Conditional probability measures on propositional systems

By I. G. KALMÁR (Debrecen)

We define the notion of conditional probability space in a general case when the propositions (events) form a not necessarily distributive orthomodular σ-lattice. Having discussed the basic properties and consequences of our axiomatic system, a chain-representation of conditional probability is given.

1. Introduction

The development of modern probability theory goes back to 1933 when A. N. KOLMOGOROV published his axiomatic system. Later on, this axiomatic system was generalized by many authors in a way that they, for example, exchanged probability for conditional probability. The most systematic foundation is due to A. Rényi who defined and examined conditional probability spaces (RÉNYI (1955), (1956)). This work was continued by Á. CSÁSZÁR (1955), P. H. KRAUSS (1968), L. E. DUBINS (1975), and others.

Rényi mentioned that his theory can be used in quantum physics. But it is known that the propositions (events) of a quantum mechanical system fail to form a Boolean algebra (BIRKHOFF (1936)), so the conditional probability space defined by Rényi is not useful in quantum mechanics. Generally the propositional system of a physical system is supposed to be an orthomodular σ-lattice. (See VARADARAJAN (1955), PIRON (1972), for example.) However, the probabilistic interpretation of quantum mechanics suggests that certain probability measures (pure states) can be interpreted on the propositional system as conditional probability measures (See PIRON (1972).)

To be consistent with the above mentioned facts we shall in this paper define the conditional probability measures on orthomodular σ-lattices. It will come to light that our system contains the conditional probability space of Rényi and the quantum mechanical propositional-state structure as special cases.

2. Basic concepts and notations

Let $\mathscr{L}$ be a partially ordered set with first and last elements $0, 1$, respectively that is closed under the complementation $a \mapsto a^\perp$ satisfying also

(i) $$(a^\perp)^\perp = a \quad \text{and}$$

(ii) $$b^\perp \leqq a^\perp \quad \text{if} \quad a \leqq b.$$

Such a complementation is called *orthocomplementation*. If the least upper bound and greatest lower bound of $a, b \in \mathscr{L}$ exist, we denote them by $a \vee b$ and $a \wedge b$, respectively.

We call $a, b \in \mathscr{L}$ to be *orthogonal* and write $a \perp b$ if $a \leqq b^{\perp}$. We say that $a, b \in \mathscr{L}$ are *compatible* and write $a \leftrightarrow b$ if there exists a Boolean sublattice B in $\mathscr{L}$ containing a and b.

If $\mathscr{L}$ is a partially ordered set with first and last elements $0, 1$ and orthocomplementation $\perp$, furthermore the l.u.b. and g.l.b. exist for all countable subsets of $\mathscr{L}$, then we say that $\mathscr{L}$ is an *orthocomplemented σ-lattice*. An orthocomplemented lattice is orthomodular if

$$b \leqq c \quad \text{implies} \quad c = b \vee (c \wedge b^{\perp}) \quad (b, c \in \mathscr{L}).$$

The following propositions are well known in an orthomodular lattice $\mathscr{L}$.

Proposition 2.1 ([9, Theorem 2.25]). *If one of the three elements a, b, c of $\mathscr{L}$ is compatible with each of the two others, then triplet (a, b, c) is distributive that is*

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

and

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$$

Proposition 2.2 ([9], [12]). *The following statements are true in $\mathscr{L}$:*

(a) $a \leqq b \Rightarrow a \leftrightarrow b$,

(b) $a \perp b \Rightarrow a \leftrightarrow b$,

(c) $a \leftrightarrow b \Rightarrow a \leftrightarrow b^{\perp}$,

(d) $a \leftrightarrow b \Leftrightarrow (a \vee b^{\perp}) \wedge b = a \wedge b$,

(e) *$\mathscr{L}$ is distributive if and only if $a \leftrightarrow b$ for all $a, b \in \mathscr{L}$,*

(f) $a \leftrightarrow b_i \Rightarrow a \leftrightarrow \bigvee_i b_i,\ a \leftrightarrow \bigwedge_i b_i$

Proposition 2.3 *For every $a, b \in \mathscr{L}$ we have*

$$(a \vee b)^{\perp} = a^{\perp} \wedge b^{\perp}$$

and

$$(a \wedge b)^{\perp} = a^{\perp} \vee b^{\perp}.$$

A *measure* μ on an orthocomplemented σ-lattice $\mathscr{L}$ is a non-negative function on $\mathscr{L}$ that is σ-additive, i.e.,

$$\mu\left(\bigvee_i a_i\right) = \sum_i \mu(a_i) \tag{2.1}$$

if the a_i's are orthogonal element of $\mathscr{L}$. If we require (2.1) only for finitely many a_i, then μ is called *finitely additive measure*.

We call the measure (finitely additive measure) *probability measure (finitely additive probability measure)* if it satisfies also $\mu(1)=1$. One can find a detailed examination of the above mentioned notions in MAEDA (1970), PIRON (1976), VARADARAJAN (1955), for example. A systematic review of the field can be found in PIRON (1976).

Note on symbols: In some cases it will be used the usual logical operations: $\Rightarrow, \Leftrightarrow$.

3. *The generalized conditional probability space and its special cases*

The starting point for our investigation is a triple $(\mathscr{L}, \mathscr{L}_c, p)$ with the following three properties:

(A) $\mathscr{L}=\mathscr{L}(\vee, \wedge, \perp, 0, 1)$ is an orthomodular σ-lattice whose elements are called events.

(B) Let $\mathscr{L}_0=\mathscr{L}\setminus\{0\}$. Assume that there exist a subset $\mathscr{L}_c$ of $\mathscr{L}_0$ and a mapping $(x, z)\rightarrow p(x|z)$ of $\mathscr{L}\times\mathscr{L}_c$ into $[0, \infty)$ such that $p(x|z)$ for every fixed $z\in\mathscr{L}_c$ is finitely additive measure on $\mathscr{L}$ and $p(z|z)=1$ also holds. $p(x|z)$ is the so-called conditional probability of the event x under the condition z.

(C) If $x, y, z\in\mathscr{L}$, $x\leftrightarrow y$ and $z, (z\vee x^\perp)\wedge x\in\mathscr{L}_c$, then

$$p(x\wedge y|z) = p(x|z)\,p\big(y|(z\vee x^\perp)\wedge x\big).$$

For brevity $(\mathscr{L}, \mathscr{L}_c, p)$ with properties (A), (B), (C) will be called *generalized conditional probability space* or GS. We shall call the function p of two variables *conditional probability*.

In some instances other properties of p are also obeyed, for example the following two:

(D) If $x\vee y\in\mathscr{L}_c$, then $p(x|x\vee y)+p(y|x\vee y)>0$.

(E) If $(z\vee x)\wedge(z\vee x^\perp)\in\mathscr{L}_c$, then $p\big(z|(z\vee x)\wedge(z\vee x^\perp)\big)>0$.

Now let us see the most important special cases of the GS. In these cases properties (D), (E) hold too.

Example 3.1. (The case of classical probability theory.)

Let $(H, \mathscr{A}, Q)$ be a probability field. Define $\mathscr{A}_c$ as the set of B's for which $Q(B)>0$. Let

$$q(A|B) = \frac{Q(AB)}{Q(B)} \quad \text{if} \quad A\in\mathscr{A}, \quad B\in\mathscr{A}_c.$$

Then $(\mathscr{A}, \mathscr{A}_c, q)$ is obviously a GS with properties (D), (E).

On the other hand let $(\mathscr{L}, \mathscr{L}_c, p)$ be a GS and assume that $\mathscr{L}$ is distributive and p is σ-additive in its first variable. Then from a theorem of Loomis (1947) it follows that there exist a measurable space $(\Omega, \mathscr{F})$ and a σ-homomorphism h from $\mathscr{F}$ onto $\mathscr{L}$. If $1\in\mathscr{L}_c$, then let

$$P(A) = p\big(h(A)|h(\Omega)\big) = p\big(h(A)|1\big), \quad A\in\mathscr{F}.$$

Now $(\Omega, \mathscr{F}, P)$ is a probability field. If $P(B)>0$, then we can define

$$P(A|B) = \frac{P(AB)}{P(B)},$$

and if $h(B)\in\mathscr{L}_c$, then

$$P(A|B) = p(h(A)|h(B)).$$

Example 3.2. (The case of conditional probability space of Rényi.)

Let $(H, \mathscr{A}_1, \mathscr{A}_2, R)$ be a conditional probability space in the sense of Rényi, i.e., H is a non empty set; $\mathscr{A}_1$ is a σ-algebra of subsets of H; $\mathscr{A}_2$ is a non empty subset of $\mathscr{A}_1$; and finally R is a function of two variables on $\mathscr{A}_1\times\mathscr{A}_2$ with the following properties:

(i) $R(A|B)\geqq 0$, if $A\in\mathscr{A}_1$ and $B\in\mathscr{A}_2$; moreover $P(B|B)=1$, if $B\in\mathscr{A}_2$.
(ii) For any fixed $B\in\mathscr{A}_2$ $R(A|B)$ is countably additive set function of $A\in\mathscr{A}_1$.
(iii) $R(A|BC)R(B|C)=R(AB|C)$, if $A, B\in\mathscr{A}_1$, $C, BC\in\mathscr{A}_2$.

Then $(\mathscr{A}_1, \mathscr{A}_2, R)$ is a GS, and properties (D), (E) also hold.

On the other hand let now $(\mathscr{L}, \mathscr{L}_i, p)$ be a GS and assume that $\mathscr{L}$ is distributive, p is σ-additive in its first variable.

Similarly as in Example 3.1, there exist a measurable space $(\Omega, \mathscr{F})$ and σ-homomorphism h of $\mathscr{F}$ onto $\mathscr{L}$. Let G denote the set of B's for which $B\in\mathscr{F}$ and $h(B)\in\mathscr{L}_c$. Let

$$S(A|B) = p(h(A)|h(B)),\quad A\in\mathscr{F},\quad B\in\mathscr{G}.$$

Then $(\Omega, \mathscr{F}, G, S)$ is a conditional probability space in the sense of Rényi.

Example 3.3. (Classical modell of quantum mechanics.)

Let $\mathscr{H}$ be a separable complex Hilbert space with dim $\mathscr{H}>2$. Let $\mathscr{P}(\mathscr{H})$ and $\mathscr{P}_c(\mathscr{H})$ be the projection lattice and the set of atomic projections of $\mathscr{H}$ respectively. Then $\mathscr{P}(\mathscr{H})$ is a complete orthomodular lattice and $\mathscr{P}_c(\mathscr{H})$ is composed of the atoms of $\mathscr{P}(\mathscr{H})$. Let $x\in\mathscr{P}(\mathscr{H})$; $z\in\mathscr{P}_c(\mathscr{H})$, then a vector φ of $\mathscr{H}$ can be chosen such that $z\varphi=\varphi$, $\|\varphi\|=1$. Define

$$q(x|z) = \langle\varphi, x\varphi\rangle,$$

where the inner product $\langle\varphi, x\varphi\rangle$ does not depend on φ ($z\varphi=\varphi$, $\|\varphi\|=1$). It is known that in this case $q(x|z)$ for fixed z is a probability measure (state) on $\mathscr{P}(\mathscr{H})$. Furthermore, the so-called pure states can be identified with the atoms z of $\mathscr{P}(\mathscr{H})$. (See Varadarajan (1955), Theorem 7.23.) Let us see the triple $(\mathscr{P}(\mathscr{H}), \mathscr{P}_c(\mathscr{H}), q)$. With notations $\mathscr{L}=\mathscr{P}(\mathscr{H})$, $\mathscr{L}_c=\mathscr{P}_c(\mathscr{H})$, $p=q$, properties (A), (B), (D) hold. But it is easy to see that $q(x|z)=0$ if and only if $x\perp z$, thus (E) also hold. It can be proved that if x is not orthogonal to z; $x\in\mathscr{P}(\mathscr{H}), z\in\mathscr{P}_c(\mathscr{H})$, then $(z\vee x^{\perp})\wedge x\in\mathscr{P}_c(\mathscr{H})$ Moreover, as it was shown by Piron (1976) property (C) is also satisfied. Combining the facts mentioned above, we arrive at the statement that $(\mathscr{P}(\mathscr{H}), \mathscr{P}_c(\mathscr{H}), q)$ is a GS and satisfy (D) and (E).

4. Some remarks on the GS

1. Analyzing a (phisical) system from a probabilistic point of view we may set out from our previous knowledge about the system. This knowledge may be the occurrence of one or more events. Example 3.3 shows that in quantum mechanics the previous information (state) concerning the system is interpretable as an atom of the proposition system.

Returning now to the axiomatic system (A), (B), (C), the quantity $p(x|z)$ must be considered as the probability of the occurrence of x under an a-priori knowledge, namely, that z certainly occurs. Thus, axiom (C) means the following. If $x \leftrightarrow y$ and we previously know the occurrence of z, moreover we perform an experiment, and x occurs, then immediately after the experiment the event $(z \vee x^{\perp}) \wedge x$ will certainly occur.

A similar argument was put forth by PIRON (1972) in the special case when x is an atom and the experiment is ideal and of the first kind. Then property (C) is easily justified by using the role of composition of probability.

2. Let us assume that $p(x|z)=p(y|z)=0$. Since generally the occurrence of $x \vee y$ does not imply the occurrence of x or y, so $p(x \vee y|z)>0$ does not follow, from the preceding condition. It seems an easier condition is more reasonable, as follows: if $x \vee y \in \mathscr{L}_c$, then $p(x|x \vee y)$ or $p(y|x \vee y)$ is positive. This is exactly property (D).

3. Property (E) is probably independent of (A), (B), (C), and (D), but in some special cases it is a consequence of the others.

We shall say that the conditional probability is *positive,* if

$$p(x|z)=0 \quad \text{if and only if} \quad x \perp z.$$

The reader can readily verify the following statements: Let $(\mathscr{L}, \mathscr{L}_c, p)$ be a GS, then

(a) if p is positive, then (E) holds,
(b) if $x \leftrightarrow z$ and $(z \vee x) \wedge (z \vee x^{\perp}) \in \mathscr{L}_c$, then $p(z)(z \vee x) \wedge (z \vee x^{\perp})>0$.
(c) if $(z \vee x) \wedge (z \vee x^{\perp})$, $1 \in \mathscr{L}_c$ and $p(z|1)>0$, then

$$p(z|1) = p\big(z|(z \vee x) \wedge (z \vee x^{\perp})\big)p\big((z \vee x) \wedge (z \vee x^{\perp}))1\big)$$

and consequently

$$p\big(z|(z \vee x) \wedge (z \vee x^{\perp})\big) > 0.$$

5. Elementary properties of conditional probability

In this section we prove fundamental properties of conditional probability that will be frequently used.

Proposition 5.1. *Let $y \in \mathscr{L}_c$. Then the following statements are true.*

(i) $p(1|y)=1$, $p(0|y)=0$.
(ii) $p(x|y) \leqq 1$ *for all* $x \in \mathscr{L}$.
(iii) $p(\cdot|y)$ *is monotone increasing.*
(iv) $p(x|y)=1$ *if* $x \geqq y$.
(v) $p(x|y)=0$ *if* $x \perp y$.

PROOF. (i) By (C)

$$p(y^{\perp}|y) = p(y^{\perp} \wedge 1|y) = p(1|y) \cdot p(y^{\perp}|y)$$

and

$$p(1|y) = p(y|y) + p(y^{\perp}|y) = 1 + p(y^{\perp}|y),$$

which imply (i).

(ii) With the help of (i) we have

$$1=p(1|y)=p(x|y)+p(x^\perp|y) \quad \text{thus} \quad p(x|y)\leqq 1.$$

(iii) Let $x_1\leqq x_2$, then $x_2=x_1\vee(x_2\wedge x_1^\perp)$, where

$$x_1\perp(x_2\wedge x_1^\perp), \text{ so } p(x_2)=p(x_1)+p(x_2\wedge x_1^\perp).$$

(iv) From the finitely additivity we get that if $x\geqq y$, then $p(x|y)=p(y|y)+$ $+p(x\wedge y^\perp|y)\geqq 1$. By (ii) $p(x|y)\leqq 1$, so $p(x|y)=1$.

(v) If $x\perp y$, then

$$1=p(x\vee y|y)=p(x|y)+p(y|y)=p(x|y)+1. \text{ Hence } p(x|y)=0. \qquad \text{Q. e. d.}$$

Proposition 5.2. *Let* $z\in\mathscr{L}_c$, *then*

(i) $p(x|z)=p\big((z\vee x^\perp)\wedge x|z\big)$ *for all* $x\in\mathscr{L}$,
(ii) *if additionally* $x\leftrightarrow z$, *then* $p(x|z)=p(x\wedge z|z)$.

PROOF. (i) By Proposition 2.3

$$\big((z\vee x^\perp)\wedge x\big)^\perp = (z^\perp\wedge x)\vee x^\perp$$

and consequently

$$p\big((z\vee x^\perp)\wedge x|z\big) = 1-p\big((z^\perp\wedge x)\vee x^\perp|z\big) = 1-p(z^\perp\wedge x|z)-p(x^\perp|z) =$$
$$= 1-p(x^\perp|z)=p(x|z).$$

(ii) If $x\leftrightarrow z$, then $(z\vee x^\perp)\wedge x=z\wedge x$, so (i) implies (ii). Q.e.d.

As concerns positivity it is obvious that p is positive if and only if

$$p(x|z)=1 \Leftrightarrow x\geqq z.$$

We have seen in Proposition 5.2 that in case of $x\leftrightarrow y$ $p(x|z)=p(x\wedge z|z)$. If p is positive, then the converse is also true as follows.

Proposition 5.3. *If* p *is positive and* $y, (y\vee x^\perp)\wedge x\in\mathscr{L}_c$, *then*

$$p(x\wedge y|y) = p(x|y) \Leftrightarrow x\leftrightarrow y.$$

PROOF. It will be sufficient to show that $p(x\wedge y|y)=p(x|y)$ implies $x\leftrightarrow y$. Since $x\leftrightarrow x\wedge y$, by (C) we have

$$p(x\wedge y|y) = p\big(x\wedge(x\wedge y)|y\big) = p(x|y)p\big(x\wedge y|(y\vee x^\perp)\wedge x\big).$$

Therefore

$$p(x|y)=0 \quad \text{or} \quad p\big(x\wedge y|(y\vee x^\perp)\wedge x\big)=1.$$

In case of $p(x|y)=0$ it follows that $p(x^\perp|y)=1$, so according to positivity $x^\perp\geqq y$, i.e., $x\perp y$. Hence $x\leftrightarrow y$. In case of $p(x\wedge y|(y\vee x^\perp)\wedge x)=1$ it follows that

$$x\wedge y\geqq (y\vee x^\perp)\wedge x, \quad \text{i.e.,} \quad x\wedge y=(y\vee x^\perp)\wedge x.$$

By Proposition 2.2 $x\wedge y=(y\vee x^\perp)\wedge x$ implies $x\leftrightarrow y$. Q.e.d.

Remark. Proposition 5.2 shows that generally $x \wedge y=0$ does not imply $p(x|y)=0$ because one can easily prove that

$$x \perp y \Leftrightarrow x \leftrightarrow y \quad \text{and} \quad x \wedge y = 0.$$

Proposition 5.4. *The following two statements are true.*

(i) *If* $x \vee y,\ x \vee y \vee z \in \mathscr{L}_c$, *then*

$$p(y|x \vee y \vee z) = p(y|x \vee y) p(x \vee y \vee z). \tag{5.1}$$

(ii) *If* $y,\ y \wedge z \in \mathscr{L}_c$ *and* $x \leqq y$, *then*

$$p(x|y \vee z) = p(x|y) p(y|y \vee z). \tag{5.2}$$

Proof. (i) Let us set $y=u$, $x \vee y=v$ and $x \vee y \vee z=w$. Then $v \leftrightarrow w$ and $(w \vee v^{\perp}) \wedge v = w \wedge v = v$, and by (C)

$$p(u \wedge v|w) = p(u|v) p(v|w).$$

(ii) Let now $x \leqq y$, then $x \leftrightarrow y$ and $((z \vee y) \vee y^{\perp}) \vee y = (y \wedge z) \wedge y = y$. Hence by (C) it follows (5.2). Q.e.d.

6. Representation of conditional probability

Now we shall investigate the problem of representation of conditional probability by an "ordered" set of measures on $\mathscr{L}$. When $\mathscr{L}$ is distributive (Boolean algebra) this was discussed by several authors, e.g. Rényi (1956), Krauss (1968). We use a natural ordering of $\mathscr{L}$ induced by the conditional probability p in a similar manner as it was introduced by Rényi (1956).

Throughout this section $(\mathscr{L}, \mathscr{L}_c, p)$ denotes a GS such that properties (D), (E) hold and it will be assumed that $\mathscr{L}_c = \mathscr{L}_0 = \{z \in \mathscr{L} | z \neq 0\}$. The last assumption is very useful because it makes the survey of the structure of conditional probability easy.

Firstly we introduce a new notion. For every $x, y \in \mathscr{L}$, $x \vee y \neq 0$ let $x \subseteqq y$ if $p(y|x \vee y) > 0$, and $x \subset y$ if $p(x|x \vee y) = 0$. For these relations we have the following statements:

Proposition 6.1. *Let* $x, y, z \in \mathscr{L}$ *and* $x \vee y \in \mathscr{L}_0$, *then*

(i) *If* $x \leqq y$, *then* $x \subseteqq y$.
(ii) *Either* $x \subseteqq y$ *or* $y \subseteqq x$.
(iii) $x \subseteqq y \Leftrightarrow x \vee y \subseteqq y$.
(iv) $x \subset y \Leftrightarrow x \subseteqq y$ *and* $y \nsubseteqq x$.
(v) *If* $x \subseteqq y,\ y \subseteqq z$, *then* $x \subseteqq z$.

Proof. Since statements (i)—(iv) are trivial, we prove only (v). Let us assume that $x \subseteqq y$ and $y \subseteqq z$. We have to distinguish two cases:
(a) If $p(y|x \vee y \vee z) = 0$, then by (5.1)

$$p|y|x \vee y \vee z) = p(y|x \vee y) p(x \vee y|x \vee y \vee z),$$

hence according to $p(y|x\vee y)>0$ we have $p(x\vee y|x\vee y\vee z)=0$. Then from property (D) we get $p(z|x\vee y\vee z)>0$. However,

$$p(z|x\vee y\vee z) = p(z|x\vee z)p(x\vee z|x\vee y\vee z),$$

and consequently $p(z|x\vee z)>0$, i.e. $x\subseteqq z$.

(b) If $p(y|x\vee y\vee z)>0$, then $p(y\vee z|x\vee y\vee z)>0$ and by (5.1)

$$p(z|x\vee y\vee z) = p(z|y\vee z)p(y\vee z|x\vee y\vee z),$$

where $p(z|y\vee z)>0$ by assumption. Thus $p(z|x\vee y\vee z)>0$; hence

$$p(z|x\vee y\vee z) = p(z|x\vee y)p(x\vee z|x\vee y\vee z)$$

implies $p(z|x\vee z)>0$, i.e., $x\subseteqq z$. Q.e.d.

In any complemented lattice L one can introduce the following operation of two variables:

If $z, x\in L$, then let $z*x=(z\vee x^{\perp})\wedge x$. Here $\perp$ is the complementation in L.

Now, for star mapping, the following statement is true. The proof is omitted because it is very simple to verify by the definition.

Proposition 6.2. *Let $\mathscr{L}$ be an orthomodular σ-lattice as before. Then the star mapping on $\mathscr{L}=\mathscr{L}$ possesses the following properties: If $x, y, z\in\mathscr{L}$ then*

(i) $z*z=z$.
(ii) $z*x=0$, *if* $z\perp x$.
(iii) $(z*x)*y=0$, *if* $x\perp y$.
(iv) $(z*x)*x=z*x$.
(v) $z*(z*x)=z*x$.
(vi) $z*z=x\wedge x \Leftrightarrow x \leftrightarrow z$.
(vii) $(z*x)*y=(z*y)*x=z*x$, *if* $x\leqq y$.
(viii) $(\bigvee_i z_i)*x=\bigvee_i (z_i*x)$, *if* $\{z_i\}$ *is countable and* $z_i\in\mathscr{L}$.

The notation of operation $*$ allows us to define an *ideal* in $\mathscr{L}$ as a non empty subset I of $\mathscr{L}$ such that

(a) if $z\in I$, then $z*x\in I$ for all $x\in\mathscr{L}$,
(b) if $z_1, z_2\in I$, then $z_1\vee z_2\in I$.

We now set

$$I(x) = \{y\in\mathscr{L}|y\subseteqq x\}$$

and

$$I^+(x) = I(x)\setminus\{\bigcup_{y\subset z} I(y)\}, \quad \text{if} \quad x\in\mathscr{L}_0.$$

For $I(x)$ we shall prove the following:

Proposition 6.3.

(i) *$I(x)$ is an ideal in $\mathscr{L}$ for every $x\in\mathscr{L}_0$.*
(ii) *$\{I(y)|y\in\mathscr{L}_0\}$ is linearly ordered by set inclusion and*

$$I(y)\subset I(z) \Leftrightarrow y\subset z,$$

where $\subset$ means the proper inclusion.

PROOF. (i) Let $y_1, y_2 \in I(x)$, then by Proposition 6.1 (ii)—(iii)

$$y_1 \subseteqq y_2 \quad \text{or} \quad y_2 \subseteqq y_1, \quad \text{i.e.,} \quad y_1 \vee y_2 \subseteqq y_1 \quad \text{or} \quad y_1 \vee y_2 \subseteqq y_2.$$

According to the transitivity of $\subseteqq$ in any case $y_1 \vee y_2 \subseteqq x$, that is, $y_1 \vee y_2 \in I(x)$. Let now $z \in I(x)$ and $y \in \mathscr{L}$. If $z=0$, then $z * x = 0 \in I(x)$ trivially. If $z \neq 0$, then $[(z \vee y^\perp) \wedge y] \vee z = (z \vee y^\perp) \wedge (z \vee y)$; hence by property (E) $p(z|(z*y) \vee z) = p(z|(z \vee y^\perp) \wedge (z \vee y)) > 0$, i.e., $z * y \subseteqq z$. By $z \subseteqq x$ it follows $z * y \in I(x)$.

(ii) This obviously follows from Proposition 6.1/(ii). Q.e.d.

Proposition 6.4. *If* $x_1, x_2 \in \mathscr{L}$, $y_1, y_2 \in \mathscr{L}_0$, $x_1 \vee x_2 \leqq y_1 \wedge y_2$ *and* $p(x_2|y_1)$, $p(x_2|y) > 0$, *then*

$$p(x_1|y_1)p(x_2|y_2) = p(x_1|y_2)p(x_2|y_1).$$

PROOF. By our conditions $x_1, x_2 \leqq y_1, y_2$, so $x_1, x_2 \subseteqq y_1, y_2$. Furthermore, $0 < p(x_2|y_1) = p(x_2|y_1 \vee x_2)$, and consequently $y_1 \subseteqq x_2$. Similarly also $y_2 \subseteqq x_2$. Then by the transitivity

$$y_1 \subseteqq y_2 \quad \text{and} \quad y_2 \subseteqq y_1.$$

Now, with the help of (5.2) we have

$$\frac{p(x_1|y_1 \vee y_2)}{p(x_2|y_1 \vee y_2)} = \frac{p(x_1|y_1)p(y_1|y_1 \vee y_2)}{p(x_2|y_1)p(y_1|y_1 \vee y_2)} = \frac{p(x_1|y_2)p(y_2|y_1 \vee y_2)}{p(x_2|y_2)p(y_2|y_1 \vee y_2)};$$

hence the statement follows. Q.e.d.

Let $x \in \mathscr{L}_0$, $y \in I(x)$, then $p(x|x \vee y) > 0$, so we can define

$$m_x(y) = \frac{p(y|x \vee y)}{p(x|x \vee y)}.$$

For fixed $x \in \mathscr{L}_0$ m_x is a function on $I(x)$. The following proposition is true:

Proposition 6.5. *For every fixed* $x \in \mathscr{L}_0$ m_x *has the following properties:*

(i) m_x *is finitely additive measure on* $I(x)$. *If* p *is* σ*-additive, then* m_x *is also* σ*-additive.*

(ii) *If* $y \in I(x)$, *then*

$$m_x(y) = 0 \Leftrightarrow I(y) \subset I(x) \quad (\text{i.e. } y \subset x).$$

(iii) *If* $x \vee y \in I^+(z)$, *then* $m_z(x) + m_z(y) > 0$.

(iv) *For all* $z \in I(x)$, $z \in \mathscr{L}_0$ *and* $y \in \mathscr{L}$

$$m_x(y \wedge z) = p(y \wedge z|z) m_x(z). \tag{6.1}$$

(v) *If* $I(x) = I(y)$ *and* $z \in I(y)$, *then*

PROOF. (i) It is sufficient to prove for (i) that m_x finitely additive. Let $y, z \in I(x)$ and assume $y \perp z$. Then

$$m_x(y \vee z) = \frac{p(y \vee z|x \vee y \vee z)}{p(x|x \vee y \vee z)} = \frac{p(y|x \vee y \vee z)}{p(x|x \vee y \vee z)} + \frac{p(z|x \vee y \vee z)}{p(x|x \vee y \vee z)} =$$

By Proposition 6.3

$$= \frac{p(y|x\vee y)}{p(x|x\vee y)} + \frac{p(z|x\vee z)}{p(x|x\vee z)} = m_x(y) + m_x(y).$$

σ-additivity of m_x can be proved in a similar manner, assumed that p is σ-additive.

(ii) Since $p(y|x\vee y)=0$ if and only if $y\subset x$, (ii) is obvious.

(iii) Let $x\vee y\in I^+(z)$, then from Proposition 5.4 it follows that

$$p(x|x\vee y\vee z) = p(x|x\vee z)p(x\vee z|x\vee y\vee z). \tag{6.2}$$

$$p(y|x\vee y\vee z) = p(y|y\vee z)p(y\vee z|x\vee y\vee z),$$

and

$$p(x|x\vee y\vee z) = p(x|x\vee y)p(x\vee y|x\vee y\vee z), \tag{6.3}$$

$$p(y|x\vee y\vee z) = p(y|x\vee y)p(x\vee y|x\vee y\vee z).$$

In according to $x\vee y\in I^+(z)$

$$m_z(x\vee y) = \frac{p(x\vee y|x\vee y\vee z)}{p(z|x\vee y\vee z)} > 0;$$

hence by (6.3) and Proposition 6.1/(ii) we have

$$p(x|x\vee y\vee z) + p(y|x\vee y\vee z) > 0.$$

Then by (6.2) we get

$$p(x|x\vee z) + p(y|y\vee z) > 0.$$

Since $m_z(x)=\frac{p(x|x\vee z)}{p(z|x\vee z)}$, $m_z(y)=\frac{p(y|y\vee z)}{p(z|y\vee z)}$, it follows

$$m_z(x) + m_z(y) > 0.$$

(iv) Let $z\in I(x)$, $z\in\mathcal{L}_0$, $y\in\mathcal{L}$. If $m_x(z)=0$, then the statement is trivial. If $m_x(z)>0$, then $p(z|x\vee z)>0$ and by Proposition 6.4

$$m_x(y\wedge z) = \frac{p\big(y\wedge z|x\vee (y\wedge z)\big)}{p\big(x|x\vee (y\wedge z)\big)} = \frac{p(y\wedge z|x\vee z)}{p(x|x\vee z)} =$$

$$= \frac{p(y\wedge z|z)\,p(z|x\vee z)}{p(x|x\vee z)} = p(y\wedge z|z)\,m_x(z),$$

where we used equality

$$p(y\wedge z)|x\vee z) = p(y\wedge z|z)p(z|x\vee z).$$

(v) Let $I(x)=I(y)$ and $z\in I(y)$. Then by Proposition 6.4

$$m_x(z) = \frac{p(z|x\vee z)}{p(x|x\vee z)} = \frac{p(z|x\vee y\vee z)}{p(x|x\vee y\vee z)}$$

and

$$m_y(z) = \frac{p(z|y\vee z)}{p(y|y\vee z)} = \frac{p(z|x\vee y\vee z)}{p(y|x\vee y\vee z)}$$

Hence

$$m_x(z) = \frac{p(y|x \vee y \vee z)}{p(x|x \vee y \vee z)} m_y(z) = \frac{p(y|x \vee y)}{p(x|x \vee y)} m_y(z) = m_x(y) m_y(z) \quad \text{Q.e.d.}$$

As we have seen in (ii) of the preceding Proposition

$$m_z(x) > 0 \iff z \in I^+(x).$$

If $z \in I^+(x)$, then by (6.1)

$$p(y \wedge z|z) = \frac{m_x(y \wedge z)}{m_x(z)}.$$

Here in general $p(y \wedge z|z) \neq p(y|z)$ (see Proposition 5.3). However, according to Proposition 5.2/(ii) the following statement is true:

Proposition 6.6. *If $y \in I(x)$, $z \in I^+(x)$ and $y \leftrightarrow z$, then*

$$p(y|z) = \frac{m_x(y \wedge z)}{m_x(z)}.$$

We should remark that m_x is generally not bounded, but in a special case, when p is positive, an important theorem is valid as follows:

Proposition 6.7. *If p is positive, then there exists a finitely additive probability measure ν on $\mathcal{L}$ that for all $y \in \mathcal{L}$, $z \in \mathcal{L}_0$,*

$$p(y|z) = \frac{\nu(y \wedge z)}{\nu(z)}$$

if and only if $y \leftrightarrow z$. If p is σ-additive, then ν is also σ-additive.

Proof. Let $z \in \mathcal{L}_0$, then from the positivity it follows $y(z) = I(1)$ and $\mathcal{L}_0 = I^+(1)$. Let $x \in \mathcal{L}$ and

$$\nu(x) = m_1(x) = \frac{p(x|1)}{p(1|1)} = p(x|1).$$

Then by (6.1) we have

$$p(y|z) = \frac{\nu(y \wedge z)}{\nu(z)}$$

if $y \leftrightarrow z$, and from Proposition 5.2/(ii) it follows that

$$p(y|z) \neq \frac{\nu(y \wedge z)}{\nu(z)}, \quad \text{if} \quad y \not\leftrightarrow z. \quad \text{Q.e.d.}$$

The most important results of this section can be summarized in one theorem which will be called the representation theorem.

Theorem 6.8. *Let $(\mathcal{L}, \mathcal{L}_c, p)$ be a* GS *with properties* (D) *and* (E). *Let us assume that $\mathcal{L}_c = \mathcal{L}_0$, then there exists $\{I_\gamma, m_\gamma, \gamma \in \Gamma\}$ such that*

(i) *For every $\gamma \in \Gamma$ I_γ is an ideal in $\mathcal{L}$ and the family $\{I_\gamma, \gamma \in \Gamma\}$ is linearly ordered by set inclusion.*

(ii) *For every* $\gamma\in\Gamma$ m_γ *is a finitely additive measure on* I_γ *and for arbitrary* $x\in I_\gamma$ $m_\gamma(x)>0$ *if and only if* $x\in I_\gamma^+$. *If* p *is* σ*-additive, then* m_γ *is* σ*-additive, too. Here*

$$I_\gamma^+ = I_\gamma \setminus \{\bigcup_{\beta:\, I_\beta\subset I_\gamma} I_\beta\}.$$

(iii) *If* $x\vee y\in I_\gamma^+$, *then* $m_\gamma(x)+m\,(y)>0$.

(iv) *For all* $x\in\mathcal{L}_0$ *there exists* $\gamma\in\Gamma$ *so that* $x\in I_\gamma^+$.

(v) *For all* $\gamma\in\Gamma$ *there exists at least one element of* I_γ^+.

(vi) *If* $I_\gamma=I_\beta$ $(\gamma,\beta\in\Gamma)$, *then there exists a positive real number* k *such that*

$$m_\gamma(x) = k m_\beta(x)$$

for all $x\in I_\gamma$.

(vii) *For every* $\gamma\in\Gamma$ *and* $z\in I_\gamma, x\in\mathcal{L}$

$$m_\gamma(x\wedge z) = p(x\wedge z|z)m_\gamma(z).$$

When $\mathcal{L}$ is a Boolean algebra an analogous theorem was proved by KRAUSS ([5], pp. 232). Following Krauss the family $\{I_\gamma, m_\gamma, \gamma\in\Gamma\}$ is said to be a *chain-representation* of p, if (i)—(vii) of Theorem 6.8 hold.

Our next object is to discuss the nature of chain-representations. Assume that there exist two chain-representations $\{I_\gamma, m_\gamma, \gamma\in\Gamma\}$ and $\{I_\omega, m_\omega, \omega\in\Omega\}$ of p such that

$$\{I_\gamma|\gamma\in\Gamma\} = \{I_\omega|\omega\in\Omega\}$$

and if $I_\gamma=I_\omega$, then there exists $c>0$ with $m_\gamma=cm_\omega$. In this case there is no essential difference between the two chain-representations. Such chain representations will be called *equivalent*.

Theorem 6.9. *Let* $(\mathcal{L},\mathcal{L}_i,p)$ *be the same as in Theorem* 6.8. *Then any two chain-representations of* p *are equivalent.*

PROOF. It is enough to prove that an arbitrary chain-representation $\{I_\gamma, m_\gamma, \gamma\in\Gamma\}$ is equivalent with the the chain representation $\{I(x), m_x, x\in\mathcal{L}_0\}$ which is induced by p.

Let $\gamma\in\Gamma$, then there exists $x\in I_\gamma^+$. We show that

$$I_\gamma = I(x). \tag{6.4}$$

If $y\in I_\gamma$, then $x\vee y\in I_\gamma$ and $x\leftrightarrow x\vee y$; hence

$$m_\gamma(x) = p(x|x\vee y)m_\gamma(x\vee y).$$

Since $m_\gamma(x)>0$, so $p(x|x\vee y)>0$, i.e., $y\in I(x)$. Hence

$$I_\gamma \subseteqq I(x). \tag{6.5}$$

When conversely $y\notin I_\gamma$, then there exists $\delta\in\Gamma$ such that $y\in I_\delta^+$ and consequently $I_\gamma\subset I_\delta$, $m_\delta(y)>0$. However, by $x\in I_\gamma$ $x\notin I_\delta^+$, i.e., $m_\delta(x)=0$. From $x\leftrightarrow x\vee y$

$$m_\delta(x) = p(x|x\vee y)m_\delta(x\vee y).$$

This implies $p(x|x\vee y)=0$, i.e., $y\notin I(x)$. Then

(6.6) $$I(x) \subseteqq I_\gamma.$$

Now (6.5) and (6.6) imply (6.4).

If $z\in\mathscr{L}_0$, then there exists $\beta\in\Gamma$ such that $z\in I_\beta^+$ and by similar arguments as before we get $I(z)=I_\beta$, i.e.,

(6.7) $$\{I(x)|x\in\mathscr{L}_0\} = \{I_\gamma|\gamma\in\Gamma\}.$$

On the other hand, if $I(x)=I_\gamma$, then in case of $w\in I(x)$ we have

$$m_\gamma(w) = p(w|x\vee w)\, m_\gamma(x\vee w),$$

$$m_\gamma(x) = p(x|x\vee w)\, m_\gamma(x\vee w)$$

and

$$m_x(w) = \frac{p(w|x\vee w)}{p(x|x\vee w)}.$$

Hence it follows that

$$m_\gamma(w) = m_\gamma(x)\, m_x(w)$$

for all $w\in I(x)=I_\gamma$. Since $0<m_\gamma(x)=c$, we have

(6.8) $$m_\gamma = c m_x.$$

(6.7) and (6.8) imply the equivalence of $\{I(x), m_x, x\in\mathscr{L}_0\}$ and $\{I_\gamma, m_\gamma, \gamma\in\Gamma\}$.

Q.e.d.

It is not very difficult to construct two such GS $(\mathscr{L}, \mathscr{L}_0, p)$ and $(\mathscr{L}, \mathscr{L}_0, r)$ with properties (D), (E) that the chain-representations generalized by p and r should be equivalent, but $p\neq r$. This means that the chain-representation does not determine exactly p.

Let now a triple $\{I_\gamma, m_\gamma, \gamma\in\Gamma\}$ be given and assume that the statements (i)—(vi) of Theorem 6.8 are true. If $z\in\mathscr{L}_0$, then there exists $\alpha\in\Gamma$ such that $z\in I_\alpha^+$, so we can define

(6.9) $$q(x|z) = \frac{m_\alpha(z*x)}{m_\alpha(z)}.$$

$q(x|z)$ is obviously independent of the choise of $\alpha\,(z\in I_\alpha^+)$. The following statement may be proved:

Proposition 6.10. *The function* $q\colon \mathscr{L}\times\mathscr{L}_0 \to [0, \infty)$ *defined by* (6.9) *possesses properties* (C), (D), (E) *and*

(6.10) $$m_\gamma(x\wedge z) = q(x\wedge z|z) m_\gamma(z)$$

holds for all $\gamma\in\Gamma$, $z\in I_\gamma$ *and* $x\in\mathscr{L}$.

Proof. Let $z\in I_\gamma^+$, then

$$q(x\wedge z|z) = \frac{m_\gamma\big(z*(x\wedge z)\big)}{m_\gamma(z)} = \frac{m_\gamma(x\wedge z)}{m_\gamma(z)},$$

i.e., (6.10) holds. If $z\in I$, but $z\notin I_\gamma^+$, then $m_\gamma(z)=m_\gamma(z\wedge x)=0$, i.e., (6.10) is true also in this case.

Concerning property (C) we have to prove that if $z, z*x \in \mathscr{L}_0$, $z \in I_\alpha^+$ and $x \leftrightarrow y$ then

$$q(x \wedge y|z) = q(x|z)q(y|z*x), \tag{6.11}$$

i.e.,

$$\frac{m_\alpha\big(z*(x\wedge y)\big)}{m_\alpha(z)} = \frac{m_\alpha(z*x)}{m_\alpha(z)}\,\frac{m\beta\big((z*x)*y\big)}{m_\delta(z*x)},$$

where $z*x \in I_\beta^+$. Since $x \leftrightarrow y^\perp$ and $x \leftrightarrow (z \vee x^\perp)$, the following calculations are correct.

$$(z*x)*y = (z\vee x^\perp \vee y^\perp)\wedge(x \vee y^\perp)\wedge y = (z \vee x^\perp \vee y^\perp)\wedge(x\wedge y) =$$

$$= \big(z\vee(x\wedge y^\perp)\big)\wedge(x\wedge y) = z*(x\wedge y).$$

Thus, setting $u = z*(x\wedge y)$, $v = z*x$, we can write (6.11) in the form

$$\frac{m_\alpha(u)}{m_\alpha(z)} = \frac{m_\alpha(v)}{m_\alpha(z)}\,\frac{m_\beta(u)}{m_\beta(v)}. \tag{6.12}$$

Since $v \in I_\alpha$, $v \in I_\beta^+$, it is clear that $I_\beta \subseteqq I_\alpha$. Now we distinguish two cases. If $v \in I_\alpha^+$, then $I_\alpha = I_\beta$ and

$$\frac{m_\beta(u)}{m_\beta(v)} = \frac{m_\alpha(u)}{m_\alpha(v)},$$

i.e., (6.12) holds. If $v \notin I_\alpha^+$, then $m_\alpha(v) = 0$ and $I_\beta \subset I_\alpha$. However, from $u \in I_\beta$ it follows $u \notin I_\alpha^+$, that is $m_\alpha(u) = 0$. This implies (6.12) and consequently (6.11).

Now we shall prove that

$$q(x)x\vee y) + q(y|x\vee y) > 0 \tag{6.13}$$

for all $x \vee y \in \mathscr{L}_0$, i.e., (D) holds for q. Let $x \vee y \in I_\gamma^+$, then

$$q(x|x\vee y) = \frac{m_\gamma\big((x\vee y)*x\big)}{m_\gamma(x\vee y)} = \frac{m_\gamma(x)}{m_\gamma(x\vee y)}.$$

$$q(y|x\vee y) = \frac{m_\gamma\big((x\vee y)*y\big)}{m_\gamma(x\vee y)} = \frac{m_\gamma(y)}{m_\gamma(x\vee y)}.$$

However, by (iii) of Theorem 6.8 we have

$$m_\gamma(x) + m_\gamma(y) > 0;$$

hence

$$q(x|x\vee y) + q(y)|x\vee y) > 0,$$

i.e., (6.13) holds.

Property (E) amounts to saying that

$$q\big(z|(z\vee x)\wedge(z\vee x^\perp\big) > 0, \tag{6.14}$$

if $z \in \mathscr{L}_0$, $x \in \mathscr{L}$. To prove (6.14) let us assume that $z \in I_\delta^+$, $x \in \mathscr{L}$. Then $z*x \in I_\delta$ and

$$(z\vee x)\wedge(z\vee x^\perp) = (z*x)\vee z \in I_\delta.$$

Furthermore, $z \leqq z \vee (z * x)$ implies $0 < m_\delta((z*x) \vee z)$. Hence

$$q(z|(z \vee x) \wedge (z \vee x^\perp)) = \frac{m_\delta(x * (z \vee (z, x)))}{m_\delta(z \vee (z * x))} = \frac{m_\delta(z)}{m_\delta(z \vee (z * x))} > 0,$$

i.e., (6.14) holds. Q.e.d

Remarks. One can easily verify the following statements:

1. In general function q defined by (6.9) is not finitely additive.
2. If $x_1 \leftrightarrow z$ and $x_1 \perp x_2$, then

$$q(x_1 \vee x_2 | z) = q(x_1|z) + q(x_2|z)$$

is also true.

3. From the preceding remark it follows that if z is in the centrum of $\mathscr{L}$, then $q(\cdot|t)$ is a finitely additive probability measure. Furthermore, if $\mathscr{L}$ is distributive, then q is obviously the only conditional probability on $\mathscr{L} \times \mathscr{L}_0$ such that $\{I_\gamma, m_\gamma, \gamma \in \Gamma\}$ is a chain-representation of p.

References

[1] A. N. Kolmogorov, Grundbegriffe der Wahrscheinlichkeitsrechnung. *Berlin,* 1933.

[2] H. Birkhoff—J. von Neumann, The logic of quantum mechanics, *Ann. of Math.* **37** (1936), 823—843.

[3] L. H. Loomis, On the representation of σ-complete Boolean algebras, *Bull. Amer. Math. Soc.* **53** (1947), 757—760.

[4] Á. Császár, Sur la structure des espaces de probabilité, *Acta Math. Acad. Sci. Hung.* **6** (1955), 337—361.

[5] A. Rényi, On a new axiomatic theory of probability, *Acta Math. Acad. Sci. Hung.* **6** (1955), 285—333.

[6] V. S. Varadarajan, Geometry of quantum theory. *Princeton,* 1955.

[7] A. Rényi, On conditional probability spaces generated by a dimensionally ordered set of measures, *Teorii veraiatnostei i ee promeniia* **1** (1956), 61—71.

[8] P. H. Krauss, Representation of conditional probability measures on Boolean algebras, *Acta Math. Acad. Sci. Hung.* **19** (1968), 229—241.

[9] F. Maeda and S. Maeda, Theory of symmetric lattices, *Berlin,* 1970.

[10] C. Piron, Survey of general quantum physics, *Found. of Phys.* **2** (1972), 287—314.

[11] L. E. Dubins, Finitely additive conditional probabilities, conglomerability and disintegrations, *The Ann. of Prob.* **3** (1975), 89—99.

[12] C. Piron, Foundation of quantum physics. *Massachusetts,* 1976.

I. G. KALMÁR
DEPT. OF MATH.
UNIVERSITY OF DEBRECEN
H-4010 DEBRECEN 10 PF. 12.
HUNGARY

(Received 8 August, 1980)