Kybernetika

Tomáš Kroupa

Many-dimensional observables on Łukasiewicz tribe: constructions, conditioning and conditional independence

Kybernetika, Vol. 41 (2005), No. 4, [451]--468

Persistent URL: `http://dml.cz/dmlcz/135669`

KYBERNETIKA — VOLUME 41 (2005), NUMBER 4, PAGES 451 – 468

MANY-DIMENSIONAL OBSERVABLES ON ŁUKASIEWICZ TRIBE: CONSTRUCTIONS, CONDITIONING AND CONDITIONAL INDEPENDENCE

Tomáš Kroupa

Probability on collections of fuzzy sets can be developed as a generalization of the classical probability on σ-algebras of sets. A Łukasiewicz tribe is a collection of fuzzy sets which is closed under the standard fuzzy complementation and under the pointwise application of the Łukasiewicz t-norm to countably many fuzzy sets. An observable is a fuzzy set-valued mapping defined on a σ-algebra of sets and satisfying some additional properties; formally, the role of an observable is in a sense analogous to that of a random variable in classical probability theory. This article aims at studying and surveying some properties of observables on a Łukasiewicz tribe of fuzzy sets with a special focus on many-dimensional observables. Namely, the definition and basic construction techniques of observables are discussed. A method for a reasonable construction and interpretation of a joint observable is proposed. Further, the contribution contains results concerning conditioning of observables. We continue in our study [3] of conditional independence in this framework and conclude that semi-graphoid properties are preserved.

Keywords: state, observable, tribe of fuzzy sets, conditional independence

AMS Subject Classification: 60B99, 06D39

1. INTRODUCTION

Probability on fuzzy sets has been developing since the publication of the paper [16] by Zadeh. Its aim is to capture both the vagueness (usually expressed by means of fuzzy set theory) and stochastic uncertainty (usually modeled by probability measures). An approach presented herein is based on MV-algebraic probability theory developed by Riečan and Mundici [10].

It is however worth mentioning that probability on fuzzy sets as a special branch of probability on MV-algebras belongs equally to a much wider context of measure theory on ordered structures such as quantum logics [8] and triangular norm based tribes [1]. After all, the terminology and some of the basic definitions (state, observable, tribe) introduced in the next section originate from both the above mentioned theories.

2. BASIC NOTIONS

This section summarizes essential constructions of probability on fuzzy sets as appearing in [10] and [2]. Let us start by recalling essential notions of measure-theoretical probability: probability space is a triple $(\Omega, \mathcal{A}, P)$, where Ω is a non-empty set, $\mathcal{A}$ is a σ-algebra of subsets of Ω and P is a non-negative measure such that $P(\Omega) = 1$. (Real) random variable ξ on $(\Omega, \mathcal{A}, P)$ is an $\mathcal{A}$-measurable mapping $\xi : \Omega \to \mathbb{R}$, where $\mathbb{R}$ is endowed with Borel σ-algebra.

2.1. Łukasiewicz tribe

Łukasiewicz tribe is a many-valued generalization of a σ-algebras of sets. Let Ω be a non-empty set and $[0,1]^\Omega$ be a family of all functions from Ω to $[0,1]$. For any $f, g \in [0,1]^\Omega$, the Łukasiewicz operations $\oplus, \otimes$ are defined pointwise for all $\omega \in \Omega$:

$$(f \oplus g)(\omega) := \min(1, f(\omega) + g(\omega)),$$

$$(f \otimes g)(\omega) := \max(0, f(\omega) + g(\omega) - 1).$$

A unary operation $\neg$ (standard complement) is further defined:

$$(\neg f)(\omega) := 1 - f(\omega), \quad \omega \in \Omega.$$

A function on Ω which is identically equal to zero (one) is denoted by $\mathbf{0}$ and $\mathbf{1}$, respectively.

A *Łukasiewicz tribe* $\mathcal{T}$ on Ω is a collection of functions $\mathcal{T} \subseteq [0,1]^\Omega$ such that

1. $\mathbf{0} \in \mathcal{T}$,
2. $f \in \mathcal{T} \Rightarrow \neg f \in \mathcal{T}$,
3. $f_n \in \mathcal{T} \Rightarrow \bigoplus_{n \in \mathbb{N}} f_n \in \mathcal{T}$, where $\bigoplus_{n \in \mathbb{N}} f_n(\omega) := \lim_{n \to \infty} \bigoplus_{k=1}^{n} f_k(\omega)$, $\omega \in \Omega$.

Observe that the pointwise limit in the last expression always exists as the sequence $f_1(\omega), f_1(\omega) \oplus f_2(\omega), \ldots$ is monotone and bounded. Elements of $\mathcal{T}$ are called *fuzzy sets* on Ω. *Boolean skeleton* $\mathcal{T}^\vee$ of $\mathcal{T}$ consists of the subsets of Ω corresponding to indicator functions in $\mathcal{T}$, i. e.

$$\mathcal{T}^\vee := \{A \subseteq \Omega : \mathbb{I}_A \in \mathcal{T}\}.$$

These are the fundamental properties [1] of a Łukasiewicz tribe $\mathcal{T}$:

1. $\mathbf{1} \in \mathcal{T}$,
2. $f_n \in \mathcal{T} \Rightarrow \bigvee_{n \in \mathbb{N}} f_n \in \mathcal{T}$, where the supremum is exactly the pointwise supremum of real-valued functions,
3. $\mathcal{T}^\vee$ is a σ-algebra of subsets of Ω,
4. each fuzzy set $f \in \mathcal{T}$ is $\mathcal{T}^\vee$-measurable.

Example 1. Let $\mathcal{A}$ be a σ-algebra of subsets of Ω. Then $\mathcal{T}_\mathcal{A} := \{\mathbb{I}_A : A \in \mathcal{A}\}$ is a Łukasiewicz tribe.

Example 2. Let $[0,1]^\Omega$ be a collection of all fuzzy sets on Ω. Trivially, $[0,1]^\Omega$ is a Łukasiewicz tribe and each fuzzy set $f \in [0,1]^\Omega$ is measurable with respect to the σ-algebra of all subsets of Ω.

Any Łukasiewicz tribe $\mathcal{T}$ is obviously endowed with a partial ordering which is just the pointwise ordering of fuzzy sets on Ω. Consequently, according to [7], $\mathcal{T}$ forms a distributive lattice with the greatest element $\mathbf{1}$ and the lowest element $\mathbf{0}$, where the supremum $\vee$ and the infimum $\wedge$ are defined pointwise. Any Łukasiewicz tribe is also *σ-complete* as a lattice, i. e. any non-empty countable subset of $\mathcal{T}$ has a supremum in $\mathcal{T}$.

Given a lattice L with supremum $\vee$ and infimum $\wedge$ whose order is $\leq$, let us made this stipulation for $a_n, a \in L$: a notation $a_n \nearrow a$ stands for '$a_1 \leq a_2 \leq \dots$ and $\bigvee_{n\in\mathbb{N}} a_n = a$.' Analogously, a notation $a_n \searrow a$ means '$a_1 \geq a_2 \geq \dots$ and $\bigwedge_{n\in\mathbb{N}} a_n = a$.'

Notice that not every Łukasiewicz tribe $\mathcal{T}$ is closed with respect to the usual pointwise multiplication of real functions. Instead of requiring that $\mathcal{T}$ be closed with respect to the multiplication we use the following purely technical simplification.

Example 3. Given a σ-algebra $\mathcal{A}$ of subsets of Ω, let $\mathcal{A}^\wedge$ be a collection of all $\mathcal{A}$-measurable $[0,1]$-valued functions on Ω. Then $\mathcal{A}^\wedge$ is a Łukasiewicz tribe (so called *full tribe*) on Ω.

2.2. State

A state on a Łukasiewicz tribe is a counterpart of a probability measure on a σ-algebra. A *state* m on a Łukasiewicz tribe $\mathcal{T}$ is a mapping

$$m : \mathcal{T} \to [0,1]$$

such that for all $f, g, f_n \in \mathcal{T}$:

1. $m(\mathbf{1}) = 1$,
2. $f \otimes g = \mathbf{0} \Rightarrow m(f \oplus g) = m(f) + m(g)$,
3. $f_n \nearrow f \Rightarrow m(f_n) \nearrow m(f)$.

The condition $f \otimes g = \mathbf{0}$ is equivalent to requiring $f + g \leq \mathbf{1}$. Any state is monotone ($f \leq g \Rightarrow m(f) \leq m(g)$) and one can also demonstrate that for any $f, g \in \mathcal{T}$:

$$m(f \oplus g) = m(f) + m(g) - m(f \otimes g).$$

Example 4. Let $\mathcal{T}$ be a Łukasiewicz tribe on Ω. Then a projection

$$\pi_\omega : f \mapsto f(\omega)$$

is a state on $\mathcal{T}$ for each $\omega \in \Omega$.

Example 5. Let $(\Omega, \mathcal{A}, P)$ be a probability space. Then $\mathcal{T}_{\mathcal{A}}$ is a Łukasiewicz tribe and a mapping $m : \mathcal{T}_{\mathcal{A}} \to [0,1]$ given by

$$m : \mathbb{I}_A \mapsto P(A)$$

is a state on $\mathcal{T}_{\mathcal{A}}$.

Example 6. For any probability space $(\Omega, \mathcal{A}, P)$, let $\mathcal{A}^{\wedge}$ be the full tribe on Ω. A mapping $m_P : \mathcal{A}^{\wedge} \to [0,1]$ given by

$$m_P : f \mapsto \int_{\Omega} f \, \mathrm{d}P$$

is a state on $\mathcal{A}^{\wedge}$.

2.3. Observable

An *observable* has an analogous role as a random variable in classical probability theory. In what follows, X denotes a set, $\mathcal{B}$ is a σ-algebra of its subsets and $\mathcal{B}^n := \sigma(\{ \times_{i=1}^{n} B_i : B_i \in \mathcal{B}\})$ is a product σ-algebra of n copies of $\mathcal{B}$.

Let $\mathcal{T}$ be a Łukasiewicz tribe. An *n-dimensional observable* x on $\mathcal{T}$ is a mapping

$$x : \mathcal{B}^n \to \mathcal{T}$$

such that for all $A, B, A_k \in \mathcal{B}^n$:

1. $x(X^n) = \mathbf{1}$,
2. $A \cap B = \emptyset \Rightarrow x(A) \otimes x(B) = \mathbf{0}$ and $x(A \cup B) = x(A) \oplus x(B)$,
3. $A_k \nearrow A \Rightarrow x(A_k) \nearrow x(A)$.

A mapping x satisfying only 1. and 2. is called a *finitely-additive n-dimensional observable.* In the sequel, 'observable' without any adjective means 'one-dimensional observable'. These are the simplest examples of observables:

Example 7. Assume that a probability space $(\Omega, \mathcal{A}, P)$ is given and let $\mathcal{T}_{\mathcal{A}}$ be the Łukasiewicz tribe on Ω as defined in Example 1. For any random variable $\xi : \Omega \to X$, an observable $x : \mathcal{B} \to \mathcal{T}_{\mathcal{A}}$ can be defined by $x(B) := \mathbb{I}_{\xi^{-1}(B)}$ for any $B \in \mathcal{B}$, where $\xi^{-1}(B) := \{\omega \in \Omega : \xi(\omega) \in B\}$. Moreover, a mapping $P_\xi : B \mapsto P(\xi^{-1}(B))$ is a probability measure on X which is called a *probability distribution* of ξ.

Example 8. Let Ω be a singleton set. Then $[0,1]^{\Omega}$ is a Łukasiewicz tribe (so called *tribe of constants*). Tribe of constants can be obviously identified with the interval $[0,1]$ equipped with the Łukasiewicz t-conorm and the standard complement. Then any observable $P : \mathcal{B} \to [0,1]^{\Omega}$ is a probability measure.

For any state and an observable on a Łukasiewicz tribe one can also find its 'distribution': let x be an n-dimensional observable on $\mathcal{T}$ and m be a state on $\mathcal{T}$. Consider a mapping $m_x : \mathcal{B}^n \to [0,1]$ such that

$$m_x(B) := m(x(B)), \quad B \in \mathcal{B}^n.$$

It straightforwardly follows from the definition of a state and an observable that m_x is a probability measure on X^n.

Remark 1. If $\mathcal{T}$ is a Łukasiewicz tribe on $\Omega = [0,1]$, then each fuzzy set $f \in \mathcal{T}$ can be interpreted as an imprecise of some uncertainty function; for example, $\mathcal{T}$ contains fuzzy sets *small_probability*, *medium_probability* etc. Then any observable $x : \mathcal{B} \to \mathcal{T}$ can be naturally conceived as some 'fuzzy set-valued probability'.

Assume there are given n observables $x_1, \dots, x_n$ on the full tribe $\mathcal{T}$. A *joint (n-dimensional) observable* of $x_1, \dots, x_n$ is any n-dimensional observable $x_{1\dots n}$ such that for all $B_1, \dots, B_n \in \mathcal{B}$:

$$x_{1\dots n}\left(\times_{i=1}^{n} B_i\right) = \prod_{i=1}^{n} x_i(B_i),$$

where the symbol $\prod$ denotes the usual pointwise multiplication of reals. A joint observable always exists [10]. Have in mind that indices in $x_{1\dots n}$ do not commute! For example, given two observables x_1, x_2, their joint observables x_{12} and x_{21} are generally different.

Remark 2. The construction of joint observable is analogous to that of random vector. Notice that neither individual components of random vectors commute: if $\xi_1, \xi_2 : \Omega \to \mathbb{R}$ are random variables, then $(\xi_1, \xi_2) \neq (\xi_2, \xi_1)$ in general.

There are many different definitions of an observable on Łukasiewicz tribes (or MV-algebras) in the literature. For example, an 'observable' according to Riečan and Mundici [10] is a mapping from Borel sets $\mathcal{B}(\mathbb{R})$ satisfying the three properties above. Pulmannová [9] considers an 'observable' to be only finitely-additive and defined on a Boolean algebra. The approach proposed in this paper is not so general from an algebraic viewpoint yet guaranteeing that some of the classical constructions with observables are possible as further explained in Section 3.

3. CONSTRUCTIONS OF OBSERVABLES

Let us consider a collection $\{f_i\}_{i \in I}$, where I is an index set and f_i is a fuzzy set on Ω. Then there is always at least one Łukasiewicz tribe $\mathcal{T}$ containing the collection $\{f_i\}_{i \in I}$: it is the full tribe $\mathcal{A}^\wedge$, where $\mathcal{A}$ is some σ-algebra of subsets of Ω. Moreover, due to Example 4, at least one state can be defined on the Łukasiewicz tribe $\mathcal{T}$. A naturally arising question is whether there can also be defined at least a finitely-additive observable $x : \mathcal{B} \to \mathcal{T}$ such that its range contains the collection $\{f_i\}_{i \in I}$. If such an observable x exists, then, using the terminology of quantum logics, we say that the collection $\{f_i\}_{i \in I}$ is *coexistent.*

Proposition 1. Any at most countable collection of fuzzy sets $\{f_i\}_{i \in I}$ is coexistent.

Proof. A proof for the countable case is to be found in [9]. Assume that I is finite. The idea of the proof is the classical construction appearing in quantum logics [13] and probability on MV-algebras [9]. For a convenience, assume that

$f_i \neq \mathbf{0}, f_i \neq \mathbf{1}$, $i \in I$. Let $\mathcal{T}$ be a Łukasiewicz tribe such that $\{f_i\}_{i\in I} \subseteq \mathcal{T}$. The proof is based on Lemma 1 in [9]: we can find a finite collection of fuzzy sets $\{g_j\}_{j\in J} \subseteq \mathcal{T}$ $(g_j \neq \mathbf{0}, g_j \neq \mathbf{1})$ such that

$$\sum_{j\in J} g_j = \mathbf{1}$$

and for any $i \in I$,

$$f_i = \sum_{j\in J_i'} g_j, \text{ for some } J_i' \subseteq J.$$

Consider a finite set of real numbers $\{a_j\}_{j\in J}$ such that $a_1 \leq a_j$ for all $j \in J$. Define a mapping $x : \mathcal{B}(\mathbb{R}) \to \mathcal{T}$ such that

$$x : B \mapsto \bigoplus_{k:a_k\in B} g_k.$$

Then $x(\mathbb{R}) = \mathbf{1}$. If $A, B \in \mathcal{B}(\mathbb{R})$ are disjoint, then it immediately follows that $x(A) \otimes x(B) = \mathbf{0}$ and $x(A \cup B) = x(A) \oplus x(B)$. The mapping x is a finitely-additive observable which is not an observable. Indeed, let $b_n \in \mathbb{R} \nearrow a_1$. Then $(-\infty, a_1] = \bigcup_{n\in\mathbb{N}}(-\infty, b_n)$ and

$$\bigvee_{n\in\mathbb{N}} x((-\infty, b_n)) = \mathbf{0} \neq g_1 = x((-\infty, a_1]).$$

Finally, it is easy to see the collection $\{f_i\}_{i\in I}$ is contained in the range of x. □

Unfortunately, in order to guarantee coexistency, we can not always restrict ourselves on observables defined on Borel sets of Euclidean space $\mathbb{R}$ or, for example, any other separable space. This can't even be done when proving the coexistency of any countable subset of a Łukasiewicz tribe. Moreover, it is a well-known fact [4] that $\operatorname{card}\mathcal{B}(X) = \mathfrak{c}$, where $\mathcal{B}(X)$ is the Borel σ-algebra of subsets of a separable space X. If $\operatorname{card}\{f_i\}_{i\in I} > \mathfrak{c}$, then there obviously doesn't exist any observable defined on $\mathcal{B}(X)$ such that $\{f_i\}_{i\in I}$ is contained in the range of x. Nevertheless, it can be proven [9] for any non-empty set Ω that $[0,1]^\Omega$ is contained in the range of an observable.

3.1. Random vectors and joint observables

Let $\xi_1, \ldots, \xi_n$ be X-valued random variables defined on Ω. Then a *random vector* (*n-dimensional random variable*) T is a mapping $T : \Omega \to X^n$ such that $T(\omega) = (\xi_1(\omega), \ldots, \xi_n(\omega))$. An image of $\omega \in \Omega$ is an n-tuple $T(\omega) = (a_1, \ldots, a_n) \in X^n$.

Let us compare the previous concept with the definition of a joint observable. For any $i = 1, \ldots, n$, let $x_i : \mathcal{B} \to \mathcal{T}$ be an observable on the full tribe $\mathcal{T}$. A joint n-dimensional observable $x_{1\ldots n}$ is also a mapping into $\mathcal{T}$. Intuitively, this is quite counterintuitive at first sight: for example, let us consider two observables *height* and *mass* each of them characterizing some real-world object in terms of fuzzy sets. These fuzzy sets are of course different and defined on different domains. It is natural to expect that a joint observable that is composed from the two observables above describes all the two properties simultaneously. On the other hand, all the observables and a joint observable share the same range $\mathcal{T}$ according to the definition. A simple technique that resolves this seeming inconvenience is thus presented in the next section.

3.2. A construction of joint observable

In the sequel, for any $i \leq n, n \in \mathbb{N}$, let $\mathcal{T}_i$ be a Łukasiewicz tribe on a non-empty given set Ω_i and let an observable $x_i' : \mathcal{B} \to \mathcal{T}_i$ be given on each $\mathcal{T}_i$. Consider further the full tribe $\mathcal{T}$ on $\Omega = \times_{i=1}^{n} \Omega_i$ and, for any $i \leq n$, a mapping $x_i : \mathcal{B} \to \mathcal{T}$ such that for any $B \in \mathcal{B}$:

$$x_i(B) : (\omega_1, \ldots, \omega_i, \ldots, \omega_n) \mapsto x_i'(B)(\omega_i), \tag{1}$$

where $x_i'(B)(\omega_i)$ is a value of $x_i'(B)$ at a point ω_i. The definition above makes sense as any $\mathcal{T}_i^{\vee}$-measurable fuzzy set (image of x_i') is a $\mathcal{T}^{\vee}$-measurable fuzzy set (image of x_i). The following proposition is an immediate consequence of (1) and the definition of an observable.

Proposition 2. x_i is an observable on the full tribe $\mathcal{T}$.

All the observables x_i' which were primarily defined on the different Łukasiewicz tribes $\mathcal{T}_i$ are now naturally extended on $\mathcal{T}$ by the formula (1): an image

$$f(\omega_1, \ldots, \omega_i, \ldots, \omega_n) = x_i(B)$$

of $B \in \mathcal{B}(\mathbb{R})$ such that $x_i'(B) = f(\omega_i)$ is only a function of ω_i, i. e.

$$f(\omega_1, \ldots, \omega_i, \ldots, \omega_n) = f(\omega_i), \quad (\omega_1, \ldots, \omega_n) \in \Omega.$$

This concept can be justified as follows: any observable x_i is one-dimensional and thus should express only a property of a single kind (e. g., either height or mass). Its values can be therefore viewed as fuzzy subsets of only the universe Ω_i disregarding other universes $\Omega_j, j \neq i$.

Consider now a joint n-dimensional observable $x_{1\ldots n}$ of $x_1, \ldots, x_n$ defined on $\mathcal{B}^n$. Then for any $B_1, \ldots, B_n \in \mathcal{B}$:

$$x_{1\ldots n}\left(\times_{i=1}^{n} B_i\right) = \prod_{i=1}^{n} x_i(B_i) = f(\omega_1, \ldots, \omega_n).$$

It follows from the definition of observables x_i that

$$f(\omega_1, \ldots, \omega_n) = \prod_{i=1}^{n} f_i(\omega_i),$$

where $f_i = x_i(B_i)$.

4. CONDITIONAL INDEPENDENCE OF OBSERVABLES

In order to demonstrate the expressional power of the presented uncertainty theory, a more complex property – so called *conditional independence* – will be defined for observables and semi-graphoid properties (see below) will be proven. There exist definitions of conditional independence based on a variety of uncertainty formalizations, e. g. classical probability theory [5], possibility theory [14]. At first, an appropriate approach to independence and conditioning must be established. For the sake of simplicity, all observables appearing in the rest of the paper are defined on the Borel sets of $\mathbb{R}$.

4.1. Independence

Let $\mathcal{T}$ be the full tribe with a state m and $x_1, \dots, x_n$ be observables on $\mathcal{T}$. Fuzzy sets $f_1, f_2 \in \mathcal{T}$ are called *independent* if

$$m(f_1 \cdot f_2) = m(f_1)m(f_2). \tag{2}$$

The use of the product t-norm $\cdot$ instead of any other t-norm can be reasonably justified [16]. We say that observables x_1, x_2 are *independent* if there exists a joint observable x_{12} such that for all $A, B \in \mathcal{B}(\mathbb{R})$:

$$m_{x_{12}}(A \times B) = m_{x_1}(A)m_{x_2}(B). \tag{3}$$

The independence of more than two observables is defined analogously. Notice that the independence of observables does not depend on their ordering.

4.2. Conditioning

The simplest way of conditioning in classical probability theory is based on two events. Generalizing this to two fuzzy sets $f_1, f_2 \in \mathcal{T}$, where $\mathcal{T}$ is the full tribe with a state m, we say that a real number $m(f|g)$ is a *conditional state* if $m(f|g)$ is a solution of the equation

$$m(g)m(f|g) = m(f \cdot g). \tag{4}$$

A more general approach to conditioning in probability theory on fuzzy sets follows the same line of reasoning as that of conditioning by a random variable in classical probability. Let us only briefly recall the definition of a conditional probability given a random variable. Let $(\Omega, \mathcal{A}, P)$ be a probability space, $A \in \mathcal{A}$ and $\xi : \Omega \longrightarrow \mathbb{R}$ be a random variable. Borel measurable function $P(A|\xi) : \mathbb{R} \longrightarrow \mathbb{R}$ is called a *version of conditional probability* of A given ξ, if for every $B \in \mathcal{B}(\mathbb{R})$:

$$\int_B P(A|\xi)\, \mathrm{d}P_\xi = P(A \cap \xi^{-1}(B)). \tag{5}$$

The following Definition 1 of a conditional state is only a minor modification of the one appearing in [15] – any n-dimensional observable can be in the condition.

Definition 1. Let $\mathcal{T}$ be the full tribe with a state m and x be an n-dimensional observable on $\mathcal{T}$. For any fuzzy set $f \in \mathcal{T}$ we say that a Borel measurable function $\kappa_f : \mathbb{R}^n \to \mathbb{R}$ is a *version of conditional state* of f given x, if for all $B \in \mathcal{B}(\mathbb{R}^n)$:

$$\int_B \kappa_f \, \mathrm{d}m_x = m(f \cdot x(B)). \tag{6}$$

The existence and m_x-a.e. uniqueness of κ_f was proven in [15] as a consequence of Radon–Nikodým Theorem. A version of conditional state κ_f of f given an n-dimensional observable x will be denoted $m(f|x)$ – bear in mind that $m(f|x)$ is a real function of n real arguments. The proposition below guarantees the desired properties of $m(f|x)$: see [15] and [3] for a proof.

Proposition 3. Any version of conditional state of f given an n-dimensional observable x satisfies these properties m_x-a.e.:

1. $m(\mathbf{0}|x) = 0$, $m(\mathbf{1}|x) = 1$,
2. $0 \leq m(f|x) \leq 1$,
3. for any $f, g \in \mathcal{T}$, $m(f \oplus g|x) = m(f|x) + m(g|x) - m(f \otimes g|x)$,
4. if $f \otimes g = \mathbf{0}$, then $m(f \oplus g|x) = m(f|x) + m(g|x)$,
5. if $f_n \nearrow f$, then $m(f_n|x) \nearrow m(f|x)$,
6. if $f \leq g$, then $m(f|x) \leq m(g|x)$.

Example 9. Consider the full tribe $\mathcal{B}(\mathbb{R})^\wedge$ on $\mathbb{R}$ and the observable x on $\mathcal{B}(\mathbb{R})^\wedge$ such that

$$x : B \mapsto \mathbb{I}_B.$$

For any $\mathbb{I}_A \in \mathcal{B}(\mathbb{R})^\wedge$, $m(\mathbb{I}_A|x) = \mathbb{I}_A$ m_x-a.e. since

$$\int_B m(\mathbb{I}_A|x) \, \mathrm{d}m_x = m(\mathbb{I}_A \cdot x(B)) = m(\mathbb{I}_{A \cap B}) = m_x(A \cap B) = \int_B \mathbb{I}_A \, \mathrm{d}m_x,$$

for any $B \in \mathcal{B}(\mathbb{R})$.

Example 10. Let x be an observable on the full tribe $\mathcal{T}$ such that x is concentrated at a point $a \in \mathbb{R}$, that is $x(\{a\}) = \mathbf{1}$. Then for any $f \in \mathcal{T}$:

$$m(f|x) = m(f) \quad m_x\text{-a.e.}$$

Indeed, for any $B \in \mathcal{B}(\mathbb{R})$:

$$\int_B m(f) \, \mathrm{d}m_x = m(f) m_x(B) = m(f \cdot x(B)) = \int_B m(f|x) \, \mathrm{d}m_x.$$

Example 11. For any $n \in \mathbb{N}$, let $\{a_i\}_{i \leq n}$ be a finite set of real numbers and $\{g_i\}_{i \leq n} \in \mathcal{T}$, where $\sum_{i \leq n} g_i = \mathbf{1}$. The last condition evidently implies

$$g_j \otimes g_k = \mathbf{0}, \quad j, k \leq n. \tag{7}$$

Now, the assignment

$$x : \{a_i\} \mapsto g_i, \quad i \leq n,$$

completely determines the finitely-additive observable x on $\mathcal{T}$. Notice that x is concentrated at the set $\{a_i\}_{i \leq n}$, that is $x(\{a_i\}_{i \leq n}) = \mathbf{1}$. Then for any $f \in \mathcal{T}$:

$$m(f|x) = \sum_{i \leq n} m(f|g_i) \mathbb{I}_{\{a_i\}} \quad m_x\text{-a.e.},$$

where $m(f|g_i) \in \mathbb{R}$ is a conditional state defined in (4). Let us verify the equality above. For any $B \in \mathcal{B}(\mathbb{R})$:

$$\int_B \sum_{i \leq n} m(f|g_i) \mathbb{I}_{\{a_i\}} \, \mathrm{d}m_x = \sum_{i \leq n} m(f|g_i) \int_{B \cap \{a_i\}} \mathrm{d}m_x. \tag{8}$$

If $B \cap \{a_i\}_{i \leq n} = \emptyset$, then $m(f \cdot x(B)) = m(f \cdot \mathbf{0}) = 0$ which is evidently equal to (8). If $B \cap \{a_i\}_{i \leq n} = \{a_{i_j}\}_{j \leq p}$ for some $p \leq n$, then (8) is further equal to

$$\sum_{j \leq p} m(f|g_{i_j}) m(g_{i_j}) = \sum_{j \leq p} m(f \cdot g_{i_j}) \tag{9}$$

Employing the property (7), the expression (9) equals

$$m\left(\bigoplus_{j \leq p} f \cdot g_{i_j} \right) = m\left(f \cdot \bigoplus_{j \leq p} g_{i_j} \right) = m(f \cdot x(B)),$$

which finishes the verification.

It is an open problem under which conditions for any $u \in \mathbb{R}^n$ a version of conditional state can be chosen within its equivalence class such that $m(.|x)(u)$ is a state on $\mathcal{T}$. In classical probability theory, this selection can be carried out only under certain topological assumptions [6]: namely, if a probability space $(\Omega, \mathcal{A}, P)$ is such that Ω is a separable complete metric space endowed with Borel σ-algebra $\mathcal{A}$, then for any real random variable ξ and any $u \in \mathbb{R}$ a version of conditional probability can be chosen within its equivalence class such that $P(.|\xi)(u)$ is a probability measure on Ω.

In the rest of this section, two important lemmas are stated: they will be utilized later when proving various properties of conditional independence. In fact, the first one is a 'weaker' definition of a conditional state.

Lemma 1. Let $\mathcal{T}$ be the full tribe with a state m and x be an n-dimensional observable on $\mathcal{T}$. Let $\mathcal{A} \subseteq \mathcal{B}(\mathbb{R}^n)$ be a family of sets containing $\mathbb{R}^n$ and closed with respect to finite intersections such that the σ-algebra generated by $\mathcal{A}$ coincides with $\mathcal{B}(\mathbb{R}^n)$, i. e. $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^n)$. Then κ_f is a version of conditional state of f given x if and only if for all $B \in \mathcal{A}$:

$$\int_B \kappa_f \,\mathrm{d}m_x = m(f \cdot x(B)). \tag{10}$$

Proof. The first implication is trivial. To prove that (10) is also the sufficient condition, let

$$\mathcal{M} := \left\{ B \in \mathcal{B}(\mathbb{R}^n) : \int_B \kappa_f \,\mathrm{d}m_x = m(f \cdot x(B)) \right\}.$$

The inclusion $\mathcal{A} \subseteq \mathcal{M}$ follows directly from the assumption. We demonstrate that $\mathcal{M}$ is an *additive system*, i. e.

1. $\mathbb{R}^n \in \mathcal{M}$,
2. $B_1, B_2 \in \mathcal{M}, B_1 \cap B_2 = \emptyset \;\Rightarrow\; B_1 \cup B_2 \in \mathcal{M}$,
3. $B_1, B_2 \in \mathcal{M}, B_1 \supseteq B_2 \;\Rightarrow\; B_1 \setminus B_2 \in \mathcal{M}$,
4. $B_n \in \mathcal{M}, B_n \nearrow B \Rightarrow B \in \mathcal{M}$.

Assume that $B_1, B_2 \in \mathcal{M}$ with $B_1 \cap B_2 = \emptyset$. It follows from basic properties of states and observables that

$$\begin{aligned} \int_{B_1 \cup B_2} \kappa_f \,\mathrm{d}m_x &= \int_{B_1} \kappa_f \,\mathrm{d}m_x + \int_{B_2} \kappa_f \,\mathrm{d}m_x = m(f \cdot x(B_1)) + m(f \cdot x(B_2)) \\ &= m(f \cdot x(B_1) + f \cdot x(B_2)) = m(f \cdot (x(B_1) + x(B_2))) = m(f \cdot x(B_1 \cup B_2)). \end{aligned}$$

Let now $B_1, B_2 \in \mathcal{M}$ and $B_1 \supseteq B_2$. We obtain

$$\begin{aligned} \int_{B_1 \setminus B_2} \kappa_f \,\mathrm{d}m_x &= \int_{B_1} \kappa_f \,\mathrm{d}m_x - \int_{B_2} \kappa_f \,\mathrm{d}m_x = m(f \cdot x(B_1)) - m(f \cdot x(B_2)) \\ &= m(f \cdot x(B_1) - f \cdot x(B_2)) = m(f \cdot (x(B_1) - x(B_2))) = m(f \cdot x(B_1 \setminus B_2)). \end{aligned}$$

Suppose finally $B_n \nearrow B$, $B_n \in \mathcal{M}$. Then

$$\begin{aligned} \int_B \kappa_f \,\mathrm{d}m_x &= \int_{\mathbb{R}} \kappa_f \mathbb{I}_{\bigcup_n B_n} \,\mathrm{d}m_x = \lim_{n\to\infty} \int_{\mathbb{R}} \kappa_f \mathbb{I}_{B_n} \,\mathrm{d}m_x \\ &= \lim_{n\to\infty} \int_{B_n} \kappa_f \,\mathrm{d}m_x = \lim_{n\to\infty} m(f \cdot x(B_n)) = m(f \cdot x(B)). \end{aligned}$$

The family $\mathcal{M}$ is thus indeed an additive system and since the least additive system containing $\mathcal{A}$ coincides with $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^n)$ (see, e. g., [12]), we have $\mathcal{M} = \mathcal{B}(\mathbb{R}^n)$. □

In the lemma below as well as in the rest of the paper we utilize the following property of Lebesgue integral. Let $\mathcal{A}_1, \mathcal{A}_2$ be a σ-algebra of subsets of Ω_1 and Ω_2, respectively, and suppose that μ is a measure on the product σ-algebra $\mathcal{A}$ of $\mathcal{A}_1$ and $\mathcal{A}_2$. Let a so called marginal measure ν be defined on $\mathcal{A}_1$ as $\nu(A) := \mu(A \times \Omega_2)$ for every $A \in \mathcal{A}_1$. Every $\mathcal{A}_1$-measurable real function h on Ω_1 can be viewed as $\mathcal{A}$-measurable real function on $\Omega_1 \times \Omega_2$ and h is integrable with respect to ν iff it is integrable with respect to μ and for any $A \in \mathcal{A}_1$:

$$\int_{A\times\Omega_2} h \,\mathrm{d}\mu = \int_A h \,\mathrm{d}\nu.$$

The second lemma states that an integral of a product of some conditional state with a measurable function g can be represented only as a certain integral of g.

Lemma 2. Let $\mathcal{T}$ be the full tribe with a state m and x, y be n-dimensional and k-dimensional observables on $\mathcal{T}$, respectively. Let further z be a joint observable of y and x, i.e. $z(E \times F) = y(E) \cdot x(F)$ for any $E \in \mathcal{B}(\mathbb{R}^k)$ and $F \in \mathcal{B}(\mathbb{R}^n)$. If $A \in \mathcal{B}(\mathbb{R}^k)$ and $\kappa_{y(A)}$ is a version of conditional state of $y(A)$ given x, then for any non-negative $\mathcal{B}(\mathbb{R}^n)$-measurable function g the following equality is satisfied for all $B \in \mathcal{B}(\mathbb{R}^n)$:

$$\int_B \kappa_{y(A)} g \,\mathrm{d}m_x = \int_{A\times B} g \,\mathrm{d}m_z. \tag{11}$$

Proof. Let us follow the idea used for a construction of Lebesgue integral.

1. Let $g = \mathbb{I}_C$, $C \in \mathcal{B}(\mathbb{R}^n)$. Then for any $B \in \mathcal{B}(\mathbb{R}^n)$:

$$\begin{aligned} \int_B \kappa_{y(A)} \mathbb{I}_C \,\mathrm{d}m_x &= \int_{B\cap C} \kappa_{y(A)} \,\mathrm{d}m_x = m(y(A) \cdot x(B \cap C)) \\ &= m(z(A \times (B \cap C)) = m_z(A \times (B \cap C)) = \int_{A\times(B\cap C)} \mathrm{d}m_z \\ &= \int_{(A\times B)\cap(\mathbb{R}^k \times C)} \mathrm{d}m_z = \int_{A\times B} \mathbb{I}_{\mathbb{R}^k\times C} \,\mathrm{d}m_z = \int_{A\times B} \mathbb{I}_C \,\mathrm{d}m_z. \end{aligned}$$

2. Let g be a non-negative simple function, that is $g = \sum_{i=1}^N \alpha_i \mathbb{I}_{C_i}$, $\alpha_i \geq 0$, $C_i \in \mathcal{B}(\mathbb{R}^n)$ and $C_i \cap C_j = \emptyset$ for $i, j \leq N, i \neq j$. Then the assertion follows from linearity of Lebesgue integral using the same technique as in the previous case.

3. Consider a non-negative $\mathcal{B}(\mathbb{R}^n)$-measurable function g. Since $g_i \nearrow g$, where g_i are non-negative simple functions, Monotone Convergence Theorem proves the general case:

$$\int_B \kappa_{y(A)} \lim_{i\to\infty} g_i \,\mathrm{d}m_x = \lim_{i\to\infty} \int_B \kappa_{y(A)} g_i \,\mathrm{d}m_x = \lim_{i\to\infty} \int_{A\times B} g_i \,\mathrm{d}m_z = \int_{A\times B} g \,\mathrm{d}m_z.$$

□

4.3. Conditional independence

Probabilistic conditional independence (see, e. g. [5], [11]) is briefly summarized at first. Let $(\Omega, \mathcal{A}, P)$ be a probability space and ξ_1, ξ_2, ξ_3 be real random variables on Ω. ξ_1 and ξ_2 are said to be *conditionally independent given* ξ_3 with respect to P, if for any $A, B \in \mathcal{B}(\mathbb{R})$:

$$P(\xi_1^{-1}(A) \cap \xi_2^{-1}(B)|\xi_3) = P(\xi_1^{-1}(A)|\xi_3)P(\xi_2^{-1}(B)|\xi_3) \quad P_{\xi_3}\text{-a.e.} \tag{12}$$

In fact, (12) is satisfied iff for any $A \in \mathcal{B}(\mathbb{R})$:

$$P(\xi_1^{-1}(A)|(\xi_2, \xi_3)) = P(\xi_1^{-1}(A)|\xi_3) \quad P_{(\xi_2,\xi_3)}\text{-a.e.} \tag{13}$$

$\xi_1 \perp\!\!\!\perp \xi_2 \mid \xi_3$ $[P]$ stands for 'ξ_1 and ξ_2 are conditionally independent given ξ_3 with respect to P' and the symbol $[P]$ is usually omitted. It is well-known that conditional independence satisfies significant properties (so called *semi-graphoid properties*) as a ternary relation of random variables. They are:

1. $\xi_1 \perp\!\!\!\perp \xi_2 \mid \xi_3 \quad \Rightarrow \quad \xi_2 \perp\!\!\!\perp \xi_1 \mid \xi_3$ (symmetry)
2. $\xi_1 \perp\!\!\!\perp (\xi_2, \xi_4) \mid \xi_3 \quad \Rightarrow \quad \xi_1 \perp\!\!\!\perp \xi_4 \mid \xi_3$ (decomposition)
3. $\xi_1 \perp\!\!\!\perp (\xi_2, \xi_4) \mid \xi_3 \quad \Rightarrow \quad \xi_1 \perp\!\!\!\perp \xi_2 \mid (\xi_4, \xi_3)$ (weak union)
4. $\xi_1 \perp\!\!\!\perp \xi_2 \mid (\xi_3, \xi_4) \quad \& \quad \xi_1 \perp\!\!\!\perp \xi_3 \mid \xi_4 \quad \Rightarrow \quad \xi_1 \perp\!\!\!\perp (\xi_2, \xi_3) \mid \xi_4$ (contraction)

The relation (12) is a starting point for definition of conditional independence for observables.

Definition 2. Let $\mathcal{T}$ be the full tribe with a state m and x_1, x_2, x_3 be one-dimensional observables on $\mathcal{T}$. Observables x_1 and x_2 are *conditionally independent given* an observable x_3 with respect to m, if for all $A, B \in \mathcal{B}(\mathbb{R})$:

$$m(x_1(A) \cdot x_2(B)|x_3) = m(x_1(A)|x_3)\, m(x_2(B)|x_3) \quad m_{x_3}\text{-a.e.} \tag{14}$$

$x_1 \perp\!\!\!\perp x_2 \mid x_3$ $[m]$ stands for 'x_1 and x_2 are conditionally independent given x_3 with respect to m' and the symbol $[m]$ is usually omitted. Notice that this definition indeed generalizes a notion of independence (3) for observables.

Proposition 4. The following three assertions are equivalent:

1. $x_1 \perp\!\!\!\perp x_2 \mid x_3$,
2. for any $A \in \mathcal{B}(\mathbb{R})$:

$$m(x_1(A)|x_{23})(u, v) = m(x_1(A)|x_3)(v) \quad m_{x_{23}}\text{-a.e.}, \tag{15}$$

3. for any $A \in \mathcal{B}(\mathbb{R})$, there exists a $\mathcal{B}(\mathbb{R})$-measurable function $\kappa(v)$ such that:

$$m(x_1(A)|x_{23})(u, v) = \kappa(v) \quad m_{x_{23}}\text{-a.e.} \tag{16}$$

Proof. Let us verify all implications:

1. $\Rightarrow$ 2.

It has to be demonstrated that $m(x_1(A)|x_3)$ is a version of $m(x_1(A)|x_{23})$ for any $A \in \mathcal{B}(\mathbb{R})$. Due to Lemma 1, it suffices to show the $m_{x_{23}}$-a.e. equality for integrals of conditional states over measurable rectangles. For all $E, F, A \in \mathcal{B}(\mathbb{R})$:

$$\begin{aligned} &\int_{E\times F} m(x_1(A)|x_{23})\,\mathrm{d}m_{x_{23}} = m(x_1(A)\cdot x_2(E)\cdot x_3(F)) \\ = &\int_F m(x_1(A)\cdot x_2(E)|x_3)\,\mathrm{d}m_{x_3} = \int_F m(x_1(A)|x_3)m(x_2(E)|x_3)\,\mathrm{d}m_{x_3}. \end{aligned}$$

Employing Lemma 2 with $g = m(x_1(A)|x_3)$, we can finally write

$$\int_F m(x_1(A)|x_3)m(x_2(E)|x_3)\,\mathrm{d}m_{x_3} = \int_{E\times F} m(x_1(A)|x_3)\,\mathrm{d}m_{x_{23}}.$$

2. $\Rightarrow$ 1.

Again due to Lemma 2, we can write for all $E, A, B \in \mathcal{B}(\mathbb{R})$:

$$\begin{aligned} &\int_E m(x_1(A)|x_3)m(x_2(B)|x_3)\,\mathrm{d}m_{x_3} = \int_{B\times E} m(x_1(A)|x_3)\,\mathrm{d}m_{x_{23}} \\ = &\int_{B\times E} m(x_1(A)|x_{23})\,\mathrm{d}m_{x_{23}} = m(x_1(A)\cdot x_2(B)\cdot x_3(E)) \\ = &\int_E m(x_1(A)\cdot x_2(B)|x_3)\,\mathrm{d}m_{x_3}. \end{aligned}$$

2. $\Rightarrow$ 3.

Put $\kappa(v) = m(x_1(A)|x_3)(v)$.

3. $\Rightarrow$ 2.

Assume that such κ exists. We will show

$$\kappa(v) = m(x_1(A)|x_3)(v) \quad m_{x_3}\text{-a.e.}.$$

For all $A, B \in \mathcal{B}(\mathbb{R})$:

$$\int_B \kappa\,\mathrm{d}m_{x_3} = \int_{\mathbb{R}\times B} \kappa\,\mathrm{d}m_{x_{23}} = m(x_1(A)\cdot x_{23}(\mathbb{R}\times B)) = m(x_1(A)\cdot x_3(B)). \qquad \square$$

The next theorem is important as it justifies the use of a notion 'conditional independence' for the ternary relation introduced above for observables. The proof of 2 – 4 already appeared in [3] but we repeat this construction in order to keep the paper self-contained.

Theorem 1. Conditional independence of observables satisfies semi-graphoid axioms:

1. $x_1 \perp\!\!\!\perp x_2 \mid x_3 \quad \Rightarrow x_2 \perp\!\!\!\perp x_1 \mid x_3$
2. $x_1 \perp\!\!\!\perp x_{24} \mid x_3 \quad \Rightarrow \quad x_1 \perp\!\!\!\perp x_4 \mid x_3$
3. $x_1 \perp\!\!\!\perp x_{24} \mid x_3 \quad \Rightarrow \quad x_1 \perp\!\!\!\perp x_2 \mid x_{43}$
4. $x_1 \perp\!\!\!\perp x_2 \mid x_{34} \quad \& \quad x_1 \perp\!\!\!\perp x_3 \mid x_4 \quad \Rightarrow \quad x_1 \perp\!\!\!\perp x_{23} \mid x_4.$

Proof. Let us verify all four assertions.

1. For any $A, B \in \mathcal{B}(\mathbb{R})$, this equality is satisfied m_{x_3}-a.e.:

$$m(x_2(B) \cdot x_1(A)|x_3) = m(x_1(A) \cdot x_2(B)|x_3) = m(x_1(A)|x_3)\, m(x_2(B)|x_3).$$

2. Since considering a Borel set is equivalent to considering any fuzzy set from the range of a given observable, a purely formal convention is made that whenever we write only $m(f|x)$, f is exactly an arbitrary (but fixed) value (fuzzy set) of the observable x_1.

We show that the assumption

$$m(f|x_{243}) = m(f|x_3) \quad m_{x_{243}}\text{-a.e.} \tag{17}$$

implies $m_{x_{43}}$–a.e. equality of $m(f|x_{43})$ and $m(f|x_3)$. For any $B \in \mathcal{B}(\mathbb{R}^2)$:

$$\int_B m(f|x_{43})\, \mathrm{d}m_{x_{43}} = m(f \cdot x_{43}(B)) = m(f \cdot x_{243}(\mathbb{R} \times B))$$
$$= \int_{\mathbb{R}\times B} m(f|x_{243})\, \mathrm{d}m_{x_{243}} = \int_{\mathbb{R}\times B} m(f|x_3)\, \mathrm{d}m_{x_{243}} = \int_B m(f|x_3)\, \mathrm{d}m_{x_{43}},$$

and therefore $x_1 \perp\!\!\!\perp x_4 \mid x_3$.

3. The assumption is equivalent to (17). We have already proved in step 2 that

$$m(f|x_{43}) = m(f|x_3) \quad m_{x_{43}}\text{-a.e.} \tag{18}$$

Comparing (17) and (18), we get the $m_{x_{243}}$-a.e. equality of $m(f|x_{243})$ and $m(f|x_{43})$. Hence $x_1 \perp\!\!\!\perp x_2 \mid x_{43}$.

4. According to the assumptions,

$$m(f|x_{234}) = m(f|x_{34}) \quad m_{x_{234}}\text{-a.e.} \tag{19}$$

and

$$m(f|x_{34}) = m(f|x_4) \quad m_{x_{34}}\text{-a.e.} \tag{20}$$

Immediately from (19) and (20), the equality

$$m(f|x_{234}) = m(f|x_4) \quad m_{x_{234}}\text{-a.e.}$$

is obtained and thus $x_1 \perp\!\!\!\perp x_{23} \mid x_4$. □

In classical probability theory, it is natural that indices in decomposition and weak union property can be interchanged so that

$$\xi_1 \perp\!\!\!\perp (\xi_2, \xi_4) \mid \xi_3 \quad \Rightarrow \quad \xi_1 \perp\!\!\!\perp \xi_2 \mid \xi_3,$$
$$\xi_1 \perp\!\!\!\perp (\xi_2, \xi_4) \mid \xi_3 \quad \Rightarrow \quad \xi_1 \perp\!\!\!\perp \xi_4 \mid (\xi_2, \xi_3)$$

are also satisfied. The proposition below states an analogous result.

Proposition 5. The following two implications hold true:

1. $x_1 \perp\!\!\!\perp x_{24} \mid x_3 \quad \Rightarrow \quad x_1 \perp\!\!\!\perp x_2 \mid x_3$
2. $x_1 \perp\!\!\!\perp x_{24} \mid x_3 \quad \Rightarrow \quad x_1 \perp\!\!\!\perp x_4 \mid x_{23}$.

Proof. Let us suppose that

$$m(f|x_{243}) = m(f|x_3) \quad m_{x_{243}}\text{-a.e.} \tag{21}$$

Due to Lemma 1, it is enough to show that

$$m(f|x_{23}) = m(f|x_3) \quad m_{x_{23}}\text{-a.e.} \tag{22}$$

holds true on measurable rectangles. For all $E, F \in \mathcal{B}(\mathbb{R})$:

$$\begin{aligned} \int_{E\times F} m(f|x_{23})\, \mathrm{d}m_{x_{23}} &= m(f \cdot x_{23}(E \times F)) = m(f \cdot x_2(E) \cdot x_4(\mathbb{R}) \cdot x_3(F)) \\ = \quad m(f \cdot x_{243}(E \times \mathbb{R} \times F)) &= \int_{E\times\mathbb{R}\times F} m(f|x_{243})\, \mathrm{d}m_{x_{243}} \\ = \quad \int_{E\times\mathbb{R}\times F} m(f|x_3)\, \mathrm{d}m_{x_{243}} &= \int_{E\times F} m(f|x_3)\, \mathrm{d}m_{x_{23}}. \end{aligned}$$

To prove the second part of the proposition, assume again

$$m(f|x_{243}) = m(f|x_3) \quad m_{x_{243}}\text{-a.e.} \tag{23}$$

Due to the first part of the proposition, we obtain

$$m(f|x_{23}) = m(f|x_3) \quad m_{x_{23}}\text{-a.e..} \tag{24}$$

From (23) and (24),

$$m(f|x_{243}) = m(f|x_{23}) \quad m_{x_{243}}\text{-a.e.} \tag{25}$$

For any $E, F, G \in \mathcal{B}(\mathbb{R})$:

$$\begin{aligned} \int_{E\times F\times G} m(f|x_{423})\, \mathrm{d}m_{x_{423}} &= m(f \cdot x_4(E) \cdot x_2(F) \cdot x_3(G)) \\ = \quad \int_{F\times E\times G} m(f|x_{243})\, \mathrm{d}m_{x_{243}} &= \int_{F\times E\times G} m(f|x_{23})\, \mathrm{d}m_{x_{243}} \\ = \quad \int_{E\times F\times G} m(f|x_{23})\, \mathrm{d}m_{x_{423}}. \end{aligned}$$

Hence $x_1 \perp\!\!\!\perp x_4 \mid x_{23}$. □

Conditional independence of observables is further demonstrated on examples below.

Example 12. Let a probability space $(\Omega, \mathcal{A}, P)$ and three real random variables ξ_1, ξ_2, ξ_3 on $(\Omega, \mathcal{A}, P)$ be given such that $\xi_1 \perp\!\!\!\perp \xi_2 \mid \xi_3\ [P]$ and let $\mathcal{T}$ be a Łukasiewicz tribe on Ω. Three observables on the Łukasiewicz tribe are defined for any $B \in \mathcal{B}(\mathbb{R})$ by the assignment

$$x_i(B) := \mathbb{I}_{\xi_i^{-1}(B)}, \quad i = 1, 2, 3. \tag{26}$$

According to Example 6, a mapping m defined for any $f \in \mathcal{T}$ by $m(f) := \int_\Omega f \, \mathrm{d}P$ is a state on the Łukasiewicz tribe. Let us assume that $x_1 = \mathbb{I}_A$, $A \in \mathcal{B}(\mathbb{R})$. Since $m_{x_3} = P_{\xi_3}$ and for any $B \in \mathcal{B}(\mathbb{R})$

$$\int_B m(\mathbb{I}_A | x_3) \, \mathrm{d}m_{x_3} = m(\mathbb{I}_A x_3(B)) = m(\mathbb{I}_{A \cap \xi_3^{-1}(B)}) = P(A \cap \xi_3^{-1}(B)),$$

we have $m(\mathbb{I}_A | x_3) = P(A | \xi_3)$ P_{ξ_3}-a.e. Analogously, the $P_{(\xi_2, \xi_3)}$-a.e. equality $m(\mathbb{I}_A | x_{23}) = P(A | (\xi_2, \xi_3))$ is obtained. Furthermore, because $m_{x_{23}} = P_{(\xi_2, \xi_3)}$ and $m(\mathbb{I}_A | x_{23}) = m(\mathbb{I}_A | x_3)$ $m_{x_{23}}$-a.e., the relation $x_1 \perp\!\!\!\perp x_2 \mid x_3\ [m]$ is finally attained.

Example 13. Assuming that random variables ξ_1, ξ_2, ξ_3 from Example 12 are not conditionally independent with respect to P, we obviously obtain observables x_1, x_2, x_3 which are not conditionally independent with respect to m.

ACKNOWLEDGEMENT

The author gratefully acknowledges the support of Grant 201/02/1540 of the Grant Agency of the Czech Republic. The author would like to express his gratitude to Prof. Ing. Mirko Navara, DrSc. and RNDr. Milan Studený, DrSc. for valuable suggestions and comments regarding the presented topic.

(Received October 15, 2003.)

REFERENCES

[1] D. Butnariu and E. P. Klement: Triangular Norm-Based Measures and Games with Fuzzy Coalitions. Kluwer Academic Publishers, Dordrecht 1993.

[2] D. Butnariu and E. P. Klement: Triangular norm-based measures. In: Handbook of Measure Theory (E. Pap, ed.), North Holland, Amsterdam 2002, pp. 947–1010.

[3] T. Kroupa: Conditional independence in probability theory on MV-algebras. Soft Computing *8* (2004), 534–538.

[4] K. Kuratowski and A. Mostowski: Set Theory. North Holland, Amsterdam 1976.

[5] S. Lauritzen: Graphical Models. Oxford Statistical Science Series, Oxford University Press, Oxford 1996.

[6] M. Loève: Probability Theory, Foundations, Random Processes. D. Van Nostrand, Toronto – New York – London 1955.

[7] R. Cignoli and I. D'Ottaviano, and D. Mundici: Algebraic Foundations of Many-valued Reasoning. (Trends in Logic, Studia Logica Library, Vol. 7.) Kluwer Academic Publishers, Dordrecht 2000.

[8] P. Pták and S. Pulmannová: Orthomodular Structures as Quantum Logics. (Fundamental Theories of Physics, Volume 44.) Kluwer Academic Publishers, Dordrecht 1991.

[9] S. Pulmannová: A note on observables on MV-algebras. Soft Computing *4* (2000), 45–48.

[10] B. Riečan and D. Mundici: Probability on MV-algebras. In: Handbook of Measure Theory (E. Pap, ed.), North Holland, Amsterdam 2002, pp. 869–909.

[11] M. Studený: Probabilistic Conditional Independence Structures. Springer, London 2005.

[12] J. Štěpán: Probability Theory. Mathematical Foundations (in Czech). Academia, Praha 1987.

[13] V. Varadarajan: Geometry of Quantum Theory, Volume 1. Van Nostrand Reinhol, Princeton, N.J. 1968.

[14] J. Vejnarová: Conditional independence relations in possibility theory. Internat. J. Uncertainty Fuzziness Knowledge-Based Systems *8* (2000), 253–269.

[15] M. Vrábelová: On the conditional probability on product MV-algebras. Soft Computing *4* (2000), 58–61.

[16] L. Zadeh: Probability measures of fuzzy events. J. Math. Anal. Appl. *23* (1968), 421–427.

Tomáš Kroupa, Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mail: kroupa@utia.cas.cz