Soft Computing (2020) 24:6521–6529
https://doi.org/10.1007/s00500-020-04762-6

FOUNDATIONS

Conditional probability on full Łukasiewicz tribes

Peter Eliaš[1] · **Roman Frič**[1]

Published online: 14 February 2020

Abstract
We study notions of conditional probability and stochastic dependence/independence in an upgraded probability model in which the space of events is modeled by a full Łukasiewicz tribe of all measurable functions from some measurable space into [0, 1]. Our study is based on properties of joint experiments and the notion of stochastic channel, a construct equivalent to the notion of Markov kernel between two measurable spaces. Using the notion of a degenerated stochastic channel, a channel transmitting no stochastic information between two spaces, we define an asymmetrical independence of random experiments. Finally, we define the notion of conditional probability on full Łukasiewicz tribes.

Keywords Random experiment · Stochastic channel · Joint experiment · Fuzzified probability theory · Observable · Probability integral · Statistical map · g-joint experiment · Independence · Conditional probability · Łukasiewicz tribe

1 Introduction

Stochastic channel is a channel through which a stochastic information is transmitted from one random experiment to another one. We study independence/dependence of two experiments based on stochastic channels and joint experiments, both for classical random experiments and for fuzzified random experiments. The latter case is developed within a categorical model of upgraded probability, where the classical outcomes are extended to fuzzified outcomes, Boolean events are extended to measurable fuzzy events (full Łukasiewicz tribes), and probability measures are extended to probability integrals. Stochastic channels are defined via observables (morphisms), and probability integrals are observables into [0,1], viewed as the trivial object. The degenerated channel transmitting no relevant stochastic information defines an asymmetrical independence. The dependence is modeled via the uniquely determined joint experiment, called g-joint, and leads to a canonical conditional probability. Each observable g can be interpreted as a conditional probability. Asymmetrical independence means that the classical outcomes of one experiment do not discriminate the classical outcomes of the other experiment.

Conditional probabilities for σ-complete MV-algebras were studied by Dvurečenskij and Pulmannová and for σ-complete MV-algebras with products by Kroupa. 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 to this special case (interpretation and the role of product).

Communicated by A. Di Nola.

✉ Peter Eliaš
elias@saske.sk

Roman Frič
fric@saske.sk

[1] Mathematical Institute, Slovak Academy of Sciences, Grešákova 6, 040 01 Košice, Slovak Republic

2 Classical stochastic channel

Stochastic channel is a channel through which a stochastic information is transmitted from one random experiment into another one. In the classical case, each random variable f models a stochastic channel from the original (sample) probability space $(\Omega, \mathbf{A}, p)$ to the one induced by f on real numbers. This way, the original stochastic information p is transmitted to the distribution of the random variable f, represented by the corresponding distribution function of f. Further, two random experiments draw the stochastic information from the same given source $(\Omega, \mathbf{A}, p)$; hence, any stochastic information between the two experiments originates in $(\Omega, \mathbf{A}, p)$ (Kolmogorov 1933).

We take a different approach. We do not assume any sample space as a source of all stochastic information. Instead, we shall consider one-way stochastic channels. From the viewpoint of category theory, a one-way stochastic channel is determined by a morphism (arrow) going from one object into another object. As we shall show, within the upgraded probability theory, this enables us to develop an asymmetrical independence of one fuzzified random experiment on another one in terms of morphisms and outcomes. Observe that this is not possible when the stochastic channel is based on classical measurable maps. Indeed, if f is a measurable and measure preserving map from $(\Omega, \mathbf{A}, p)$ into another probability space $(\Xi, \mathbf{B}, q)$, then the "stochastics of the latter space is determined by the stochastics of the former one."

We start with some simple observations. Each classical probability space $(\Omega, \mathbf{A}, p)$ has an "algebraic" component: a measurable space $(\Omega, \mathbf{A})$ (an event space) and a "random" component: a probability measure p. Here, Ω represents experimental outcomes, $\mathbf{A}$ represents random events—measurable sets of outcomes; we identify $A \in \mathbf{A}$ and its indicator function $\chi_A : \Omega \to \{0, 1\}$, $\chi_A(\omega) = 1$ if $\omega \in A$ and $\chi_A(\omega) = 0$ otherwise, and we consider A as the truth set of propositional function "outcome ω supports event A." In what follows, we shall assume that if $\omega_1, \omega_2 \in \Omega$, $\omega_1 \neq \omega_2$, then there exists $A \in \mathbf{A}$ such that $\omega_1 \in A$, $\omega_2 \in (\Omega \setminus A)$ and all singletons are events, i.e., $\{\omega\} \in \mathbf{A}$, $\omega \in \Omega$. Further, p represents the choice of one probability measure (from all possible) the one that "describes the randomness of experiment in the best possible way" and $p(A)$ tells us "how p-big" $A \in \mathbf{A}$ is when compared with the sure event $\Omega \in \mathbf{A}$. In what follows, classical probability spaces will be called **classical** (random) **experiments**.

Let $(\Omega, \mathbf{A})$ and $(\Xi, \mathbf{B})$ be measurable spaces and let $f : \Omega \to \Xi$ be a measurable map . Then, the preimage map $f^{\leftarrow} : \mathbf{B} \to \mathbf{A}$, $f^{\leftarrow}(B) = \{\omega \in \Omega;\ f(\omega) \in B\}$, $B \in \mathbf{B}$, is a sequentially continuous (with respect to the pointwise convergence of indicator functions) Boolean homomorphism and it defines the so-called push-forward map D_f on 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 $\mathbf{B}$, $D_f(t) = t \circ f^{\leftarrow}$, $t \in \mathcal{P}(\mathbf{A})$. As a rule, we identify each outcome o and the corresponding Dirac probability measure δ_o concentrated at the outcome o and then the restriction of D_f to Ω, sending δ_ω to $\delta_{f(\omega)}$, coincides with f. The relationship between $f^{\leftarrow}$ and D_f yields a duality (dual equivalence) and leads to a better understanding of stochastic channels. Without going into details, by a stochastic information we understand information related to outcomes, events, and probabilities of events. Indeed, if $p \in \mathcal{P}(\mathbf{A})$ and $q = D_f(p) = p \circ f^{\leftarrow}$, then all stochastic information about $(\Xi, \mathbf{B}, q)$ is contained in $(\Omega, \mathbf{A}, p)$. The information about outcomes $\xi \in \Xi$ is transmitted via $f : \Omega \to \Xi$, the information about events $B \in \mathbf{B}$ is transmitted via $f^{\leftarrow} : \mathbf{B} \to \mathbf{A}$, and the information about probability measure $p \in \mathcal{P}(\mathbf{A})$ is transmitted via the push-forward map $D_f : \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{B})$. Since $D_f(\delta_\omega) = \delta_\omega \circ f^{\leftarrow} = \delta_{f(\delta_\omega)}$, the transmission of stochastic information is completely determined by $f^{\leftarrow}$. Push-forward maps in the fuzzified probability theory are called statistical maps.

If $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ are classical experiments and $f : \Omega \to \Xi$ is a measurable map such that $D_f(p) = q$, then an outcome $\xi \in \Xi$ appears iff some outcome $\omega \in \Omega$ appears, an event $B \in \mathbf{B}$ appears iff the event $f^{\leftarrow}(B) \in \mathbf{A}$ appears, and $q(B) = p(f^{\leftarrow}(B))$. It is natural to say that $(f^{\leftarrow}, D_f)$ is a **classical stochastic channel** from $(\Omega, \mathbf{A}, p)$ to $(\Xi, \mathbf{B}, q)$.

In the next sections, we outline the theory of one-way stochastic channels, joint experiments, and asymmetrical independence/dependence within upgraded (fuzzified) probability theory (Zadeh 1968; Mesiar 1992; Dvurečenskij and Pulmannová 2000; Foulis and Bennett 1994; Papčo 2004, 2005, 2008; Frič and Papčo 2015, 2017a, b; Babicová 2018). The outline is based on results contained in Babicová and Frič (2019).

3 Classical joint experiment

Intuitively, a joint experiment is an experiment which "from one side looks like one constituent and from the other side looks like the other one." Of course, each viewpoint represents a stochastic channel.

In this section, we analyze what stochastic information is transmitted from one constituent of the joint experiment to the other one. We show that classical stochastic channels do not model such one-way transmissions (via a measurable map) and that the transition from classical probability to fuzzified probability is exactly what is needed. Recall that the transition from $(\Omega, \mathbf{A}, p)$ to $(\Omega, \mathcal{M}(\mathbf{A}), \int(\cdot)\, \mathrm{d}p)$, where $\mathcal{M}(\mathbf{A})$ is the full Łukasiewicz tribe of all measurable fuzzy events, can be best described in terms of an epireflection (cf. Babicová 2018). Further, the transition from transmission of stochastic information from classical joint experiment to the transmission of stochastic information from fuzzified joint experiment is canonical in the sense that the involved classical stochastic channels are fuzzified via the corresponding epireflections. Hence, the classical transmission is uniquely extended to fuzzified transmission. The added value of fuzzification is that the one-way **transmission of no relevant stochastic information** from one fuzzified constituent to the other fuzzified constituent can be explicitly formalized via a **degenerated observable**, i.e., via a degenerated fuzzified stochastic channel.

Let $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ be classical random experiments. Then, the product $\Omega \times \Xi$ models the outcomes (ω, ξ) of joint experiment, the product σ-field $\mathbf{A} \times \mathbf{B}$ models the events of joint experiment, and a probability measure $r \in \mathcal{P}(\mathbf{A} \times \mathbf{B})$ describes the "randomness related to the joint events." Let $pr_1 \colon \Omega \times \Xi \to \Omega$, $pr_1(\omega, \xi) = \omega$, $pr_2 \colon \Omega \times \Xi \to \Xi$, $pr_2(\omega, \xi) = \xi$, be the usual projections. Then, the preimage maps $pr_1^{\leftarrow} \colon \mathbf{A} \to \mathbf{A} \times \mathbf{B}$, $pr_1^{\leftarrow}(A) = A \times \Xi$, $pr_2^{\leftarrow} \colon \mathbf{B} \to \mathbf{A} \times \mathbf{B}$, $pr_2^{\leftarrow}(B) = \Omega \times B$, are sequentially continuous Boolean homomorphisms and define stochastic channels through which $r \in \mathcal{P}(\mathbf{A} \times \mathbf{B})$ is transmitted to lateral (marginal) probabilities $L_1(r) \in \mathcal{P}(\mathbf{A})$, $(L_1(r))(A) = r(pr_1^{\leftarrow}(A)) = r(A \times \Xi)$, $A \in \mathbf{A}$, resp. $L_2(r) \in \mathcal{P}(\mathbf{B})$, $(L_2(r))(B) = r(pr_2^{\leftarrow}(B)) = r(\Omega \times B)$, $B \in \mathbf{B}$. Denote $L_1 \colon \mathcal{P}(\mathbf{A} \times \mathbf{B}) \to \mathcal{P}(\mathbf{A})$, resp. $L_2 \colon \mathcal{P}(\mathbf{A} \times \mathbf{B}) \to \mathcal{P}(\mathbf{B})$, the corresponding lateral statistical maps. Schematically,

$$\Omega \xleftarrow{pr_1} \Omega \times \Xi \xrightarrow{pr_2} \Xi,$$
$$\mathbf{A} \xrightarrow{pr_1^{\leftarrow}} \mathbf{A} \times \mathbf{B} \xleftarrow{pr_1^{\leftarrow}} \mathbf{B},$$
$$\mathcal{P}(\mathbf{A}) \xleftarrow{L_1} \mathcal{P}(\mathbf{A} \times \mathbf{B}) \xrightarrow{L_2} \mathcal{P}(\mathbf{B}).$$

Provided that $L_1(r) = p$ and $L_2(r) = q$, it is natural to call $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ a **classical joint experiment**. Indeed, $(pr_1^{\leftarrow}, L_1)$ yields a stochastic channel from $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ to $(\Omega, \mathbf{A}, p)$, resp. $(pr_2^{\leftarrow}, L_2)$ yields a stochastic channel from $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ to $(\Xi, \mathbf{B}, q)$, and any joint experiment provides (via the corresponding stochastic channel) complete stochastic information about each of the two constituents. Naturally, as we shall see, each particular $r \in \mathcal{P}(\mathbf{A} \times \mathbf{B})$ provides additional stochastic information transmitted between the constituent experiments.

Let us denote $\mathcal{J}(p, q) = \{r \in \mathcal{P}(\mathbf{A} \times \mathbf{B});\ L_1(r) = p,\ L_2(r) = q\}$. Clearly, $p \times q \in \mathcal{J}(p, q)$. If $r = p \times q$, then $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$ are said to be **stochastically independent** in $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$.

Question. Let $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ be a joint experiment of $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$. What stochastic information about $(\Omega, \mathbf{A}, p)$ is transmitted from $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ to $(\Xi, \mathbf{B}, q)$ via the stochastic channel $(pr_2^{\leftarrow}, L_2)$? Similarly, what stochastic information about $(\Xi, \mathbf{B}, q)$ is transmitted from $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ to $(\Omega, \mathbf{A}, p)$ via the stochastic channel $(pr_1^{\leftarrow}, L_1)$?

Observations. 1. The occurrence of an event $B \in \mathbf{B}$ in $(\Xi, \mathbf{B}, q)$ amounts to the occurrence of $pr_2^{\leftarrow}(B) = \Omega \times B$ in $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ and from $L_2(r) = q$ we get $q(B) = r(pr_2^{\leftarrow}(B)) = r(\Omega \times B)$. Hence, with respect to $(pr_2^{\leftarrow}, L_2)$, $q(B)$ does not depend on p. Even more, $q(B)$ will be the same if, instead of $(\Omega, \mathbf{A}, p)$ some other constituent in the joint experiment is considered.

2. If an outcome $\xi \in \Xi$ occurs in $(\Xi, \mathbf{B}, q)$, then the occurrence yields via $(pr_2^{\leftarrow}, L_2)$ no information about outcomes $\omega \in \Omega$ in $(\Omega, \mathbf{A}, p)$, meaning that $pr_2(\omega, \xi) = \xi$ for all $\omega \in \Omega$.

3. If an event $B \in \mathbf{B}$ occurs in $(\Xi, \mathbf{B}, q)$, i.e., if some outcome $\xi \in B$ occurs, then the occurrence yields via $(pr_2^{\leftarrow}, L_2)$ no information about outcomes $\omega \in \Omega$ in $(\Omega, \mathbf{A}, p)$ and hence, without additional assumptions, about events $A \in \mathbf{A}$ in $(\Omega, \mathbf{A}, p)$.

4. Clearly, analogous observations hold about the channel $(pr_1^{\leftarrow}, L_1)$. To sum up, a classical joint experiment $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ contains all stochastic information about both constituent experiments but, without additional assumptions, the constituent experiments do not influence each other. Moreover, simple example of a joint experiment (of two discrete probability spaces having different number of outcomes) shows that in general no measure preserving map of one constituent experiment to the other one exists. As we shall show, the situation in the fuzzified probability theory is different (cf. Frič and Papčo 2010c).

Answer. For each joint experiment $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ of $(\Omega, \mathbf{A}, t)$ and $(\Xi, \mathbf{B}, q)$, $t \in \mathcal{P}(\mathbf{A})$, no relevant information about $(\Omega, \mathbf{A}, t)$ is transmitted from the joint experiment to $(\Xi, \mathbf{B}, q)$ via $(pr_2^{\leftarrow}, L_2)$ and the information about q obtained from the joint experiment via $(pr_2^{\leftarrow}, L_2)$ is the same (does not depend on $t \in \mathcal{P}(\mathbf{A})$): $pr_2^{\leftarrow}(B) = \Omega \times B$, $L_2(r) = q$, $q(B) = r(pr_2^{\leftarrow}(B)) = r(\Omega \times B)$, $B \in \mathbf{B}$. Vice versa, for each joint experiment $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ of $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, s)$, $s \in \mathcal{P}(\mathbf{B})$, no relevant information about $(\Xi, \mathbf{B}, q)$ is transmitted from the joint experiment to $(\Omega, \mathbf{A}, p)$ via $(pr_1^{\leftarrow}, L_1)$ and the information about p obtained from the joint experiment via $(pr_1^{\leftarrow}, L_1)$ is the same (does not depend on $s \in \mathcal{P}(\mathbf{B})$): $pr_1^{\leftarrow}(A) = A \times \Xi$, $L_1(r) = p$, $p(A) = r(pr_1^{\leftarrow}(A)) = r(A \times \Xi)$, $A \in \mathbf{A}$.

Let us note that in classical probability an exceptional situation occurs if $f \colon \Omega \to \Xi$, resp. $g \colon \Xi \to \Omega$, is a degenerated measurable map, i.e., f maps each $\omega \in \Omega$ to a fixed $\xi \in \Xi$, resp. g maps each $\xi \in \Xi$ to a fixed $\omega \in \Omega$. Then, $f^{\leftarrow}(B) = \Omega$ for $\xi \in B$ and $f^{\leftarrow}(B) = \emptyset$ for $\xi \notin B$, $B \in \mathbf{B}$, resp. $g^{\leftarrow}(A) = \Xi$ for $\omega \in A$ and $g^{\leftarrow}(A) = \emptyset$ for $\omega \notin A$, $A \in \mathbf{A}$; the corresponding indicator functions $\chi_{f^{\leftarrow}(B)} \in \mathcal{M}(\mathbf{A})$ and $\chi_{g^{\leftarrow}(A)} \in \mathcal{M}(\mathbf{B})$ are constant functions. But degenerated measurable maps describe deterministic (not stochastic) classical experiments: all outcomes in one experiment are transmitted to a fixed outcome in the other experiment.

4 Asymmetrical independence/dependence via *g*-joint experiment

In this section, we briefly recall (cf. Frič and Papčo 2010a, b, c, 2011, 2015, 2016; Babicová 2018; Papčo 2013; Babicová and Frič 2019) the transition from classical to fuzzified (random) experiments and show that in the fuzzified probability theory an experiment can transmit "no relevant stochastic information" to another experiment via a fuzzified stochastic channel. Our main goal is to discuss the notion of g-joint experiment, a mathematical construct on which the asymmetrical independence and conditional probability on full Łukasiewicz tribes is based.

Let $(\Omega, \mathbf{A})$ be a measurable space. Denote $\mathcal{M}(\mathbf{A})$ the system of all measurable functions $f: \Omega \to [0,1]$, carrying the usual partial order and pointwise sequential convergence of functions. Consider $\mathbf{A}$ as a subset of $\mathcal{M}(\mathbf{A})$. It is known that $\mathbf{A}$ and $\mathcal{M}(\mathbf{A})$ can be viewed as Łukasiewicz tribes (where the Łukasiewicz sum of two functions is defined via $\min\{1, a+b\}$ and the Łukasiewicz product of two functions is defined via $\max\{0, a+b-1\}$). If $\mathbf{A}$ is the trivial field of sets $\mathbf{T} = \{\emptyset, \{\omega\}\}$, then $[0,1]$ can be viewed as $\mathcal{M}(\mathbf{T})$. Denote ELT the category in which objects are of the form $\mathbf{A}$ or $\mathcal{M}(\mathbf{A})$ (extremal Łukasiewicz tribes) and morphisms are sequentially continuous maps preserving partial order, top and bottom constant functions, and partial addition ($a+b$ is defined for $b \le 1-a$); morphisms are called **observables**. Let $(\Xi, \mathbf{B})$ be another measurable space. Then:

- Observables from $\mathbf{B}$ to $\mathbf{A}$ are exactly Boolean homomorphisms from $\mathbf{B}$ to $\mathbf{A}$;
- Observables from $\mathbf{A}$ to $\mathcal{M}(\mathbf{T})$ are exactly probability measures on $\mathbf{A}$;
- Observables from $\mathcal{M}(\mathbf{A})$ to $\mathcal{M}(\mathbf{T})$ are exactly probability integrals on $\mathcal{M}(\mathbf{A})$ ($\overline{p} = \int (\cdot)\, \mathrm{d}p$);
- Each observable from $\mathbf{B}$ to $\mathbf{A}$ (from $\mathbf{B}$ to $\mathcal{M}(\mathbf{A})$) can be uniquely extended to an observable from $\mathcal{M}(\mathbf{B})$ to $\mathcal{M}(\mathbf{A})$.

Accordingly (cf. Babicová 2018), the (full) subcategory FLT of full Łukasiewicz tribes (objects of the form $\mathcal{M}(\mathbf{A})$) is epireflective in ELT, $\mathcal{M}(\mathbf{A})$ is the epireflection of $\mathbf{A}$, $\overline{p} = \int (\cdot)\, \mathrm{d}p$ is the epireflection of $p: \mathbf{A} \to \mathcal{M}(\mathbf{T})$.

Let $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B})$ be the product of two measurable spaces $(\Omega, \mathbf{A})$ and $(\Xi, \mathbf{B})$. Clearly, the canonical embeddings (constant prolongations) $e_1: \mathcal{M}(\mathbf{A}) \to \mathcal{M}(\mathbf{A} \times \mathbf{B})$ and $e_2: \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A} \times \mathbf{B})$ 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 and the unique extensions of coprojections $pr_1^{\leftarrow}$ and $pr_2^{\leftarrow}$, respectively.

Definition 1 Let $(\Omega, \mathbf{A}, p)$ be a classical random experiment. Then, $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ is said to be an **experiment**; it will be called the **fuzzification** of $(\Omega, \mathbf{A}, p)$. Let $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r)$ be a classical joint experiment of $(\Omega, \mathbf{A}, p)$ and $(\Xi, \mathbf{B}, q)$. 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 $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$.

Let $(\Omega, \mathcal{M}(\mathbf{A}), \int (\cdot)\, \mathrm{d}p)$ be an experiment. Usually, the probability integral $\overline{p} = \int (\cdot)\, \mathrm{d}p$ is called a state. In what follows, the elements of $\mathcal{M}(\mathbf{A})$ will be called fuzzy events, or simply events, probability integral will be condensed to probability, and $\overline{p}(v) = \int (v)\, \mathrm{d}p$ will be called the probability of event v.

Let $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ be experiments, and let $g: \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ be an observable. The corresponding **statistical map** $T_g: \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{B})$ is defined as follows: for $t \in \mathcal{P}(\mathbf{A})$ put $T_g(t) = s$, where $\overline{s} = \overline{t} \circ g$.

Observation. It is known (cf. Frič 2005; Bugajski 2001a, b; Gudder 1998) that a statistical map is uniquely determined by the restriction to point (Dirac) measures. More precisely, the following holds. Let $g: \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$, $f: \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ be observables, and let $T_g: \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{B})$ and $T_f: \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{B})$ be corresponding statistical maps. If $T_g(\delta_\omega) = T_f(\delta_\omega)$ for each $\omega \in \Omega$, then $g = f$.

For $u \in \mathcal{M}(\mathbf{B})$ and $\omega \in \Omega$, put $(g(u))(\omega) = \int u\, \mathrm{d}q$. This defines a mapping g of $\mathcal{M}(\mathbf{B})$ into $[0,1]^\Omega$. It is known (cf. Babicová and Frič 2019) that g is an observable (into $\mathcal{M}(\mathbf{A})$) and $T_g(s) = q$ for all $s \in \mathcal{P}(\mathbf{A})$. Further, if $h: \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ is an observable such that each $h(u) \in \mathcal{M}(\mathbf{A})$, $u \in \mathcal{M}(\mathbf{B})$, is a constant function, then T_h maps each $t \in \mathcal{P}(\mathbf{A})$ to $q_h \in \mathcal{P}(\mathbf{B})$, where for $B \in \mathbf{B}$ we have $q_h(B) = (h(\chi_B))(\omega)$, $\omega \in \Omega$. Such observables and statistical maps are called **degenerated**.

Definition 2 Let $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ be experiments. Let $g: \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ be an observable, and let T_g be the corresponding statistical map. Then, (g, T_g) is said to be a **stochastic channel**. If $T_g(p) = q$, then (g, T_g) is said to be a stochastic channel from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. If g and T_g are degenerated, then (g, T_g) is said to be **degenerated**.

In this sense, the interpretation of a degenerated stochastic channel is that the classical outcomes of second experiment **do not discriminate** the classical outcomes of first experiment.

A construction of conditional probability for full Łukasiewicz tribes using stochastic channels has been developed in Babicová and Frič (2019). Important assertions, on which the construction is based, can be reformulated as follows (cf. Proposition 2.8, Proposition 2.10, and Corollary 2.11 in Babicová and Frič 2019), see Fig. 1.

Theorem 3 *Let $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ and $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ be experiments. Let $id: \mathcal{M}(\mathbf{A}) \to \mathcal{M}(\mathbf{A})$ be the identity observable, let $g: \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ be an observable, and let $T_{id}: \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{A})$ and $T_g: \mathcal{P}(\mathbf{A}) \to \mathcal{P}(\mathbf{B})$ be the corresponding statistical maps.*

(1) *There exists a unique observable $h: \mathcal{M}(\mathbf{A} \times \mathbf{B}) \to \mathcal{M}(\mathbf{A})$ such that $h \circ ev_1 = id$ and $h \circ ev_2 = g$, where h is equal to the product $id \otimes g$ of observables id and g defined as*

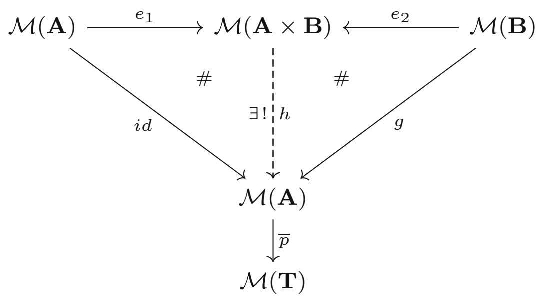

Fig. 1 Commuting observables in a g-joint experiment

$$((id \otimes g)(u))(\omega) = \int u \,\mathrm{d}(T_{id}(\delta_\omega) \times T_g(\delta_\omega)) \qquad (\otimes)$$

for all $\omega \in \Omega$, $u \in \mathcal{M}(\mathbf{A} \times \mathbf{B})$.

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

(3) *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 $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r})$, where $r \in \mathcal{J}(p, q) \subseteq \mathcal{P}(\mathbf{A} \times \mathbf{B})$, 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 $g : \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ be an observable such that $\overline{q} = \overline{p} \circ g$. From Theorem 3, it follows that there exists a unique joint experiment $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r_p})$, $\overline{r_p} = \overline{p} \circ (id \otimes g)$, which is, in some sense, "the best" of all joint experiments taking into account g. It is determined by the observable $h : \mathcal{M}(\mathbf{A} \times \mathbf{B}) \to \mathcal{M}(\mathbf{A})$ and $h = id \otimes g$, and h is the unique observable satisfying two conditions: $h \circ ev_1 = id$ and $h \circ ev_2 = g$. The first condition guarantees that, for each $v \in \mathcal{M}(\mathbf{A})$, h does not distort the stochastic information about $e_1(v) \in \mathcal{M}(\mathbf{A} \times \mathbf{B})$ (the same as the stochastic information about $v = id(v)$). The second condition guarantees that, for each $u \in \mathcal{M}(\mathbf{B})$, h transmits the same stochastic information about $e_2(u)$ (the same stochastic information as about u) as g transmits about u: for each for each $u \in \mathcal{M}(\mathbf{B})$, $g(u) \in \mathcal{M}(\mathbf{A})$ provides stochastic information about u, $e_2(u) \in \mathcal{M}(\mathbf{A} \times \mathbf{B})$ provides stochastic information about u, and $h(e_2(u)) \in \mathcal{M}(\mathbf{A})$ provides stochastic information about u. For $h = id \otimes g$, we have $g(u) = h(e_2(u))$.

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

Intuitively, the g-joint experiment is "the best" joint experiment which reflects the stochastic information transmitted via channel (g, T_g) from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ and, moreover, if (g, T_g) is degenerated, then $r_p = p \times q$. The g-joint experiment is an important auxiliary mathematical tool used in the construction of conditional probability on full Łukasiewicz tribes.

Definition 5 Let (g, T_g) be a stochastic channel from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. If (g, T_g) is degenerated, then $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ is said to be *independently joined by g* to $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$.

Consider the g-joint experiment of $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$, $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. Observe (see (2) in Proposition 2.8 in Babicová and Frič (2019)) that for each $\omega \in \Omega$, $T_h(\delta_\omega) = T_{id \otimes g}(\delta_\omega)$ is a product measure of the form $\delta_\omega \times T_g(\delta_\omega)$. Next, we calculate $T_{id \otimes g}(t)$ for an arbitrary $t \in \mathcal{P}(\mathbf{A})$.

Lemma 6 *For $t \in \mathcal{P}(\mathbf{A})$ put $\overline{r_t} = \overline{t} \circ (id \otimes g)$. Let $A \in \mathbf{A}$ and $B \in \mathbf{B}$. Then, $r_t(A \times B) = \int \chi_A \cdot \chi_B \,\mathrm{d}r_t = \int \chi_A \cdot g(\chi_B)\,\mathrm{d}t$. If g is degenerated, then $r_t = p \times q$.*

Proof From $\overline{r_t} = \overline{t} \circ (id \otimes g)$, it follows that

$$\int \chi_A \cdot \chi_B \,\mathrm{d}r_t = \int (id \otimes g)(\chi_A \cdot \chi_B)\,\mathrm{d}t.$$

From $(\otimes)$, for $\omega \in \Omega$ we get

$$((id \otimes g)(\chi_A \cdot \chi_B))(\omega) = \int \chi_A \cdot \chi_B \,\mathrm{d}(T_{id}(\delta_\omega) \times T_g(\delta_\omega))$$

and, by Fubini theorem,

$$((id \otimes g)(\chi_A \cdot \chi_B))(\omega) = \chi_A(\omega) \cdot \int \chi_B \,\mathrm{d}(T_g(\delta_\omega)).$$

But

$$\int \chi_B \,\mathrm{d}(T_g(\delta_\omega)) = \int g(\chi_B)\,\mathrm{d}(\delta_\omega).$$

Hence,

$$((id \otimes g)(\chi_A \cdot \chi_B))(\omega) = (\chi_A \cdot g(\chi_B))(\omega)$$

and

$$r_t(A \times B) = \int \chi_A \cdot \chi_B \,\mathrm{d}r_t = \int \chi_A \cdot g(\chi_B)\,\mathrm{d}t.$$

The last assertion follows from the fact that if g is degenerated, then $g(\chi_B)$ is a constant function the value of which is $q(B)$. Indeed, then $r_t(A \times B) = p(A) \cdot q(B)$, and hence, $r_t = p \times q$. □

Proposition 7 *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* $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$.

(1) *If* $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ *is independently joined by g to* $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$*, then* $r_p = p \times q$.
(2) *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* $p(A_B) = 0$.

Proof (1) The assertion follows from Lemma 6.

(2) Let $r_p = p \times q$. From Lemma 6, it follows that for each $A \in \mathbf{A}$, $B \in \mathbf{B}$ we have $r_p(A \times B) = p(A) \cdot q(B) = \int \chi_A \cdot g(\chi_B)\,\mathrm{d}p$. Denote $A_1 = \{\omega \in \Omega\,;\ (g(\chi_B))(\omega) = q(B)\}$, $A_2 = \{\omega \in \Omega\,;\ (g(\chi_B))(\omega) < q(B)\}$, and $A_3 = \{\omega \in \Omega\,;\ (g(\chi_B))(\omega) > q(B)\}$. 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 $(\Omega, \mathbf{A}, p)$, $(\Xi, \mathbf{B}, q)$ 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 "independence" of their fuzzifications (cf. Proposition 3.7 in Babicová and Frič 2019): in the broader context of g-joint experiment, the asymmetrical "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 "independence" of $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ on $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$.

Proposition 8 *Consider the g-joint of* $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ *and* $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. *Let* $w \in \mathcal{M}(\mathbf{A} \times \mathbf{B})$. *For* $\omega \in \Omega$*, denote* $w_\omega \in \mathcal{M}(\mathbf{B})$ *the* ω*-cut of w,* $w_\omega(\xi) = w(\omega, \xi), \xi \in \Xi$. *Then,*

(1) $((id \otimes g)(w))(\omega) = \int w_\omega\,\mathrm{d}(T_g(\delta_\omega))$;
(2) $\int w\,\mathrm{d}r_p = \int (\int w_\omega\,\mathrm{d}(T_g(\delta_\omega)))\,\mathrm{d}p$;
(3) *Let* $A \in \mathbf{A}$ *and* $B \in \mathbf{B}$. *Then,* $r_p(A \times B) = \int (\int (\chi_{A\times B})_\omega\,\mathrm{d}(T_g(\delta_\omega)))\,\mathrm{d}p$.

Proof (1) From (⊗), for $\omega \in \Omega$ we get

$$((id \otimes g)(w))(\omega) = \int w\,\mathrm{d}(T_{id}(\delta_\omega) \times T_g(\delta_\omega))$$

and, by Fubini theorem,

$$((id \otimes g)(w))(\omega) = \int w(\omega, \xi)\,\mathrm{d}(T_g(\delta_\omega)).$$

(2) From $\overline{r_p} = \overline{p} \circ (id \otimes g)$, it follows that

$$\int w\,\mathrm{d}r_p = \int (id \otimes g)(w)\,\mathrm{d}p.$$

Now (2) follows from (1).

(3) Clearly, (3) is a special case of (2). □

Corollary 9 *Let* $u \in \mathcal{M}(\mathbf{B})$. *Then, in the preceding proposition, for* $w = e_2(u)$ *we have* $((id \otimes g)(w))(\omega) = \int u\,\mathrm{d}(T_g(\delta_\omega))$ *and, moreover, if g is degenerated, then* $g(u) = (id \otimes g)(e_2(u))$ *is a constant function the value of which is* $\int u\,\mathrm{d}q$.

Observe that, in the view of Corollary 9, the asymmetrical independence of $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ on $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ in their g-joint experiment can be restated as follows: "**the classical outcomes of** $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ **do not discriminate the classical outcomes of** $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$." Indeed, if an outcome $\xi \in \Xi$ does occur then the degenerated observable $g\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ maps each $u \in \mathcal{M}(\mathbf{B})$ to a constant function and so the occurrence of ξ "does not discriminate the outcomes of Ω." Vice versa, the degenerated stochastic channel from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ is the unique one that "transmits no relevant stochastic information about outcomes."

5 Conditional probability on full Łukasiewicz tribes

Conditional probabilities have been defined for σ-complete MV-algebras with products by Kroupa (2005b) and for σ-complete MV-algebras by Dvurečenskij and Pulmannová (2005); 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 to this special case. We start with a stochastic channel (g, T_g) from $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$ to $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$, then 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.

Additional information on probability and conditional probability can be found in Jurečková (2001), Kalina and Nánásiová (2006), Kroupa (2005a), Navara (2005), Riečan (1999), Riečan and Mundici (2002), Vrábelová (2000).

We discuss how $\overline{r_p}$ reflects the "dependence/independence" of $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$ 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 will show that this leads to a natural construction of $R(u|v)$. For $w \in \mathcal{M}(\mathbf{A} \times \mathbf{B})$, each pair $((\omega, \xi), a), 0 < a \leq w(\omega, \xi)$, can be considered as a "fuzzy outcome supporting w" and the set $M_w = \{((\omega, \xi), a)\,;\ 0 < a \leq w(\omega, \xi)\}$ can be considered as the fuzzy event in $\mathcal{M}(\mathbf{A} \times \mathbf{B})$ 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 $M_{\chi_B \cdot v} = M_{\widetilde{\chi_B} \cdot \tilde{v}} = M_{\widetilde{\chi_B}} \cap M_{\tilde{v}} = \{((\omega, \xi), a)\,;\ 0 < a \leq \tilde{v}(\omega, \xi), \omega \in B\}$ 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

$$R(\chi_B|v) = \frac{\int \widetilde{\chi_B} \cdot \tilde{v} \, \mathrm{d}r_p}{\int \tilde{v} \, \mathrm{d}r_p}$$

Clearly, for each $v \in \mathcal{M}(\mathbf{A}), 0 < \int \tilde{v} \, \mathrm{d}r_p = \int v \, \mathrm{d}p$, $R(\chi_B|v)$ defines a probability measure $R(\cdot|v)$ on $\mathbf{B}$. Finally, $R(\cdot|v)$ is an observable into $\mathcal{M}(\mathbf{T})$, and hence, it can be uniquely extended to an observable over $\mathcal{M}(\mathbf{B})$ into $\mathcal{M}(\mathbf{T})$. Then,

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

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 $(\Omega \times \Xi, \mathbf{A} \times \mathbf{B}, r_p)$.

Lemma 10 *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*

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

Proof First, from $\overline{r_p} = \overline{p} \circ (id \otimes g)$ we get $\int (\tilde{v} \cdot \tilde{u}) \, \mathrm{d}r_p = \int (id \otimes g)(\tilde{v} \cdot \tilde{u}) \, \mathrm{d}p$. Second, from $(\otimes)$ we get $(id \otimes g)(\tilde{v} \cdot \tilde{u}) = v \cdot g(u)$. 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 $L_1(r_p) = p$. □

Definition 11 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** $v \in \mathcal{M}(\mathbf{A})$.

Let $(\Lambda, \mathbf{C}, P)$ be a classical random experiment (i.e., a 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. Loève 1963) 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

$$\int_D \left(E^{\mathbf{D}}w\right) \mathrm{d}P_{\mathbf{D}} = \int_D w \, \mathrm{d}P, \quad D \in \mathbf{D};$$

function $E^{\mathbf{D}}w$ 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 D** and denoted $P^{\mathbf{D}}C = E^{\mathbf{D}}_{\chi_C}$.

We shall deal with a special case $\Lambda = \Omega \times \Xi, \mathbf{C} = \mathbf{A} \times \mathbf{B}$, $\mathbf{D} = \mathbf{A} \times \{\emptyset, \Xi\}$, $P = r_p$, 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 $E^{\mathbf{D}}$.

Proposition 12 *Let $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r_p})$ be the g-joint experiment of $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$, $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. 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)*

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

Proof Let $u \in \mathcal{M}(\mathbf{B})$. Since

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

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

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

Since $\overline{r_p} = \overline{p} \circ (id \otimes g)$, we get

$$\int_{A \times \Xi} \tilde{u} \, \mathrm{d}r_p = \int \chi_{A \times \Xi} \cdot \tilde{u} \, \mathrm{d}r_p = \int (id \circ g)(\chi_A \cdot u) \, \mathrm{d}p.$$

Using $(\otimes)$, for all $\omega \in \Omega$ we get

$$((id \otimes g)(\chi_A \cdot u))(\omega) = \int \chi_A \cdot u \, \mathrm{d}(\delta_\omega \times T_g(\delta_\omega))$$
$$= \chi_A(\omega) \cdot \int g(u) \, \mathrm{d}(\delta_\omega) = (\chi_A \cdot g(u))(\omega),$$

and hence

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

Thus $\widetilde{g(u)} = E^{\mathbf{D}}\tilde{u}$ (up to $P_{\mathbf{D}}$-equivalence). □

Definition 13 Let $(\Omega \times \Xi, \mathcal{M}(\mathbf{A} \times \mathbf{B}), \overline{r_p})$ be the g-joint experiment of $(\Omega, \mathcal{M}(\mathbf{A}), \overline{p})$, $(\Xi, \mathcal{M}(\mathbf{B}), \overline{q})$. The observable $g\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ is then said to be the **conditional probability on $\mathcal{M}(\mathbf{B})$ given $\mathcal{M}(\mathbf{A})$**.

Observe that $E^{\mathbf{D}}$ is defined in terms of crisp sets ($\mathbf{D}$), hence the observable $g\colon \mathcal{M}(\mathbf{B}) \to \mathcal{M}(\mathbf{A})$ might be called the conditional probability on $\mathcal{M}(\mathbf{B})$ given $\mathbf{A}$ (cf. p. 380 in Kroupa 2005b).

For generalized probability domains (MV-algebras, Łukasiewicz tribes, D-posets, ...), an additional binary operation "product" has been studied primarily in connection with joint observables, stochastic independence, conditional expectation, and conditional probability, see Riečan (1999), Di Nola and Dvurečenskij (2001), Vrábelová (2000), Jurečková (2001), Riečan and Mundici (2002), Kroupa (2005a, b), Dvurečenskij and Pulmannová (2005), Kalina and Nánásiová (2006), Chovanec et al. (2014), Kôpka (2008). It is known (Riečan and Mundici 2002; Kroupa 2005b) that in a full Łukasiewicz tribe the "product" reduces to the usual pointwise product of functions.

Observe that the construction of generalized conditional probability for MV-algebras and D-posets is based on the operation of product. In Kroupa (2005b), Dvurečenskij and Pulmannová (2005), for $u, v \in \mathcal{M}(\mathbf{A})$, $0 < \int v \, \mathrm{d}p$, $P(u|v)$ is defined via

$$\frac{\int v \cdot u \, \mathrm{d}p}{\int v \, \mathrm{d}p}.$$

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. Let us 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 $g \equiv id$ 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 $v = \chi_B, a = \chi_A$, $A, B \in \mathbf{A}$, $p(A) > 0$ we get

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

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

Acknowledgements First author gratefully acknowledge the support by the grant of the Slovak Scientific Grant Agency VEGA No. 2/0097/20. Second author gratefully acknowledge the support by the Grant of the Slovak Research and Development Agency under contract APVV-16-0073.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Babicová D (2018) Probability integral as a linearization. Tatra Mt Math Publ 72:1–15
Babicová D, Frič R (2019) Real functions in stochastic dependence. Tatra Mt Math Publ 74:47–56
Bugajski S (2001a) Statistical maps I. Basic properties. Math Slovaca 51:321–342
Bugajski S (2001b) Statistical maps II. Operational random variables. Math Slovaca 51:343–361
Chovanec F, Drobná E, Kôpka F, Nánásiová O (2014) Conditional states and independence in D-posets. Soft Comput 14:1027–1034
Di Nola A, Dvurečenskij A (2001) Product MV-algebras. Multi Val Logic 6:193–215
Dvurečenskij A, Pulmannová S (2000) New trends in quantum structures. Kluwer Academic Publ. and Ister Science, Dordrecht and Bratislava
Dvurečenskij A, Pulmannová S (2005) Conditional probability on σ-MV-algebras. Fuzzy Sets Syst 155:102–118
Foulis DJ, Bennett MK (1994) Effect algebras and unsharp quantum logics. Found Phys 24:1331–1352
Frič R (2005) Remarks on statistical maps and fuzzy (operational) random variables. Tatra Mt Math Publ 30:21–34
Frič R, Papčo M (2010a) On probability domains. Internat J Theoret Phys 49:3092–3100
Frič R, Papčo M (2010b) A categorical approach to probability. Stud Log 94:215–230
Frič R, Papčo M (2010c) Fuzzification of crisp domains. Kybernetika 46:1009–1024
Frič R, Papčo M (2011) On probability domains II. Internat J Theoret Phys 50:3778–3786
Frič R, Papčo M (2015) On probability domains III. Internat J Theoret Phys 54:4237–4246
Frič R, Papčo M (2016) Upgrading probability via fractions of events. Commun Math 24:29–41
Frič R, Papčo M (2017a) Probability: from classical to fuzzy. Fuzzy Sets Syst 326:106–114
Frič R, Papčo M (2017b) On probability domains IV. Internat J Theoret Phys 56:4084–4091
Gudder S (1998) Fuzzy probability theory. Demonstr Math 31:235–254
Jurečková M (2001) On the conditional expectation on probability MV-algebras with product. Soft Comput 5:381–385
Kalina M, Nánásiová O (2006) Conditional states and joint distributions on MV-algebras. Kybernetika 42:129–142

Kolmogorov AN (1933) Grundbegriffe der wahrscheinlichkeitsrechnung. Springer, Berlin

Kôpka F (2008) Quasi product on Boolean D-posets. Int J Theor Phys 47:26–35

Kroupa T (2005a) Many-dimensional observables on Łukasiewicz tribe: constructions, conditioning and conditional independence. Kybernetika 41:451–468

Kroupa T (2005b) Conditional probability on MV-algebras. Fuzzy Sets Syst 149:369–381

Loève M (1963) Probability theory. D. Van Nostrand Inc, Princeton

Mesiar R (1992) Fuzzy sets and probability theory. Tatra Mt Math Publ 1:105–123

Navara M (2005) Probability theory of fuzzy events. In: Montseny E, Sobrevilla P (eds.) Fourth Conference of the European society for fuzzy logic and technology (EUSFLAT 2005) and Eleventh Rencontres Francophones sur la Logique Floue et ses Applications, Barcelona, Spain, pp 325–329

Papčo M (2004) On measurable spaces and measurable maps. Tatra Mt Math Publ 28:125–140

Papčo M (2005) On fuzzy random variables: examples and generalizations. Tatra Mt Math Publ 30:175–185

Papčo M (2008) On effect algebras. Soft Comput 12:373–379

Papčo M (2013) Fuzzification of probabilistic objects. In: 8th conference of the European society for fuzzy logic and technology (EUSFLAT 2013), pp 67–71. https://doi.org/10.2991/eusat.2013.10

Riečan B (1999) On the product MV-algebras. Tatra Mt Math Publ 16:143–149

Riečan B, Mundici D (2002) Probability on MV-algebras. In: Pap E (ed) Handbook of measure theory, vol II. North-Holland, Amsterdam, pp 869–910

Vrábelová M (2000) A note on the conditional probability on product MV-algebras. Soft Comput 4:58–61

Zadeh LA (1968) Probability measures of fuzzy events. J Math Anal Appl 23:421–27