The Logical Concept of Probability and Statistical Inference

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

Abstract

A consistent concept of logical probability affords the employment of interval probability. Such a concept which attributes probability to arguments consisting of premise and conclusion, can be used to generate a system of axioms for statistical inference.

Keywords. The theory of interval probability, logical probability, statistical inference, inductive probability.

1 Introduction

Traditionally the two most important of the fundamental problems in statistics are the question about the nature of probability and an appropriate theory of statistical inference. The close connection between both problems is evident: The subjectivist — or the objective Bayesian — accepts the idea of ever available knowledge providing prior probability, and relies on Bayes' theorem in order to learn from experience. If, however, the concept of trustworthy prior probability is not accepted, statistical inference is restricted to methods, which do not exhaust the total information contained in a given sample.

A logical concept of probability several times in history was seen as a possible solution of both problems, but never a comprehensive mathematical treatment of probability as a two-place function succeeded. The reason is obvious: If probability consistently is seen as a function of arguments, the dependence upon the premise and the dependence upon the conclusion have to be described. In classical probability theory, however, no tool can be found allowing for an addition formula with respect to unions of premises.

The situation becomes different, if the theory of interval probability can be relied on: In this theory the concept of F-probability plays a central role, and the union of F-probability-fields on a measurable space always constitutes an F-probability-field on this space. Therefore by means of the theory of interval probability a consistent logical concept of probability can be created (see also [3]). Since this concept is purely epistemic, it is acceptable from the subjectivistic view on probability as well as from the objectivistic one. According to this concept every probability statement is related to an argument, consisting of a premise and of a conclusion; probabilities of events — if existing at all — are not subject of the theory.

The radical switch from probability as a one-place-function to probability as a two-place-function produces many insights. It might be regarded the most important gain, that statistical inference can be conceived as an integral branch of probability theory — without engaging any prior probability. Based on fundamental ideas of statistical reasoning — which are to be found in the concepts of significance and confidence — a duality of logical probability-fields can be established by axiomatic method, allowing for probability statements of statistical inference in clearly defined situations. Since a frequency interpretation of all logical probability statements is available, the conceptual quality is the same for primal and for dual probability-fields.

While the fundamental ideas of the resulting Symmetric Theory — together with a few prominent examples of application — are represented in this article, the main task of developing the methodology and exhibiting its boundaries remains to be done. Furthermore the paper does not contain the necessary comparison of the Symmetric Theory with preceding logical theories of probability from J.M. Keynes to H. Kyburg and I. Levi. It will also be fruitful to review Bayesian and Non-Bayesian approaches of statistical inference (e.g. R.A. Fisher's Fiducial Probability, D.A.S. Fraser's Structural Inference and A. Birnbaum's Likelihood Theory) from the position of the Symmetric Theory. So far the present paper remains preliminary and should be understood as an appeal for cooperation.

The description of the theory of interval probability in Section 2 is reduced to a minimum in order to avoid unnecessary repetitions. The introduction of the logical concept of probability in Section 3 concentrates on the mathematical aspects of the construction of W-fields containing all information which is the basic prerequisite for application of this concept. Section 4 describes the axiomatics of S-models producing duality of W-fields and characterizes the three quality-levels of duality. Some examples of application demonstrate in Section 5 statistical inference for constellations to be found in classical statistics. Section 6 gives an outlook to further results allowing the employment of the theory in more general situations.

2 The Theory of Interval Probability

This theory [2] is based on Kolmogorov's system of axioms. Let $(\Omega; \mathcal{A})$ be a measure-space, then $p(.)$ is nominated *K-function*, if it obeys the Kolmogorov-axioms.

The concept of interval probability is introduced in two steps. *R-probability* $P(A)$ on $(\Omega; \mathcal{A})$ is given if additionally to Kolmogorov's axioms T IV and T V hold.

T IV: $P(A) = [L(A); U(A)] \subseteq [0; 1]$

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.

$\mathcal{M}$ is nominated the *structure* of the R-probability-field $\mathcal{R}$.

The higher quality of interval probability is named *F-probability* and is given iff additionally

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

is valid.

A consequence of T VI is:

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

In the case of an F-probability-field $\mathcal{F} = (\Omega; \mathcal{A}; L(.))$ the structure is sufficient to describe all interval-limits. Any subset of the structure $\mathcal{M}$ which is sufficient to describe $\mathcal{M}$ and therefore the interval-limits of F-probability, is called a *prestructure* of the F-field.

Two partitions $\mathcal{C}_1$ and $\mathcal{C}_2$ of the sample space Ω of an F-probability-field are *mutually independent*, if a subset of the structure consisting of K-functions which are independent on $\mathcal{C}_1$ and $\mathcal{C}_2$, serves as prestructure. This definition allows the concept of an independent identically distributed (i.i.d.) sample, for which it can be proved: In a very large sample it is practically sure that the relative frequency of an event A with $P(A) = [L(A); U(A)]$ lies between $L(A)$ and $U(A)$. The model employed contains, however, no information about the behaviour of the relative frequency inside $[L(A); U(A)]$. This result allows a frequency-interpretation of interval probability.

In this theory some algebraic operations with probability-fields are possible, e.g. the *union* of F-probability-fields with identical $(\Omega; \mathcal{A})$: Let $\mathcal{F}_i = (\Omega; \mathcal{A}; L_i(.))$, $i \in I$, $\mathcal{F}_0 = (\Omega; \mathcal{A}; L_0(.))$, then $\mathcal{F}_0 = \bigcup_{i\in I} \mathcal{F}_i \quad\Leftrightarrow\quad L_0(A) = \inf_{i\in I} L_i(A),\ \forall A \in \mathcal{A}$. $\mathcal{F}_0$ is an F-probability-field with prestructure $\bigcup_{i\in I} \mathcal{M}(\mathcal{F}_i)$.

3 The Logical Concept of Probability

As it is understood here, this concept attaches components of probability exclusively to *arguments*, understood as *pairs* $(A||B)$ *of propositions*: the *premise* B and the *conclusion* A.

Therefore it never produces probability statements about propositions or events! All evidence concerning the conclusion A is relative to information contained in the premise B.

So far this concept is clearly distinct from any *objectivistic* approach.

On the other hand it contains no *personalistic* element and allows nothing which could be interpreted as "belief". Consequently it is clearly distinct from any subjectivistic approach.

The mathematical model employed by this concept is that of interval probability, especially the theory of F-probability-fields.

Definition 1 *Let $(\Omega_A; \mathcal{A})$ and $(\Omega_B; \mathcal{B})$ be two measurable spaces with $\{x\} \in \mathcal{A},\ \forall x \in \Omega_A$, $\{y\} \in \mathcal{B},\ \forall y \in \Omega_B$. Then a* W-field $\mathcal{W} = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L(.||.))$ *is created, iff the following axioms are valid:*

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

L II: *Let $I \neq \emptyset$ a set of indices, $B_0 \in \mathcal{B}^+$, $B_i \in \mathcal{B}^+$, $i \in I$, with $B_0 = \bigcup_{i\in I} B_i$. Then $\mathcal{F}(B_0) = \bigcup_{i\in I} \mathcal{F}(B_i)$ holds.*

L III: *Let*

$$\Omega_A = \Omega_A^{(1)} \times \Omega_A^{(2)}, \quad \Omega_B = \Omega_B^{(1)} \times \Omega_B^{(2)},$$
$$\mathcal{A}^{(r)} \subseteq \mathcal{P}ot(\Omega_A^{(r)}), \quad \mathcal{B}^{(r)} \subseteq \mathcal{P}ot(\Omega_B^{(r)}),$$

$r = 1, 2$, $\mathcal{A}^{(r)}$, $\mathcal{B}^{(r)}$ σ-fields, $r = 1, 2$;

$\{A^{(1)} \times A^{(2)} | A^{(r)} \in \mathcal{A}^{(r)}, r = 1, 2\}$ is a generating set of $\mathcal{A}$;

$\{B^{(1)} \times B^{(2)} | B^{(r)} \in \mathcal{B}^{(r)}, r = 1, 2\}$ is a generating set of $\mathcal{B}$.

Whenever

$$L(A^{(1)} \times \Omega_A^{(2)} || B^{(1)} \times B^{(2)}) = L(A^{(1)} \times \Omega_A^{(2)} || B^{(1)} \times \Omega_B^{(2)}),$$

$$L(\Omega_A^{(1)} \times A^{(2)} || B^{(1)} \times B^{(2)}) = L(\Omega_A^{(1)} \times A^{(2)} || \Omega_B^{(1)} \times B^{(2)}),$$

$\forall A^{(1)} \in \mathcal{A}^{(1)}$, $\forall A^{(2)} \in \mathcal{A}^{(2)}$, $\forall B^{(1)} \in \mathcal{B}^{(1)}$, $\forall B^{(2)} \in \mathcal{B}^{(2)}$,

then

$$L(A^{(1)} \times A^{(2)} || B^{(1)} \times B^{(2)}) = L(A^{(1)} \times \Omega_A^{(2)} || B^{(1)} \times \Omega_B^{(2)}) \cdot L(\Omega_A^{(1)} \times A^{(2)} || \Omega_B^{(1)} \times B^{(2)})$$

$\forall A^{(1)} \in \mathcal{A}^{(1)}$, $\forall A^{(2)} \in \mathcal{A}^{(2)}$, $\forall B^{(1)} \in \mathcal{B}^{(1)}$, $\forall B^{(2)} \in \mathcal{B}^{(2)}$,

holds. □

The elements of $\mathcal{A}$ and of $\mathcal{B}^+$ are understood as representing propositions. A pair $(A||B)$, $A \in \mathcal{A}$, $B \in \mathcal{B}$, is interpreted as an argument with premise B and conclusion A. The probability $P(A)$ in the F-field $\mathcal{F}(B)$ is interpreted as probability of the argument $(A||B)$ and designated by $P(A||B)$.

Axiom L I states that in evaluating arguments the higher quality-level of interval probability (defined in axiom T VI) has to be employed. Axiom L II describes the method to calculate $P(A||B_1 \text{ or } B_2)$, if $P(A||B_1)$ and $P(A||B_2)$ are known. This axiom clearly distinguishes the concept from all types of Bayesian approach. Axiom L III is formulated for a W-field $\mathcal{W}$ with two projection-fields $\mathcal{W}^{(1)}$ and $\mathcal{W}^{(2)}$ but by induction it can be generalized to any finite number of projection-fields $\mathcal{W}^{(i)}$, $i = 1, ..., n$, if the premise in each $\mathcal{W}^{(i)}$ is relevant only for the conclusion in the same $\mathcal{W}^{(i)}$ but totally irrelevant for all other conclusions. Thus *strong independence* of arguments is characterized. Due to Axiom L III such a situation produces mutual independent F-fields, and consequently allows for a frequency-interpretation of the logical concept:

If a large number of strongly independent arguments are evaluated by $P(A_i||B_i) = [L; U]$, $i = 1, ..., n \gg 1$, this is equivalent to: $P(A^T||B^T) = [1 - \varepsilon; 1]$, where

$$A^T = \bigcup_{L-\delta \le \frac{|I|}{n} \le U+\delta} \left[\bigcap_{i \in I} A_i \cap \bigcap_{j \in \neg I} (\neg A_j) \right],$$

$$B^T = \bigcap_{i=1}^{n} B_i;$$

$I \subseteq \{1, 2, ..., n\}$, $\varepsilon \ll 1$, $\delta \ll 1$.

This result can be interpreted as follows: From total evidence B^T, created by all B_i, with probability very close to 1 the proposition A^T may be concluded, stating that the proportion of true propositions in the set $\{A_i | i = 1, ..., n\}$ lies between $L - \delta$ and $U + \delta$, where δ is very small. If an argument $(A||B)$ with $P(A||B) = [L; U]$ is conceived as one out of an infinite set of mutually strongly independent arguments with $P(A_i||B_i) = [L; U]$, where all premises B_i are true, it is understood from L III, that the proportion of true conclusions A_i in this set lies between L and U.

Four definitions are important in employing this concept for statistical reasoning.

1. In the W-field $\mathcal{W} = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L(.||.))$ the set $\mathcal{Y}(I) = \{(A_i||B_i) | i \in I\}$ is nominated a *W-support of* $\mathcal{W}$ if all $L(.||.)$ of $\mathcal{W}$ can be calculated, provided that L_i and U_i out of $P(A_i||B_i) = [L_i; U_i]$, $\forall i \in I$, are known. □

2. In the W-field $\mathcal{W} = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L(.||.))$ the conclusions A_1 and A_2 are designated *mutually probabilistic equivalent*, $A_1 \sim A_2$, if $P(A_1||B) = P(A_2||B)$, $\forall B \in \mathcal{B}^+$; the premises B_1 and B_2 are designated *mutually probabilistic equivalent*, $B_1 \sim B_2$, if $P(A||B_1) = P(A||B_2)$, $\forall A \in \mathcal{A}$. □

3. For $\mathcal{A}_0 \subseteq \mathcal{A}$, $\mathcal{B}_0 \subseteq \mathcal{B}$ let $(\mathcal{A}_0||\mathcal{B}_0) := \{(A||B) | (A, B) \in \mathcal{A}_0 \times \mathcal{B}_0\}$. The W-support $\mathcal{Y} = (\mathcal{Y}_A||\mathcal{Y}_B) \cup (\overline{\mathcal{Y}_A}||\overline{\mathcal{Y}_B})$ of $\mathcal{W}$ is *regular W-support of* $\mathcal{W}$ iff
   (a) $\mathcal{Y}_A, \overline{\mathcal{Y}_A} \subseteq \mathcal{A}' := \mathcal{A} \setminus \{\emptyset, \Omega_A\}$, $\mathcal{Y}_A \cap \overline{\mathcal{Y}_A} = \emptyset$
   $\mathcal{Y}_B, \overline{\mathcal{Y}_B} \subseteq \mathcal{B}' := \mathcal{B} \setminus \{\emptyset, \Omega_B\}$, $\mathcal{Y}_B \cap \overline{\mathcal{Y}_B} = \emptyset$
   (b) $\forall A \in \mathcal{Y}_A \; \exists \overline{A} \in \overline{\mathcal{Y}_A} : \overline{A} \sim \neg A$,
   $\forall B \in \mathcal{Y}_B \; \exists \overline{B} \in \overline{\mathcal{Y}_B} : \overline{B} \sim \neg B$,
   (c) $\forall \overline{A} \in \overline{\mathcal{Y}_A} \; \exists A \in \mathcal{Y}_A : A \sim \neg \overline{A}$,
   $\forall \overline{B} \in \overline{\mathcal{Y}_B} \; \exists B \in \mathcal{Y}_B : B \sim \neg \overline{B}$. □

4. If in the W-field $\mathcal{W} = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L(.||.))$ to every $y \in \Omega_B$ a K-function $p(.||y)$ is corresponding — instead of a general F-field $\mathcal{F}(\{y\}) = (\Omega_A; \mathcal{A}; L(.||\{y\}))$ — $\mathcal{W}$ is nominated a *classical W-field.* □

Example 1α

a) $\mathcal{W} = (\Omega_A;\, \mathcal{A};\, \Omega_B;\, \mathcal{B};\, L(.||.))$ with

$$\begin{array}{rclrcl} \Omega_A & = & \mathbb{R}^1, & \mathcal{A} & = & \mathcal{B}or(\Omega_A) \\ \Omega_B & = & \mathbb{R}^1, & \mathcal{B} & = & \mathcal{B}or(\Omega_B) \end{array}$$

$$L(A||B) = \inf_{y \in B} \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-y)^2} dt$$

implying

$$U(A||B) = \sup_{y \in B} \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-y)^2} dt$$

is a classical W-field, since for every $y \in \Omega_B$, $p(A||y) = \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-y)^2} dt$ describes a classical probability field $(\Omega_A;\, \mathcal{A};\, p(.||y))$.

b) $\mathcal{Y}_0 = \{(]-\infty;\, x]||y)|\ x \in \Omega_A, y \in \Omega_B\}$ is a W-support of $\mathcal{W}$, since $p(]-\infty;\, x]||y) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-y)^2} dt$ is sufficient information to describe $\mathcal{W}$.

c) $]-\infty;\, x]$ and $]-\infty;\, x[$ are probabilistic equivalent since $F(x;\, y) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-y)^2} dt$ is everywhere continuous in x.

$]-\infty;\, y]$ and $]-\infty;\, y[$ are probabilistic equivalent since $F(x;\, y)$ is everywhere continuous in y.

d) While $\mathcal{Y}_0$ is not a regular support of $\mathcal{W}$, $\mathcal{Y}_1 = (\mathcal{Y}_A||\mathcal{Y}_B) \cup (\overline{\mathcal{Y}_A}||\overline{\mathcal{Y}_B})$ with

$$\begin{array}{rcl} \mathcal{Y}_A & = & \{]-\infty;\, x]|\, x \in \Omega_A\} \\ \overline{\mathcal{Y}_A} & = & \{[x;\, +\infty[|\, x \in \Omega_A\} \\ \mathcal{Y}_B & = & \{[y;\, +\infty[|\, y \in \Omega_B\} \\ \overline{\mathcal{Y}_B} & = & \{]-\infty;\, y]|\, y \in \Omega_B\} \end{array}$$

is a regular support of $\mathcal{W}$.

e) $\mathcal{W}^* = (\Omega_A;\, \mathcal{A};\, \Omega_B;\, \mathcal{B};\, L^*(.||.))$ with

$$\begin{array}{rclrcl} \Omega_A & = & \mathbb{R}^1, & \mathcal{A} & = & \mathcal{B}or(\Omega_A) \\ \Omega_B & = & \mathbb{R}^1, & \mathcal{B} & = & \mathcal{B}or(\Omega_B) \end{array}$$

$$L^*(A||B) = \inf_{\substack{y \in B \\ 1 \le \sigma \le 2}} \int_A \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(t-y)^2} dy$$

is a W-field, but not a classical one. □

4 The Symmetric Theory of Probability

Definition 2 *Let $\mathcal{W} = (\Omega_A;\, \mathcal{A};\, \Omega_B;\, \mathcal{B};\, L(.||.))$ be a W-field. An argument $(A||B)$, $(A,\, B) \in \mathcal{A}' \times \mathcal{B}'$, is* concordant in $\mathcal{W}$ *iff:*

a) $P(A||B) = [0;\, \alpha]$

b) $P(\neg A||\neg B) = [0;\, 1-\alpha]$,

$(A||B)$ *then is* α-concordant in $\mathcal{W}$, $(\neg A||\neg B)$ *is* $(1-\alpha)$*-concordant in* $\mathcal{W}$. □

The search for a set of concordant arguments in a W-field is the key task in employing the Symmetric Theory.

Example 1β In the case of W-field $\mathcal{W}$ according to Example 1α

$$\begin{array}{rcl} P(]-\infty;\, x]\,||\,[y;\, +\infty[) & = & \left[0;\, \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-y)^2} dt\right] \\ P(]x;\, +\infty]\,||\,[-\infty;\, y[) & = & \left[0;\, \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-y)^2} dt\right] \end{array}$$

Therefore every $(A_x||B_y)$, where $A_x =]-\infty;\, x]$, $B_y = [y;\, +\infty[$, is a concordant argument with $\alpha = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-y)^2} dt$. □

Definition 3 *Let* $\mathcal{W}_1 = (\Omega_A;\, \mathcal{A};\, \Omega_B;\, \mathcal{B};\, L_1(.||.))$ *and* $\mathcal{W}_2 = (\Omega_B;\, \mathcal{A};\, \Omega_A;\, \mathcal{B};\, L_2(.||.))$ *be W-fields and* $\mathcal{N} \subseteq \mathcal{A}' \times \mathcal{B}'$*. Then* $\mathcal{S} := (\mathcal{W}_1;\, \mathcal{N};\, \mathcal{W}_2)$ *is a* model of Symmetric Probability *or* S-Model, *iff Axiom S I is valid.*

S I: *a)* $\mathcal{N}_1 := \{(A||B)|(A,\, B) \in \mathcal{N}\}$
is a W-support of $\mathcal{W}_1$*, and its elements are concordant in* $\mathcal{W}_1$*.*

b) $\mathcal{N}_2 := \{(B||A)|(A,\, B) \in \mathcal{N}\}$
is a W-support of $\mathcal{W}_2$*, and its elements are concordant in* $\mathcal{W}_2$*.*

c) $\forall A,\, B \in \mathcal{N}$*:* $P_1(A||B) = P_2(B||A)$*.*

$\mathcal{N}$ *is designated the* nomenclature of $\mathcal{S}$*, the W-fields* $\mathcal{W}_1$ *and* $\mathcal{W}_2$ *are* dual with respect to the nomenclature $\mathcal{N}$*.* □

Example 1γ In case of primal W-field $\mathcal{W}_1 = \mathcal{W}$ (Example 1α) according to d) $\mathcal{Y}_1$ is a regular support of $\mathcal{W}_1$.

The W-field $\mathcal{W}_2 = (\Omega_B;\, \mathcal{B};\, \Omega_A;\, \mathcal{A};\, L_2(.||.))$, where

$$L_2(B||A) = \inf_{x \in A} \int_B \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} dt$$

implies

$$U_2(B||A) = \sup_{x \in A} \int_B \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} dt,$$

together with $\mathcal{W}_1$ obeys to rule S Ic) for all $(A, B) \in \mathcal{Y}_1$:

$$\begin{aligned} P_1(]-\infty; x] \,||\, [y; +\infty[) &= \left[0; \int_{-\infty}^{x-y} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}} d\tau\right] \\ &= \left[0; \int_{y-x}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}} d\tau\right] \\ &= P_2([y; +\infty[\,||\,]-\infty; x]). \end{aligned}$$

Since $\mathcal{Y}_2 = (\mathcal{Y}_B||\mathcal{Y}_A) \cup (\overline{\mathcal{Y}_B}||\overline{\mathcal{Y}_A})$ is a regular support of $\mathcal{W}_2$, $\mathcal{N} := (\mathcal{Y}_A \times \mathcal{Y}_B) \cup (\overline{\mathcal{Y}_A} \times \overline{\mathcal{Y}_B})$ represents the nomenclature of the S-model $(\mathcal{W}_1; \mathcal{N}; \mathcal{W}_2)$, and $\mathcal{W}_1$ and $\mathcal{W}_2$ are dual with respect to $\mathcal{N}$. □

The relation between two dual fields characterizes statistical inference. Axiom S I may be understood as adoption of basic concepts in classical non-Bayesian inference: The ideas of confidence and significance are easily recognizable, especially in S I c[1]. S-models according to Definition 3 define the elementary level of the methodology, where duality depends on the choice of nomenclature. It is possible that for the same W-field $\mathcal{W}_1$ with two different nomenclatures duality to two different W-fields $\mathcal{W}_2$ is established.

Definition 4 *Let $\mathcal{S} := (\mathcal{W}_1; \mathcal{N}; \mathcal{W}_2)$ be an S-model. It is* regular*, iff Axiom S II holds:*

S II: *a)* $\mathcal{N} = (\mathcal{N}_A \times \mathcal{N}_B) \cup (\overline{\mathcal{N}_A} \times \overline{\mathcal{N}_B})$.
b) $\mathcal{N}_1 = (\mathcal{N}_A||\mathcal{N}_B) \cup (\overline{\mathcal{N}_A}||\overline{\mathcal{N}_B})$ *is a regular W-support of* $\mathcal{W}_1$.
c) $\mathcal{N}_2 = (\mathcal{N}_B||\mathcal{N}_A) \cup (\overline{\mathcal{N}_B}||\overline{\mathcal{N}_A})$ *is a regular W-support of* $\mathcal{W}_2$.

In this case $\mathcal{W}_1$ and $\mathcal{W}_2$ are regularly dual with respect to $\mathcal{N}$. □

Theorem 1 *Let $\mathcal{S} := (\mathcal{W}_1; \mathcal{N}; \mathcal{W}_2)$ be an S-model. Then $\mathcal{N}$ defines a linear order on Ω_A as well as a linear order on Ω_B, so that equivalence classes consist of probabilistic equivalent elements only.* □

Vice versa $\mathcal{N}$ and the regular duality of $\mathcal{W}_1$ and $\mathcal{W}_2$ are uniquely determined, if besides the W-fields $\mathcal{W}_1$ and $\mathcal{W}_2$ also linear orders on Ω_A and Ω_B are given (as usual in one-dimensional classical probability, at least if a one-parametric family of distributions is given): Regular duality represents a higher quality-level of duality.

Example 1δ The nomenclature $\mathcal{N} := (\mathcal{Y}_A \times \mathcal{Y}_B) \cup (\overline{\mathcal{Y}_A} \times \overline{\mathcal{Y}_B})$ employed in the S-model $\mathcal{S} = (\mathcal{W}_1; \mathcal{N}; \mathcal{W}_2)$ of Example 1γ obviously is in accordance with Axiom S II: $\mathcal{W}_1$ and $\mathcal{W}_2$ are regularly dual with respect to $\mathcal{N}$. Due to Theorem 1 regular nomenclature $\mathcal{N}$ is determined by the linear orders on Ω_A and Ω_B and so is regular duality of $\mathcal{W}_1$ and $\mathcal{W}_2$. □

The highest quality-level of duality can be employed if probability ratios are monotone.

Definition 5 *Let $\mathcal{W}_1$ and $\mathcal{W}_2$ be classical W-fields, $\mathcal{S} := (\mathcal{W}_1; \mathcal{N}; \mathcal{W}_2)$ be a regular S-model,*

$$\mathcal{Z}_A: \text{ “}x_1 \prec x_2\text{” and } \mathcal{Z}_B: \text{ “}y_1 \prec y_2\text{”}$$

be the two linear orders defined by $\mathcal{N}$ on Ω_A and Ω_B. Additionally let

$$[x_1; x_2] := \{x \in \Omega_A | x_1 \precsim x \precsim x_2\}, \forall x_1, x_2 \in \Omega_A : x_1 \precsim x_2,$$

$$[y_1; y_2] := \{y \in \Omega_B | y_1 \precsim y \precsim y_2\}, \forall y_1, y_2 \in \Omega_B : y_1 \precsim y_2.$$

Then $\mathcal{S}$ is a perfect S-model*, if the following axiom holds:*

S III: *a) Let $x_i \in \Omega_A$, $i = 1, 2, 3, 4$:*
$x_1 \prec x_2 \prec x_3 \prec x_4$,
$y_i \in \Omega_B$, $i = 1, 2$: $y_1 \prec y_2$;
then:
$$p_1([x_1; x_2]||y_1) \cdot p_1([x_3; x_4]||y_2) \le p_1([x_1; x_2]||y_2) \cdot p_1([x_3; x_4]||y_1);$$
b) Let $y_i \in \Omega_B$, $i = 1, 2, 3, 4$:
$y_1 \prec y_2 \prec y_3 \prec y_4$,
$x_i \in \Omega_A$, $i = 1, 2$: $x_1 \prec x_2$;
then:
$$p_2([y_1; y_2]||x_1) \cdot p_2([y_3; y_4]||x_2) \le p_2([y_1; y_2]||x_2) \cdot p_2([y_3; y_4]||x_1).$$ □

Definition 6 *If $\mathcal{S} = (\mathcal{W}_1; \mathcal{N}; \mathcal{W}_2)$ is a perfect S-model, the W-fields $\mathcal{W}_1$ and $\mathcal{W}_2$ are* mutually perfectly dual. □

If Axiom S III holds, then its validity is restricted to the linear orders $\mathcal{Z}_A$ and $\mathcal{Z}_B$, except for the simultaneous inversion of both orders, which in fact is only a change of symbols.

Perfect duality, if it can be established, depends only on the W-fields $\mathcal{W}_1$ and $\mathcal{W}_2$: the perfect dual W-field $\mathcal{W}_2$ is the perfect answer to the request for statistical inference if $\mathcal{W}_1$ is given. This applies to many problems which arise in statistical reasoning with models employing classical probability.

Example 1ε In the case of the regular S-model $\mathcal{S} = (\mathcal{W}_1; \mathcal{N}; \mathcal{W}_2)$ according to Example 1δ $\mathcal{W}_1$ as

[1]If for a one-sided test of significance the critical region A and the hypothesis B are compared with conclusion A and premise B of a concordant argument employed in the nomenclature of an S-model.

well as $\mathcal{W}_2$ are classical W-fields defined by the Normal Law. Since this law is distinguished by monotone density-ratios, the conditions of Axioms S III are given. Therefore $\mathcal{W}_1$ and $\mathcal{W}_2$ are perfectly dual, and this relation is unique: No other perfect S-model with participation of one of the W-fields $\mathcal{W}_1$ or $\mathcal{W}_2$ is possible. □

5 Some applications in classical statistics

1) The L-constellation: A one-dimensional continuous distribution function is given.

Let: $-\infty \leq x_L < x_U \leq +\infty,$
$-\infty \leq y_L < y_U \leq +\infty,$

and

$0 \leq F(x;\, y) \leq 1,\ \forall x:\ x_L < x < x_U,$
$\forall y:\ y_L < y < y_U,$

$F(x_2;\, y) \geq F(x_1;\, y),\ \forall x_1,\, x_2 : x_L < x_1 < x_2 < x_U,$
$\forall y:\ y_L < y < y_U,$

$F(x;\, y_1) \geq F(x;\, y_2),\ \forall x:\ x_L < x < x_U,$
$\forall y_1,\, y_2 : y_L < y_1 < y_2 < y_U,$

$\lim_{x \to x_U} F(x;\, y) = 1,\ \forall y:\ y_L < y < y_U,$

$\lim_{x \to x_L} F(x;\, y) = 0,\ \forall y:\ y_L < y < y_U,$

$\lim_{y \to y_U} F(x;\, y) = 0,\ \forall x:\ x_L < x < x_U,$

$\lim_{y \to y_L} F(x;\, y) = 1,\ \forall x:\ x_L < x < x_U.$

$F(x;\, y)$ continuous in x and y, $\forall x:\ x_L < x < x_U$, $\forall y:\ y_L < y < y_U$: describing a *stochastically ordered* family of classical continuous probability distributions with parameter y:

$$p_1(X \leq x || y) = F(x;\, y)$$

This constellation produces a W-field $\mathcal{W}_1 = (\Omega_A;\, \mathcal{A};\, \Omega_B;\, \mathcal{B};\, L_1(.||.))$ with

$$\begin{array}{lcll} \Omega_A &=& \{x | x_L < x < x_U\}, & \mathcal{A} = \mathcal{B}or(\Omega_A), \\ \Omega_B &=& \{y | y_L < y < y_U\}, & \mathcal{B} = \mathcal{B}or(\Omega_B), \end{array}$$

$$L_1(]x_L;\, x] || \{y\}) = F(x;\, y).$$

A regular nomenclature is created by

$$\begin{array}{lcllcl} A_x &=&]x_L;\, x], & \overline{A_x} &=& [x;\, x_U[\\ B_y &=& [y;\, y_U[, & \overline{B_y} &=&]y_L;\, y]. \end{array}$$

$$\begin{array}{rcl} P_1(A_x || B_y) &=& [0;\, F(x;\, y)], \\ && \forall x \in \Omega_A,\, \forall y \in \Omega_B, \\ P_1(\overline{A_x} || \overline{B_y}) &=& P_1(\neg A_x || \neg B_y) \\ &=& [0;\, 1 - F(x;\, y)], \\ && \forall x \in \Omega_A,\, \forall y \in \Omega_B. \end{array}$$

$$\begin{array}{ll} \mathcal{N}_A = \{A_x | x \in \Omega_A\}, & \overline{\mathcal{N}_A} = \{\overline{A_x} | x \in \Omega_A\}, \\ \mathcal{N}_B = \{B_y | y \in \Omega_B\}, & \overline{\mathcal{N}_B} = \{\overline{B_y} | y \in \Omega_B\}, \end{array}$$

$$\mathcal{N} = (\mathcal{N}_A \times \mathcal{N}_B) \cup (\overline{\mathcal{N}_A} \cup \overline{\mathcal{N}_B}).$$

The dual W-field $\mathcal{W}_2 = (\Omega_B;\, \mathcal{B};\, \Omega_A;\, \mathcal{A};\, L_2(.||.))$ is determined by $P_1(A||B) = P_2(B||A)$, $\forall (A,\, B) \in \mathcal{N}$:

$$\begin{array}{rcl} P_2(B_y || A_x) &=& [0;\, F(x;\, y)] \\ && \forall y \in \Omega_B,\, \forall x \in \Omega_A, \\ P_2(\neg B_y || \neg A_x) &=& P_2(\overline{B_y} || \overline{A_x}) \\ &=& [0;\, 1 - F(x;\, y)], \\ && \forall y \in \Omega_B,\, \forall x \in \Omega_A. \end{array}$$

$$\begin{array}{rcl} P_2([y;\, y_U[\, || \,]x_L;\, x]) &=& [0;\, F(x;\, y)] \\ P_2(]y_L;\, y]\, || \,[x;\, x_U[) &=& [0;\, 1 - F(x;\, y)] \end{array}$$

Due to Axiom L II it can be concluded:

$$L_2(]y_L;\, y] || \{x\}) = 1 - F(x;\, y),$$

and $\mathcal{W}_2$ describes an ordered family of one-dimensional probability distributions as well, this time the parameter is x:

$$p_2(Y \leq y || x) = 1 - F(x;\, y).$$

$\mathcal{W}_1$ and $\mathcal{W}_2$ are regularly dual with respect to the linear orders on Ω_A and Ω_B.

If additionally Axiom S III is fulfilled, the quality of duality is perfect, a level which cannot be reached by any other linear order of Ω_A or Ω_B.

Perfect duality is given for instance in the case of the normal law

$$F(x,\, y) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(t-y)^2}\, dt$$

for any fixed $\sigma \in]0;\, \infty[$, but not in the case of the Cauchy law:

$$F(x;\, y) = \int_{-\infty}^{x} \frac{1}{\pi} \frac{1}{1 + (t-y)^2}\, dt.$$

2) The DC-constellation: the distribution function is discontinuous in x, but continuous in y.

Let $-\infty \leq y_L < y_U \leq \infty$,

$$\begin{array}{rcl} \Omega_A^+ &=& \{x_0,\, ...,\, x_i < x_{i+1},\, ...,\, x_N\}, \\ \mathcal{A}^+ &=& \mathcal{P}ot(\Omega_A^+), \\ \Omega_A &=& \Omega_A^+ \setminus \{x_0\}, \\ \mathcal{A} &=& \mathcal{P}ot(\Omega_A), \\ \Omega_B &=& \{y | y_L < y < y_U\}, \\ \mathcal{B} &=& \mathcal{B}or(\Omega_B), \end{array}$$

and:

$0 \leq F(x;\, y) \leq 1,\ \forall x \in \Omega_A,\ \forall y \in \Omega_B,$

$F(x_{i+1};\, y) \geq F(x_i;\, y),\ \forall x_i \in \Omega_A \setminus \{x_N\},\ \forall y \in \Omega_B,$

$F(x_i;\, y_1) \geq F(x_i;\, y_2),\ \forall x_i \in \Omega_A,\ \forall y_1,\, y_2 \in \Omega_B : y_1 < y_2,$

$F(x;\, y)$ continuous in y, $\forall x_i \in \Omega_A$, $\forall y \in \Omega_B$,

$$\begin{aligned} F(x_0;\, y) &= 0,\ \forall y \in \Omega_B, \\ F(x_N;\, y) &= 1,\ \forall y \in \Omega_B, \\ F(x_i;\, y_L) &= 1,\ \forall x_i \in \Omega_A, \\ F(x_i;\, y_U) &= 0,\ \forall x_i \in \Omega_A^+ \setminus \{x_N\}. \end{aligned}$$

The stochastically ordered family of discrete probability distributions

$$p_1(X \leq x_i || y) = F(x_i;\, y)$$

with continuous parameter y is described by the W-field $\mathcal{W}_1 = (\Omega_A;\, \mathcal{A};\, \Omega_B;\, \mathcal{B};\, L_1(.||.))$ with $L_1(\{x_i\}||\{y\}) = F(x_i;\, y) - F(x_{i-1};\, y)$, $\forall x_i \in \Omega_A$, $\forall y \in \Omega_B$.

A regular nomenclature is given by $\mathcal{N} = (\mathcal{N}_A \times \mathcal{N}_B) \cup (\overline{\mathcal{N}_A} \times \overline{\mathcal{N}_B})$ with

$$\begin{aligned} \mathcal{N}_A &= \{A_i | i = 1, ..., N-1\}, \\ A_i &= \{x_1, ..., x_i\},\ i = 1, ..., N-1, \\ \overline{\mathcal{N}_A} &= \{\overline{A_i} | i = 1, ..., N-1\}, \\ \overline{A_i} &= \{x_{i+1}, ..., x_N\},\ i = 1, ..., N-1, \\ \mathcal{N}_B &= \{B_y | y_L \leq y \leq y_U\}, \\ B_y &= [y;\, y_U], \\ \overline{\mathcal{N}_B} &= \{\overline{B_y} | y_L \leq y \leq y_U\}, \\ \overline{B_y} &= [y_L;\, y]. \end{aligned}$$

Then it can be shown, that $\mathcal{W}_1$ and the W-field $\mathcal{W}_2 = (\Omega_B;\, \mathcal{B};\, \Omega_A;\, \mathcal{A};\, L_2(.||.))$ with

$$L_2([y;\, y_U]||\{x_i\}) = F(x_{i-1};\, y), \forall x_i \in \Omega_A, \forall y \in \Omega_B,$$

$$L_2([y_L;\, y]||\{x_i\}) = 1 - F(x_i;\, y), \forall x_i \in \Omega_A, \forall y \in \Omega_B,$$

are regularly dual with respect to the linear orders given on Ω_A and on Ω_B.

The W-field $\mathcal{W}_2$ describes a stochastically ordered family of interval probability distributions — not classical ones — distinguished by:

$$P_2([y;\, y_U]\, ||\{x_i\}) = [F(x_{i-1};\, y);\, F(x_i;\, y)]\,,\quad \forall x_i \in \Omega_A,\ \forall y \in \Omega_B,$$

$$P_2([y_L;\, y]\, ||\{x_i\}) = [1 - F(x_i;\, y);\, 1 - F(x_{i-1};\, y)],\quad \forall x_i \in \Omega_A,\ \forall y \in \Omega_B.$$

From this result it can be concluded, that

$$\begin{aligned} P_2([y_1;\, y_2]\, ||\{x_i\}) &= [\max(0;\, F(x_{i-1};\, y_1) - F(x_i;\, y_2));\\ &\quad F(x_i;\, y_1) - F(x_{i-1};\, y_2)],\\ &\quad \forall x_i \in \Omega_A,\ \forall y_1 < y_2 : y_1,\, y_2 \in \Omega_B, \end{aligned}$$

and

$$P_2(\{y\}||\{x_i\}) = [0;\, F(x_i;\, y) - F(x_{i-1},\, y)],\quad \forall x_i \in \Omega_A,\ \forall y \in \Omega_B.$$

In the case of the classical Binomial law with parameters n and y determining $\mathcal{W}_1$, the dual W-field $\mathcal{W}_2$ is given by

$$P_2([0;\, y]||\{i\}) = \left[\sum_{j=i+1}^{n} \binom{n}{j} y^j (1-y)^{n-j};\sum_{j=i}^{n} \binom{n}{j} y^j (1-y)^{n-j}\right] \tag{1}$$

$$P_2([y;\, 1]||\{i\}) = \left[\sum_{j=0}^{i-1} \binom{n}{j} y^j (1-y)^{n-j};\sum_{j=0}^{i} \binom{n}{j} y^j (1-y)^{n-j}\right]. \tag{2}$$

It produces

$$P_2(\{y\}||\{i\}) = \left[0;\, \binom{n}{i} y^i (1-y)^{n-i}\right]$$

and allows an answer to the famous question about the probability for success in the "next" trial: If the premise is, that n trials produce i successes, the argument for success in the $(n+1)$. trial is given by

$$P_2(x_{n+1} = 1||\{i\}) = \left[\frac{i}{n+1};\, \frac{i+1}{n+1}\right].$$

6 Some further problems

6.1 Restricting the range of the parameter

In certain situations the statistician has to combine information about a parameter y which is given in advance with the information stemming from the realisation of a random variable. Mostly this means explicitly that there exist restrictions concerning the possible values of y.

In such a situation the Symmetric Theory is applied regardless of the restriction to the 'natural' set of values y: Ω_B and, by means of duality, $\mathcal{W}_2$ are defined. In a second step the restrictions are considered by means of *conditional probability.* If $\mathcal{W}_2$ is a classical W-field, the classical concept of conditional probability can be employed, provided that for the restricted range of y: Ω_B^R the condition $P_2(\Omega_B^R||\{x\}) > 0$ holds. If Ω_B^R is of the same dimension as Ω_B — f.i. a non-degenerate interval in $\Omega_B = \mathbb{R}^1$ — but $P_2(\Omega_B^R||\{x\}) = 0$, it must be concluded, that the outcome of x and the restriction Ω_B^R of the values y are contradicting each other.

If on the other hand $P_2(\Omega_B^R||\{x\}) > 0$, then

$$P_2^R(B||\{x\}) := \frac{P_2(B||\{x\})}{P_2(\Omega_B^R||\{x\})},\ \forall B \in \mathcal{B} : B \subseteq \Omega_B^R,$$

defines a classical probability field representing the answer to the statistician's request.

Example 1ι The W-field $\mathcal{W}_2$ of Example 1γ according to 1ε is a classical one. Since the range of positive density for the argument $(\{y\}||\{x\})$ consists of all $(x, y) \in \mathbb{R}^1 \times \mathbb{R}^1$, the condition $P_2(\Omega_B^R||\{x\}) > 0$ holds for every non-degenerate interval $\Omega_B^R \subset \mathbb{R}^1$ and every $x \in \Omega_A$. If e.g. $\Omega_B^R := [0; +\infty[$, for every $B \in \mathcal{B} : B \subseteq [0; +\infty[$ and for every $x \in \mathbb{R}^1$ the conditional probability:

$$P_2(B|\Omega_B^R||\{x\}) = \frac{\int_B \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-x)^2} dt}{\int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-x)^2} dt} = P_2^R(B||\{x\})$$

describes statistical inference in the situation of prior information that $y \geq 0$. If, however, for an outcome x, $\int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-x)^2} dt$ is very small, this is a problem of interpreting the result, not of formal methods. □

Example 2 If a family of classical rectangular distributions with range $[y - \frac{1}{2}; y + \frac{1}{2}]$ is given, $\mathcal{Y}_1$ as defined in Example 1αd) is a regular nomenclature, $\mathcal{W}_1 = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L_1(.||.))$ and $\mathcal{W}_2 = (\Omega_B; \mathcal{B}; \Omega_A; \mathcal{A}; L_2(.||.))$ with $\Omega_A = \mathbb{R}^1$, $\mathcal{A} = \mathcal{B}or(\Omega_A)$, $\Omega_B = \mathbb{R}^1$, $\mathcal{B} = \mathcal{B}or(\Omega_B)$,

$$\begin{aligned} P_1(]-\infty; x]\,||y) &= P_2([y; +\infty[\,||x) \\ &= \begin{cases} 0, & x \leq y - \frac{1}{2} \\ x-y+\frac{1}{2}, & y-\frac{1}{2}<x<y+\frac{1}{2} \\ 1, & x \geq y + \frac{1}{2} \end{cases} \end{aligned}$$

are mutually regularly dual classical W-fields with respect to the linear orders on Ω_A and Ω_B. If f.i. $\Omega_B^R = [0; +\infty[$, an outcome $x \leq -\frac{1}{2}$ means contradiction, an outcome $x : -\frac{1}{2} < x < +\frac{1}{2}$ leads to

$$\begin{aligned} P_2^R(B||\{x\}) &= P_2(B|\Omega_B^R||\{x\}) \\ &= \frac{P_2(B||\{x\})}{P_2(\Omega_B^R||\{x\})} \\ &= \frac{P_2(B||\{x\})}{x + \frac{1}{2}},\ \forall B \in \mathcal{B} : B \subseteq [0; \infty[. \end{aligned}$$

In case of $x \geq \frac{1}{2}$ the restriction becomes irrelevant:

$$\begin{aligned} P_2^R(B||\{x\}) &= \frac{P_2(B||\{x\})}{1} \\ &= P_2(B||\{x\}),\ \forall B \in \mathcal{B} : B \subseteq [0; \infty[. \end{aligned}$$ □

If $\mathcal{W}_2$ is not a classical W-field, the *intuitive concept of conditional interval-probability* [4] has to be employed instead of the classical one: the structure of $iP(B|C)$ consisting of the classical conditional probability $p(B|C)$ for all structural elements of the F-field.

This applies at first to the question of a possible contradiction of restriction and outcome: The criterion in this situation is, whether there exists at least one element $p^*(.||x)$ of the structure $\mathcal{M}(\mathcal{W}_2)$, so that $p^*(\Omega_B^R||x) > 0$ holds. If $p(\Omega_B^R||x) = 0$, $\forall p(.||.) \in \mathcal{M}(\mathcal{W}_2)$, this has to be interpreted as contradiction.

Otherwise, restriction and outcome are compatible, but — if necessary — the structure has to be reduced to

$$\mathcal{M}^R(\mathcal{W}_2) = \{p(.||.) \in \mathcal{M}(\mathcal{W}_2) : p(\Omega_B^R||x) > 0\}.$$

The restricted probability for every $B \in \mathcal{B} : B \subseteq \Omega_B^R$, then is given by

$$\begin{aligned} &P_2^R(B||\{x\}) = iP_2(B|\Omega_B^R||\{x\}) \\ &= \left[\inf_{p \in \mathcal{M}^R(\mathcal{W}_2)} \frac{p(B||x)}{p(\Omega_B^R||x)};\ \sup_{p \in \mathcal{M}^R(\mathcal{W}_2)} \frac{p(B||x)}{p(\Omega_B^R||x)} \right]. \end{aligned}$$

Example 3 If the primal W-field $\mathcal{W}_1$ is determined by the Binomial law for any fixed n, $\mathcal{W}_2$ is given by (1) and (2). Let a restriction Ω_B^R of Ω_B be defined by $y \geq \eta$ with $0 < \eta < 1$. Concerning a possible contradiction of restriction and outcome, (2) reveals that in the case of $\eta = \frac{1}{2}$

$$U_2\left(\left[\tfrac{1}{2}; 1\right] ||\{i\}\right) > 0,\ \forall i = 0, 1, ..., n.$$

There is never a contradiction between Ω_B^R and $\{i\}$!

If $B = [y; 1]$ with $y > \eta$ — so that $B \subset \Omega_B^R$ — according to (2) one arrives at

$$P_2^R(B||\{i\}) = \left[L_2^R(B||\{i\}); U_2^R(B||\{i\})\right]$$

with

$$\begin{aligned} L_2^R(B||\{i\}) &= \frac{\sum_{j=0}^{i-1} \binom{n}{j} y^j (1-y)^{n-j}}{\sum_{j=0}^{i} \binom{n}{j} \eta^j (1-\eta)^{n-j}} \\ U_2^R(B||\{i\}) &= \min\left(1,\ \frac{\sum_{j=0}^{i} \binom{n}{j} y^j (1-y)^{n-j}}{\sum_{j=0}^{i-1} \binom{n}{j} \eta^j (1-\eta)^{n-j}}\right) \end{aligned}$$

Since $P_2^R(\Omega_B^R) = [1]$, the result shows that $L_2^R(.)$ as a function of y is discontinuous for $y = \eta$. □

6.2 I.i.d. sample

Some of the concepts described above are useful in a constellation of highest importance: The statistician is given the realisation of an i.i.d. sample $\vec{x}$ of the size $n > 1$ out of a stochastically ordered family of classical one-dimensional probability distributions with parameter y. Instead of relying on a sufficient estimator, the statistician wants to employ the sample as a whole for the premise of an argument about the parameter y.

In such a situation the statistician in a first step should neglect his information that the parameter y is the same for all elements of the sample. On the contrary he attributes a parameter $y^{(r)}$ to each $x^{(r)}$ and takes it as granted that $y^{(r)}$ influences only $x^{(r)}$ and no other element of the sample. In this way he establishes strong independence of the arguments $(A^{(r)}||B^{(r)})$ according to Axiom L III.

Let $\mathcal{W}_1^{(r)}$ be the W-field describing arguments with premise about $y^{(r)}$ and conclusion about $x^{(r)}$, then

$$\mathcal{W}_1 = \bigtimes_{r=1}^{n} \mathcal{W}_1^{(r)}$$

constitutes a W-field, for which each $\mathcal{W}_1^{(r)}$ is a projection-field. A fundamental theorem of the Symmetric Theory says that due to strong independence

$$\mathcal{W}_2 = \bigtimes_{r=1}^{n} \mathcal{W}_2^{(r)}$$

holds together with

$$\begin{aligned} \mathcal{N} \;=\; \bigcup_{r=1}^{n} \Big\{ \Big(& \Omega_A^{(1)} \times ... \times A^{(r)} \times ... \times \Omega_A^{(n)}, \\ & \Omega_B^{(1)} \times ... \times B^{(r)} \times ... \times \Omega_B^{(n)} \Big) \\ & \Big| \Big(A^{(r)}, B^{(r)} \Big) \in \mathcal{N}^{(r)} \Big\}, \end{aligned} \qquad (3)$$

and that $\mathcal{S} = (\mathcal{W}_1; \mathcal{N}; \mathcal{W}_2)$ is an S-model, provided that $\mathcal{S}^{(r)} = (\mathcal{W}_1^{(r)}; \mathcal{N}^{(r)}; \mathcal{W}_2^{(r)})$, $r = 1, ..., n$, are S-models.

The W-field $\mathcal{W}_2$ defines the probability of an argument from a premise about $\vec{x}$ to the n-dimensional $\vec{y}$. The additional information that $y^{(1)} = ... = y^{(n)} = y$ distinguishes a one-dimensional subset Ω_B^R of values y in the space Ω_B of conclusions. If $\mathcal{W}_2$ is a classical W-field, a contradiction between the outcome $\vec{x}$ as a premise and Ω_B^R as a conclusion has to be admitted if $P_2(\Omega_B^R||\{\vec{x}\}) = 0$ for a discrete probability-field or $\int_{\Omega_B^R} f_2(y, .., y||\{\vec{x}\})dy = 0$ in the case of an existing n-dimensional density $f_2(y^{(1)}, ..., y^{(n)}||\{\vec{x}\})$.

If otherwise $P_2(\Omega_B^R||\{\vec{x}\}) > 0$, for any $B \in \mathcal{B} : B \subseteq \Omega_B^R$ classical conditional probability

$$\begin{aligned} P_2(B|\Omega_B^R||\{\vec{x}\}) &= \frac{P_2(B||\{\vec{x}\})}{P_2(\Omega_B^R||\{\vec{x}\})} \\ &\overset{\text{resp.}}{=} \frac{\int_B f_2(y, ..., y||\{\vec{x}\})dy}{\int_{\Omega_B^R} f_2(y, ..., y||\{\vec{x}\})dy} \end{aligned}$$

can be understood as the probability of B if all $y^{(r)}$ all equal, and is designated $P_2^R(B||\{\vec{x}\})$.

Example 1ζ Let $\mathcal{W}_2$ be given by $N(x; 1)$ according to Example 1γ and $\bar{x} = \frac{1}{n}\sum_{r=1}^{n} x^{(r)}$. Then

$$f_2(y^{(1)}, ..., y^{(n)}||\{\bar{x}\}) = \prod_{r=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y^{(r)}-x^{(r)})^2}.$$

$$\begin{aligned} f_2(y, ..., y||\{\vec{x}\}) &= \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{r=1}^{n}(y-x^{(r)})^2} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\left[n(y-\bar{x})^2 + \sum_{r=1}^{n} x^{(r)2} - n\bar{x}^2\right]}; \end{aligned}$$

$$\begin{aligned} &\int_{-\infty}^{+\infty} f_2(y, ..., y||\{\vec{x}\})dy \\ &\quad = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\left[\sum_{r=1}^{n} x^{(r)2} - n\bar{x}^2\right]} \cdot \int_{-\infty}^{+\infty} e^{-\frac{n}{2}t^2} dt; \end{aligned}$$

$$\begin{aligned} \frac{f_2(y, ..., y||\{\vec{x}\})}{\int_{-\infty}^{+\infty} f_2(y, ..., y||\{\vec{x}\})dy} &= \frac{e^{-\frac{n}{2}(y-\bar{x})^2}}{\sqrt{2\pi}\frac{1}{\sqrt{n}}} \\ &= \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{n}{2}(y-\bar{x})^2} \\ &= f_2^R(y||\{\vec{x}\}). \end{aligned}$$

The density — and therefore the probability — for the argument with premise $\vec{x}$ and conclusion y is obviously the one which is also created by the perfectly dual $\mathcal{W}_2^*$, if $\mathcal{W}_1^*$ describes the probability of the argument with premise $\{y\}$ and conclusion $\{\bar{x}\}$.

If $\mathcal{W}_2$ is not a classical W-field the intuitive concept of conditional interval-probability has to be applied instead of the classical conditional probability. □

Example 4: Let $\mathcal{W}_1^{(r)} = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L_1^{(r)}(.||.))$, $\Omega_A = \{0; 1\}$, $\mathcal{A} = \mathcal{P}ot(\Omega_A)$, $\Omega_B = [0; 1]$, $\mathcal{B} = \mathcal{B}or(\Omega_B)$, $L_1^{(r)}(\{0\}||y^{(r)}) = 1 - y^{(r)}$, the simplest case of a DC-constellation. According to (1) and (2) for $\mathcal{W}_2^{(r)} = (\Omega_B; \mathcal{B}; \Omega_A; \mathcal{A}; L_2^{(r)}(.||.))$ it follows for the regular S-model $\mathcal{S}^{(r)} = (\mathcal{W}_1^{(r)}; \mathcal{N}^{(r)}; \mathcal{W}_2^{(r)})$ with $\mathcal{N}^{(r)}$ like $\mathcal{N}$ as described in 5.2:

$$\begin{aligned} P_2^{(r)}([0; y^{(r)}]||\{0\}) &= [y^{(r)}; 1] \\ P_2^{(r)}([0; y^{(r)}]||\{1\}) &= [0; y^{(r)}] \end{aligned}$$

for $r = 1, ..., n$.

Let $\mathcal{W}_1 = \bigtimes_{r=1}^{n} \mathcal{W}_1^{(r)}$, $\mathcal{W}_2 = \bigtimes_{r=1}^{n} \mathcal{W}_2^{(r)}$ and $\mathcal{N}$ according to (3).

Then

$$P_2([0;\, y^{(1)}] \times ... \times [0;\, y^{(n)}] || \{x^{(1)}, .., x^{(n)}\}) =$$

$$= \begin{cases} \left[\prod_{r=1}^{n} y^{(r)};\, 1\right], \text{if } x^{(1)} = ... = x^{(n)} = 0; \\ \left[0;\, \prod_{r \in I} y^{(r)}\right] \quad \text{with } I := \{r \in \{1, ..., n\} | x^{(r)} = 1\}, \\ \qquad\qquad \text{if } 1 \in \{x^{(1)}, ..., x^{(n)}\}. \end{cases}$$

The calculation of conditional interval-probability according to the intuitive concept relies on the extreme values of the classical conditional probabilities for structural elements.

α) $x^{(1)} = ... = x^{(n)} = 0$: For every $0 < Y < 1$ the smallest value out of the structure for

$$\frac{p_2([0;\, Y] \times ... \times [0;\, Y] || \{0, ..., 0\})}{p_2([0;\, 1] \times ... \times [0;\, 1] || \{0, ..., 0\})}$$
$$= p_2([0;\, Y] \times ... \times [0;\, Y] || \{0, ..., 0\})$$

is found for

$$p_2([0;\, y^{(1)}] \times ... \times [0;\, y^{(n)}] || \{0, ..., 0\}) = \prod_{r=1}^{n} y^{(r)}.$$

For this n-dimensional classical probability the density $f_2(y^{(1)}, ..., y^{(n)} || \{0, ..., 0\}) = 1$. Therefore

$$\begin{aligned} \int_0^Y f_2(y, ..., y || \{0, ..., 0\}) dy &= Y, \\ \frac{\int_0^Y f_2(y, ..., y || \{0, ..., 0\}) dy}{\int_0^1 f_2(y, ..., y || \{0, ..., 0\}) dy} &= \frac{Y}{1} = Y, \\ L_2([0;\, Y] || \{0, ..., 0\}) &= Y. \end{aligned}$$

The largest value out of the structure for $p_2([0;\, Y] \times ... \times [0;\, Y] || \{0, ..., 0\})$, $Y > 0$, obviously is 1.

Therefore: $P_2([0;\, Y] || \{0, ..., 0\}) = [Y;\, 1]$, independent of n (!).

β) $x^{(r)} = 1 : r \in I \neq \emptyset$; $x^{(r)} = 0 : r \notin I$; $\vec{x}(I)$

The smallest value for $p_2([0;\, Y] \times ... \times [0;\, Y] || \{\vec{x}(I)\})$, $Y < 1$, obviously is 0; the largest one is created by the structural element with $p_2([0;\, Y] || \{0\}) = 1$, $p_2([0;\, Y] || \{1\}) = Y$. $f_2(y^{(1)}, ..., y^{(n)} || \vec{x}(I))$ is independent of $y^{(r)} : r \notin I$, and in the subspace generated by $y^{(r)}$, $r \in I$, the density with dimension $|I|$ has the value 1.

This produces — in analogy to the lower limit in the case $\{0, ..., 0\}$: $U_2([0;\, Y] || \{\vec{x}(I)\}) = Y$ and $P_2([0;\, Y] || \{\vec{x}(I)\}) = [0;\, Y]$ independent of n and of $|I|$ (!).

The result says, that increasing the sample does not produce more accuracy in evaluating arguments with conclusions about y. Of course a substantial gain in evidence is achieved by using the Binomial model as described in 5.2. But Example 4 — describing indeed a special case — demonstrates that the impression generated by Example 1ζ must not be generalized. □

6.3 Updating

It must always be borne in mind, that the subject of the Symmetric Theory are arguments, not events or propositions. If an S-model presents dual probability for arguments based on an outcome $\{x_1\}$, this premise remains unchanged, whatever manipulations are applied in the field of conclusions. This is evidentially true, even if updating by means of Bayes' Theorem considers any further outcome $\{x_2\}$.

If, however, an S-model produces dual probability based on a premise about $\{x_1, x_2\}$, this represents a totally different result — even if in certain cases the probabilities may be equal for mathematical reasons.

This aspect is a fundamental difference between the Symmetric Theory and those approaches of statistical inference which produce probability statements about parameters (see T. Seidenfeld [1], pp. 139–140).

REFERENCES

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

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

[3] K. Weichselberger. The theory of interval probability, the logical concept of probability and statistical reasoning, in: ISI2005, 55th Session of the International Statistical Institute (CD).

[4] K. Weichselberger and T. Augustin. On the Symbiosis of two Concepts of Conditional Interval Probability, in: J.–M. Bernard, T. Seidenfeld, M. Zaffalon (eds.), ISIPTA03, Proceedings in Informatics 18, Carleton Scientific, 608 – 629, 2003.