CHAPTER XIII

Probability sheaves

13.1. Introduction.

We generalise the *probability* between two *events* related by a binary logical operation, to any numbers, which can include different values than 1 for certain and 0 for impossible. To do this I use multilinear interpolation of values of truth tables.

We extend these considerations to truth tables with results of operations which are themselves probabilities. The case for multilinear probabilities is subsumed under the case for polynomials in more than one variable.

We introduce as further developments *hyperintricate probabilities*, *exponentiated probability* and *probability sheaves*.

13.2. Multilinear probabilities from truth tables.

Instead of considering 1 as the number for an absolutely certain event, choose the real number τ to represent this, and instead of 0 for an impossible event, choose the real number υ, $\tau \neq \upsilon$.

If there are n occurrences of an event, where the complete set of possible occurrences is N, this gives an allocation of a probability measure

$$(n/N)\tau + [1 - (n/N)]\upsilon$$

for the event. As linear values these occupy a unique straight line through τ and υ.

A model for the bijection between set theory and logic is to introduce an event $x \in A$, where A is a set.

Suppose the probability of an event C is $c\tau + (1 - c)\upsilon$. For the following truth tables for the logical operations & and OR, we can determine the linear probability of C & D and C OR D, where D is the event with probability $d\tau + (1 - d)\upsilon$. The axiomatic basis of our calculations will be based on interpolating multilinear relations like $r = \alpha + \beta c + \gamma d + \delta cd$ from their values given in truth tables.

C & D	C	D
τ	τ	τ
υ	τ	υ
υ	υ	τ
υ	υ	υ

C OR D	C	D
τ	τ	τ
τ	τ	υ
τ	υ	τ
υ	υ	υ

For the probability of C & D, if we designate this as P(C & D), we have

$$P(C \,\&\, D) = cd\tau + (1 - cd)\upsilon. \tag{1}$$

The linear probability of C OR D can be obtained from C & D using the relations

$$C \text{ OR } D = \text{NOT}(\text{NOT } C \,\&\, \text{NOT } D)$$

and

$$P(\text{NOT } C) = (1 - c)\tau + c\upsilon,$$

so the probability is

$$P(C \text{ OR } D) = (c + d - cd)\tau + (1 - c - d + cd)\upsilon. \tag{2}$$

The logic of *exclusive* OR (XOR) is given by the relation

$$C \text{ XOR } D = (C \text{ OR } D) \,\&\, \text{NOT}(C \,\&\, D).$$

The linear interpolation from the truth table turns out to be

$$P(C \text{ XOR } D) = r\tau + (1 - r)\upsilon = (c + d - 2cd)\tau + (1 - c - d + 2cd)\upsilon.$$

To show how I obtained this, let us look at the truth table for C XOR D:

C XOR D	C	D
υ	τ	τ
τ	τ	υ
τ	υ	τ
υ	υ	υ

So to interpolate, considering $r = \alpha + \beta c + \gamma d + \delta cd$, we have the following equations:

$$0 = \alpha + \beta + \gamma + \delta$$
$$1 = \alpha + \beta$$
$$1 = \alpha \quad\quad + \gamma$$
$$0 = \alpha.$$

The relation *if and only if*, given by the symbol $\Leftrightarrow$, is defined as

$\Leftrightarrow$ is equivalent to NOT XOR.

The logical relations $\Rightarrow$ (implies) and $\Leftarrow$ (is implied by) are defined by

$\Rightarrow$ is equivalent to (NOT C) OR D

$\Leftarrow$ is equivalent to C OR (NOT D),

so for example

$$P(C \Rightarrow D) = (1 - c + cd)\tau + (c - cd)\upsilon.$$

However, although

$$C \Leftrightarrow D = (C \Rightarrow D) \,\&\, (C \Leftarrow D),$$

$(C \Rightarrow D)$ and $(C \Leftarrow D)$ are not independent events, so a substitution does not give a linear probability. Nevertheless, we *can* construct a relation between $\Leftrightarrow$, $\Rightarrow$ and $\Leftarrow$ probabilities:

$$P(C \Leftrightarrow D) = P(C \Rightarrow D) + P(C \Leftarrow D) - \tau.$$

To complete the logical set of relationships, define the seemingly trivial relations

$(C \to D)$ "C gives D"

as resulting in D and

$(C \leftarrow D)$ "C is given by D"

as resulting in C. There is also the relation which gives identically τ for all values.

Together with the NOT relation, which interchanges τ and υ, this gives a complete set of 16 binary logical relations between C and D, and the corresponding linear probabilities.

As a further extension, consider an operation * that acts on C and D probabilistically. We can determine the formula for C * D from the truth table

C * D	C	D
ε_1	τ	τ
ε_2	τ	υ
ε_3	υ	τ
ε_4	υ	υ

where ε_m is the probability $e_m\tau + (1 - e_m)\upsilon$, with $m = 1, \dots 4$, so

$$P(C * D) = X\tau + (1 - X)\upsilon$$

where

$$X = e_4 + (e_2 - e_4)c + (e_3 - e_4)d + (e_1 - e_2 - e_3 + e_4)cd.$$

The * operation can be represented as a linear sum, with non-negative coefficients together summing to 1, of the 16 possible binary logical operations.

13.3. Multipolynomial probabilities from truth tables.

Rather than have probabilities defined under a multilinear relation, we now extend this to multipolynomials, that is, polynomials in more than one variable.

For the unary NOT operation

NOT A	A
υ	τ
τ	υ

we consider the polynomial

$$[1 - \Sigma(i = 0 \text{ to } n)\theta_i c^i]\tau + [\Sigma(i = 0 \text{ to } n)\theta_i c^i]\upsilon,$$

so

$$0 = \Sigma(i = 0 \text{ to } n)\theta_i c^i,$$
$$1 = \theta_0.$$

For a general unary operation

$\sim$ A	A
g_1	τ
g_2	υ

similar considerations apply.

For the binary C * D operation,

$$X = [\Sigma(i = 0 \text{ to } n)\lambda_i c^i][\Sigma(j = 0 \text{ to } n)\mu_j d^j]$$
$$= \lambda_0\mu_0 + \mu_0\Sigma(i = 1 \text{ to } n)\lambda_i c^i + \lambda_0\Sigma(j = 1 \text{ to } n)\mu_j d^j$$
$$+ [\Sigma(i = 1 \text{ to } n)\lambda_i c^i][\Sigma(j = 1 \text{ to } n)\mu_j d^j],$$

with

$$e_4 = \lambda_0\mu_0$$
$$(e_2 - e_4) = \mu_0\Sigma(i = 1 \text{ to } n)\lambda_i c^i$$
$$(e_3 - e_4) = \lambda_0\Sigma(j = 1 \text{ to } n)\mu_j d^j$$
$$(e_1 - e_2 - e_3 + e_4) = [\Sigma(i = 1 \text{ to } n)\lambda_i c^i][\Sigma(j = 1 \text{ to } n)\mu_j d^j].$$

It is possible to consider multipolynomial probability logics where

$$P(C * D) = X\tau + (1 - Y)\upsilon$$

and Y is also a multipolynomial, not necessarily identical to X except at the τ, υ boundary for C * D. These multipolynomial probability logics are generally non-commutative.

The point of view of our work when it was conceived 35 years ago was that τ and υ are themselves in polynomial format.

We can also consider the case with λ_i, c^i, μ_j and d^j matrices for X and Y, or indeed for τ and υ. In the latter case, τ and υ may be thought of as split into their hyperintricate basis components, so we are dealing with multi-valued logics.

13.4. Hyperintricate multivalued probabilities.

We can consider vectors τ and υ. For instance, if we have two components for τ, given by τ_A and τ_B and two components for υ, namely υ_A and υ_B, then for two events C and D we have an isomorphism with a truth table in four events – J, K, L and M. We demonstrate this with portions of the two truth tables below.

An example (with multiplication) is where τ_A and υ_A are real and τ_B and υ_B are imaginary.

We can extend this idea further to matrix τ's and υ's, via the hyperintricate representation. In particular, for a unique representation of 2 × 2 matrices we obtain the intricate representation $\tau = \tau_A + \tau_B i + \tau_\alpha\alpha + \tau_\phi\phi$.

C		D	
τ_A	τ_B	τ_A	τ_B
τ_A	τ_B	τ_A	υ_B
τ_A	τ_B	υ_A	τ_B
τ_A	τ_B	υ_A	υ_B
τ_A	υ_B	τ_A	τ_B
τ_A	υ_B	τ_A	υ_B
τ_A	υ_B	υ_A	τ_B
τ_A	υ_B	υ_A	υ_B

J	K	L	M
τ	τ	τ	τ
τ	τ	τ	υ
τ	τ	υ	τ
τ	τ	υ	υ
τ	υ	τ	τ
τ	υ	τ	υ
τ	υ	υ	τ
τ	υ	υ	υ

13.5. Exponentiated probabilities and the exponential map.

Let B: A $\rightarrow$ P(A) be a mapping. We allow the formation of $P^2(A)$, the "probability derived from a probability", and $P^n(A)$. The NOT operation for linear probabilities

$$a\tau \rightarrow a\upsilon, \qquad (1 - a)\upsilon \rightarrow (1 - a)\tau$$

satisfies NOT NOT = id, where id is the identity operation, giving

$$P[NOT(NOT\ A)] = P(A).$$

We will denote this NOT mapping by B^{op}. There is also the crossover transformation, $B^c = B^{cop}$, which swaps the order of the pair (τ, υ). For linear probabilities this is

$$a\tau \rightarrow (1 - a)\upsilon$$

and

$$(1 - a)\upsilon \rightarrow a\tau,$$

or equivalently

$$a\upsilon \rightarrow (1 - a)\tau.$$

For multipolynomial probabilities the solution set of P(NOT A) can range over various values for different mappings of NOT. In particular, for two such maps $P_p(NOT\ A) \equiv P(NOT_p\ A)$ and $P_q(NOT\ A) \equiv P(NOT_q\ A)$, in general we have outside the (τ, υ) boundary

$$P_q(NOT_q(P_p(NOT_p\ A))) \equiv P(NOT_q(NOT_p\ A)) \neq P(A).$$

Thus on the (τ, υ) boundary the polynomial logic is Boolean, and outside of this it is intuitionalistic, as defined in *What is mathematics* of the Prologue.

Consider the commutative diagram

$$\begin{tikzcd}
C \times D \arrow[r] \arrow[d] & P(C\ OR\ D) \arrow[d] \arrow[dr, "NOT"] & \\
C^{op} \times D^{op} \arrow[r] & P(C^{op}\ OR\ D^{op}) \arrow[r] & \equiv P(NOT_c C\ OR\ NOT_d D)
\end{tikzcd}$$

We have introduced a mapping of maps, called a functor, whose domain mapping is the OR operation and whose codomain mapping is the NOT & operation.

For linear probabilities

$$P(\text{NOT } C \text{ OR NOT } D) = P(\text{NOT } (C \,\&\, D)),$$

and we are mapping

$$P(C) + P(D) - P(C)P(D) \xrightarrow{\text{OR}} 1 - P(C)P(D) \xrightarrow{\text{NOT \&}} P(C)P(D),$$

whereas for polynomial probabilities we define

$$P(\text{NOT}_c\ C\ \text{OR}_{c,d}\ \text{NOT}_d\ D) = P(\text{NOT}_{c,d}\ (C\ \&_{c,d}\ D)),$$

this idea being extendible to multiple events C, D, ... E.

It is then possible to form the sum of P(C) and P(D), which is an operation of an abelian group that is not a probability within the limits $c\tau + (1 - c)\upsilon$, $0 \le c \le 1$, that is, as

$$P(C \text{ OR } D) + P(C \,\&\, D).$$

For the mapping

$$P(C \text{ OR } D) + P(C \,\&\, D) \rightarrow P(C \,\&\, D), \qquad (1)$$

the left hand side deals with the sum, so “inside range” the intersection of C and D is limited to be within the values c + d = 0 to 1, but now that we can deal with addition directly in terms of logical connectives “outside range”, the above mapping of (1) transforms the additive kernel 0 on the left to the multiplicative kernel 1 of P(C & D) on the right, *and the sum does not have to be disjoint or within range*.

In certain circumstances, in particular when P(C) and P(D) are complex numbers, it is possible to obtain again a multiplicative function from the sum by an exponential map f from

$$P(C) + P(D) \rightarrow f[P(C)P(D)]$$

given by

$$e^{P(A) + P(B)} = e^{P(A)}e^{P(B)},$$

because an abelian exponential map is of the form

$$(G + H) \rightarrow e{\uparrow}(G + H) = (e{\uparrow}G)(e{\uparrow}H)$$

and

$$(JK) \rightarrow e{\uparrow}(JK) = (e{\uparrow}J){\uparrow}K.$$

When this function is non-commutative, we do not have these equalities on the right hand side. A consistent evaluation is then given in chapter XIV.

This enables us to evaluate an exponential map from P(C * D).

13.6. The hyperintricate probability sheaf.

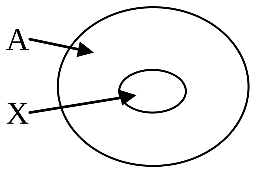
A *sheaf* is a tool for systematically tracking locally defined data attached to the open sets of a topological space. The data can be restricted to smaller open sets, and the data assigned to an open set is equivalent to all collections of compatible data assigned to collections of smaller open sets covering the original one. For example, such data can consist of the rings of continuous or infinitely differentiable Eudoxus-valued functions defined on each open set.

Sheaves are by design quite general and abstract objects, and their correct definition is rather technical. There exist several varieties such as sheaves of sets or sheaves of rings, depending on the type of data assigned to open sets.

There are also maps from one sheaf to another.

Due to their general nature and versatility, sheaves have several applications in topology and especially in algebraic and differential geometry. First, geometric structures such as that of a differentiable manifold can be expressed in terms of a sheaf of rings on the space. In such contexts several geometric constructions such as vector bundles or divisors are naturally specified in terms of sheaves. Second, sheaves provide the framework for a very general cohomology theory, which encompasses also topological cohomology theories. Especially in algebraic geometry and the theory of complex manifolds, sheaf cohomology provides a powerful link between topological and geometric properties of spaces.

It is usual in category theory to replace $x \in A$, where x is an element of a set A, by an arrow. If we replace x by the set X, then in mZFC if sets X and A are not empty, then the logic operator $X \Leftarrow A$ is mapped to the existence of an $X \subset A$ (its only everywhere false value, which corresponds to the empty set $\varnothing$, can be shown from the inclusion diagram



for $X \subset A$).

The ideas of logic and of sets can be conflated. We replace each logical operation $C * D$ bijectively by the statement $(x \in s(C)) * (x \in s(D))$, where there is a bijection between the statement C and the set s(C). The interpretation for sets of * is given directly by the above relation. Thus C & D maps to $s(C) \cap s(D)$ and $C \Rightarrow D$ is mapped to $s(C) \supset s(D)$.

We now have

$$s(C) * s(D) \leftrightarrow C * D \leftrightarrow P(C * D).$$

Thus the operations + and $\times$ of a noncommutative ring map directly to probability logic operations taken outside of range for OR and & respectively given by (1) and (2) of section 2, and their respective set operations for $\cup$ and $\cap$, with NOT mapping to set complement. We can introduce the inverse polynomial in the abelian case, which in the extended case of hyperintricate multipolynomials is a minimal model for associative noncommutative rings. By this means all set operations are transformed to operations in arithmetic and extensions of this idea.

Since we now have a complement and an intersection operation identified with a polynomial ring, which can be hyperintricate, we can introduce noncommutative sheaves.

In order to consider mappings of sets

$$\begin{array}{ccc} A & \rightarrow & B \\ \downarrow & & \downarrow \\ C & \rightarrow & D \end{array}$$

it is convenient to stipulate the maximal partition of $A \cup B \cup C \cup D$ in terms of derived subsets.

Theorem 12.6.1. *The intersection between each of the following* 16 *subsets in the partition of* $A \cup B \cup C \cup D$ *is* $\varnothing$, *and the union of the* 16 *is* $A \cup B \cup C \cup D$.

(1) $\varnothing$
(2) $A \setminus (B \cup C \cup D)$
(3) $B \setminus (A \cup C \cup D)$
(4) $C \setminus (A \cup B \cup D)$
(5) $D \setminus (A \cup B \cup C)$
(6) $(A \cap B) \setminus (C \cup D)$
(7) $(A \cap C) \setminus (B \cup D)$
(8) $(A \cap D) \setminus (B \cup C)$
(9) $(B \cap C) \setminus (A \cup D)$
(10) $(B \cap D) \setminus (A \cup C)$
(11) $(C \cap D) \setminus (A \cup B)$
(12) $(A \cap B \cap C) \setminus D$
(13) $(A \cap B \cap D) \setminus C$
(14) $(A \cap C \cap D) \setminus B$
(15) $(B \cap C \cap D) \setminus A$
(16) $(A \cap B \cap C \cap D)$. □

The extension to more than four sets is obvious.

It follows that each stable mapping is an automorphism of one or more of these partitioned subsets.

13.7. Exercises.

(A) *A nuclear power plant has two processes. The probability of the first process failing is* $A = a\tau + (1 - a)\upsilon$, *and of the second failing is* $B = b\tau + (1 - b)\upsilon$. *There is a* $1/5^{th}$ *certainty that if* A *fails then* B *will. What is the probability that* A *and* B *will happen together*?