The Theory of Interval-Probability as a Unifying Concept for Uncertainty

Kurt Weichselberger
Ludwig-Maximilians University, Munich
weichsel@stat.uni-muenchen.de

Abstract

The concept of interval-probability is motivated by the goal to generalize classical probability so that it can be used for describing uncertainty in general. The foundations of the theory are based on a system of three axioms — in addition to Kolmogorov's axioms — and definitions of independence as well as of conditional probability. The resulting theory does not depend upon interpretations of the probability concept. As an example of generalizing classical results the Bayes' rule is described — other theorems are only mentioned.

Keywords. Interval-probability, uncertainty, conditional probability, Bayes' Rule.

1 The Scope of the Theory

The theory of interval-probability, as developed in Munich over many years, is motivated by the following goals:

1. Different kinds of uncertainty should be treated by the same concept.

   This applies to

   - imprecise probability and uncertain knowledge
   - imprecise data
   - the use of capacities
   - the concept of ambiguity and its employment in decision theory
   - belief functions and related concepts
   - interpretation of interval-estimates in classical theory
   - the study of experiments with possibly diverging relative frequencies
   - non-additive measures (fuzzy measures)

2. As a special case, classical probability must fit into this theory.
3. A simple system of axioms must describe the fundamentals of the theory.
4. All statements of the theory must be derivable from the given axioms and appropriate definitions.
5. The domain of application must neither be limited to purely formal aspects nor be bound by a certain interpretation of probability.

In classical probability one system of axioms exists not being restricted to a certain type of interpretation: Kolmogorov's axioms. Therefore, the concept of interval-probability is directly related to this system of axioms.

There is one obvious limitation for any theory of interval-probability: Only those assessments assigning intervals to random events qualify as genuine subjects of the theory. The benefits and the power of the theory are due to the duality between a set of interval-limits and the corresponding set of classical probabilities. These qualities distinguish the approach described in the following chapters from those admitting more general types of probability assignments, e.g. Kofler and Menges[1]or Peter Walley[2].

Since the theory of interval-probability is independent on the kind of interpretation it suits for fields of application, where probability is understood as means of argumentation without relation either to betting or to large series of random experiments.

Also it produces freedom in describing behaviour in a very general way: Ellsberg's remark that everyone will switch to a favourable event with probability [0; 1] instead of an equally favourable with probability p, provided that p is small enough, can be taken into consideration adequately.

Altogether, theory of interval-probability comes nearer to the classical understanding of probability

[1] [3].
[2] [4].

assignment than those approaches relying on more general types of assessment.

2 Basic Concepts

2.1 The Axioms

In a slightly specialized version of the axioms all closed intervals in [0;1] are admitted as components of interval-probability. In this case the following definitions may be understood as describing the system of axioms for interval-probability.

Definition 2.1 Given a sample space Ω and a σ-field $\mathcal{A}$ of random events in Ω, a set function $p(\cdot)$ defined on $\mathcal{A}$ is named a *K-function*, if it obeys the axioms of Kolmogorov (I–III).

Since K-functions have the same properties as classical probabilities, sometimes they are named *K-probabilities*.

Definition 2.2 An interval-valued set function $P(\cdot)$ on $\mathcal{A}$ is called an *R-probability* if it obeys the following two axioms.

IV. $$P(A) = [L(A); U(A)], \quad \forall A \in \mathcal{A}, \tag{1}$$

with $$0 \leq L(A) \leq U(A) \leq 1, \quad \forall A \in \mathcal{A}. \tag{2}$$

V. The set $\mathcal{M}$ of K-functions $p(\cdot)$ on $\mathcal{A}$ with

$$L(A) \leq p(A) \leq U(A), \quad \forall A \in \mathcal{A}, \tag{3}$$

is not empty.

The name “R-probability” may be related to the word “reasonable”. A quadruple consisting of a sample space Ω, a σ-field $\mathcal{A}$ of random events and a certain R-probability on $(\Omega; \mathcal{A})$ will be called an R-(probability) field $(\Omega; \mathcal{A}; L(\cdot), U(\cdot))$. The set $\mathcal{M}$ of K-probabilities being in accordance with (3) is named the *structure* of the R-probability field. Therefore the existence of a non-empty structure is the only condition for any R-field. It is obvious that

$$L(\emptyset) = 0, \quad U(\Omega) = 1 \tag{4}$$

are among the necessary conditions for R-probability.

Definition 2.3 An R-probability obeying the following axiom is named an *F-probability*.

VI. $$\left.\begin{array}{l} \inf_{p \in \mathcal{M}} p(A) = L(A) \\ \sup_{p \in \mathcal{M}} p(A) = U(A) \end{array}\right\} \forall A \in \mathcal{A}. \tag{5}$$

The letter F may be connected with the word “feasible”. In any F-probability field none of the limits $L(\cdot)$ and $U(\cdot)$ are too wide, while this may be the case for an R-field. Furthermore the property of F-probability implies the validity of

$$U(A) = 1 - L(\neg A), \ \forall A \in \mathcal{A}, \tag{6}$$

and of

$$U(\emptyset) = 0, \quad L(\Omega) = 1, \tag{7}$$

which by use of the symbol

$$[a] := [a; a] \tag{8}$$

together with (4) read as

$$P(\emptyset) = [0], \quad P(\Omega) = [1]. \tag{9}$$

A triple consisting of a sample-space Ω, a σ-field $\mathcal{A}$ of random events and a given F-probability is understood to be an F-(probability) field $(\Omega; \mathcal{A}; L(\cdot))$.

The concept of structure is fundamental for the theory of interval-probability. Most definitions and proofs are directly or indirectly related to it. Another important concept is that of *prestructure*.

Definition 2.4 Let $(\Omega; \mathcal{A}; L(.))$ be an F-probability field. Then any set $\mathcal{J}$ of K-probabilities on $(\Omega; \mathcal{A})$ is named a *prestructure* of $\mathcal{F}$ if the following relations hold:

$$\inf_{p \in \mathcal{J}} p(A) = L(A), \quad \forall A \in \mathcal{A}. \tag{10}$$

According to that, every subset of the structure producing the same lower — and as a set of K-probabilities also the same upper — limits constitutes a prestructure. As long as it does not contain all K-functions in accordance with (10), $\mathcal{J}$ is different from $\mathcal{M}$.

The given system of axioms may be applied to finite sample spaces and to infinite ones, but in the case of F-probability it proves useful to distinguish continuous F-probability, since Axioms I-VI do not guarantee the continuity of the set functions $L(\cdot)$ and $U(\cdot)$.

Definition 2.5 An F-probability is called *continuous* if for any decreasing sequence of events of $\mathcal{A}$:

$$A_1 \supset A_2 \supset \ldots \supset A_n \ldots, \tag{11}$$

for which

$$\bigcap_{i=1}^{\infty} A_i =: A \tag{12}$$

is valid, the following equation holds:

$$\lim_{n \to \infty} U(A_n) = U(A). \tag{13}$$

2.2 R-Probability and F-Probability

R-probability may be interpreted as "not contradictory, but not necessarily perfect", since on the one hand it allows the existence of a structure, but on the other hand some of the limits may be not narrow enough with respect to this structure. The concept is used by Huber[3] and materially it is related to Walley's concept of "avoiding sure loss"[4]. For R-probability fields which do not possess the F-property one may use the expression "redundant R-probability fields".

F-probability may be interpreted as a perfect generalization of classical probability to an interval-valued one. The structure and the set of interval-limits imply each other. Huber[5] calls probability of this nature "representable", materially it corresponds to Walley's "coherent probability"[6].

Since the probabilist must expect to be confronted with a redundant R-probability, he should be prepared to "improve" such an assessment. There are two possible standpoints concerning his attitude towards a certain redundant R-probability field.

1. He may use the interval-limits to derive the structure of the R-probability field and pass over to the limits of that F-probability field, which is in accordance with this structure. In this way an F-probability field can uniquely be *derived* from every redundant R-probability without violating any of the interval-limits. This is called the *rigid standpoint*: It reduces the original interval-length for every redundant R-probability field.

2. It may – at least in some cases – be argued that, after adjustment to (9), any of the remaining limits should necessarily describe the outcome of the probability component $p(A)$ for at least one element $p(\cdot)$ of the structure. None of the values contained in such an interval therefore must be excluded: The structure has to be enlarged in order to include at least one K-probability $p(\cdot)$ for which $p(A) = L(A)$ is true and one $p'(\cdot)$ for which $p'(A) = U(A)$ holds. There is no unique way of enlarging the structure for this purpose, but there exist criteria to distinguish "minimum enlargements". If no information in favour of a certain kind of minimum enlargement is provided, the union of all F-probability fields produced in this way may be used. It is an F-probability field itself and is named the *F-cover* of the given redundant R-probability field. The standpoint producing this type of procedure may be called the *cautious standpoint*, because it leads to larger intervals and therefore weaker statements.

[3] [2], p. 257.
[4] [4], pp. 67-72, 135.
[5] [2], p. 255.
[6] [4], pp. 72-86, 135.

2.3 Partially Determinate Probability

Definition 2.6 Let Ω be a sample space and $\mathcal{A}$ be the σ-field of random events. Let furthermore be

$$\mathcal{A}' = \mathcal{A} \setminus \{\Omega, \emptyset\} \tag{14}$$

and

$$\mathcal{A}_L \subseteq \mathcal{A}',\ \mathcal{A}_U \subseteq \mathcal{A}'. \tag{15}$$

Then an assessment is called a *partially determinate R-probability* if (9) holds, and for each $A \in \mathcal{A}_L$ a lower limit $L(A)$ is given, as well as for each $A \in \mathcal{A}_U$ an upper limit $U(A)$, so that there exists a non-empty structure $\mathcal{M}$ of K-probabilities $p(\cdot)$, for which the following inequalities hold:

$$\left.\begin{array}{ll} L(A) \leq p(A), & \forall A \in \mathcal{A}_L \\ p(A) \leq U(A), & \forall A \in \mathcal{A}_U \end{array}\right\} \tag{16}$$

Definition 2.7 If for a partially determinate R-probability the conditions

$$\inf_{p \in \mathcal{M}} p(A) = L(A), \quad \forall A \in \mathcal{A}_L, \tag{17}$$

$$\sup_{p \in \mathcal{M}} p(A) = U(A), \quad \forall A \in \mathcal{A}_U, \tag{18}$$

are fulfilled, it is called a *partially determinate F-probability*.

For a partially determinate F-probability there exists a rather simple way of constructing the complete F-probability field: The use of (5) produces all originally lacking limits. This procedure is called *normal completion*.

Definition 2.8 Let $\mathcal{F} = (\Omega; \mathcal{A}; L(\cdot))$ be an F-probability field. Then $(\mathcal{A}_L, \mathcal{A}_U)$ is named a *support* of $\mathcal{F}$ if there exists a partially determinate F-probability according to (16) together with (17) and (18) which produces $\mathcal{F}$ via normal completion.

Interpretation of this concept is obvious: Information about limits $L(A), \forall A \in \mathcal{A}_L$, and $U(A), \forall A \in \mathcal{A}_U$, is sufficient for constructing $\mathcal{F}$.

3 Conditional Probability

3.1 General Remarks

Following Kolmogorov's procedure the system of axioms has to be completed by two definitions: the definition of conditional probability and the definition of independence. Defining conditional probability affords a series of considerations both on principle and of a technical kind and can be referred to here only in a highly abridged version.

The concepts of conditional probability are applied to F-probability fields only, since it may be assumed that redundant R-probability fields are transferred into F-probability fields by one of the ways described in Chapter 2.

Generally conditional probability affords the existence of a partition $\mathcal{C}$ of Ω:

$$\left.\begin{aligned} &\mathcal{C} = \{C_1, C_2, \ldots, C_r\} \\ &C_i \in \mathcal{A}', \; C_i \cap C_k = \emptyset, \; i \neq k \\ &\bigcup_{i=1}^{r} C_i = \Omega \end{aligned}\right\} \qquad (19)$$

It produces an assessment of conditional K-probability:

$$p_{\mathcal{C}}(A|C), \quad \forall A \in \mathcal{A}, \; \forall C \in \mathcal{C}. \qquad (20)$$

It should be stressed that (20) is applied to all conditioning events in $\mathcal{C}$, but not to conditioning events which belong to the field produced by $\mathcal{C}$ and not to $\mathcal{C}$ itself: With respect to K-probability this restriction serves to avoid paradoxical results.

3.2 The Intuitive Concept

Concerning the transfer to F-probability there is one first concept promising a simple solution. It is described by the following definitions:

Definition 3.1 Let $\mathcal{M}$ be the structure of the F-probability field $(\Omega; \mathcal{A}; L(\cdot))$ and $C \in \mathcal{A}$. Then

$$\mathcal{M}_C := \{p(\cdot) | p(\cdot) \in \mathcal{M} : p(C) > 0\} \qquad (21)$$

is called the *C-docked structure.*

Definition 3.2 Under the requirements of Definition 3.1

$$iL_{\mathcal{C}}(A|C) := \inf_{p \in \mathcal{M}_C} \frac{p(A \cap C)}{p(C)}, \qquad \forall A \in \mathcal{A}, \; \forall C \in \mathcal{C}, \qquad (22)$$

together with

$$\begin{aligned} iU_{\mathcal{C}}(A|C) &:= 1 - iL_{\mathcal{C}}(\neg A|C) = 1 - \inf_{p \in \mathcal{M}_C} \frac{p(\neg A \cap C)}{p(C)} \\ &= \sup_{p \in \mathcal{M}_C} \frac{p(A \cap C)}{p(C)} \end{aligned} \qquad (23)$$

produces the *intuitive concept of conditional probability.*

This concept has a considerable number of pleasant properties. It is easy to understand and easy to use. Nevertheless it is of limited interest, because it does not allow the reconstruction of the F-probability field from which it is gained.

This may be demonstrated by the following two examples.

Example 3.1 The following assessment produces an F-probability field on the sample-space $\Omega = E_1 \cup E_2 \cup E_3$:

$$\begin{array}{ll} P(E_1) = [0.10; 0.25] & P(E_1 \cup E_2) = [0.40; 0.60] \\ P(E_2) = [0.20; 0.40] & P(E_1 \cup E_3) = [0.60; 0.80] \\ P(E_3) = [0.40; 0.60] & P(E_2 \cup E_3) = [0.75; 0.90]. \end{array}$$

A partition $\mathcal{C}$ of Ω be given by

$$\mathcal{C} = \{C_1, C_2\}, \; C_1 = E_1 \cup E_2, \; C_2 = E_3$$

$$\Rightarrow \; P(C_1) = [0.40; 0.60] \qquad P(C_2) = [0.40; 0.60].$$

Application of (22) leads to:

$$iL_{\mathcal{C}}(E_1|C_1) = \inf_{p \in \mathcal{M}} \frac{p(E_1)}{p(C_1)} = \frac{0.10}{0.10 + 0.40} = 0.20$$

$$iL_{\mathcal{C}}(E_2|C_1) = \inf_{p \in \mathcal{M}} \frac{p(E_2)}{p(C_1)} = \frac{0.20}{0.20 + 0.25} = 0.4\dot{4}.$$

Therefore the intuitive concept generates the following assessment:

$$\begin{array}{ll} iP_{\mathcal{C}}(E_1|C_1) = [0.20; 0.5\dot{5}] & iP_{\mathcal{C}}(E_1|C_2) = [0] \\ iP_{\mathcal{C}}(E_2|C_1) = [0.4\dot{4}; 0.80] & iP_{\mathcal{C}}(E_2|C_2) = [0] \\ iP_{\mathcal{C}}(E_3|C_1) = [0] & iP_{\mathcal{C}}(E_3|C_2) = [1]. \end{array}$$

Example 3.2 Another F-probability field:

$$\begin{array}{ll} P(E_1) = [0.11; 0.225] & P(E_1 \cup E_2) = [0.40; 0.60] \\ P(E_2) = [0.18; 0.44] & P(E_1 \cup E_3) = [0.56; 0.82] \\ P(E_3) = [0.40; 0.60] & P(E_2 \cup E_3) = [0.775; 0.89] \end{array}$$

with the same partition as in Example 3.1 and the same marginal probability produces:

$$iL_{\mathcal{C}}(E_1|C_1) = \frac{0.11}{0.11 + 0.44} = 0.20$$

$$iL_{\mathcal{C}}(E_2|C_1) = \frac{0.18}{0.225 + 0.18} = 0.4\dot{4}$$

$$\Rightarrow \; \begin{array}{ll} iP_{\mathcal{C}}(E_1|C_1) = [0.20; 0.5\dot{5}] & iP_{\mathcal{C}}(E_1|C_2) = [0] \\ iP_{\mathcal{C}}(E_2|C_1) = [0.4\dot{4}; 0.80] & iP_{\mathcal{C}}(E_2|C_2) = [0] \\ iP_{\mathcal{C}}(E_3|C_1) = [0] & iP_{\mathcal{C}}(E_3|C_2) = [1]. \end{array}$$

As examples 3.1 and 3.2 demonstrate, it is possible that two different F-probability fields lead to the same marginal probability and to the same conditional probability, if the intuitive concept is used. For that reason it is impossible to reconstruct the given F-probability field from marginal probability together with conditional probability. Therefore conditional probability according to this concept can never contain the type of information one wants to transfer from one model to the other.

This failure rules out the use of the intuitive concept as the only concept of conditional probability. Nevertheless it remains useful as a means of describing and characterizing the phenomenon of conditional probability.

As far as analysis is concerned one is led to the canonical concept.

3.3 The Canonical Concept

Definition 3.3 The subfields with respect to elements of the partition $\mathcal{C}$ are denoted as $\mathcal{A}(C)$:

$$\mathcal{A}(C) := \{C \cap A \,|\, A \in \mathcal{A}\}, \quad C \in \mathcal{C}. \tag{24}$$

Definition 3.4 An F-probability field $\mathcal{F} = (\Omega; \mathcal{A}; L(\cdot))$ together with a partition $\mathcal{C}$ of Ω is called a *laminar constellation* $(\mathcal{F}, \mathcal{C})$ if a support $(\mathcal{A}_L, \mathcal{A}_U)$ exists such that for all $A \in \mathcal{A}_L \cup \mathcal{A}_U$ the following holds:

$$A \in \bigcup_{C \in \mathcal{C}} \mathcal{A}(C). \tag{25}$$

The definition of laminar constellation distinguishes constellations in which all information about interval-limits is given by the assessment $P(A)$ for those random events A which are contained in one single element of the partition. The reason for this definition is the following: Information with respect to a random event which does not obey (25), can be contained neither in marginal probability nor in conditional probability.

In this article the construction of conditional probability for a laminar constellation is described for the case only, that the conditions

$$L(C) > 0, \quad \forall C \in \mathcal{C}, \tag{26}$$

are fulfilled. The concept then requires the calculation of

$$L_{\mathcal{C}}(A|C) := \frac{L(A)}{L(C)}, \quad \forall A \in \mathcal{A}(C),\ \forall C \in \mathcal{C}, \tag{27}$$

$$U_{\mathcal{C}}(A|C) := \frac{U(A)}{U(C)}, \quad \forall A \in \mathcal{A}(C),\ \forall C \in \mathcal{C}. \tag{28}$$

Concerning this assessment a decisive distinction has to be made.

Definition 3.5 If for all $C \in \mathcal{C}$ the assessment described by (27) and (28) constitutes an F-probability field[7]

$$\mathcal{F}_C := (C; \mathcal{A}(C), L_{\mathcal{C}}(\cdot|C)), \tag{29}$$

then $(\mathcal{F}, \mathcal{C})$ is called an *F-laminar constellation*. In this situation each $\mathcal{F}_C$ represents the *conditional F-probability* with respect to the condition C.

Knowledge of conditional probability and of marginal probability allows reconstruction of the given F-probability field. (27) and (28) are converted to

$$L(A) = L_{\mathcal{C}}(A|C) \cdot L(C), \quad \forall A \in \mathcal{A}(C),\ \forall C \in \mathcal{C}, \tag{30}$$

$$U(A) = U_{\mathcal{C}}(A|C) \cdot U(C), \quad \forall A \in \mathcal{A}(C),\ \forall C \in \mathcal{C}, \tag{31}$$

and because of laminarity all the rest of the F-probability field may be reconstructed by use of the limits defined by (30) and (31). Furthermore it may be shown that combination of conditional probability according to (30) and (31) with any marginal F-probability produces an F-probability field. Therefore (30) together with (31) may be used in order to transfer information to any comparable model.

Example 3.3 The following assessment produces an F-probability field:

$$\begin{array}{ll} P(E_1) = [0.10; 0.30] & P(E_1 \cup E_2) = [0.40; 0.60] \\ P(E_2) = [0.20; 0.45] & P(E_1 \cup E_3) = [0.55; 0.80] \\ P(E_3) = [0.40; 0.60] & P(E_2 \cup E_3) = [0.70; 0.90]. \end{array}$$

With regard to the partition

$$\mathcal{C} = \{C_1, C_2\},\ C_1 = E_1 \cup E_2,\ C_2 = E_3$$

$$\Rightarrow P(C_1) = [0.40; 0.60] \qquad P(C_2) = [0.40; 0.60],$$

the effect of combining in a new model the conditional probability with another marginal probability, namely

$$P'(C_1) = [0.60; 0.80] \qquad P'(C_2) = [0.20; 0.40]$$

is to be determined.

Conditional probability according to (27) and (28):

$$L_{\mathcal{C}}(E_1|C_1) = \frac{L(E_1)}{L(C_1)} = \frac{0.10}{0.40} = 0.25$$

$$L_{\mathcal{C}}(E_2|C_1) = \frac{L(E_2)}{L(C_1)} = \frac{0.20}{0.40} = 0.50$$

$$L_{\mathcal{C}}(E_3|C_2) = \frac{L(E_3)}{L(C_2)} = \frac{0.40}{0.40} = 1$$

[7] which means among others that (6) must hold for $L_{\mathcal{C}}(\cdot\,|\,C)$ and $U_{\mathcal{C}}(\cdot\,|\,C)$.

$$U_{\mathcal{C}}(E_1|C_1) = \frac{U(E_1)}{U(C_1)} = \frac{0.30}{0.60} = 0.50$$
$$U_{\mathcal{C}}(E_2|C_1) = \frac{U(E_2)}{U(C_1)} = \frac{0.45}{0.60} = 0.75$$
$$U_{\mathcal{C}}(E_3|C_2) = \frac{U(E_3)}{U(C_2)} = \frac{0.60}{0.60} = 1$$

$$\Rightarrow \begin{array}{ll} P_{\mathcal{C}}(E_1|C_1) = [0.25; 0.50] & P_{\mathcal{C}}(E_1|C_2) = [0] \\ P_{\mathcal{C}}(E_2|C_1) = [0.50; 0.75] & P_{\mathcal{C}}(E_2|C_2) = [0] \\ P_{\mathcal{C}}(E_3|C_1) = [0] & P_{\mathcal{C}}(E_3|C_2) = [1] \end{array}$$

proves to constitute an F-probability field as well for C_1 as for C_2; therefore both assessments represent conditional F-probability. The transfer of $P_{\mathcal{C}}(\cdot|C_1)$ to the alternative marginal probability produces by analogy to (30) and (31):

$$L'(E_1) = 0.25 \cdot 0.60 = 0.15$$
$$L'(E_2) = 0.50 \cdot 0.60 = 0.30$$

$$U'(E_1) = 0.50 \cdot 0.80 = 0.40$$
$$U'(E_2) = 0.75 \cdot 0.80 = 0.60$$

$$\Rightarrow \begin{array}{ll} P'(E_1) = [0.15; 0.40] & P'(E_1 \cup E_2) = [0.60; 0.80] \\ P'(E_2) = [0.30; 0.60] & P'(E_1 \cup E_3) = [0.40; 0.70] \\ P'(E_3) = [0.20; 0.40] & P'(E_2 \cup E_3) = [0.60; 0.85]. \end{array}$$

That this assessment represents an F-probability field, is easily proven by the fact that $p_1(\cdot)$ with

$$p_1(E_1) = 0.15,\ p_1(E_2) = 0.60,\ p_1(E_3) = 0.25$$

reaches $L'(E_1)$, $U'(E_2)$, $L'(E_1 \cup E_3)$, $U'(E_2 \cup E_3)$, $p_2(\cdot)$ with

$$p_2(E_1) = 0.40,\ p_2(E_2) = 0.40,\ p_2(E_3) = 0.20$$

reaches $U'(E_1)$, $L'(E_3)$, $U'(E_1 \cup E_2)$, $L'(E_2 \cup E_3)$ and $p_3(\cdot)$ with

$$p_3(E_1) = 0.30,\ p_3(E_2) = 0.30,\ p_3(E_3) = 0.40$$

reaches $L'(E_2)$, $U'(E_3)$, $L'(E_1 \cup E_2)$, $U'(E_1 \cup E_3)$. Therefore Axioms I-VI are fulfilled.

Definition 3.6 If there exists $C \in \mathcal{C}$, so that (27) and (28) violate at least one of the Axioms IV and V, then $(\mathcal{F}, \mathcal{C})$ is called a *0-laminar constellation.*

In this case application of the concept of conditional probability according to the canonical concept is not possible.

Example 3.4 For the F-probability field

$$\begin{array}{ll} P(E_1) = [0.16; 0.21] & P(E_1 \cup E_2) = [0.40; 0.60] \\ P(E_2) = [0.22; 0.42] & P(E_1 \cup E_3) = [0.58; 0.78] \\ P(E_3) = [0.40; 0.60] & P(E_2 \cup E_3) = [0.79; 0.84] \end{array}$$

and the partition

$$\mathcal{C} = \{C_1, C_2\},\ C_1 = E_1 \cup E_2,\ C_2 = E_3,$$

the result

$$L_{\mathcal{C}}(E_1|C_1) = \frac{0.16}{0.40} = 0.40$$
$$U_{\mathcal{C}}(E_1|C_1) = \frac{0.21}{0.60} = 0.35$$

violates Axiom IV: The relative length of the interval $P(E_1)$ is too small compared with $P(C_1)$. Consequently $P(E_1)$ can never be produced by a procedure of the type defined by (30) and (31).

Between F-laminarity and 0-laminarity lies what is called *R-laminarity.*

Definition 3.7 If for all $C \in \mathcal{C}$ the assessment created by Equations (27) and (28) constitutes R-probability, then $(\mathcal{F}, \mathcal{C})$ is called an *R-laminar constellation.*

R-laminar constellation which is not F-laminar may be described as *redundant R-laminar.* While the reconstruction of the original F-probability field can be achieved through (30) and (31) also in these situations, the proper way of transferring information contained in (27) and (28) to another model requires a bunch of considerations and decisions far above the scope of this article.

Example 3.5 The constellation described in Example 3.1 proves to be redundant R-laminar.

$$L_{\mathcal{C}}(E_1|C_1) = \frac{0.10}{0.40} = 0.25$$
$$L_{\mathcal{C}}(E_2|C_1) = \frac{0.20}{0.40} = 0.50$$
$$U_{\mathcal{C}}(E_1|C_1) = \frac{0.25}{0.60} = 0.41\dot{6}$$
$$U_{\mathcal{C}}(E_2|C_1) = \frac{0.40}{0.60} = 0.6\dot{6}$$

defines a redundant R-probability field: $p(E_1|C_1) = 0.40$, $p(E_2|C_1) = 0.60$ is an element of the structure, therefore $P(\cdot|C_1)$ produces an R-field. Since $p(E_1|C_1) = L_{\mathcal{C}}(E_1|C_1) = 0.25$ is not possible, this R-probability field is redundant.

3.4 Bayes' Rule

If both concepts of conditional probability are employed in their specific roles, the Bayes' rule for interval-probability can be derived. It is reported here without proof.

Let $\mathcal{F} = (\Omega; \mathcal{A}; L(.))$ be an F-Field and $\mathcal{C}$ a partition of Ω, so that $(\mathcal{F}, \mathcal{C})$ is an F-laminar constellation. Then the following information allows the reconstruction of $\mathcal{F}$:

$P(+\cap D)$	=	$[0.12; 0.32]$	$P(-\cap D)$	=	$[0.04; 0.16]$	$P(D)$	=	$[0.2; 0.4]$
$P(+\cap S)$	=	$[0.12; 0.24]$	$P(-\cap S)$	=	$[0.42; 0.64]$	$P(S)$	=	$[0.6; 0.8]$
$P(+)$	=	$[0.24; 0.54]$	$P(-)$	=	$[0.46; 0.76]$		$[1]$	

Table 1: The F-field of Example 3.6; components relevant for the intuitive conditional probability with respect to the outcome of the test.

1. $F_{\mathcal{C}} = (\mathcal{C}; \mathcal{P}(\mathcal{C}); L(.))$ is the marginal F-field with respect to the partition $\mathcal{C}$: the "prior probability".

2. $\{\mathcal{F}_C = (C; \mathcal{A}(C); L_{\mathcal{C}}(.\,|\,C)) \,|\, C \in \mathcal{C}\}$ is the set of conditional F-probability fields with respect to the canonical concept.

One should remember that because of laminarity each of the — originally not known — true interval-limits $L(.)$ und $U(.)$ of the field $\mathcal{F}$ is produced either directly via

$$P(A) = [L(A\,|\,C)\cdot L(C);\; U(A\,|\,C)\cdot U(C)] \qquad (32)$$

if $A \subseteq C$, or through normal completion and so is the structure $\mathcal{M}$. For each $B \in \mathcal{A}$ the intuitive concept of conditional probability creates an F-field

$$iP(A\,|\,B) = \left[\inf_{p\in\mathcal{M}} p(A\,|\,B);\; \sup_{p\in\mathcal{M}} p(A\,|\,B)\right]. \qquad (33)$$

By definition each component $iP(A\,|\,B)$ represents the set of all posterior K-probabilities calculated for elements of the structure $\mathcal{M}$. As a consequence $iP(A\,|\,B)$ as posterior F-probability given B possesses the properties of classical posterior probability.

Example 3.6 The (prior) probability of having a certain disease D is known to be between 0.2 and 0.4. A test producing either positive or negative results can be characterized by (D: disease, S: soundness)

$$P(+\,|\,D) = [0.6; 0.8] \qquad P(-\,|\,D) = [0.2; 0.4]$$
$$P(+\,|\,S) = [0.2; 0.3] \qquad P(-\,|\,S) = [0.7; 0.8].$$

If this information is understood to be conditional probability due to the canonical concept, application of (30) and (31) and normal completion produce the F-field $\mathcal{F}$, the relevant components of which — support and marginal probabilities — are given in Table 1.

Intuitive conditional probabilities with respect to the outcome of the test can be calculated as

$$iP(D\,|\,+) = [0.33; 0.73] \qquad iP(D\,|\,-) = [0.06; 0.28]$$
$$iP(S\,|\,+) = [0.27; 0.67] \qquad iP(S\,|\,-) = [0.72; 0.94].$$

Dependent upon the outcome of the test, the posterior F-probability may be used as prior probability for another test.

4 Further Aspects

The second complement of the system of axioms is produced by the definition of independence. It is reported here in an abridged version, since it materially produces the concept already used by P. Walley and T. Fine in 1982.[8] Let for a sample space of four elements,

$$\Omega = E_{11} \cup E_{12} \cup E_{21} \cup E_{22}, \qquad (34)$$

two partitions consisting of dichotomies be given:

$$\mathcal{C}_A = \{C_{1\cdot}, C_{2\cdot}\} \qquad \mathcal{C}_B = \{C_{\cdot 1}, C_{\cdot 2}\} \qquad (35)$$

with

$$\begin{array}{ll} C_{1\cdot} = E_{11} \cup E_{12} & C_{2\cdot} = E_{21} \cup E_{22} \\ C_{\cdot 1} = E_{11} \cup E_{21} & C_{\cdot 2} = E_{12} \cup E_{22}. \end{array} \qquad (36)$$

This can be represented in a four-fold table:

E_{11}	E_{12}	$C_{1\cdot}$
E_{21}	E_{22}	$C_{2\cdot}$
$C_{\cdot 1}$	$C_{\cdot 2}$	Ω

Definition 4.1 A partially determinate F-probability on $(\Omega; \mathcal{P}(\Omega))$ according to (34) is named *marginal probability on the four-fold table* if

$$\mathcal{A}_L = \mathcal{A}_U = \{C_{1\cdot}, C_{\cdot 1}\}. \qquad (37)$$

Let

$$P(C_{1\cdot}) := [L_1; U_1] \qquad P(C_{2\cdot}) = [1-U_1; 1-L_1]$$
$$P(C_{\cdot 1}) := [L_2; U_2] \qquad P(C_{\cdot 2}) = [1-U_2; 1-L_2].$$

[8] [5], p. 745.

Normal completion produces the following lower interval-limits:

$$
\begin{aligned}
&L(E_{11}) = \text{Max}(0, L_1 + L_2 - 1)\\
&L(E_{12}) = \text{Max}(0, L_1 - U_2)\\
&L(E_{21}) = \text{Max}(0, L_2 - U_1)\\
&L(E_{22}) = \text{Max}(0, 1 - U_1 - U_2)\\
&L(E_{11} \cup E_{12}) = L_1\\
&L(E_{11} \cup E_{21}) = L_2\\
&L(E_{11} \cup E_{22}) = \text{Max}(0, 1 - U_1 - U_2, L_1 + L_2 - 1)\\
&L(E_{12} \cup E_{21}) = \text{Max}(0, L_2 - U_1, L_1 - U_2)\\
&L(E_{12} \cup E_{22}) = 1 - U_2\\
&L(E_{21} \cup E_{22}) = 1 - U_1\\
&L(E_{11} \cup E_{12} \cup E_{21}) = \text{Max}(L_1, L_2)\\
&L(E_{11} \cup E_{12} \cup E_{22}) = \text{Max}(L_1, 1 - U_2)\\
&L(E_{11} \cup E_{21} \cup E_{22}) = \text{Max}(1 - U_1, L_2)\\
&L(E_{12} \cup E_{21} \cup E_{22}) = \text{Max}(1 - U_1, 1 - U_2)
\end{aligned} \tag{38}
$$

and the set of conjugate upper interval-limits as defined by (6).

The structure of this F-probability field is denominated by $\mathcal{M}_M$.

Using the concept of prestructure one may define independence of the two partitions as a property of a certain F-probability field on $(\Omega; \mathcal{P}(\Omega))$ conforming to the marginal probability.

Definition 4.2 If $\mathcal{F} = (\Omega; \mathcal{A}; L(\cdot))$ is an F-probability on the sample-space (34) and $\mathcal{M}_M$ is the structure of the marginal F-probability according to Definition 4.1, then the partitions $\mathcal{C}_A$ and $\mathcal{C}_B$ are *mutually independent*, provided that the set

$$\mathcal{M}_I := \{p(.) \in \mathcal{M}_M | p(E_{ij}) = p(C_{i\cdot})p(C_{\cdot j});\ i, j = 1, 2\} \tag{39}$$

serves as a prestructure of $\mathcal{F}$.

This definition requires that for the F-field $\mathcal{F}$ with independence of $\mathcal{C}_A$ and $\mathcal{C}_B$ all interval-limits have to be just wide enough to include all K-functions which qualify for $\mathcal{M}_I$ by

1. as well being in accordance with the given marginal probability (i.e. being elements of $\mathcal{M}_M$) as
2. obeying the classical rule of independence.

The lower interval-limits defined by this prestructure are the following:

$$
\begin{aligned}
&L(E_{11}) = L_1 L_2\\
&L(E_{12}) = L_1(1 - U_2)\\
&L(E_{21}) = (1 - U_1)L_2\\
&L(E_{22}) = (1 - U_1)(1 - U_2)\\
&L(E_{11} \cup E_{12}) = L_1\\
&L(E_{11} \cup E_{21}) = L_2\\
&L(E_{11} \cup E_{22}) = \text{Min}[U_1U_2 + (1 - U_1)(1 - U_2),\\
&\qquad U_1L_2 + (1 - U_1)(1 - L_2),\\
&\qquad L_1U_2 + (1 - L_1)(1 - U_2),\\
&\qquad L_1L_2 + (1 - L_1)(1 - L_2)]\\
&L(E_{12} \cup E_{21}) = \text{Min}[L_1(1 - L_2) + L_2(1 - L_1),\\
&\qquad L_1(1 - U_2) + U_2(1 - L_1),\\
&\qquad U_1(1 - L_2) + L_2(1 - U_1),\\
&\qquad U_1(1 - U_2) + U_2(1 - U_1)]\\
&L(E_{12} \cup E_{22}) = 1 - U_2\\
&L(E_{21} \cup E_{22}) = 1 - U_1\\
&L(E_{11} \cup E_{12} \cup E_{21}) = 1 - (1 - L_1)(1 - L_2)\\
&L(E_{11} \cup E_{12} \cup E_{22}) = 1 - (1 - L_1)U_2\\
&L(E_{11} \cup E_{21} \cup E_{22}) = 1 - U_1(1 - L_2)\\
&L(E_{12} \cup E_{21} \cup E_{22}) = 1 - U_1U_2.
\end{aligned} \tag{40}
$$

Again the corresponding upper interval-limits are given by equation (6).

It must be remembered that $\mathcal{M}_I$ is a prestructure, defining the interval-limits, but in most cases is not the total structure $\mathcal{M}$ of the F-field $\mathcal{F}$. If $\mathcal{M}_I$ contains more than one K-function, $\mathcal{M} \setminus \mathcal{M}_I$ is not empty. Therefore $\mathcal{M}$ includes elements in accordance with (40) but violating the classical multiplication rule for independent K-probabilities: If $\mathcal{F}$ deviates from classical probability the interval-limits tolerate K-functions with slight dependence of $\mathcal{C}_A$ and $\mathcal{C}_B$.

On the other hand all kinds of deviation from the limits given by (40) would violate the concept of independence: either the interval-limits would exclude elements of $\mathcal{M}_I$ or they would include too many dependent K-functions.

Example 4.1 Marginal probability on a four-fold-table is determinate by the following assessment:

$$
\begin{aligned}
&P(C_{1\cdot}) = [0.3; 0.5] \qquad P(C_{2\cdot}) = [0.5; 0.7]\\
&P(C_{\cdot 1}) = [0.2; 0.4] \qquad P(C_{\cdot 2}) = [0.6; 0.8]
\end{aligned}
$$

With normal completion the marginal F-probability field is derived:

$$\begin{aligned}P(E_{11}) &= [0.0; 0.4]\\ P(E_{12}) &= [0.0; 0.5]\\ P(E_{21}) &= [0.0; 0.4]\\ P(E_{22}) &= [0.1; 0.7]\end{aligned}$$

$$\begin{aligned}P(E_{11} \cup E_{12}) &= [0.3; 0.5]\\ P(E_{11} \cup E_{21}) &= [0.2; 0.4]\\ P(E_{11} \cup E_{22}) &= [0.1; 1]\\ P(E_{12} \cup E_{21}) &= [0.0; 0.9]\\ P(E_{12} \cup E_{22}) &= [0.6; 0.8]\\ P(E_{21} \cup E_{22}) &= [0.5; 0.7]\end{aligned}$$

$$\begin{aligned}P(E_{11} \cup E_{12} \cup E_{21}) &= [0.3; 0.9]\\ P(E_{11} \cup E_{12} \cup E_{22}) &= [0.6; 1]\\ P(E_{11} \cup E_{21} \cup E_{22}) &= [0.5; 1]\\ P(E_{12} \cup E_{21} \cup E_{22}) &= [0.6; 1].\end{aligned}$$

According to (40) and (6) partitions $\mathcal{C}_A$ and $\mathcal{C}_B$ are independent iff

$$\begin{aligned}P(E_{11}) &= [0.06; 0.20]\\ P(E_{12}) &= [0.18; 0.40]\\ P(E_{21}) &= [0.10; 0.28]\\ P(E_{22}) &= [0.30; 0.56]\end{aligned}$$

$$\begin{aligned}P(E_{11} \cup E_{12}) &= [0.30; 0.50]\\ P(E_{11} \cup E_{21}) &= [0.20; 0.40]\\ P(E_{11} \cup E_{22}) &= [0.50; 0.62]\\ P(E_{12} \cup E_{21}) &= [0.38; 0.50]\\ P(E_{12} \cup E_{22}) &= [0.60; 0.80]\\ P(E_{21} \cup E_{22}) &= [0.50; 0.70]\end{aligned}$$

$$\begin{aligned}P(E_{11} \cup E_{12} \cup E_{21}) &= [0.44; 0.70]\\ P(E_{11} \cup E_{12} \cup E_{22}) &= [0.72; 0.90]\\ P(E_{11} \cup E_{21} \cup E_{22}) &= [0.60; 0.82]\\ P(E_{12} \cup E_{21} \cup E_{22}) &= [0.80; 0.94]\end{aligned}$$

holds.

The comparison of marginal probability and independent probability shows that a remarkable sharpening of the intervals is caused by independence.

For instance:

$$p_1(.) \text{ with } \quad \begin{aligned}&p_1(E_{11}) = 0.00,\ p_1(E_{12}) = 0.40,\\ &p_1(E_{21}) = 0.30,\ p_1(E_{22}) = 0.30\end{aligned}$$

is an element of $\mathcal{M}_M$, but not of $\mathcal{M}$, while

$$p_2(.) \text{ with } \quad \begin{aligned}&p_2(E_{11}) = 0.12,\ p_2(E_{12}) = 0.28,\\ &p_2(E_{21}) = 0.18,\ p_2(E_{22}) = 0.42\end{aligned}$$

represents an independent K-probability in $\mathcal{M}_M$, therefore being an element of $\mathcal{M}_I$ and consequently of $\mathcal{M}$. On the other hand

$$p_3(.) \text{ with } \quad \begin{aligned}&p_3(E_{11}) = 0.10,\ p_3(E_{12}) = 0.20,\\ &p_3(E_{21}) = 0.20,\ p_3(E_{22}) = 0.50\end{aligned}$$

shows no independence between lines and columns, consequently being not an element of $\mathcal{M}_I$, but belonging to $\mathcal{M}$, since $p_3(.)$ is in accordance with all of the limits (40).

In the theory of interval-probability the concept of mutually independent partitions in an F-field among other aspects provides the fundamentals for a *Weak Law of Large Numbers*.

At first it serves to define independently identically F-distributed (i.i.F-d.) samples. This is demonstrated by means of a simple model sufficient for the purpose of studying relative frequencies.

Definition 4.3 Let $\mathcal{F}_n = (\Omega^n; \mathcal{P}(\Omega^n); L_n(.))$ be an F-field with $\Omega^n = \times_{i=1}^n \Omega_i$, $\Omega_i = E_{i,1} \cup E_{i,2}$, $i = 1, \ldots, n$. Partitions $\mathcal{C}_i$ are given by $\mathcal{C}_i = \{C_{i,1}, C_{i,2}\}$, $i = 1, \ldots, n$, with

$$C_{i,r} = \Omega_1 \times \ldots \times \Omega_{i-1} \times E_{i,r} \times \Omega_{i+1} \times \ldots \times \Omega_n, \quad (41)$$

$r = 1, 2$.

Then $\mathcal{F}_n$ describes an i.i.F-d. sample of size n, provided that the marginal probabilities are:

1. $\left.\begin{aligned}P(C_{i,1}) &= [L; U]\\ P(C_{i,2}) &= [1-U; 1-L]\end{aligned}\right\} i = 1, \ldots, n, \quad (42)$
2. $\mathcal{C}_i$ and $\mathcal{C}_{i'}$ are mutually independent, $i, i' = 1, \ldots, n,\ i \neq i'$.

The relative frequency of $E_{.1}$ is defined by the $\mathcal{F}_n$-random variable

$$T^{(n)} = \frac{1}{n}\sum_{i=1}^{n} T_i, \text{ where } T_i(E_{i,1}) = 1,\ T_i(E_{i,2}) = 0. \quad (43)$$

In order to arrive at a Weak Law of Large Numbers the concept of convergence in F-probability has to be introduced.

Definition 4.4 With respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$ let $(X_n)_{n\in\mathbb{N}}$ be $\mathcal{F}_n$-random variables in $\mathbb{R}^1$. For $-\infty < \alpha \leq \beta < +\infty$ the random event $A_n(\epsilon)$ is defined by:

$$A_n(\epsilon) = \bigcup\{E \subseteq \Omega_n \mid \alpha - \epsilon \leq X_n(E) \leq \beta + \epsilon\}. \quad (44)$$

Then $(X_n)_{n\in\mathbb{N}}$ is *convergent in F-probability into* $[\alpha, \beta]$ iff for every (ϵ, δ), $\epsilon > 0$, $0 < \delta < 1$, there exists a $N(\epsilon, \delta)$, such that for all $n \geq N(\epsilon, s)$

$$L_n(A_N(\epsilon)) \geq 1 - \delta \quad (45)$$

holds.

Using the concepts of Definitions 4.3 and 4.4 the following statement can be proven:

If for all $n \in \mathbb{N}$ the F-field $\mathcal{F}_n$ describes an i.i.F-d. sample of size n with marginal probabilities given by (42), the $\mathcal{F}_n$-random variable $T^{(n)}$ defined by (43) is convergent in F-probability into $[L; U]$.

One proof can be derived from [5][9].

This generalization of Bernoulli's theorem allows a frequency interpretation of interval-probability: For a long i.i.F-d. sample with marginal F-probability $P(\text{"success"}) = [L; U]$ the relative frequency of successes at last almost surely will be found in $[L; U]$. If $L < U$ it is not possible to know in which part of $[L; U]$ the relative frequency will be found and whether the sequence will be convergent in the classical sense.

Despite all differences between the schools concerning the meaning of probability assessments, it might be useful for everybody to consider these logical implications of an assignment containing interval-probability.

A comprehensive study of the theory is in progress, the first volume will be published in 1999 [10].

An important aspect of the theory not mentioned in the present article is the use of Linear Optimization to solve fundamental problems on finite sample spaces. Some results of this type are already reported in [8]. The concept of uniform F-probability and its use in describing sampling is briefly described in [7] — together with consequences concerning an improvement of the principle of insufficient reason.

Among those aspects not mentioned in the foregoing sections is that of decision theory. A general approach to decision problems with respect to behavioural viewpoints is made possible by the theory. Ellsberg's results and their consequences can be respected. Behaviour under ambiguity can be analyzed and classified. A preliminary report is found in [9]. One of the many problems concerning statistical methodology under interval-probability has yet been studied thoroughly: testing statistical hypotheses. Fundamental results are given in [1].

Altogether the unifying concept for uncertainty contained in the theory of interval-probability produces a great number of aspects which deserve intensive research and will create many chances for methodological improvements.

Acknowledgements

I like to thank A. Wallner and Th. Augustin for their valuable cooperation.

[9] [5], Theorem 4.1, pp. 747-748.

References

[1] T. Augustin. Optimale Tests bei Intervallwahrscheinlichkeit. Vandenhoeck & Ruprecht, Göttingen, 1998.

[2] P.J. Huber. Robust Statistics. Wiley. New York et al, 1981.

[3] E. Kofler and G. Menges. Entscheidungen bei unvollständiger Information, Lecture Notes, *Economics and Mathematical Systems*, 136, Springer, Berlin, Heidelberg, New York, 1976.

[4] P. Walley. Statistical Reasoning with Imprecise Probabilities. Chapman and Hall. London, New York, 1991.

[5] P. Walley and T.L. Fine. Towards a frequentist theory of upper and lower probability, *AMS*, 10, 741-761, 1982.

[6] K. Weichselberger. Axiomatic foundations of the theory of interval-probability. In: V. Mammitzsch, H. Schneeweiß (eds.): *Proceedings of the Second Gauß Symposion, Section B.* De Gruyter, Berlin; 47-64, 1995A.

[7] K. Weichselberger. Stichproben und Intervallwahrscheinlichkeit, *ifo Studien* 41; 653-676, 1995B.

[8] K. Weichselberger. Interval-probability on finite sample-spaces. In: H. Rieder (ed.): *Robust Statistics, Data Analysis and Computer Intensive Methods.* Springer, New York (Lecture Notes in Statistics, 109); 391-409, 1996.

[9] K. Weichselberger, T. Augustin. Analysing Ellsberg's paradox by means of interval-probability. In: R. Galata, H. Küchenhoff (eds.): *Econometrics in Theory and Practice.* Physika-Verlag, Heidelberg, 1998.

[10] K. Weichselberger. Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung. Bd. 1. In preparation, 1999.