The Logical Concept of Probability: Foundation and Interpretation

Kurt Weichselberger
(In cooperation with Anton Wallner)
Ludwig-Maximilians-University Munich, Germany,
Kurt.Weichselberger@stat.uni-muenchen.de

Abstract

The Logical concept of probability, introduced to ISIPTA 2005 in a tutorial ([15]), is based on the theory of Interval probability. Since the main feature of the Logical concept is given by the evaluation of arguments consisting of premises and conclusions, it proves necessary to define exactly which kinds of propositions can be employed hereby. If this is done, the analysis allows the definition of independent arguments by examination of the contents of premises and conclusions. If Interval probability is attributed to arguments according to the relevant axioms, a frequency interpretation becomes feasible which decisively relies on the autonomous concept of independence.

Keywords. Interval probability, evaluation of arguments, concept of independence, frequency interpretation, Symmetric theory of probability.

1 Introduction

The motivation to develop the Logical concept of probability and the main features of this concept are thoroughly described in [15]. While in this paper the emphasis was laid on the merits of this concept and of the Symmetric theory of probability which is based on this concept, the present article concentrates on the understanding of the elements which constitute the Logical concept.

Since by the Logical concept probability exclusively is attributed to arguments, the question must be answered, which kind of arguments are suitable for such an evaluation: It is easy to find counter-examples. In Section 2 the concept of P-argument is introduced, meant to cover all situations of interest in statistical analysis and reasoning. This approach allows the definition of mutually independent P-arguments, based on the contents of the respective premises and conclusions. Irrelevance of one proposition with respect to another one and mutual independence of P-arguments constitute the most important aspects of this approach — prior to any kind of evaluating the arguments.

The axioms governing the introduction of interval probability via the establishment of W-fields must be based on this conceptual foundation: Axioms L III and L IV produce the result that mutual independence has the effect of multiplicativity — but not vice versa[1] (Section 3).

A generalization of the classical Binomial law produces a Weak law of large numbers based on mutual independence of arguments. Its result is described by a proposed new expression: weak invergence (Section 4).

The frequency interpretation of the Logical concept which arises out of these results (Section 5) is not liable to objections of circular reasoning. It is seen as the basis of understanding the statements of the Logical concept and of the Symmetric theory of probability.

Section 6 contains a few historical remarks and a comparison of some different approaches to the combined aspects of probability and independence. It also outlines the importance of these results with respect to the Symmetric theory.

2 P-Arguments

Let A and B be propositions describing contingent facts, which may be right or wrong. The propositions A and B are therefore neither tautologies nor antinomies. Only propositions of this kind will be considered in this article. Concerning the ordered pairs (A, B) the question arises what kind of consequences can be drawn from B concerning A.

[1] The paper of ISIPTA'05 does not cover the aspects described in Section 2 of the present article. As a consequence the axioms of W-fields in 2005 are partially different from that in Section 3 of the present article.

Four categories of such pairs (A, B) may be distinguished. They are characterized by:

$$\begin{aligned} K(A, B) &= +1, \\ K(A, B) &= -1, \\ K(A, B) &= 0, \\ K(A, B) &= P. \end{aligned}$$

Note that here $+1$, -1, 0, P are mere symbols and don't represent numbers!

Definition 1 *$K(A, B) = +1$ is meant to describe pairs (A, B), where A can be derived logically from B.* □

Definition 2 *$K(A, B) = -1$ describes pairs (A, B) where $\neg A$ logically can be derived from B.* □

Definition 3 *$K(A, B) = 0$ distinguishes such pairs (A, B), where absolutely no consequences about A can be drawn from B, in particular that for all potential premises B_1 (where $B \wedge B_1$ is not an antinomy)*

$$K(A, B_1) = K(A, B \wedge B_1)$$

holds and for all potential conclusions A_1 (provided that $A \vee A_1$ is not a tautology)

$$K(A_1, B) = K(A \vee A_1, B)$$

holds: B is irrelevant with respect to A. □

Corollary 1 *With respect to the meaning of "irrelevance" it can be concluded that the following implications hold:*

$$K(A, B) = 0 \;\Rightarrow\; K(\neg A, B) = 0;$$

$$K(A, B_1) = 0,\, K(A, B_2) = 0$$
$$\Rightarrow K(A, B_1 \vee B_2) = 0,\, K(A, B_1 \wedge B_2) = 0;$$

$$K(A_1, B) = 0,\, K(A_2, B) = 0$$
$$\Rightarrow K(A_1 \vee A_2, B) = 0,\, K(A_1 \wedge A_2, B) = 0.$$ □

The attachment $K(A, B) = 0$ is determined by the contents of A and of B, and its meaning is generally unquestioned. It is, however, possible that persons with different background disagree with respect to the eventuality of consequences which the facts described by the proposition B can have for the facts described by the proposition A.

It may be expected that a parapsychologist or a supporter of Chaos-theory refuse attachments $K(A, B) = 0$, which are selfunderstanding for other scientists. Concerning characteristic types of (A, B) with religious background there will be an influence of creed.

On the other hand distinction of pairs (A, B) with $K(A, B) = 0$ constitutes fundamental prerequisites in most scientific disciplines as far as empirical research is concerned. Consequently these attachments are inevitable tools of statistical modeling.

Historically the idea that the circumstances of one game of chance must have no consequence whatever for the following games, was prior and fundamental to the idea of introducing probability in analyzing the results of games of chance.

Definition 4 *All ordered pairs (A, B) in consideration which do not belong to the categories $+1$, -1 or 0, are attached to category P. A pair (A, B) belonging to this category is named a* partial argument *or* P-argument *$(A||B)$. B is named the* premise*, A is named the* conclusion *of the P-argument $(A||B)$.* □

It must be agreed that generally P-arguments are the most important tools of learning and therefore are the means of evidential reasoning. The class of P-arguments is huge and extremely heterogeneous.

For clearness it must be pointed out that the category of P-arguments contains pairs (A, B), which at first sight could be expected to belong to category 0:

A pair of propositions (A, B) where $B = R \vee M$ and $K(A, R) = +1$, $K(A, M) = -1$ does not qualify for $K(A, B) = 0$ since it violates a criterion of this category. This aspect may be demonstrated by

Example 1 *Let*

$$\begin{aligned} A &= \text{"It is freezing"}, \\ R_1 &= \text{"The temperature is } -3°C\text{"}, \\ M_1 &= \text{"The temperature is } +2°C\text{"} \end{aligned}$$

and $B_1 = R_1 \vee M_1$. Obviously $K(A, B_1) \in \{0, P\}$. Now let

$$\begin{aligned} R_2 &= \text{"The temperature is } -1°C\text{"}, \\ M_2 &= \text{"The temperature is between } 0°C \text{ and } +3°C\text{"}. \end{aligned}$$

$B_2 = R_2 \vee M_2$ produces $K(A, B_2) \in \{0, P\}$.

$B_1 \wedge B_2 = (R_1 \vee M_1) \wedge (R_2 \vee M_2) =$

$$M_1 = \text{"The temperature is } +2°C\text{"}.$$

Therefore: $K(A, B_1 \wedge B_2) = -1$. Comparison with Definition 3 reveals that $K(A, B_1) = 0$ as well as $K(A, B_2) = 0$ would be in contradiction with the requirements of this definition. Accordingly $K(A, B_1) = P$ and $K(A, B_2) = P$ must hold true and: $(A||B_1)$ as well as $(A||B_2)$ are P-arguments. □

This example points at the possibility of pairs (A, B), where B is not informative directly with respect to A, but nevertheless (A, B) does not belong to category 0, because B contains information which can be relevant with respect to A, if combined with some complementary information.

On the other hand it reveals the existence of P-arguments where the premises are not informative with respect to the conclusions — as long as both of them stand alone.

This possibility contrasts sharply to another type of P-arguments describing reliable empirical knowledge which may be classified as “practically sure”.

Altogether the kinds of treatment with P-arguments are very different in different fields of application: in daily life, in court, in science or in humanities. It is, however, possible to define generally a relation between P-arguments which is of special importance for establishing concepts to evaluate P-arguments.

Definition 5 *The P-arguments* $(A_1||B_1)$ *and* $(A_2||B_2)$ *are* independent of each other, *iff* $K(A_1, B_2) = 0$ *and* $K(A_2, B_1) = 0$. $\square$

Mutual independence of P-arguments is distinguished, therefore, solely by the reciprocal irrelevance of premises with respect to the conclusion of the other P-argument. This definition is prior to all attempts to introduce the concept of probability and in the context it is seen as a prerequisite for establishing a suitable theory.

Generalizations of Definition 5 seemingly can be established in two different ways.

Definition 6 *Let* $(A_i||B_i)$, $i = 1, ..., r$, *be P-arguments. Iff*

$$K(A_i, B_j) = 0, \forall\, i, j \in \{1, ..., r\}, i \neq j,$$

holds, the P-arguments $(A_1||B_1)$, ..., $(A_r||B_r)$ *are* pairwise independent. $\square$

Definition 7 *Iff, under the assumptions of Definition 6,*

$$K\left(\bigwedge_{i\in I_1} A_i,\ \bigwedge_{j\in I_2} B_j\right) = 0,$$
$$\forall \emptyset \subsetneqq I_1,\ I_2 \subsetneqq \{1, ..., r\},\ I_1 \cap I_2 = \emptyset,$$

holds, the P-arguments $(A_1||B_1)$, ..., $(A_r||B_r)$ *are* totally independent from each other. $\square$

Obviously P-arguments which are totally independent from each other are pairwise independent, too. But additionally the following Lemma holds:

Lemma 1 *If* $(A_i||B_i)$, $i = 1, ..., r$, *are pairwise independent P-arguments, then they are totally independent from each other.* $\square$

The *proof* of this Lemma is based on Corollary 1 by induction on r. Obviously Definition 6 and 7 coincide for $r = 2$. It is now presupposed that the assertion of Lemma 1 holds for $r \geq 2$.

Therefore, if $(A_i||B_i)$, $i = 1, ..., r + 1$, are taken as pairwise independent, for every $I_0 \subseteq \{1, ..., r + 1\}$ with $|I_0| \leq r$, the relation

$$K\left(\bigwedge_{i\in I_1} A_i,\ \bigwedge_{j\in I_2} B_j\right) = 0,$$
$$\forall \emptyset \subsetneqq I_1,\ I_2,\ I_1 \cup I_2 \subseteq I_0,\ I_1 \cap I_2 = \emptyset,$$

must be valid.

Now let $\emptyset \subsetneqq I_1,\ I_2 \subsetneqq \{1, ..., r + 1\}$ and $I_1 \cap I_2 = \emptyset$. If there exists $I_0 \subseteq \{1, ..., r + 1\}$, $|I_0| \leq r$, and $I_1 \cup I_2 \subseteq I_0$, according to the assumption

$$K\left(\bigwedge_{i\in I_1} A_i,\ \bigwedge_{j\in I_2} B_j\right) = 0$$

must hold.

If, however, $I_1 \cup I_2 = \{1, ..., r+1\}$, no such I_0 exists. Now two cases have to be distinguished:

a) If $|I_1| < r$, $|I_2| \geq 2$, let

$$I_2 = I_2' \cup I_2'',\ \emptyset \subsetneqq I_2',\ I_2'',\ I_2' \cap I_2'' = \emptyset.$$

Therefore $|I_1 \cup I_2'| \leq r$, $|I_1 \cup I_2''| \leq r$.

Due to the assumption

$$K\left(\bigwedge_{i\in I_1} A_i,\ \bigwedge_{j\in I_2'} B_j\right) = K\left(\bigwedge_{i\in I_1} A_i,\ \bigwedge_{j\in I_2''} B_j\right) = 0,$$

and Corollary 1 produces

$$K\left(\bigwedge_{i\in I_1} A_i,\ \bigwedge_{j\in I_2} B_j\right) =$$
$$= K\left(\bigwedge_{i\in I_1} A_i,\ \bigwedge_{j\in I_2'} B_j \wedge \bigwedge_{j\in I_2''} B_j\right) = 0.$$

b) If $|I_1| = r$, $|I_2| = 1$, let

$$I_1 = I_1' \cup I_1'',\ \emptyset \subsetneqq I_1',\ I_1'',\ I_1' \cap I_1'' = \emptyset.$$

$|I_1' \cup I_2| \leq r$, $|I_1'' \cup I_2| \leq r$,

$$K\left(\bigwedge_{i\in I_1'} A_i,\ \bigwedge_{j\in I_2} B_j\right) = K\left(\bigwedge_{i\in I_1''} A_i,\ \bigwedge_{j\in I_2} B_j\right) = 0,$$

and due to Corollary 1:

$$K\left(\bigwedge_{i\in I_1} A_i,\ \bigwedge_{j\in I_2} B_j\right) =$$

$$= K\left(\bigwedge_{i\in I_1'} A_i \wedge \bigwedge_{i\in I_1''} A_i,\ \bigwedge_{j\in I_2} B_j\right) = 0.$$

In both cases the conditions for total independence of $(A_i||B_i)$, $i = 1, ..., r$, are satisfied. □

Consequently in the following only the concept of r mutual independent P-arguments has to be taken into consideration.

This result characterizes the difference between the concept of independence in the theory employed here and in the Classical theory: Independence is a more demanding relation in the theory of the Logical concept.

3 W-Fields

The most advanced method of evaluating P-arguments is that of attaching interval-probability. It affords the selection of two sets of propositions: $\mathcal{A}_P$, $\mathcal{B}_P$ with $\mathcal{A}_P \cap \mathcal{B}_P = \emptyset$, so that

$$K(A, B) \in \{0, P\},\ \forall A \in \mathcal{A}_P,\ B \in \mathcal{B}_P.$$

$\mathcal{A}_P$ as well as $\mathcal{B}_P$, in the second step, have to be completed, if necessary, to generate sets $\mathcal{A}_P^*$ and $\mathcal{B}_P^*$, which are closed under the logical operations $\vee$, $\wedge$, and $\setminus$ ("logical difference"). Additional potential conclusions A and additional potential premises B may produce additional P-arguments, but ordered pairs (A, B) of category $+1$, -1 or 0 as well. It must be secured that all assignments are in concordance with the definitions of $K(A, B)$.

According to the tradition and the actual practice of probability theory conclusions as well as premises should be described by sets. Therefore the elements of $\mathcal{A}_P^*$ and $\mathcal{B}_P^*$ must be represented by sets in a way guaranteeing that logical operations on $\mathcal{A}_P^*$ and on $\mathcal{B}_P^*$ are transformed to the corresponding set operations on the representing sets $\mathcal{A}$ and $\mathcal{B}$ with $\mathcal{A} \cap \mathcal{B} = \emptyset$.

Obviously this representation is by no means uniquely determined. It always must be borne in mind that the tools of representation must not influence the decisive probabilistic reasoning.

Then any assignment of interval-probability is produced by

$$P(A||B) = [L(A||B),\ U(A||B)],\ \forall A \in \mathcal{A},\ B \in \mathcal{B}.$$

It may be understood as the result of evaluating P-arguments completed by the following attachments:

$L(A||B) = 1$, $U(A||B) = 1$, if for the corresponding propositions $K(A, B) = 1$ holds;

$L(A||B) = 0$, $U(A||B) = 0$, if $K(A, B) = -1$ holds;

$L(A||B) = 0$, $U(A||B) = 1$, if $K(A, B) = 0$ holds.

The probability of any P-argument $(A||B)$ determines the interval-limits $L(A||B)$ and $U(A||B)$ for the representing sets A and B. The rules governing this assessment are given in a three-level hierarchy:

1) Classical theory of probability:

Any function $p(.)$ on a measure space $(\Omega; \mathcal{A})$ which obeys Kolmogorov's three axioms is called a *K-function.*

2) Theory of interval probability (see [14])**:**

An *F-(probability-)field* $\mathcal{F} = (\Omega; \mathcal{A}; L(.))$ is given, iff the following three axioms hold[2]:

T IV: $P(A) = [L(A); U(A)] \subseteq [0; 1],\ \forall A \in \mathcal{A}$.

T V: The set $\mathcal{M}$ of K-functions $p(.)$ on $(\Omega; \mathcal{A})$ with $L(A) \leq p(A) \leq U(A)$, $\forall A \in \mathcal{A}$, is not empty.

T VI: $\inf_{p(.)\in\mathcal{M}} p(A) = L(A)$, $\sup_{p(.)\in\mathcal{M}} p(A) = U(A)$, $\forall A \in \mathcal{A}$.

3) Logical concept of probability:

Let $(\Omega_A; \mathcal{A})$ and $(\Omega_B; \mathcal{B})$, $\Omega_A \cap \Omega_B = \emptyset$, be two measure spaces, where $\{x\} \in \mathcal{A}$, $\forall x \in \Omega_A$, $\{y\} \in \mathcal{B}$, $\forall y \in \Omega_B$.

A *W-field* $\mathcal{W} = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L(.||.))$ is given, iff the following four axioms hold:

L I: To each $B \in \mathcal{B}^+ := \mathcal{B} \setminus \{\emptyset\}$ an F-field $\mathcal{F}(B) = (\Omega_A; \mathcal{A}; L(.||B))$ is attached.

L II: Let $I \neq \emptyset$ be an index set, $B_0 \in \mathcal{B}^+$, $B_i \in \mathcal{B}^+$, $i \in I$, and

$$B_0 = \bigcup_{i\in I} B_i.$$

Then[3]:

$$\mathcal{F}(B_0) = \bigcup_{i\in I} \mathcal{F}(B_i).$$

[2]According to Axioms T IV–T VI the function $U(.)$ is conjugate to $L(.)$: $U(A) = 1 - L(\neg A)$, $\forall A \in \mathcal{A}$.

[3]The union $\mathcal{F} = \cup_{i\in I}\mathcal{F}_i = (\Omega_A; \mathcal{A}; L(.))$ of F-fields $\mathcal{F}_i = (\Omega_A; \mathcal{A}; L_i(.))$, $i \in I$, is defined by $L(.) := \inf_{i\in I} L_i(.)$. Hence $U(.) = \sup_{i\in I} U_i(.)$, and $\mathcal{F}$ is an F-field too. The employment of this procedure in assigning probability of arguments characterizes the Logical concept in contrast to the Bayesian approach.

L III: Let $A \in \mathcal{A}$, $B_1 \in \mathcal{B}^+$ irrelevant for A. Then:

$$L(A||B_1 \cap B_2) = L(A||B_2),\ \forall B_2 \in \mathcal{B}^+.$$

L IV: Let $A_i \in \mathcal{A}^+$, $B_i \in \mathcal{B}^+$, $i = 1, 2$, $(A_1||B_1)$ and $(A_2||B_2)$ independent from each other. Then:

$$L(A_1 \cap A_2||B_1 \cap B_2) = \\ = L(A_1||B_1 \cap B_2) \cdot L(A_2||B_1 \cap B_2)$$

$$U(A_1 \cap A_2||B_1 \cap B_2) = \\ = U(A_1||B_1 \cap B_2) \cdot U(A_2||B_1 \cap B_2).$$

The Logical concept of probability defined by Axioms L I–L IV as a general principle employs probability as a two-place-function: $P(A||B)$ is to be interpreted as probability of the argument with premise B and with conclusion A and never must be mistaken as conditional probability. According to this concept $P(A)$ and $P(B)$ do not exist and therefore $P(A|B)$ never exists either. (On the other hand $P((A_1|A_2)||B)$ is a valuable information in many situations.) Axiom L II characterizes the distinction of the Logical concept and any kind of Bayesian concept.

The fact that $\Omega_A \cap \Omega_B = \emptyset$ and therefore $\mathcal{A}$ and $\mathcal{B}$ always are disjoint, demonstrates the basic distinction between W-fields and Popper-spaces (cf. [12] and [11]). This does not prevent the idea of combining both aspects — but the success of such a program cannot be foreseen.

On the other hand there is no relationship of the Logical concept with approaches of Default reasoning (cf. [7] and [10]) or of Plausibility measures and Possibility measures. The Logical Concept does not extend the field of application beyond that of classical probability: Its main goal is to improve the methodology of statistical reasoning by introducing duality between appropriate W-fields and hereby allowing the employment of probability to describe results of statistical inference. With respect to the intention there is a relationship to approaches by R.A. Fisher ([5]), D.A.S. Fraser ([6]), A. Dempster ([4]), A. Birnbaum ([2]), and I. Hacking ([8]), but there exist fundamental differences in methodology[4].

A survey of the resulting Symmetric theory of probability was given in the ISIPTA 05 paper, a short survey of duality in statistical inference can be found in a report for the 56th Session of ISI in Lisboa, 2007 ([16]).

[4]A review of these approaches is given by T. Seidenfeld ([13]).

4 Independence and Multiplicativity

Axioms L I–L IV allow to establish a corpus of definitions and statements constituting the theory of the Logical concept of probability. With one important exception the theory of the classical concept can be regarded as a special case of this theory. The difference between the two approaches with respect to the concept of independence is characterized by the results of this section.

Corollary 2 *From Axiom L III and Corollary 1 it follows that under the conditions for L III:*

$$U(A||B_1 \cap B_2) = U(A||B_2),\ \forall\, B_2 \in \mathcal{B}^+,$$

holds. □

From Axioms L III and L IV together with Corollary 2, it may be concluded:

Corollary 3 *If $(A_1||B_1)$ and $(A_2||B_2)$ are mutually independent,*

$$\begin{aligned} L(A_1 \cap A_2||B_1 \cap B_2) &= L(A_1||B_1) \cdot L(A_2||B_2) \\ U(A_1 \cap A_2||B_1 \cap B_2) &= U(A_1||B_1) \cdot U(A_2||B_2) \end{aligned}$$

hold. □

This result says that, according to the Logical Concept, mutual independence of P-arguments produces total multiplicativity of probabilities. However, on the other hand, it is not possible in this theory to infer mutual independence of P-arguments from multiplicativity of probabilities. This is a decisive difference to the objectivistic view of classical theory, where independence of events is defined by means of multiplicativity of probabilities. It should be emphasized that mutual independence of P-arguments can only be understood as the fact that each premise is irrelevant to the conclusion of the other P-argument.

Now let $(A_i||B_i)$ with $P(A_i||B_i) = [L; U]$, $i = 1, 2, ...$, be a potentially infinite series of mutually independent P-arguments and $r \in \mathbb{N}$.

Due to independence, the probability for the combined P-argument $(\vec{A}||\vec{B})$ with $\vec{A} = A_1^* \times ... \times A_r^*$ where $A_i^* \in \{A_i, \neg A_i\}$, $\vec{B} = B_1 \times ... \times B_r$ is multiplicative:

$$P^{[r]}(\vec{A}||\vec{B}) = \left[\prod_{i=1}^{r} L(A_i^*||B_i); \prod_{i=1}^{r} U(A_i^*||B_i) \right].$$

Let $I \subseteq \{1, ..., r\}$, $A_i^* = A_i$, $\forall\, i \in I$, $A_i^* = \neg A_i$, $\forall\, i \notin I$. Then I describes the conclusion $\vec{A} =: \vec{A}(I)$ uniquely. Because of

$$\begin{aligned} P^{[r]}(A_i^*||B_i) &= [L; U],\ \forall\, i \in I, \\ P^{[r]}(A_i^*||B_i) &= [1 - U; 1 - L],\ \forall\, i \notin I, \end{aligned}$$

one arrives at

$$\begin{aligned} P^{[r]}(I||\vec{B}) &:= P(\vec{A}(I)||\vec{B}) \\ &= \left[L^{|I|}\cdot(1-U)^{r-|I|};\ U^{|I|}\cdot(1-L)^{r-|I|}\right]. \end{aligned}$$

Let $\rho := |I|$. In order to calculate the probability of an argument with a conclusion of very few $\neg A_i$ and almost all A_i,

$$P^{[r]}(\rho \geq \rho_0||\vec{B}) =$$
$$= \left[\inf_{\substack{L\leq p_i\leq U\\ i\in\{1,\,...,\,r\}}} \sum_{|I|\geq\rho_0} \prod_{i\in I} p_i \prod_{i\in\{1,\,...,\,r\}\setminus I} (1-p_i);\ \sup_{\substack{L\leq p_i\leq U\\ i\in\{1,\,...,\,r\}}} \sum_{|I|\geq\rho_0} \prod_{i\in I} p_i \prod_{i\in\{1,\,...,\,r\}\setminus I} (1-p_i)\right]$$

has to be calculated. Due to the monotonicity of the function

$$p^{[r]}(\rho \geq \rho_0||\vec{B}) = \sum_{|I|\geq\rho_0} \prod_{i\in I} p_i \prod_{i\in\{1,\,...,\,r\}\setminus I} (1-p_i)$$

in each of the $p_i \in [L;\, U]$, $i = 1,\, ...,\, r$,

$$P^{[r]}(\rho \geq \rho_0||\vec{B}) =$$
$$= \left[\sum_{\rho\geq\rho_0} \binom{r}{\rho} L^\rho (1-L)^{r-\rho};\ \sum_{\rho\geq\rho_0} \binom{r}{\rho} U^\rho (1-U)^{r-\rho}\right]$$

holds.

On the other hand:

$$P^{[r]}(\rho \leq \rho_0||\vec{B}) =$$
$$= \left[\sum_{\rho\leq\rho_0} \binom{r}{\rho} U^\rho (1-U)^{r-\rho};\sum_{\rho\leq\rho_0} \binom{r}{\rho} L^\rho (1-L)^{r-\rho}\right].$$

Therefore the probabilities of arguments with conclusions of extremely many or extremely few factors A_i can be calculated employing classical Binomial law and Tschebysheff's inequality:

Let

$$\rho_0 = rU + r\delta: \quad U^{[r]}(\rho \geq \rho_0||\vec{B}) \leq \frac{U(1-U)}{r\delta^2},$$

let

$$\rho_0^* = rL - r\delta: \quad U^{[r]}(\rho \leq \rho_0^*||\vec{B}) \leq \frac{L(1-L)}{r\delta^2}.$$

As a consequence:

$$L^{[r]}\left(L-\delta < \frac{\rho}{r} < U+\delta\right) \geq 1 - \frac{L(1-L)+U(1-U)}{r\delta^2}$$

and

$$\begin{aligned} &L^{[r]}(L-\delta < \frac{\rho}{r} < U+\delta) \geq 1-\varepsilon, \\ &\qquad \text{if } r \geq \frac{L(1-L)+U(1-U)}{\varepsilon\delta^2}. \end{aligned} \tag{1}$$

This result can be interpreted by means of appropriate concepts of converging sequences of W-fields:

Definition 8 *Let* $\mathcal{W}^{[r]} = (\Omega_A;\, \mathcal{A};\, \Omega_B^{[r]};\, \mathcal{B}^{[r]};\, L^{[r]}(.||.))$, $r \in \mathbb{N}$, *be a sequence of W-fields and* $\mathcal{Z} \in \mathcal{A}$ *be a non-empty conclusion. If for* $\vec{B}_0^{[r]} = B_0^{[1]} \times ... \times B_0^{[r]}$, $r \in \mathbb{N}$, *and for every* $\mathcal{Z}^* \in \mathcal{A}$ *with* $\mathcal{Z}^* \supsetneqq \mathcal{Z}$ *there exists a function* $N(\mathcal{Z}^*,\, \varepsilon) \in \mathbb{N}$, *so that*

$$L^{[r]}(\mathcal{Z}^*||B_0^{[r]}) \geq 1-\varepsilon,\ \forall\, r \geq N(\mathcal{Z}^*,\, \varepsilon), \tag{2}$$

then with respect to the arguments $(\mathcal{Z}||\vec{B}_0^{[r]})$, $r \in \mathbb{N}$, *the sequence* $\mathcal{W}^{[r]}$ *is named* stochastically convergent *to a sequence* $\overline{\mathcal{W}}^{[r]}$ *of W-fields,* $\overline{\mathcal{W}}^{[r]} = (\Omega_A;\, \mathcal{A};\, \Omega_B^{[r]};\, \mathcal{B}^{[r]};\, \overline{L}^{[r]}(.||.))$, $r \in \mathbb{N}$, *with* $\overline{P}^{[r]}(\mathcal{Z}||B_0^{[r]}) = [1;\, 1] =: [1]$. $\square$

According to Definition 8 the result (1) can be utilized to formulate

Corollary 4 *Let the sequence of P-arguments* $(A_0^{(i)}||B_0^{(i)})$, $i \in \mathbb{N}$, *with* $P^{(i)}(A_0^{(i)}||B_0^{(i)}) = [L;\, U]$ *be mutually independent. For* $r \in \mathbb{N}$ *let*

$$\begin{aligned} \vec{B}_0^{[r]} &= B_0^{(1)} \times ... \times B_0^{(r)}, \\ t &= \frac{\rho}{r},\ \rho \in \{0,\, ...,\, r\}, \\ A_0^{[r]}(t) &:= \bigcup_{\substack{I\subseteq\{0,\,...,\,r\}\\ |I|=r\cdot t}} \left[\bigcap_{i\in I} A_0^{(i)} \cap \bigcap_{i\in\{0,\,...,\,r\}\setminus I} \neg A_0^{(i)}\right]. \end{aligned}$$

Let $\mathcal{W}^{[r]} = (\Omega_A^{[r]};\, \mathcal{A}^{[r]};\, \Omega_B^{[r]};\, \mathcal{B}^{[r]};\, L^{[r]}(.||.))$, $r \in \mathbb{N}$, *be W-fields containing the probability of arguments with premise* $\vec{B}_0^{[r]}$ *and conclusions of the kind* $A_0^{[r]}(J) = \bigcup_{r\cdot t\in J} A_0^{[r]}(t)$, $J \subseteq \{0,\, ...,\, r\}$, *so that*

$$\begin{aligned} \Omega_A^{[r]} &= \{0, \tfrac{1}{r}, \tfrac{2}{r}, ...,\, 1\}, \\ \mathcal{A}^{[r]} &= \mathcal{P}ot(\Omega_A^{[r]}), \\ \vec{B}_0^{[r]} &\in \mathcal{B}^{[r]}. \end{aligned}$$

The sequence $\mathcal{W}^{[r]}$, $r \in \mathbb{N}$, *is then with respect to the arguments* $(A_0^{[r]}(t)||\vec{B}_0^{[r]})$ *stochastically convergent to the sequence* $\overline{\mathcal{W}}^{[r]}$, $r \in \mathbb{N}$, *of W-fields*

$$\overline{\mathcal{W}}^{[r]} = (\overline{\Omega}_A^{[r]};\, \overline{\mathcal{A}}^{[r]};\, \Omega_B^{[r]};\, \mathcal{B}^{[r]};\, \overline{L}^{[r]}(.||.))$$

with

$$\begin{aligned}\overline{\Omega}_A^{[r]} &= [0;\,1],\\ \overline{\mathcal{A}}^{[r]} &= \mathcal{B}or(\overline{\Omega}_A^{[r]}),\\ \overline{L}^{[r]}(A||\vec{B}_0^{[r]}) &= \begin{cases}[1], & A \supseteq [L;\,U]\\ [0], & A \cap [L;\,U] = \emptyset\\ [0;\,1], & \textit{else.}\end{cases}\end{aligned}\quad\square$$

Corollary 4 is the obvious consequence of (1) if applied to a sequence of mutually independent arguments with probability $[L;\,U]$.

Introducing abbreviations, the result (1) may be expressed by the statement

$$\lim_{r\to\infty} P^{[r]}(L \le t_r \le U||\vec{B}_0^{[r]}) = [1] \qquad (3)$$

and may be interpreted in a way similar to that concerning convergence of a sequence of variables in classical statistics — if the decisive difference is seen, that $[L;\,U]$ is an interval and normally not a single number. From the facts given, it is not possible to describe the result by more information than, what may be characterized by "finally: $L \le t_r \le U$". Neither convergence nor divergence of the sequence t_r, $r = 1, \ldots$, can be excluded as a possible conclusion. As an appropriate new expression to denote results of this type the sentence "The sequence t_r inverges the set $[L;\,U]$" is proposed.

5 Frequency Interpretation

As a consequence of these results a frequency interpretation of the Logical concept of probability is available.

If $(A||B)$ is a P-argument with $P(A||B) = [L;\,U]$ in a kind of thought experiment, then $(A||B)$ may be conceived as one out of a potentially infinite sequence of mutually independent P-arguments $(A_i||B_i)$ with exactly the same probability assessment:

$$P(A_i||B_i) = [L;\,U],\ i = 1,\,2,\,\ldots$$

Then the P-argument $(\hat{A}^{[r]}||\hat{B}^{[r]})$ is considered, where

$$\hat{B}^{[r]} := \bigcap_{i=1}^{r} B_i$$

is the conjunction of all single premises and

$$\hat{A}^{[r]} := \bigcup_{\substack{I\subseteq\{1,\,\ldots,\,r\}\\ rL\le|I|\le rU}} \left[\bigcap_{i\in I} A_i \cap \bigcap_{i\in\{1,\,\ldots,\,r\}\setminus I} \neg A_i\right]$$

is the adjunction of all combined conclusions, for which the proportion of A_i lies between L and U.

Due to (2) and (3)

$$\lim_{r\to\infty} P^{[r]}(\hat{A}^{[r]}||\hat{B}^{[r]}) = [1]$$

holds.

If only r is large enough, the conclusion $\hat{A}^{[r]}$ can be derived from the premise $\hat{B}^{[r]}$ with practical surety.

Therefore the P-argument $(A||B)$ with $P(A||B) = [L;\,U]$ can be interpreted as if it was one out of a huge set of mutual independent P-arguments $(A_i||B_i)$, for which the proportion of successful arguments $(A_i||B_i)$ — producing unsuccessful arguments $(\neg A_i||B_i)$ — lies between L and U, the proportion of unsuccessful arguments $(A_i||B_i)$ — and therefore successful arguments $(\neg A_i||B_i)$ — lies between $1-U$ and $1-L$.

The conceptual basis of this kind of procedure is given by Cournot's Lemma, which was formulated with reference to the objectivistic view on probability, and can be transferred to the Logical concept in the following way:

Cournot's Lemma: *If $L(A||B) = 1-\varepsilon$ and ε is extremely small, the P-argument $(A||B)$ may be understood as if $P(A||B) = [1]$.* $\square$

The validity of this interpretation is founded on the fact that in a set of mutually independent P-arguments, for which only $P(A_i||B_i) = [L;\,U]$ is known, obviously no subset can be identified, for which an additional information about the proportion of successful arguments $(A_i||B_i)$ would be possible.

It must be pointed to the fact, that the value of this frequency-interpretation of the Logical concept of probability in the first line depends on the concept of independent P-arguments: A set of mutual independent P-arguments is defined by contents of premises and conclusions. If additionally all of them are evaluated by the same $P(A_i||B_i)$ according to the axioms of the Logical concept, and if the set is large enough, inference about the proportion of successful ones in the set is possible.

The availability of an unassailable frequency interpretation of the Logical concept is of high importance with respect to the Symmetric theory (see [15]): In this theory probabilistic statements about arguments are employed not only to describe statistical modeling but also — by means of dual W-fields — for statistical inference. The concept of imaging any evaluated argument as one out of a potentially infinite sequence of mutually independent arguments with the same probability guarantees the uniformity in understanding the assessments employed in modeling as well as in inference.

Finally it must be mentioned that it is possible to improve the result of weak invergence: If a more general concept of W-field is introduced, the concept of strong invergence can be defined and it can be proven, that t_r inverges $[L; U]$ with probability [1] according to this concept. However, this result does not influence the understanding of a frequency interpretation of the Logical concept.

6 Conclusions and Prospect

When early in the 17th century the concept of probability arose from the study of games of chance, the possibility of repeating any game was a fundamental idea, comprising the concept of mutual independence of the repetitions. Multiplicativity of probability under this supposition was accepted as intuition resulting in the close relationship between frequency and probability ([14], pp. 42 ff.).

When the theory of probability developed in the following centuries it was obvious that mutual independence of events is a relation which cannot be defined without employment of exogenous concepts.

It was A. N. Kolmogorov who produced a solution by defining mutual independence via multiplicativity of probabilities ([9], pp. 8–12 of the English version), hereby accepting some border cases which are more or less counter-intuitive.

Employment of imprecise probabilities, as it is propagated by Peter Walley must rely on definitions as they are given for precise probabilities. Owing to a behaviouristic approach the concept of irrelevance is near at hand ([3]). As long as probability is seen as a one-place-function — attributed to events or statements — this remains conditional irrelevance of one event with respect to another one. Considerations of this kind come near to a justification of multiplicativity, but fail to explain independence without using the concept of probability ([3]). An additional aspect is provided by the question which type of conditional interval probability is to be employed in such consideration ([1]). Altogether: The difference between the concept of independence as employed in the present approach and other concepts of independence introduced in methodology of imprecise probability is fundamental: Independence of P-arguments is defined by the contents of propositions employed, while independence of events with imprecise probability is defined by relations of — total or conditional — probabilities.

This remains the situation of classical probability theory according to the objectivistic view. By means of the weak law of large numbers a frequency interpretation of classical probability can be derived, based on the concept of a sequence of mutual independent events with the same probability.

Criticism of the objective view stresses that probability being the conceptual basis defining independence of events, it should not be interpreted by a characteristic of a sequence of mutual independent events.

If probability is attributed to arguments — instead of events — the situation is different, since independence of two arguments can be identified with irrelevance of both premises with respect to the conclusion of the other argument.

The ISIPTA'05 paper ([15]) defines mutual independence of arguments by means of the probability assessment: It is in fact Axiom L III of the present paper which is employed in [15] to distinguish mutual independent arguments without denoting this procedure explicitly. Simultaneously multiplicativity of probability for independent arguments is required: Thus Axiom L III of the '05-paper combines two different aspects in one equation.

This procedure can be criticized not only because of its complexity, but also with respect to the detail that it encourages objections against a frequency interpretation of probability employing a sequence of independent arguments because of the role of probability in defining independence.

The present paper relies on definitions of irrelevance and independence through the contents of the propositions involved. Axiom L III contains the demand that irrelevance always is reflected by the probability assessment, and Axiom L IV insists on multiplicativity with respect to conclusions in case of mutual independence — in analogy to multiplicativity in classical probability.

Consequently the Logical concept of probability according to Section 5 is characterized by a frequency interpretation employing an autonomous concept of independence — a feature not to be found elsewhere.

This result characterizes the Symmetric theory of probability, which relies decisively on the Logical concept. The establishment of duality between W-fields generates a methodology of statistical inference employing the concept of probability in the same way as in statistical modeling attached to P-arguments.

Any statement of probability arising from applied Symmetric theory therefore is of the same quality and should be understood by means of the frequency interpretation: one out of a very large series of assessments with the same $[L; U]$ concerning mutually independent P-arguments. The proportion of "successful" arguments in this series lies between L and U.

Present research into Symmetric theory aims at the range of possible applications. Which types of problems in classical statistics can be solved by means of the concept of duality?! A comprehensive report [17] is in preparation.

Acknowledgements

I am grateful to Joseph Kadane, with whom I had a short but very impressive talk together with filet Mignon at Andy Warhol-Museum in Pittsburgh on July 21, 2005, I am grateful to Thomas Augustin whose remarks upon the draft were very instructive, and I am grateful to Anton Wallner who supports me through many years and corrects me — hopefully always when it is necessary!

References

[1] T. Augustin, K. Weichselberger. On the symbiosis of two concepts of conditional interval probability, in: J.-M. Bernard, T. Seidenfeld, M. Zaffalon (eds.), *ISIPTA'03, Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, Carleton Scientific, 608–629, 2003.

[2] A. Birnbaum. Concepts of statistical evidence, in: S. Morgenbesser, P. Suppes, M. White (eds.), *Philosophy, Science, and Method*, St. Martin's Press, New York, 1969.

[3] I. Couso, S. Moral, P. Walley. A survey of concepts of independence for imprecise probabilities, in: *Risk, Decision and Policy*, vol. 5, 165–181, 2000.

[4] A.P. Dempster. Upper and lower probabilities generated by a random closed interval, in: *Annals of Mathematical Statistics* 39, 957–966, 1968.

[5] R.A. Fisher. *Statistical Methods and Scientific Inference*, 3rd ed., Hafner, New York, 1973.

[6] D.A.S. Fraser. *The Structure of Inference*, Wiley, New York, 1968.

[7] N. Friedman, J.Y. Halpern, D. Koller. First-order conditional logic for default reasoning revisited, in: *ACM Transactions on Computational Logic*, Vol. 1, No. 2, 175–207, 2000.

[8] I. Hacking. *Logic of Statistical Inference*, Cambridge University Press, Cambridge, 1965.

[9] A. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin, Heidelberg, New York, 1933. English version: *Foundations of the Theory of Probability*, Chelsea Publishing Company, New York, 1950.

[10] J. Pearl. *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Mateo, CA, 1988.

[11] K.R. Popper. *The Logic of Scientific Discovery (revised edition)*, Hutchison, London, 1968. The first version of this book appeared as *Logik der Forschung*, 1934.

[12] A. Rényi. On conditional probability spaces generated by a dimensionally ordered set of measures, in: *Theory of Probability and its Applications 1*, 61–71, 1956. Reprinted as paper 120 in *Selected Papers of Alfred Rényi, I: 1948–1956*, Akadémia Kiadó, 554–557, 1976.

[13] T. Seidenfeld. *Philosophical Problems of Statistical Inference*, Reidel, Dordrecht, 1979.

[14] K. Weichselberger. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I — Intervallwahrscheinlichkeit als umfassendes Konzept*, in cooperation with T. Augustin and A. Wallner, Physica, Heidelberg, 2001.

[15] K. Weichselberger. The logical concept of probability and statistical inference, in: F.G. Cozman, R. Nau, T. Seidenfeld (eds.), *ISIPTA'05, Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications*, published by SIPTA, 396–405, 2005.

[16] K. Weichselberger. Applying symmetric probability theory, in: *Proceedings of the 56th Session of International Statistical Institute ISI 2007*, CD, Internet.

[17] K. Weichselberger. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung II — Symmetrische Wahrscheinlichkeitstheorie*, in cooperation with A. Wallner, in preparation.