full iff it is divisible. Observe that σ-fields of sets are exactly $\{0,1\}$-valued Łukasiewicz tribes.

3. Top-down

Taking into account previous arguments, in this section we outline our categorical approach to probability.

3.1. Probability domains

Fields of random events and their generalizations are called *probability domains.* In plain words, a probability domain can be described as follows.

- Start with a "system $\mathcal{A}$ of events";
- Choose a separating set X of "characteristic properties" of the events;
- Choose an "evaluator" E – a suitable set for evaluating the properties (e.g., $\{0,1\}$ in case of Boolean events, or $[0,1]$ in case of fuzzy events);
- Represent each event $a \in \mathcal{A}$ via the "evaluation" of $\mathcal{A}$ into the set E^X of all maps of X into E sending $a \in \mathcal{A}$ to $ev_X(a) \in E^X$, $ev_X(a) \equiv \{ev_x(a);\ x \in X\}$;
- Equip E^X with a suitable algebraic structure and form the minimal "subalgebra" containing $\{a_X;\ a \in \mathcal{A}\}$ so that the resulting object D, called probability domain, has nice categorical properties.

Extending a σ-field $\mathbf{A}$ of Boolean events to $\mathcal{M}(\mathbf{A})$ we get a probability domain which is divisible, $\mathcal{M}(\mathbf{T})$ is the evaluator, and each probability integral on $\mathcal{M}(\mathbf{A})$ is a morphism into the evaluator $\mathcal{M}(\mathbf{T})$.

The Boolean logic can be extended to fuzzy events in many ways. In particular, via the Łukasiewicz logic. The transition from $\mathbf{A}$ to $\mathcal{M}(\mathbf{A})$ guarantees that "no stochastic information is lost": $\mathbf{A}$ is "dense" in $\mathcal{M}(\mathbf{A})$ (the embedding is an epimorphism) and $\mathcal{M}(\mathbf{A})$ is an epireflection of $\mathbf{A}$ (cf. [1]).

3.2. From extremal to full Łukasiewicz tribes

Definition 3.1. Let $\mathcal{X} \subseteq [0,1]^\Omega$ be a Łukasiewicz tribe. If $\mathcal{X} = \mathbf{A}_\mathcal{X}$ or $\mathcal{X} = \mathcal{M}(\mathbf{A}_\mathcal{X})$, then $\mathcal{X}$ is said to be *extremal.* Let h be a map of an extremal Łukasiewicz tribe $\mathcal{Y} \subseteq [0,1]^\Xi$ into an extremal Łukasiewicz tribe $\mathcal{X} \subseteq [0,1]^\Omega$. If $h(0_\Xi) = 0_\Omega$, $h(1_\Xi) = 1_\Omega$, h is subtractive, sequentially continuous with respect to the pointwise convergence of sequences and preserves order, then h is said to be an *observable.* If for each $u \in \mathcal{Y}$, $h(u) \in \mathcal{X}$ is a constant function, then h is said to be *degenerated.*

Clearly, the identity map is an observable and the composition of two observables is an observable. Denote $\mathbb{EL}$ the category consisting of extremal Łukasiewicz tribes as objects and observables as morphisms. Let $\mathbb{FEL}$ be the subcategory of $\mathbb{EL}$ the objects of which are full Łukasiewicz tribes. Then, according to Theorem 2.3, every observable $h_\mathbf{B} \colon \mathbf{B} \to \mathcal{M}(\mathbf{A})$ can be uniquely extended to an observable $h \colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$. Consequently (cf. [1]), if two observables on $\mathcal{M}(\mathbf{B})$ to $\mathcal{M}(\mathbf{A})$ coincide on $\mathbf{B}$, then they are equal, i.e., the embedding of $\mathbf{B}$ into $\mathcal{M}(\mathbf{B})$ is an epimorphism. Hence, for each $q \in \mathcal{P}(\mathbf{B})$, the probability integral $\int(\cdot)\,\mathrm{d}q$ is the unique observable into $\mathcal{M}(\mathbf{T})$ which extends q. This fact is very useful in some calculations related to stochastic independence/dependence.

As a corollary we get the following theorem (cf. [1]).

Theorem 3.2. *$\mathbb{FEL}$ is an epireflective subcategory of $\mathbb{EL}$.*

3.3. Random experiment and stochastic channel

Definition 3.3. Let $\mathbf{A}$ be a σ-algebra of subsets of a set Ω and let $\mathcal{M}(\mathbf{A})$ be the corresponding full Łukasiewicz tribe of measurable functions into [0,1]. An observable on $\mathcal{M}(\mathbf{A})$ into $\mathcal{M}(\mathbf{T})$ is said to be a *state* on $\mathcal{M}(\mathbf{A})$.

For a full Łukasiewicz tribe $\mathcal{M}(\mathbf{A})$, denote $\mathcal{S}(\mathbf{A})$ the set of all states on $\mathcal{M}(\mathbf{A})$. Since states are exactly probability integrals (Theorem 2.2), there is a one-to-one correspondence between $\mathcal{S}(\mathbf{A})$ and $\mathcal{P}(\mathbf{A})$, sending $\int(\cdot)\,\mathrm{d}s$ to s; *denote*

$$\int(\cdot)\,\mathrm{d}s = \overline{s}.$$

Let $(\Omega, \mathbf{A})$, $(\Xi, \mathbf{B})$ be measurable spaces, let $\mathcal{M}(\mathbf{A})$, $\mathcal{M}(\mathbf{B})$ be the corresponding full Łukasiewicz tribes, and let $h\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ be an observable. For each state $\overline{s} \in \mathcal{S}(\mathbf{A})$, the composition $\overline{s}\circ h$ is a state on $\mathcal{M}(\mathbf{B})$. Consequently, h defines the dual map $T_h\colon \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{B})$, $T_h(s) = t$, where for $u \in \mathcal{M}(\mathbf{B})$ we have $\int u\,\mathrm{d}t = \int h(u)\,\mathrm{d}s$; it will be called a *statistical map.* If $g, h\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ are distinct observables, i.e., $(g(u))(\omega) \neq (h(u))(\omega)$ for some $u \in \mathcal{M}(\mathbf{B})$ and $\omega \in \Omega$, then $\int g(u)\,\mathrm{d}\delta_\omega \neq \int h(u)\,\mathrm{d}\delta_\omega$ implies $\int u\,\mathrm{d}(T_g(\delta_\omega)) \neq \int u\,\mathrm{d}(T_h(\delta_\omega))$ and hence $T_g \neq T_h$. In the opposite direction, if $T_g(\delta_\omega) = T_h(\delta_\omega)$ for all $\omega \in \Omega$, then from $\int u\,\mathrm{d}(T_g(\delta_\omega)) = \int u\,\mathrm{d}(T_h(\delta_\omega))$ we get $(g(u))(\omega) = \int g(u)\,\mathrm{d}\delta_\omega = \int h(u)\,\mathrm{d}\delta_\omega = (h(u))(\omega)$, $u \in \mathcal{M}(\mathbf{B})$, and hence $g = h$.

Observe that statistical maps and their products can be defined in terms of Markov kernels ([3, 4, 15]) and the definitions are equivalent to the corresponding notions defined as dual maps to observables ([2, 7, 8]).

Definition 3.4. Let $(\Omega, \mathbf{A})$ be a measurable space and let $p \in \mathcal{P}(\mathbf{A})$. Then $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ is said to be a *random experiment* and $v \in \mathcal{M}(\mathbf{A})$ is said to be a *random event.* Let $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ be another random experiment and let $h\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ be an observable such that $T_h(p) = q$. Then (h, T_h) is said to be a *stochastic channel* from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. If h is degenerated, then (h, T_h) is said to be *degenerated.*

Let $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ be a random experiment. Having in mind "quantum" aspects of "divisible" probability theory, we will view $\mathcal{P}(\mathbf{A})$ as the (generalized) outcomes and Ω, considered as a subset of $\mathcal{P}(\mathbf{A})$, it will represent the classical outcomes. Having in mind "fuzzy" aspects of "divisible" probability theory, we will view pairs (ω, a), $\omega \in \Omega$, $a \in (0, 1]$, as fuzzified classical outcomes; accordingly, $\omega \in \Omega$ will be viewed as $(\omega, 1)$. For $v \in \mathcal{M}(\mathbf{A})$ the set $S_v = \{(\omega, a);\, \omega \in \Omega, 0 < a \leq v(\omega)\}$ will be viewed as fuzzified outcomes supporting the event v and $\int v\,\mathrm{d}p$ measures how "big" the set S_v is.

Let (h, T_h) be a stochastic channel from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. Then the stochastic information is transmitted from the former to the latter experiment via (h, T_h). First, for $u \in \mathcal{M}(\mathbf{B})$ we get $\int u\,\mathrm{d}q$ by sending u to $h(u)$ and measuring $h(u)$ by $\overline{p}$. Second, to each outcome $s \in \mathcal{P}(\mathbf{A})$ of the former experiment we assign the corresponding outcome $t \in \mathcal{P}(\mathbf{B})$, where $\overline{t} = \overline{s} \circ h \in \mathcal{S}(\mathbf{B})$. If h is degenerated, then each $h(u) \in \mathcal{M}(\mathbf{A})$, $u \in \mathcal{M}(\mathbf{A})$, is a constant function the value of which is $\int u\,\mathrm{d}q$. Since for each $s \in \mathcal{P}(\mathbf{A})$ we have $\int h(u)\,\mathrm{d}s = \int u\,\mathrm{d}q$, T_h maps all outcomes $s \in \mathcal{P}(\mathbf{A})$ to q.

Let $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ be an experiment. Recall that a fuzzified outcome (ω, a), $\omega \in \Omega$, $a \in (0, 1]$, supports a random event $v \in \mathcal{M}(\mathbf{A})$ whenever $a \leq v(\omega)$. Let $v \in \mathcal{M}(\mathbf{A})$, $v = c\chi_\Omega$, $c \in [0, 1]$, be a constant function. Then the random event v does not discriminate the outcomes $\omega \in \Omega$ in the following sense: if a fuzzified outcome (ω, a), $\omega \in \Omega$, $a \in (0, 1]$, supports v, then every fuzzified outcome $(\omega', a), \omega' \in \Omega$, supports v and, if (ω, a), $\omega \in \Omega$, $a \in (0, 1]$, does not support v, then no (ω', a), $\omega' \in \Omega$, supports v. Let (h, T_h) be the degenerated stochastic channel from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to an experiment $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. Denote $\mathcal{M}_\xi(\mathbf{B}) = \{u \in \mathcal{M}(\mathbf{B}); 0 < u(\xi)\}$ the set of all random events in $\mathcal{M}(\mathbf{B})$ which are supported by some $(\xi, b), b \in (0, 1]$. Then no $v \in h[\mathcal{M}_\xi(\mathbf{B})] = \{h(u) \in \mathcal{M}(\mathbf{A});\, u \in \mathcal{M}_\xi(\mathbf{B})\}$ discriminates the outcomes $\omega \in \Omega$. In this sense, *the interpretation of a degenerated stochastic channel is that the classical outcomes of the second experiment (points of Ξ) are stochastically independent of the classical outcomes (points of Ω) of the first experiment.*

4. Stochastic dependence via joint experiments

Let $(\Omega, \mathbf{A})$ and $(\Xi, \mathbf{B})$ be measurable spaces and let $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B})$ be their product. Denote L_1 and L_2 the lateral (marginal) projections of $\mathcal{P}(\mathbf{A} \times \mathbf{B})$ into $\mathcal{P}(\mathbf{A})$ and $\mathcal{P}(\mathbf{B})$, respectively, i.e., for $r \in \mathcal{P}(\mathbf{A} \times \mathbf{B})$ and $A \in \mathbf{A}$, $B \in \mathbf{B}$, put $(L_1(r))(A) = r(A \times \Xi)$ and $(L_2(r))(B) = r(\Omega \times B)$.

Let $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ be classical probability spaces and let $r \in \mathcal{P}(\mathbf{A} \times \mathbf{B})$. If $L_1(r) = p$ and $L_2(r) = q$, then we usually identify $A \in \mathbf{A}$ and $A \times \Xi$, resp. $B \in \mathbf{B}$ and $\Omega \times B$, and then we consider $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ as a joint probability space. Denote $\mathcal{J}(p, q) = \{r \in \mathcal{P}(\mathbf{A} \times \mathbf{B}); L_1(r) = p, L_2(r) = q\}$. Then $p \times q \in \mathcal{J}(p, q)$ and if $r = p \times q$, then $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ are said to be *stochastically independent* in $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$.

Clearly, the canonical embeddings (constant prolongations) $e_1\colon \mathcal{M}(\mathbf{A}) \to \mathcal{M}(\mathbf{A} \times \mathbf{B})$ and $e_2\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A} \times \mathbf{B})$ defined by $\tilde{v}(\omega, \xi) = (e_1(v))(\omega, \xi) = v(\omega)$, $\omega \in \Omega, \xi \in \Xi$, $v \in \mathcal{M}(\mathbf{A})$, resp. $\tilde{u}(\omega, \xi) = (e_2(u))(\omega, \xi) = u(\xi)$, $\omega \in \Omega, \xi \in \Xi$, $u \in \mathcal{M}(\mathbf{B})$, are observables.

Lemma 4.1. *L_1 and L_2 are the statistical maps corresponding to e_1 and e_2, respectively.*

We shall show that the classical construction of a conditional expectation (cf. [18]) applied to a joint probability space $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ reveals the relationships between $r \in \mathcal{P}(\mathbf{A} \times \mathbf{B})$ and the asymmetrical stochastic dependence of $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ on $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ in terms of a stochastic channel. In particular, to each stochastic channel (g, T_g) from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ there corresponds a unique $r_p \in \mathcal{P}(\mathbf{A} \times \mathbf{B})$ satisfying certain canonical conditions and g can be interpreted as the conditional probability in the joint probability space $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r_p)$ of $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$.

Definition 4.2. Let $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ be classical probability spaces and let r be a probability measure on $\mathbf{A} \times \mathbf{B}$ such that $L_1(r) = p$ and $L_2(r) = q$. Then $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r})$ is said to be a *joint experiment* of $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$.

Next, we summarize (cf. [2: Proposition 2.8, Proposition 2.10, and Corollary 2.11]) some properties of a joint experiment needed in the sequel.

Theorem 4.3. *Let $(\Omega, \mathbf{A})$ and $(\Xi, \mathbf{B})$ be measurable spaces, let $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B})$ be their product, let $e_1\colon \mathcal{M}(\mathbf{A}) \to \mathcal{M}(\mathbf{A} \times \mathbf{B})$ and $e_2\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A} \times \mathbf{B})$ be the corresponding embeddings, and let L_1, L_2 be the corresponding lateral maps. Let $\mathrm{id}\colon \mathcal{M}(\mathbf{A}) \to \mathcal{M}(\mathbf{A})$ be the identity observable, let $g\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ be an observable, let $T_{\mathrm{id}}\colon \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{A})$ and $T_g\colon \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{B})$ be the corresponding statistical maps.*

(i) *There exists a unique observable $h\colon \mathcal{M}(\mathbf{A} \times \mathbf{B}) \to \mathcal{M}(\mathbf{A})$ such that $h \circ e_1 = \mathrm{id}$ and $h \circ e_2 = g$, where h is equal to the product $\mathrm{id} \otimes g$ of observables id and g defined as follows*

$$\big((\mathrm{id} \otimes g)(u)\big)(\omega) = \int u \,\mathrm{d}(T_{\mathrm{id}}(\delta_\omega) \times T_g(\delta_\omega)), \quad \omega \in \Omega,\ u \in \mathcal{M}(\mathbf{A} \times \mathbf{B}). \tag{$\otimes$}$$

(ii) *Let $T_h\colon \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{A} \times \mathbf{B})$ be the statistical map defined by h. Then $T_h(\delta_\omega) = T_{\mathrm{id} \otimes g}(\delta_\omega) = \delta_\omega \times T_g(\delta_\omega)$ and $L_1 \circ T_h = T_{\mathrm{id}}$, $L_2 \circ T_h = T_g$.*

(iii) *If (g, T_g) is a degenerated stochastic channel from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$, i.e., $T_g(t) = q$ for all $t \in \mathcal{P}(\mathbf{A})$, then $T_h(t) = t \times q$ for all $t \in \mathcal{P}(\mathbf{A})$ and, in particular, $T_h(p) = p \times q$.*

A joint experiment $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r})$, $r \in \mathcal{J}(p, q) \subseteq \mathcal{P}(\mathbf{A} \times \mathbf{B})$, is characterized by the requirement that it contains all stochastic information about its constituents transmitted via the lateral stochastic channels (e_1, L_1) and (e_2, L_2), respectively. Let $g\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ be an observable such that $\overline{q} = \overline{p} \circ g$. From Theorem 4.3 it follows that there exists a unique joint experiment $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r_p})$, $\overline{r_p} = \overline{p} \circ (\mathrm{id} \otimes g)$, which is "the best" of all joint experiments taking into account g. It is determined by the observable $h\colon \mathcal{M}(\mathbf{A}\times\mathbf{B}) \to \mathcal{M}(\mathbf{A})$, $h = \mathrm{id}\otimes g$, and h is the unique observable satisfying two conditions: (i) $h \circ ev_1 = \mathrm{id}$ and (ii) $h \circ ev_2 = g$. Condition (i) guarantees that, for each $v \in \mathcal{M}(\mathbf{A})$, h does not distort the stochastic information $\overline{r_p}(e_1(v)) = \overline{p}(v)$ about $e_1(v) \in \mathcal{M}(\mathbf{A}\times\mathbf{B})$, i.e., $\overline{r_p}(e_1(v)) = \overline{p}(h(e_1(v)) = \overline{p}(v)$. Condition (ii) guarantees that, for each $u \in \mathcal{M}(\mathbf{B})$, h conveys the same stochastic information $\overline{r_p}(e_2(u)) = \overline{p}(h(e_2(u))) = \overline{q}(u) = \overline{p}(g(u))$ about $e_2(u) \in \mathcal{M}(\mathbf{A}\times\mathbf{B})$ as g conveys about u.

Definition 4.4. Let (g, T_g) be a stochastic channel from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. Then $(\Omega\times\Xi, \mathcal{M}(\mathbf{A}\times\mathbf{B}), \overline{r_p})$, $\overline{r_p} = \overline{p}\circ(\mathrm{id}\otimes g)$, is said to be the *$g$-joint experiment* of $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$.

The g-joint experiment is "the best" of all joint experiments reflecting the stochastic information transmitted via (g, T_g) from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. In particular, if (g, T_g) is degenerated, then $r_p = p \times q$. Intuitively, $h = id \otimes g$ is a "prism inside the g-joint experiment changing its outlook": if we look at $(\Omega\times\Xi, \mathcal{M}(\mathbf{A}\times\mathbf{B}), \overline{r_p})$, then "we see $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and within it we see the g-image of $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$".

The g-joint experiment is an important auxiliary mathematical tool used in the construction of conditional probability on full Łukasiewicz tribes.

Definition 4.5. Let (g, T_g) be a stochastic channel from an experiment $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to an experiment $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. Let g be degenerated. Then we say that $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ is *stochastically independently joined to* $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ is *stochastically independent on* $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ *in their g-joint experiment* $(\Omega\times\Xi, \mathcal{M}(\mathbf{A}\times\mathbf{B}), \overline{r_p})$.

Proposition 4.6. *Let* $(\Omega\times\Xi, \mathcal{M}(\mathbf{A}\times\mathbf{B}), \overline{r_p})$, $\overline{r_p} = \overline{p}\circ(\mathrm{id}\otimes g)$, *be the g-joint experiment of* $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ *and* $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$.

(i) *If* $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ *is stochastically independently joined to* $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$, *then* $r_p = p \times q$.

(ii) *For* $B \in \mathbf{B}$ *denote* $A_B = \{\omega \in \Omega;\ (g(\chi_B)(\omega) \neq q(B)\}$. *If* $r_p = p\times q$, *then* $p(A_B) = 0$.

P r o o f. (i) The assertion follows from the preceding lemma.

(ii) Let $r_p = p \times q$. From the previous lemma it follows that for each $A \in \mathbf{A}$ and $B \in \mathbf{B}$ we have $r_p(A\times B) = p(A)\cdot q(B) = \int \chi_A \cdot g(\chi_B)\,\mathrm{d}p$. Denote $A_1 = \{\omega \in \Omega;\ (g(\chi_B))(\omega) = q(B)\}$, $A_2 = \{\omega \in \Omega;\ (g(\chi_B))(\omega) < q(B)\}$, $A_3 = \{\omega \in \Omega;\ (g(\chi_B))(\omega) > q(B)\}$. Clearly, $A_i \in \mathcal{M}(\mathbf{A})$, $i = 1,2,3$. From $p(A_i)\cdot q(B) = \int \chi_{A_i}\cdot g(\chi_B)\,\mathrm{d}p$, $i = 2,3$, we get $p(A_2) = p(A_3) = 0$. Hence $p(A_B) = p(A_2 \cup A_3) = 0$. ☐

Recall that $r = p \times q$ means that the classical experiments $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ are (symmetrically) stochastically independent in their joint experiment $(\Omega\times\Xi, \mathbf{A}\times\mathbf{B}, r)$. The previous proposition provides an explicit description of the relationships between the (symmetrical) stochastic independence of classical experiments and the asymmetrical stochastic independence of their fuzzifications (cf. [2: Proposition 3.7]): in the broader context of g-joint experiment, in one direction, the asymmetrical stochastic independence of $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ on $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ implies the symmetrical stochastic independence of $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ and, conversely, the symmetrical stochastic independence of $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ implies "p-almost" asymmetrical stochastic independence of $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ on $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$.

5. Conditional probability on full Łukasiewicz tribes

Answering a problem posed in [22], conditional probabilities have been defined for σ-complete MV-algebras by A. Dvurečenskij and S. Pulmannová ([6]) and for σ-complete MV-algebras with products by T. Kroupa ([17]) and the operation of product plays an important role. As a special case, this yields a definition of conditional probability for Łukasiewicz tribes, where the operation of product coincides with the usual product of functions. Our construction of conditional probability sheds light on this special case. We start with a stochastic channel (g, T_g) from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$, form the g-joint experiment $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r_p})$ uniquely determined by (g, T_g) and construct the conditional probability on full Łukasiewicz tribes (of fuzzy random events in $\mathcal{M}(\mathbf{B})$ conditioned by fuzzy random events in $\mathcal{M}(\mathbf{A})$ having positive probability) as the canonical fuzzification of classical conditional probability related to the g-joint experiment. The product of fuzzy events extends the intersection of classical (crisp) random events.

We discuss how $\overline{r_p}$ reflects the "stochastic dependence/independence" of experiment $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ on $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$. In particular, we are interested in the construction of a "conditional probability $R(u|v)$ of $u \in \mathcal{M}(\mathbf{B})$ given $v \in \mathcal{M}(\mathbf{A})$". Using the embeddings e_1, e_2, we consider the conditional event $u|v$ as the event $e_2(u)|e_1(v)$ in the joint experiment $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r_p})$ and we show that this leads to a natural construction of $R(u|v)$.

For $u \in \mathcal{M}(\mathbf{B})$ and $v \in \mathcal{M}(\mathbf{A})$ denote $\tilde{u} = e_2(u) \in \mathcal{M}(\mathbf{A} \times \mathbf{B})$ and $\tilde{v} = e_1(v) \in \mathcal{M}(\mathbf{A} \times \mathbf{B})$ the corresponding constant prolongations. For $w \in \mathcal{M}(\mathbf{A} \times \mathbf{B})$, each pair $\big((\omega, \xi), a\big)$, $0 < a \le w(\omega, \xi)$, can be considered as a "fuzzy outcome supporting w". Then the fuzzy event w can be considered as set $M_w = \{\big((\omega, \xi), a\big); 0 < a \le w(\omega, \xi)\}$ of all fuzzy outcomes supporting w and $\int w \, \mathrm{d}r_p$ measures "how big" the set M_w is.

For $B \in \mathbf{B}$, put $\widetilde{\chi_B} = e_2(\chi_B) = \chi_{B \times \Omega} \in \mathcal{M}(\mathbf{A} \times \mathbf{B})$. Then the set $M_{\chi_B \cdot v} = M_{\widetilde{\chi_B} \cdot \tilde{v}} = M_{\widetilde{\chi_B}} \cap M_{\tilde{v}} = \{\big((\omega, \xi), a\big); 0 < a \le \tilde{v}(\omega, \xi), \omega \in B\}$ can be considered as "the set of all fuzzy outcomes supporting $\widetilde{\chi_B}$ given $\tilde{v}$". For $0 < \int \tilde{v} \, \mathrm{d}r_p = \int v \, \mathrm{d}p$, put

$$R_{\mathbf{B}}(\chi_B|v) = \frac{\int \widetilde{\chi_B} \cdot \tilde{v} \, \mathrm{d}r_p}{\int \tilde{v} \, \mathrm{d}r_p}.$$

Clearly, for each $v \in \mathcal{M}(\mathbf{A}), 0 < \int \tilde{v} \, \mathrm{d}r_p = \int v \, \mathrm{d}p$, $R_{\mathbf{B}}(\chi_B|v)$ defines a probability measure $R_{\mathbf{B}}(\cdot|v)$ on $\mathbf{B}$. Consequently, the measure can be uniquely extended to a probability integral, i.e., to an observable over $\mathcal{M}(\mathbf{B})$ into $\mathcal{M}(\mathbf{T})$. Then

$$R(u|v) = \frac{\int \tilde{u} \cdot \tilde{v} \, \mathrm{d}r_p}{\int \tilde{v} \, \mathrm{d}r_p}, \quad u \in \mathcal{M}(\mathbf{B}), \tag{$*$}$$

is the unique extension and it yields the only natural definition of generalized conditional probability based on the stochastic channel (g, T_g) and the corresponding g-joint experiment. For $v = \chi_A$, $A \in \mathbf{A}$, $p(A) > 0$, and $u = \chi_B$, $B \in \mathbf{B}$, $(*)$ reduces to the classical conditional probability for crisp events in $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r_p)$.

Definition 5.1. Let $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r_p})$ be the g-joint experiment of $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. Let $v \in \mathcal{M}(\mathbf{A}), 0 < \int \tilde{v} \, \mathrm{d}r_p = \int v \, \mathrm{d}p$. Then, the observable $R(\cdot|v) \colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{T})$ is said to be the *conditional probability on* $\mathcal{M}(\mathbf{B})$ *given* $v \in \mathcal{M}(\mathbf{A})$.

Lemma 5.2. *Let* $R(\cdot|v) \colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{T})$, $v \in \mathcal{M}(\mathbf{A}), 0 < \int \tilde{v} \, \mathrm{d}r_p = \int v \, \mathrm{d}p$, *be the observable defined by* $(*)$. *Then, for each* $u \in \mathcal{M}(\mathbf{B})$ *we have* $\int \tilde{u} \cdot \tilde{v} \, \mathrm{d}r_p = \int v \cdot g(u) \, \mathrm{d}p$, $\int \tilde{v} \, \mathrm{d}r_p = \int v \, \mathrm{d}p$, *and*

$$R(u|v) = \frac{\int v \cdot g(u) \, \mathrm{d}p}{\int v \, \mathrm{d}p}, \quad u \in \mathcal{M}(\mathbf{B}). \tag{$**$}$$

Proof. First, from $\overline{r_p} = \overline{p} \circ (\mathrm{id} \otimes g)$ we get $\int \tilde{v} \cdot \tilde{u} \, \mathrm{d}r_p = \int (\mathrm{id} \otimes g)(\tilde{v} \cdot \tilde{u}) \, \mathrm{d}p$. Second, from $(\otimes)$ we get $(\mathrm{id} \otimes g)(\tilde{v} \cdot \tilde{u}) = v \cdot g(u)$. Thus, $\int \tilde{v} \cdot \tilde{u} \, \mathrm{d}r_p = \int v \cdot g(u) \, \mathrm{d}p$. Now, the other assertion follows from the fact that $L_1(r_p) = p$. □

From $(**)$ it follows that the observable $g\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ can be viewed as a "global" conditional expectation on $\mathcal{M}(\mathbf{B})$ given $\mathcal{M}(\mathbf{A})$ via the conditional expectation theory.

Let $(\Lambda, \mathbf{C}, P)$ be a classical random experiment (probability space), let $\mathbf{D}$ be a σ-field contained in $\mathbf{C}$, and let $P_{\mathbf{D}}$ be the restriction of P to $\mathbf{D}$. Let $\mathcal{E}$ be the family of all integrable $\mathbf{C}$-measurable functions. Clearly $\mathcal{M}(\mathbf{C}) \subset \mathcal{E}$. Then (cf. [18]) for each $w \in \mathcal{E}$ there exists a $\mathbf{D}$-measurable function $E^{\mathbf{D}}w$, defined up to $P_{\mathbf{D}}$-equivalence by

$$\int_D (E^{\mathbf{D}}w)\,\mathrm{d}P_{\mathbf{D}} = \int_D w\,\mathrm{d}P, \qquad D \in \mathbf{D}. \tag{CE$\mathbf{D}$}$$

$E^{\mathbf{D}}w$ is called the *conditional expectation of w given* $\mathbf{D}$. The restriction of $E^{\mathbf{D}}$ to indicator functions $\chi_C, C \in \mathbf{C}$, is called *conditional probability given* $\mathbf{D}$ and denoted $P^{\mathbf{D}}C = E^{\mathbf{D}}_{\chi_C}$.

We shall deal with a special case, where $\Lambda = \Omega \times \Xi$, $\mathbf{C} = \mathbf{A} \times \mathbf{B}$, $\mathbf{D} = \mathbf{A} \times \{\emptyset, \Xi\}$, $P = r_p$, and $w = \tilde{u} = e_2(u)$, $u \in \mathcal{M}(\mathbf{B})$. Our goal is to describe the relationship between an observable $g\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ and the conditional expectation $E^{\mathbf{D}}$. In [11] the following proposition has been proved.

Proposition 5.3. *Let $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r_p})$ be the g-joint experiment of $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. For $\Lambda = \Omega \times \Xi$, $\mathbf{C} = \mathbf{A} \times \mathbf{B}$, $\mathbf{D} = \mathbf{A} \times \{\emptyset, \Xi\}$, $P = r_p$, let $E^{\mathbf{D}}$ be the corresponding conditional expectation given $\mathbf{D}$. Then (up to $P_{\mathbf{D}}$-equivalence)*

$$E^{\mathbf{D}}\tilde{u} = \widetilde{g(u)} = e_1(g(u)), \quad \tilde{u} = e_2(u), \quad u \in \mathcal{M}(\mathbf{B}).$$

Sketch of the proof. Let $u \in \mathcal{M}(\mathbf{B})$. Since

$$\int_{A\times\Xi} (E^{\mathbf{D}}\tilde{u})\mathrm{d}P_{\mathbf{D}} = \int_{A\times\Xi} \tilde{u}\mathrm{d}P, \quad A \in \mathbf{A},$$

it suffices to prove that for each $A \in \mathbf{A}$ we have

$$\int_{A\times\Xi} \widetilde{g(u)}\mathrm{d}P_{\mathbf{D}} = \int_{A\times\Xi} \tilde{u}\mathrm{d}r_p.$$

Utilizing $(\otimes)$ and Fubini theorem, we get

$$\int_{A\times\Xi} \widetilde{g(u)}\mathrm{d}P_{\mathbf{D}} = \int \chi_A \cdot g(u)\mathrm{d}p = \int_{A\times\Xi} \tilde{u}\mathrm{d}r_p.$$

Since $\tilde{u}$ and $\widetilde{g(u)}$ are canonical embeddings of the events $u \in \mathcal{M}(\mathbf{B})$ and $g(u) \in \mathcal{M}(\mathbf{A})$ into $\mathcal{M}(\mathbf{A} \times \mathbf{B})$, the next definition is quite natural.

Definition 5.4. Let $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r_p})$ be the g-joint experiment of $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. Then, the observable $g\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ is said to be the *conditional probability on $\mathcal{M}(\mathbf{B})$ given $\mathcal{M}(\mathbf{A})$*.

Remark 5.5. From the viewpoint of conditional expectation, for each classical random variable, the corresponding preimage map can be interpreted as a conditional probability. Indeed, let $(\Omega, \mathbf{A}, p)$ be a classical probability space and let $f\colon \Omega \to R$ be a classical random variable. Then the preimage map $f^{\leftarrow}$ is a sequentially continuous Boolean homomorphism mapping real Borel sets $\mathbf{B}$ into $\mathbf{A}$ and it can be uniquely extended to an observable $g\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$. Accordingly, for $\Xi = R$, $B \in \mathbf{B}$, $A \in \mathbf{A}$, $0 < p(A)$, condition $(**)$ gives

$$R(B|A) = \frac{\int \chi_A \cdot \chi_{f^{\leftarrow}(B)}\,\mathrm{d}p}{\int \chi_A\,\mathrm{d}p} = \frac{p(A \cap f^{\leftarrow}(B))}{p(A)}.$$

In view of the previous definition, $f^{\leftarrow}$ as the restriction of g to $\mathbf{B}$ can be interpreted as the *conditional probability on* $\mathbf{B}$ *given* $\mathbf{A}$.

In [17] and [6], a generalized conditional probability for MV-algebras with product is constructed. In particular, for $u, v \in \mathcal{M}(\mathbf{A}), 0 < \int v \, \mathrm{d}p$, $P(u|v)$ is defined via $(\int v \cdot u \, \mathrm{d}p)/(\int v \, \mathrm{d}p)$. Our construction fully supports "conditioning via product" and, what is more important, we claim that for full Łukasiewicz tribes the "conditioning via product" is canonical.

The following special case might be of interest. Consider a g-joint experiment of $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$, where the two experiments are identical and $g \equiv \mathrm{id}$ is the identity observable, or $\mathbf{B} \subset \mathbf{A}$ and g is an embedding. Let $v \in \mathcal{M}(\mathbf{A}), 0 < \int v \, \mathrm{d}p$. Then, for each $u \in \mathcal{M}(\mathbf{A})$ we have

$$R(u|v) = \frac{\int v \cdot u \, \mathrm{d}p}{\int v \, \mathrm{d}p}$$

and for $u = \chi_B$, $v = \chi_A$, $A, B \in \mathbf{A}$, $p(A) > 0$, we get

$$R(u|v) = \frac{\int v \cdot u \, \mathrm{d}p}{\int v \, \mathrm{d}p} = \frac{p(A \cap B)}{p(A)}.$$

As suggested by the referee, the case when v is a constant function is interesting, too. Indeed, then $R(u|v)$ does not depend on v.

Finally, observe that the usual approach to independence via conditional probability is compatible with our approach via stochastic channels. Namely, from $R(v|u) = R(v)$ it follows that $r_p = p \times q$ and "g is p-almost degenerated".

Acknowledgement. The authors are indebted to the referee for valuable suggestions to improve the original text.

REFERENCES

[1] BABICOVÁ, D.: *Probability integral as a linearization*, Tatra Mt. Math. Publ. **72** (2018), 1–15.
[2] BABICOVÁ, D.—FRIČ, R.: *Real functions in stochastic dependence*, Tatra Mt. Math. Publ. **74** (2019), 17–34.
[3] BUGAJSKI, S.: *Statistical maps I. Basic properties*, Math. Slovaca **51** (2001), 321–342.
[4] BUGAJSKI, S.: *Statistical maps II. Operational random variables*, Math. Slovaca **51** (2001), 343–361.
[5] DVUREČENSKIJ, A.—PULMANNOVÁ, S.: *New Trends in Quantum Structures*, Kluwer Academic Publ. and Ister Science, Dordrecht and Bratislava, 2000.
[6] DVUREČENSKIJ, A.—PULMANNOVÁ, S.: *Conditional probability on σ-MV-algebras*, Fuzzy Sets and Systems **155** (2005), 102–118.
[7] ELIAŠ, P.—FRIČ, R.: *Factorization of observables*, Internat. J. Theoret. Phys. **56** (2017), 4073–4083.
[8] FRIČ, R.: *Remarks on statistical maps and fuzzy (operational) random variables*, Tatra Mt. Math. Publ. **30** (2005), 21–34.
[9] FRIČ, R.: *Extension of domains of states*, Soft Comput. **13** (2009), 63–70.
[10] FRIČ, R.: *On D-posets of fuzzy sets*, Math. Slovaca **64** (2014), 545–554.
[11] FRIČ, R.: *Product of measurable spaces and applications*, Tatra Mt. Math. Publ. **74** (2019), 47–56.
[12] FRIČ, R.—PAPČO, M.: *A categorical approach to probability*, Studia Logica **94** (2010), 215–230.
[13] FRIČ, R.—PAPČO, M.: *On probability domains*, Internat. J. Theoret. Phys. **49** (2010), 3092–3100.
[14] FRIČ, R.—PAPČO, M.: *On probability domains IV*, Internat. J. Theoret. Phys. **56** (2017), 4084–4091.
[15] GUDDER, S.: *Fuzzy probability theory*, Demonstratio Math. **31** (1998), 235–254.
[16] KOLMOGOROV, A. N.: *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin, 1933.
[17] KROUPA, T.: *Conditional probability on MV-algebras*, Fuzzy Sets and Systems **149** (2005), 369–381.
[18] LOÈVE, M.: *Probability Theory*, D. Van Nostrand, Inc., Princeton, New Jersey, 1963.
[19] MESIAR, R.: *Fuzzy sets and probability theory*, Tatra Mt. Math. Publ. **1** (1992), 105–123.

[20] NAVARA, M.: *Probability theory of fuzzy events*. In: Fourth Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2005) and Eleventh Rencontres Francophones sur la Logique Floue et ses Applications (E. Montseny, P. Sobrevilla, eds.), Barcelona, Spain, 2005, pp. 325–329.
[21] PAPČO, M.: *On measurable spaces and measurable maps*, Tatra Mt. Math. Publ. **28** (2004), 125–140.
[22] RIEČAN, B.—MUNDICI, D.: *Probability on MV-algebras*. In: Handbook of Measure Theory (E. Pap., ed.), Vol. II, North-Holland, Amsterdam, 2002, pp. 869–910.
[23] ZADEH, L. A.: *Probability measures of fuzzy events*, J. Math. Anal. Appl. **23** (1968), 421–427.

Received 28. 6. 2019
Accepted 22. 4. 2020

* *Mathematical Institute*
Slovak Academy of Sciences
Grešákova 6
SK–040 01 Košice
SLOVAKIA
E-mail: fric@saske.sk
elias@saske.sk

** *Catholic University in Ružomberok*
Hrabovská cesta 1
SK–034 01 Ružomberok
SLOVAKIA
Mathematical Institute
Slovak Academy of Sciences
Štefánikova 49
SK–814 73 Bratislava
SLOVAKIA
E-mail: papco@ruzomberok.sk