Symmetric Probability Theory

Kurt Weichselberger, Munich

Kurt.Weichselberger@stat.uni-muenchen.de

I. The Project.

Symmetric Probability Theory attaches probability to arguments (and not to events!) understood as ordered pairs of propositions.
Argument $A\|B$: A is the conclusion, B is the premise.

Symmetric Probability Theory employs probability statements in statistical inference. A duality of two probability-fields with inverse roles of premises and conclusions produces evaluations of probability $P(B\|A)$ of the same quality as $P(A\|B)$.

Symmetric Probability Theory rests on the *Logical Concept of Probability*.

The Logical Concept of Probability is based on the *Theory of Interval Probability*.

Therefore the Theory of Interval Probability is vital for Symmetric Probability Theory.

Aims of the theory: A comprehensive methodology of probabilistic modeling and statistical reasoning, which makes possible hierarchical modeling with information gained from empirical data.

To achieve the goals of Bayesian approach — but without the pre-requisite of an assumed prior probability.

II. The Theory of Interval Probability

Axioms of Interval Probability

T I - T III = axioms of Kolmogorov

A function obeying Kolmogorov's axioms is named

$$\text{"K-function"}: p(.) \in \mathcal{K}(\Omega_A; \mathcal{A})$$

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

T V:

$$\mathcal{M} = \{p(.) \in \mathcal{K}(\Omega_A; \mathcal{A}) \mid L(A) \leq p(A) \leq U(A), \forall A \in \mathcal{A}\} \neq \emptyset$$

R-probability obeys T IV and T V.

$\mathcal{M}$ is named "the structure" of this R-probability.

T VI:

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

F-probability obeys T IV, T V and T VI.

In case of F-probability: $U(A) = 1 - L(\neg A)$, $\forall\, A \in \mathcal{A}$.

F-probability-field: $\mathcal{F} = (\Omega;\, \mathcal{A};\, L(.))$

Union of F-fields:

Let $\mathcal{F}_i = (\Omega_A;\, \mathcal{A};\, L_i(.))\ i \in I \neq \emptyset$

The union of $\mathcal{F}_i$ is defined as

$$\bigcup_{i\in I} \mathcal{F}_i = \mathcal{F} = (\Omega_A;\, \mathcal{A};\, L(.)) \quad \text{with} \quad \left.\begin{array}{rcl} L(A) & = & \inf_{i\in I} L_i(A) \\ U(A) & = & \sup_{i\in I} U_i(A) \end{array}\right\} \forall\, A \in \mathcal{A}$$

The class of F-fields is closed under the operation "union".

III. The Logical Concept of Probability

F-probability is assigned to certain *pairs of propositions*:

Propositions are elements of epistemic variables. An epistemic variable is the set of sensible answers to a certain question concerning the reality. (Compare with random variables in Classical Theory!!)

An epistemic variable $\mathcal{B}$ is irrelevant for the epistemic variable $\mathcal{A}$, if none of its elements contains any information concerning an element of $\mathcal{A}$.

If epistemic variables $\mathcal{B}_1$, ..., $\mathcal{B}_r$ are irrelevant for $\mathcal{A}$, so is the combined epistemic variable $(\mathcal{B}_1, \ldots, \mathcal{B}_r)$.

Generally relevance of an epistemic variable $\mathcal{B}$ for an epistemic variable $\mathcal{A}$ is a question of *ideology* — in our case it is the question of the *statistical model in use*.

If the proposition B is an element of $\mathcal{B}$ and the proposition A is an element of $\mathcal{A}$ and if $\mathcal{B}$ is relevant for A, then B is relevant for A and $(A\|B)$ is an argument:

A is the conclusion, B is the premise.

Independence of Arguments:

Two arguments $(A_1\|B_1)$ and $(A_2\|B_2)$ are *independent of each other*, iff

$A_1 \in \mathcal{A}_1$, $B_1 \in \mathcal{B}_1$, $A_2 \in \mathcal{A}_2$, $B_2 \in \mathcal{B}_2$ and

the epistemic variable $\mathcal{B}_2$ is irrelevant for the epistemic variable $\mathcal{A}_1$ and

the epistemic variable $\mathcal{B}_1$ is irrelevant for the epistemic variable $\mathcal{A}_2$.

Independence of more than two arguments $(A_i\|B_i)$, $i \in \{1, 2, ..., r\}$:

α) pairwise independence: $(A_i\|B_i)$ and $(A_j\|B_j)$, $i \neq j$, are independent of each other, iff $j \in \{1, ..., r\}$;

β) total independence: the combined epistemic variable $\bigwedge_{j\in I_2}\mathcal{B}_j$ is irrelevant for the combined epistemic variable $\bigwedge_{i\in I_1}\mathcal{A}_i$, if $I_1 \cap I_2 = \emptyset$, $I_1,\ I_2 \subset \{1,\ ...,\ r\}$: the arguments $\left(\bigwedge_{i\in I_1} A_i \middle\| \bigwedge_{i\in I_1} B_i\right)$ and $\left(\bigwedge_{j\in I_2} A_j \middle\| \bigwedge_{j\in I_2} B_j\right)$ are always independent of each other.

It can be shown that pairwise independence produces total independence.

Under these conditions F-probability can be applied to evaluate arguments.

W-fields:

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

$(\Omega_A;\ \mathcal{A})$ is a space of conclusions

$(\Omega_B;\ \mathcal{B})$ is a space of premises, $\mathcal{B}^+ = \mathcal{B} \setminus \{\emptyset\}$.

Axioms

L I: To each $B \in \mathcal{B}^+$ the F-probability-field $\mathcal{F}(B) = (\Omega_A;\, \mathcal{A};\, L(.\|B))$ is attributed.

$\mathcal{M}(B)$ is the structure of $\mathcal{F}(B)$.

L II: Let $B_0 \in \mathcal{B}^+$, $B_i \in \mathcal{B}^+$, $i \in I \neq \emptyset$ with $B_0 = \bigcup_{i \in I} B_i$,

then

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

L III: Let $A_1,\, A_2 \in \mathcal{A}$, $B_1,\, B_2 \in \mathcal{B}$,

$(A_1 \| B_1)$ and $(A_2 \| B_2)$ independent from each other.

Then

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

Multiplicativity is generated by independence — and *not* vice versa.

For the interpretation:

It can be proven: Let the arguments $(A_i \| B_i)$, $i \in \{1, 2, ...\}$ be totally independent, $P(A_i \| B_i) = [L; U]$, $\forall\, i = 1, 2, ...$ (the same probability component for each of the arguments).

Let $\mathcal{C} = (C_1, C_2, ...)$ with $C_i \in \{A_i, \neg A_i\}$, $i = 1, 2, ...$ be a possible sequence of conclusions.

$t(r)$ is the number of A_i among $\{C_1, ..., C_r\}$

$r - t(r)$ is the number of $\neg A_i$ among $\{C_1, ..., C_r\}$.

$$A^{[r]}(\delta) := \left\{ L - \delta \leq \frac{t(r)}{r} \leq U + \delta \right\};\ \delta > 0;\ \ B^{[r]} = \bigcap_{i=1}^{r} B_i.$$

$A^{[r]}(\delta)$ is the union of all conclusions C_i in accordance with the condition for $t(r)$;

$B^{[r]}$ is the combined premise: $L\left(A^{[r]}(\delta) \,\middle\|\, B^{[r]}\right)$ characterizes what can be derived from the premise about $t(r)$.

Then: $\lim_{r\to\infty} L(A^{[r]}(\delta)\|B^{[r]}) = 1, \forall\, \delta > 0.$

Frequency interpretation of the Logical Concept:

Every $P(A\|B) = [L; U]$ may be interpreted, as if $(A\|B)$ was one out of an infinite series of totally independent arguments $(A_i\|B_i)$ where the proportion of arguments for which A came true, is between L and U.

W-field

The set of F-fields generated by singletons out of Ω_B according to Axiom L I, may be viewed as a *family of F-fields* with parameters $y \in \Omega_B$.

Each of the remaining F-fields generated by compound elements B of $\mathcal{B}$, is the union of F-fields generated by singletons.

Classical W-field

Any W-field $\mathcal{W} = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L(.||.))$, for which $\mathcal{M}(y) = \{p_y(.) \in \mathcal{K}(\Omega_A; \mathcal{A})\}, \forall\, y \in \Omega_B$, is named *classical W-field*. Each elementary premise in such W-field produces a classical probability-field. Due to this aspect any classical W-field can be employed as a model of a *family of classical probability-distributions*.

W-support

Let $\mathcal{W} = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L(.||.))$ be a W-field. Any set Y of arguments $(A_i \| B_i)$, $A_i \in \mathcal{A}$, $B_i \in \mathcal{B}^+$, large enough that the information contained in $\{P(A_i \| B_i), \forall\, (A_i \| B_i) \in Y\}$, is sufficient for reconstruction of $\mathcal{W}$, is named *W-support of* $\mathcal{W}$.

IV. Inference

Perfect inference for an ordered family of classical distributions is created through the concept of "dual W-field": If $\mathcal{W}_1$ is a model of the family $\mathcal{F}_1$, the dual W-field $\mathcal{W}_2$ is a model of the family $\mathcal{F}_2$ which is perfect inference to $\mathcal{F}_1$.

Let $\mathcal{F}_1$ be an ordered family of classical distributions with one-dimensional variable and one-dimensional parameter.

1-Concordance

Let $\mathcal{W} = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L(.||.))$ be a classical W-field. Let $\emptyset \neq \mathcal{Z}_A \subsetneqq \mathcal{A}$ be a monotone system of conclusions in $\mathcal{W}$ and $\overline{\mathcal{Z}}_A = \{\neg A |\, A \in \mathcal{Z}_A\}$.

Let $\emptyset \neq \mathcal{Z}_B \subsetneqq \mathcal{B}^+$ be a monotone system of premises in $\mathcal{W}$ and $\overline{\mathcal{Z}}_B = \{\neg B |\, B \in \mathcal{Z}_B\}$, so that for each pair $(A,\, B) \in \mathcal{Z}_A \times \mathcal{Z}_B$ there exists $\alpha_U \in \,]0;\, 1[$ with

$$B = \{y \in \Omega_B |\, p(A\|y) \geq \alpha_U\} \text{ and } \inf_{y \in B} p(A\|y) = \alpha_U.$$

Because of $\neg B = \{y \in \Omega_B |\, p(A\|y) < \alpha_U\}$,

there exists $\alpha_L \leq \alpha_U$ with $\sup_{y \in \neg B} p(A\|y) = \alpha_L$.

Then $(\mathcal{Z}_A \| \mathcal{Z}_B) = \big\{(A\|B)|\; A \in \mathcal{Z}_A \cup \overline{\mathcal{Z}}_A,\, B \in \mathcal{Z}_B \cup \overline{\mathcal{Z}}_B\big\}$ is called *1-Concordance in* $\mathcal{W}$.

Symmetrically one-dimensional W-field

Let $\mathcal{W}$ be a classical W-field. If there exists a 1-Concordance $(\mathcal{Z}_A \| \mathcal{Z}_B)$ in $\mathcal{W}$, which is a W-support of $\mathcal{W}$, then $\mathcal{W}$ is named *symmetrically one-dimensional W-field*.

Any field of this type contains the model of a family of classical probability-distributions with

one-dimensional variable, one-dimensional parameter
and a distribution function $F(x; y)$ monotonically decreasing in the parameter y:
ordered family of one-dimensional distributions.

Such family is named *complete family of type 1*, if the set of parameter-values is distinguished only by the probability law engaged;
it is named *truncated family of type 1*, if the set of parameter-values is distinguished by additional information.

Duality

Let $\mathcal{W}_1 = (\Omega_A;\, \mathcal{A};\, \Omega_B;\, \mathcal{B};\, L_1(.||.))$ be a symmetrically one-dimensional W-field and $\mathcal{W}_2 = (\Omega_B;\, \mathcal{B};\, \Omega_A;\, \mathcal{A};\, L_2(.||.))$ be a W-field.

If there exists a 1-concordance $(\mathcal{Z}_A \| \mathcal{Z}_B)$ in $\mathcal{W}_1$ which is a W-support of $\mathcal{W}_1$

and $(\mathcal{Z}_B \| \mathcal{Z}_A) := \{(B\|A) |\, B \in \mathcal{Z}_B,\, A \in \mathcal{Z}_A\}$ is a W-support of $\mathcal{W}_2$,

where

$$B = \{y \in \Omega_B |\, p_1(A\|y) \geq \alpha\} \text{ and } P_1(A\|B) = [\alpha;\, 1]$$

produces

$$A = \{x \in \Omega_A |\, p_2(B\|x) \geq 1 - \alpha\} \text{ and } P_2(B\|A) = [1 - \alpha;\, 1],$$

then $\mathcal{W}_2$ is the **dual W-field to $\mathcal{W}_1$**.

In the dual W-field $\mathcal{W}_2$ the roles of $(\Omega_A; \mathcal{A})$ and of $(\Omega_B; \mathcal{B})$ are exchanged: $\mathcal{W}_2$ contains probability of inference.

It can be shown:

Let $A \in \mathcal{Z}_A$, $B \in \mathcal{Z}_B$ and $P_1(A\|B) = [\alpha; 1]$.

Then according to Classical Theory, $[1-\alpha; 1]$ is the set of possible *confidence coefficients* for the confidence interval B, if any one of the outcomes $A' \in \mathcal{Z}_A$, $A' \subseteq A$ is employed. Therefore, as far as the elements of the 1-concordance are concerned, dual probability for the argument with conclusion B and premise A in Classical Theory represents the probability that B contains the true value of y generating the information $A' \subseteq A$, $A' \in \mathcal{Z}_A$.

Axiom S I: Let $\mathcal{F}_1$ be a complete family of type 1, represented by the symmetrically one-dimensional W-field $\mathcal{W}_1 = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L_1(.||.))$.
If the W-field $\mathcal{W}_2 = (\Omega_B; \mathcal{B}; \Omega_A; \mathcal{A}; L_2(.||.))$ is dual to $\mathcal{W}_1$, the family of F-probability $\mathcal{F}_2$, generated in $\mathcal{W}_2$ by the elements of Ω_A, yields the *perfect inference* to $\mathcal{F}_1$. □

Since the dual W-field $\mathcal{W}_2$ can be derived from $\mathcal{W}_1$ uniquely, perfect inference $\mathcal{F}_2$ of a family $\mathcal{F}_1$ can be described by means of the respective distribution functions $F_1(x;\, y)$ and $F_2(y;\, x)$.

Let

$$F_1(x_+;\, y_-) = \lim_{h \to 0_+} F_1(x+h;\, y-h); F_1(x_-;\, y_+) = \lim_{h \to 0_+} F_1(x-h;\, y+h)$$

Then:

$$\boxed{F_2(y;\, x) = [1 - F_1(x_+;\, y_-);\, 1 - F_1(x_-;\, y_+)] = [L_2(y;\, x);\, U_2(y;\, x)]}$$

This formula in the general case describes a special type of partially determined F-probability, *standardized cumulative F-probability*: Only the interval-limits

$$\left.\begin{array}{rcl} L_2(y;\, x) &=& L_2\left(]y_L;\, y]\,\|x\right) \\ U_2(y;\, x) &=& U_2\left(]y_L;\, y]\,\|x\right) \end{array}\right\} \forall\, y \in \Omega_B,\, x \in \Omega_A$$

are given.

The evaluation of the information available is determined by two results:

1)

$$
\begin{aligned}
L_2([y_1;\, y_2] \| x) &= \max(0;\, L_2(y_2;\, x) - U_2(y_1;\, x)) \\
U_2([y_1;\, y_2] \| x) &= U_2(y_2;\, x) - L_2(y_1;\, x))
\end{aligned}
$$

2) Let A be a union of pairwise disjoint intervals in Ω_B:

$$
A = \bigcup_{i=1}^{r} A^{(i)},\ A^{(i)} = [l_i;\, u_i],
$$

$$
y_L < l_i < u_i < l_{i+1} < y_U.
$$

Then:

$$
\begin{aligned}
L_2(A \| x) &= \textstyle\sum_{i=1}^{r} L_2\left(A^{(i)} \middle\| \, x\right) \\
U_2(A \| x) &= \textstyle\sum_{i=1}^{r} U_2\left(A^{(i)} \middle\| \, x\right) - \sum_{i=1}^{r-1} \max(0;\, U_2(u_i;\, x) - L_2(l_{i+1};\, x)).
\end{aligned}
$$

(See: Weichselberger, 2001, pp. 416-424)

Continuous Distributions-functions.

If $F_1(x;\, y)$ is continuous as well in x as in y, then:

$$F_2(y;\, x) = 1 - F_1(x;\, y).$$

Perfect inference to the complete family of type 1 $\mathcal{F}_1$, is given by $\mathcal{F}_2$, a complete family of type 1. As long as the premise is a singleton x, evaluation is achieved by classical probability. Compare approaches by R.A. Fisher, D.A.S. Fraser, A. Dempster and Podobnik and Živko.

For a compound premise — created by an imprecise observation — perfect inference is described by cumulative F-probability.

Examples

A. Continuous Distribution-functions

A1 Exact observation: elementary premise for $\mathcal{W}_2$

If a density exists:

$$f_1(x\|\{y\}) = \frac{\partial F_1(x;\, y)}{\partial x}; \quad f_2(y\|\{x\}) = \frac{-\partial F_1(x;\, y)}{\partial y}$$

(Fisher's definition of Fiducial Probability)

If $x - y$ is a pivotal quantity, this leads to

$$f_1(x\|\{y\}) = f_2(y\|\{x\}).$$

(Normal distribution, Cauchy distribution, rectangular distribution)

A2 Interval observation $[x_1, x_2]$
Perfect inference to $\mathcal{F}_1$ is described by a family of standardized cumulative F-probability:

$$P_2\left(]y_L;\, y] \| \, [x_1;\, x_2]\right) = [1 - F_1(x_2;\, y);\, 1 - F_1(x_1;\, y)]$$

Normal distribution with variance σ^2:

$$P_2\left(]-\infty;\, y] \| \, [x_1;\, x_2]\right) =$$
$$= \left[\frac{1}{\sigma\sqrt{2\pi}} \int_{x_2}^{\infty} e^{-\dfrac{(t-y)^2}{2\sigma^2}} dt;\, \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{\infty} e^{-\dfrac{(t-y)^2}{2\sigma^2}} dt\right]$$

$$P_2\left(]y_1;\, y_2] \| \left[x_1;\, x_2\right]\right) =$$

$$\left[\max\left(0; \frac{1}{\sigma\sqrt{2\pi}}\int_{x_2}^{\infty} e^{-\frac{(t-y_2)^2}{2\sigma^2}}dt - \frac{1}{\sigma\sqrt{2\pi}}\int_{x_1}^{\infty} e^{-\frac{(t-y_1)^2}{2\sigma^2}}dt\right);\right.$$

$$\left.\frac{1}{\sigma\sqrt{2\pi}}\int_{x_1}^{\infty} e^{-\frac{(t-y_2)^2}{2\sigma^2}}dt - \frac{1}{\sigma\sqrt{2\pi}}\int_{x_2}^{\infty} e^{-\frac{(t-y_1)^2}{2\sigma^2}}dt\right]$$

B. Discontinuous Distribution-functions

Perfect inference is described by a family of standardized cumulative F-probability

Binomial distribution:

$$F_1(i,\, y) = \sum_{r=1}^{i} \binom{n}{r} y^r (1-y)^{n-r}$$

Exact observation $\{i\}$:

$$\begin{aligned} P_2\left([0;\, y] \| \{i\}\right) &= \left[1 - \sum_{r=1}^{i} \binom{n}{r} y^r (1-y)^{n-r};\; 1 - \sum_{r=1}^{i-1} \binom{n}{r} y^r (1-y)^{n-r}\right] = \\ &= \left[\sum_{r=i+1}^{n} \binom{n}{r} y^r (1-y)^{n-r};\; \sum_{r=i}^{n} \binom{n}{r} y^r (1-y)^{n-r}\right] \end{aligned}$$

$$P_2\left([y_1;\,y_2]\|\{i\}\right) =$$

$$= \left[\max\left(0;\ \sum_{r=i+1}^{n}\binom{n}{r}y_2^r(1-y_2)^{n-r} - \sum_{r=i}^{n}\binom{n}{r}y_1^r(1-y_1)^{n-r}\right);\right.$$

$$\left.\sum_{r=i}^{n}\binom{n}{r}y_2^r(1-y_2)^{n-r} - \sum_{r=i+1}^{n}\binom{n}{r}y_1^r(1-y_1)^{n-r}\right]$$

Interval observation $[i_1;\, i_2]$

$$P_2\left([0;\, y] \| [i_1;\, i_2]\right) =$$

$$= \left[\sum_{r=i_2+1}^{n} \binom{n}{r} y^r (1-y)^{n-r};\, \sum_{r=i_1}^{n} \binom{n}{r} y^r (1-y)^{n-r} \right]$$

$$P_2\left([y_1;\, y_2] \| [i_1;\, i_2]\right) =$$

$$= \left[\max\left(0;\, \sum_{r=i_2+1}^{n} \binom{n}{r} y_2^r (1-y_2)^{n-r} - \sum_{r=i_2}^{n} \binom{n}{r} y_1^r (1-y_1)^{n-r}\right);\right.$$

$$\left. \sum_{r=i_1}^{n} \binom{n}{r} y_2^r (1-y_2)^{n-r} - \sum_{r=i_1+1}^{n} \binom{n}{r} y_1^r (1-y_1)^{n-r} \right]$$

Hypergeometric distribution:
$y =$ number of red balls in the universe of size N
$n =$ sample size

$$F_1(i,\, y) = \sum_{r=1}^{i} \frac{\binom{y}{i}\binom{N-y}{n-i}}{\binom{N}{n}}$$

Exact observation $\{i\}$

$$\begin{aligned} P_2\left([0;\, y]\,\|\{i\}\right) \quad &= \quad \left[1 - \sum_{r=1}^{i} \frac{\binom{y-1}{i}\binom{N-y+1}{n-i}}{\binom{N}{n}}; 1 - \sum_{r=1}^{i-1} \frac{\binom{y}{i}\binom{N-y}{n-i}}{\binom{N}{n}}\right] = \\ &= \quad \left[\sum_{r=i+1}^{\min(y,n)} \frac{\binom{y-1}{i}\binom{N-y+1}{n-i}}{\binom{N}{n}}; \sum_{r=i}^{\min(y,n)} \frac{\binom{y}{i}\binom{N-y}{n-i}}{\binom{N}{n}}\right] \end{aligned}$$

$$P_2\left([y_1;\, y_2] \| \{i\}\right) =$$

$$= \left[\max\left(0;\ \sum_{r=i+1}^{\min(y_2,\,n)} \frac{\binom{y_2-1}{i}\binom{N-y_2+1}{n-i}}{\binom{N}{n}} - \sum_{r=i}^{\min(y_1,\,n)} \frac{\binom{y_1}{i}\binom{N-y_1}{n-i}}{\binom{N}{n}}\right);\right.$$

$$\left.\sum_{r=i}^{\min(y_2,\,n)} \frac{\binom{y_2}{i}\binom{N-y_2}{n-i}}{\binom{N}{n}} - \sum_{r=i+1}^{\min(y_1,\,n)} \frac{\binom{y_1-1}{i}\binom{N-y_1+1}{n-i}}{\binom{N}{n}}\right]$$

Interval observation $[i_1;\, i_2]$

$$P_2\left([0;\, y] \| [i_1;\, i_2]\right) = \left[\sum_{r=i_2+1}^{\min(y,n)} \frac{\binom{y-1}{i}\binom{N-y+1}{n-i}}{\binom{N}{n}};\ \sum_{r=i_1+1}^{\min(y,n)} \frac{\binom{y-1}{i}\binom{N-y+1}{n-i}}{\binom{N}{n}}\right]$$

$$P_2\left([y_1;\, y_2] \| [i_1;\, i_2]\right) =$$

$$= \left[\max\left(0;\ \sum_{r=i_2+1}^{\min(y_2,n)} \frac{\binom{y_2-1}{i}\binom{N-y_2+1}{n-i}}{\binom{N}{n}} - \sum_{r=i_1+1}^{\min(y_1,n)} \frac{\binom{y_1}{i}\binom{N-y_1+1}{n-i}}{\binom{N}{n}}\right);\right.$$

$$\left.\sum_{r=i_1+1}^{\min(y_2,n)} \frac{\binom{y_2-1}{i}\binom{N-y_2+1}{n-i}}{\binom{N}{n}} - \sum_{r=i_2+1}^{\min(y_1,n)} \frac{\binom{y_1-1}{i}\binom{N-y_1+1}{n-i}}{\binom{N}{n}}\right]$$

Truncated Families

Axiom S II:

Let $\mathcal{F}_1$ represented by $\mathcal{W}_1 = (\Omega_A; \mathcal{A}; \Omega_B; \mathcal{B}; L_1(.||.))$, be a complete family of type 1 and $\mathcal{F}_1^*$ be the truncated family produced by the additional statement "$y \in \Omega_B^*$".

Let $\mathcal{F}_2$ represented by $\mathcal{W}_2 = (\Omega_B; \mathcal{B}; \Omega_A; \mathcal{A}; L_2(.||.))$, be the perfect inference to $\mathcal{F}_1$ and

$$\Omega_A^* = \{x \,|\, U_2\,(\Omega_B^*|\; x) > 0\}\,.$$

Perfect inference $\mathcal{F}_2^*$ to $\mathcal{F}_1^*$ is produced by the conditional probability according to $\mathcal{F}_2$ of the argument $(B\|x)$ relative to the argument $(\Omega_B^*\|x)$ for all $B \in \mathcal{B}$, $B \subseteq \Omega_B^*$ and for all $x \in \Omega_A^*$. □

In case that the distribution-function $F_1(x;\, y)$ is continuous in x and y, as long as the premise A is given by an exact observation $\{x\}$, the conditional probability defining $\mathcal{F}_2^*$ is classical too.

It can be shown that in all other cases this axiom requires the employment of the *intuitive concept of conditional F-probability* defined in the following way:

Let $\mathcal{W} = (\Omega_A;\ \mathcal{A};\ \Omega_B;\ \mathcal{B};\ L(.||.))$ be a W-field.

$$\Omega_B^* \subset \Omega_B,\ \mathcal{B}^* = \{B \,|\, B \subseteq \Omega_B^*\} \cap \mathcal{B}.$$
$$\Omega_A^* = \{x \in \Omega_A \,|\, U\left(\Omega_B^* \| x\right) > 0\},\ \mathcal{A}^* = \mathcal{A} \cap \Omega_A^*.$$

Then $\mathcal{W}^* = (\Omega_B^*;\ \mathcal{B}^*;\ \Omega_A^*;\ \mathcal{A}^*;\ L^*(.||.))$ is the conditional W-field with respect to Ω_B^* according to the *intuitive concept of conditional F-probability*, if $\forall\, B \in \mathcal{B}^*$, $\forall\, A \in \mathcal{A}^*$:

$$L^*(B\|A) := L(B|\,\Omega_B^*\|A) = \inf_{p(.\|A) \in \mathcal{M}(\Omega_B^*\|A)} \frac{p(B\|A)}{p(\Omega_B^*\|A)}$$

with

$$\begin{aligned} \mathcal{M}(A) &= \{p(.\|A) \,|\, p(B\|A) \geq L(B\|A),\ \forall\, B \in \mathcal{B}\} \\ \mathcal{M}(\Omega_B^*\|A) &= \{p(.\|A) \in \mathcal{M}(A) |\ p(\Omega_B^*\|A) > 0\} \end{aligned}$$

□

Summarizing: The employment of Axiom S I and Axiom S II produces the perfect inference for all ordered families of classical distributions with one-dimensional conclusions and one-dimensional parameters.

Symmetric Theory of Probability is work in progress. Perfect inference in the case of multi-dimensional models is the next step to be taken.

Preliminary results show that principles can be generalized to higher dimensions and results of familiar types can be expected.

The definite goal is a comprehensive approach
incorporating the methodology of inference into probability theory,
but avoiding explicitly subjectivistic elements like a-priori-distributions.