PROBABILITIES, OBSERVATIONS AND PREDICTIONS

A. H. COPELAND, Sr.
THE UNIVERSITY OF MICHIGAN

1. A formalization of statistical reasoning

The purpose of this paper is to formalize statistical reasoning, or, more generally, to formalize the type of reasoning employed in experimental science. It is natural to question whether or not such formalization has already been developed in modern logic. The answer, in the opinion of the author, is no. An essential part of the reasoning in experimental science is concerned with the making of predictions and deciding whether certain experimental evidence confirms these predictions, is at variance with them, or is indecisive with regard to them. Experimental confirmation is not even expressible in present day logic.

It can happen that some evidence confirms a given prediction whereas other evidence is at variance with it. Thus there is always uncertainty attached to our predictions even after they have been experimentally verified. This uncertainty suggests the advisability of introducing probabilities into our formalization. It has been popular in recent years to define probability as a measure. That is, we consider a nonnegative measure function defined over a field of subsets of some space such that the measure of the entire space is unity. A probability is then a measure assigned to one of the sets. There are of course infinitely many measure functions which can be chosen except in the most trivial case, and this concept of probability does not suggest any criterion for preferring one choice to another. In experimental science, on the other hand, the choice of measure is of vital importance, and our formalization must take this fact into account.

2. Predictions and observations

Since a prediction is a sentence we shall let sentences be the elements of our formal system, and shall denote them by small *italic* letters with or without subscripts. We shall combine sentences in the usual manner by the Boolean operators. We shall denote the operators "and," "or," "not" respectively by $\wedge$, $\vee$, $\sim$. We shall denote the complete disjunctive operator by $+$. It is defined by the equation $x + y = (x \wedge \sim y) \vee (y \wedge \sim x)$. That is, $x + y$ is interpreted to mean "x or y, but not both." Thus the basis of our system is a Boolean algebra (or Boolean ring), B. We shall denote the zero element of B by 0 and the unit element by 1.

We shall not regard all of the elements of B as predictions. In particular, we shall never predict the element 0. We shall predict only those elements which are sufficiently probable. An assignment of probabilities to the elements of B together with a decision as to the confidence level will automatically determine which of the elements are predictions.

It remains to indicate how observations can be formalized. Consider two predictions x, y and suppose that after these predictions have been made, x is observed to have occurred, that is, x has been verified. We should then reappraise y. That is, we should now concern ourselves with "y on the assumption that x has been verified." This is, of course, the conditional. All of the elements of B should be reappraised in this manner, that is, they should be replaced by the corresponding conditionals. Thus an observation is a mapping of B into a new Boolean algebra B' whose elements are conditionals. In particular, x maps into 1.

3. Conditionals

Consider two sentences x, y which can be stated in the form of predictions. Then if $x \neq 0$ it is legitimate to make the prediction "y if x" and hence "y if x" should be an element of B if x and y are elements of B and x is not 0. We shall let $\subset$ denote the operator "if" and $y \subset x$ denote the sentence "y if x." Then $y \subset x$ is a conditional and, in fact, in computing its probability it is customary to regard it as meaning the same thing as the sentence "y on the assumption that x has been verified." We shall identify these two meanings of the conditional.

If an element x is verified by observation then an element y is mapped into $y \subset x$. The new Boolean algebra which is produced by this mapping is denoted by $B \subset x$. Thus the verification of x is the mapping $B \to B \subset x$. An observation on x is either a verification of x or a verification of $\sim x$. Let us investigate what happens when two successive verifications occur on two elements x, y of B. After the verification $B \to B \subset x$ the element y becomes $y \subset x$ and the verification of y becomes the verification of $y \subset x$, that is, the mapping $B \subset x \to (B \subset x) \subset (y \subset x)$. The resultant of the two verifications maps an element z into $(z \subset x) \subset (y \subset x)$. On the other hand the resultant of the two verifications is equivalent to a verification of $x \wedge y$ and hence z maps into $z \subset (x \wedge y)$. Thus we must have the equation

$$(z \subset x) \subset (y \subset x) = z \subset (x \wedge y)\,. \tag{3.1}$$

The resultant of the two mappings is the mapping $B \to B \subset (x \wedge y)$.

We shall require the conditional to be such that the sentences x and $y \subset x$ determine the sentence $y \wedge x$. Hence if $y \subset x = z \subset x$ then we must have $y \wedge x = z \wedge x$. Finally, if x, y are elements of B and $x \neq 0$ then we require the existence of an element z such that $z \subset x = y$. This existence requirement is not essential for the present discussion but it is harmless and does result in a somewhat more elegant set of properties of the conditional. The following are the postulates for the conditional.

P3.1. $x, y \in B$ and $x \neq 0 \Rightarrow y \subset x \in B$.

P3.2. $x \neq 0 \Rightarrow x \subset x = 1$.

P3.3. $(\sim y) \subset x = \sim(y \subset x)$ if $x \neq 0$.

P3.4. $(y \subset x) \wedge (z \subset x) = (y \wedge z) \subset x$ if $x \neq 0$.

P3.5. $z \subset (x \wedge y) = (z \subset x) \subset (y \subset x)$ if $x \wedge y \neq 0$.

P3.6. $y \subset x = z \subset x$ and $x \neq 0 \Rightarrow y \wedge x = z \wedge x$.

P3.7. $x, y \in B$ and $x \neq 0 \Rightarrow$ there exists z such that $z \subset x = y$.

We can interpret $y \subset x$ as the sentence "y is implied by x" or as "x implies y." However, $y \subset x$ is not equivalent either to material or strict implication. In fact the operator $\subset$ cannot be defined in terms of the Boolean operators. A Boolean algebra which contains an operator $\subset$ satisfying the above seven postulates is said to be implicative.

Henceforth whenever we write an expression such as $y \subset x$ it will be understood that $x \neq 0$. The following five theorems are easily proved.

T3.1. $(y \subset x) \vee (z \subset x) = (y \vee z) \subset x$.

T3.2. $(y \subset x) + (z \subset x) = (y + z) \subset x$.

T3.3. $(y \subset x) = (x \wedge y) \subset x$.

T3.4. $1 \subset x = 1$.

T3.5. $x \subset 1 = x$.

For $x \subset 1 = x \subset (1 \wedge 1) = (x \subset 1) \subset (1 \subset 1) = (x \subset 1) \subset 1$ and hence $x = x \subset 1$.

The following definition introduces a new operator denoted by $\times$.

D3.1. (a) $z \subset x = y$ *and* $z \wedge x = z \Leftrightarrow z = x \times y$ *when* $x \neq 0$.

(b) $x \times y = 0$ *when* $x = 0$.

T3.6. $x \times y$ *is unique.*

T3.7. $x \times (y \subset x) = x \wedge y$.

Theorem 3.7 enables us to interpret $\times$ as the operator by means of which we can combine the sentences x and $y \subset x$ to produce the sentence $x \wedge y$.

T3.8. $x \wedge (x \times y) = x \times y$.

T3.9. $x \times y = x \times z$ *and* $x \neq 0 \Rightarrow y = z$.

For let $u \subset x = y, v \subset x = z$. Then $x \wedge u = x \times y = x \times z = x \wedge v$ and $y = u \subset x = (x \wedge u) \subset x = (x \wedge v) \subset x = v \subset x = z$.

T3.10. $x \times (y + z) = (x \times y) + (x \times z)$.

T3.11. $x \times (y \wedge z) = (x \times y) \wedge (x \times z)$.

T3.12. $x \times 0 = 0$.

T3.13. $x \times 1 = 1 \times x = x$.

The equations of T3.13 are obvious when $x = 0$. If $x \neq 0$ then $x \times 1 = x \times (x \subset x) = x \wedge x = x$. Moreover $1 \times x = z$ where $z = z \subset 1 = x$.

T3.14. $x \times y = 0 \Rightarrow x = 0$ *or* $y = 0$.

T3.15. $(x \subset y) \subset z = x \subset (y \times z)$.

For let $u \subset y = z$. Then $u \wedge y = y \times z$ and $x \subset (y \times z) = x \subset (u \wedge y) = (x \subset y) \subset (u \subset y) = (x \subset y) \subset z$.

T3.16. $x \times (y \times z) = (x \times y) \times z$.

For let $u \subset (x \times y) = z = (u \subset x) \subset y$. Then $u \wedge (x \times y) = (x \times y) \times z$ and $(u \subset x) \wedge y = y \times z$. Hence $x \times (y \times z) = x \times ((u \subset x) \wedge y) = (x \times (u \subset x)) \wedge (x \times y) = x \wedge u \wedge (x \times y) = u \wedge (x \times y) = (x \times y) \times z$.

4. The ring-like character of implicative Boolean algebra

We shall give an alternative set of postulates for implicative Boolean algebra using only the operators $+$ and $\times$. In terms of these two operators we shall then define the remaining Boolean operators and the operator $\subset$. These postulates will display the ring-like character of implicative Boolean algebra.

P4.1. B is an Abelian group with respect to 0.

P4.2. $x \in B \Rightarrow x + x = 0$.

P4.3. B is a semigroup with respect to $\times$, 1.

P4.4. $x \times y = x \times z$ and $x \neq 0 \Rightarrow y = z$.

P4.5. $x \times (y + z) = (x \times y) + (x \times z)$.

P4.6. $0 \times x = 0$.

P4.7. $x \times y = 1 \Rightarrow y = 1$.

P4.8. $x, y \in B \Rightarrow$ there exist $x', y', z \in B$ such that $x \times y' = y \times x'$, $(1 + x) \times z = y \times (1 + x')$.

P4.9. $x, y, x', y', z, x'', y'' \in B$ and $x \neq 0$ and $x \times y' = y \times x'$ and $(1 + x) \times z = y \times (1 + x')$ and $x \times y'' = y \times x'' \Rightarrow$ there exists u such that $y'' = y' \times u$.

These postulates can be satisfied by a system in which $1 = 0$, but in this case all of the elements are 0. In the following theorems it will be assumed that this trivial case is excluded. In P4.3 it is understood that 1 acts as a unit when operating on the right.

T4.1. $1 \times x = x$ *for* $1 \times x = 1 \times (1 \times x)$ *and hence* $x = 1 \times x$.

T4.2. $x \times y = 1 \Rightarrow x = y = 1$.

T4.3. $x, y \in B \Rightarrow$ *there exist unique* x', y' *such that*

(a) $x \times y' = y \times x'$.

(b) $x \times y'' = y \times x'' \Rightarrow$ *there exist* u, v *such that* $x'' = x' \times u$, $y'' = y' \times v$.

In the proof we consider separately three different cases.

$$\begin{aligned} \text{I.}\quad & x = y = 0 \Rightarrow x' = y' = 1. \\ \text{II.}\quad & x = 0, y \neq 0 \Rightarrow x' = 0, y' = 1. \\ & x \neq 0, y = 0 \Rightarrow x' = 1, y' = 0. \\ \text{III.}\quad & x \neq 0, y \neq 0. \end{aligned}$$

The existence of x', y' follows from P4.8, P4.9. To prove the uniqueness suppose that there exist a pair x', y' and another pair x'', y'' both satisfying (a) and (b). Then there exist u, u' such that $x'' = x' \times u$, $x' = x'' \times u' = x' \times (u \times u') = x' \times 1$. Then $x' = 0 \Rightarrow x'' = 0 = x'$ and $x' \neq 0 \Rightarrow u \times u' = 1 \Rightarrow u = u' = 1 \Rightarrow x'' = x'$. Similarly $y'' = y'$.

We have the following definitions.

D4.1. $x \wedge y = x \times y' = y \times x'$ *where* x', y' *are defined as in* T4.3.

D4.2. $x \subset y = x'$ *if* $y \neq 0$, $y \subset x = y'$ *if* $x \neq 0$ *where* x', y' *are defined as in* T4.3.

T4.4. $x \wedge x = x$.

T4.5. $x \wedge y = y \wedge x$.

T4.6. $(x \wedge y) \wedge z = x \wedge (y \wedge z)$.

For let

(a) $x \wedge y = x \times y' = y \times x'$,

(b) $(x \wedge y) \wedge z = x \times y' \times z' = y \times x' \times z' = z \times u$,

(c) $y \wedge z = y \times z'' = z \times y''$,

(d) $x \wedge (y \wedge z) = x \times v = y \times z'' \times x'' = z \times y'' \times x''$.

Then (b), (c) $\Rightarrow$ there exists x''' such that $u = y'' \times x'''$. Hence

(e) $x \times y' \times z' = z \times y'' \times x'''$.

But (d), (e) $\Rightarrow$ there exists s such that $x''' = x'' \times s$. Hence

(f) $(x \wedge y) \wedge z = (x \wedge (y \wedge z)) \times s$.

But $x \wedge (y \wedge z) = (y \wedge z) \wedge x$ and $(x \wedge y) \wedge z = z \wedge (x \wedge y)$. Therefore, there must exist t such that

(g) $x \wedge (y \wedge z) = ((x \wedge y) \wedge z) \times t$.

Theorem 4.6 is then a simple consequence of (f), (g).

T4.7. $x \times (y \wedge z) = (x \times y) \wedge (x \times z)$.

For if $x = 0$ then $x \times (y \wedge z) = 0 = (x \times y) \wedge (x \times z)$. If $x \neq 0$ let $y \wedge z = y \times z' = z \times y'$. Then $x \times (y \wedge z) = x \times y \times z' = x \times z \times y'$. Moreover $x \times y \times z'' = x \times z \times y'' \Rightarrow y \times z'' = z \times y'' \Rightarrow$ there exist u, v such that $y'' = y' \times u$, $z'' = z' \times v$. Therefore $x \times (y \wedge z) = (x \times y) \wedge (x \times z)$.

D4.3. $x \vee y = 1 + (1 + x) \wedge (1 + y)$.

T4.8. $x \vee y = x + y + (x \wedge y)$.

T4.9. $x \times (y \vee z) = (x \times y) \vee (x \times z)$.

For $x \times (y \vee z) = x \times (y + z + (y \wedge z)) = (x \times y) + (x \times z) + ((x \times y) \wedge (x \times z)) = (x \times y) \vee (x \times z)$.

T4.10. $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$.

For $x \wedge (y \vee z) = x \wedge (1 + (1 + y) \wedge (1 + z)) = x + (x \wedge (1 + y) \wedge (1 + z)) = (x \wedge y) + (x \wedge z) + (x \wedge y \wedge z) = (x \wedge y) \vee (x \wedge z)$.

T4.11. $x + y = (x \wedge (1 + y)) \vee (y \wedge (1 + x))$.

T4.12. $x \wedge (y + z) = (x \wedge y) + (x \wedge z)$.

For $x \wedge (y + z) = x \wedge ((y \wedge (1 + z)) \vee (z \wedge (1 + y))) = (x \wedge y \wedge (1 + z)) \vee (x \wedge z \wedge (1 + y)) = (x \wedge y \wedge (1 + (x \vee z))) \vee (x \wedge z \wedge (1 + (x \wedge y))) = (x \wedge y) + (x \wedge z)$.

T4.13. $x \wedge (x \times y) = x \times y$.

For $(x \times y) \times 1 = x \times y$.

T4.14. $x \times x = x \Rightarrow x = 0$ *or* $x = 1$.

With the aid of these theorems it is not difficult to verify that a system B satisfying the above nine postulates is a Boolean algebra provided $\wedge$, $\vee$ are defined by D4.1, D4.3 and $\sim x$ is defined to be $1 + x$. Moreover, the operator $\subset$ defined by D4.2 satisfies postulates P3.1 to P3.7.

From T4.13, T4.14 it follows that a nontrivial implicative Boolean algebra is atomless. In fact every element is infinitely divisible. Since there exist atomic Boolean algebras, the operator $\subset$ cannot be defined in terms of the Boolean operators. On the other hand, it is known that an arbitrary Boolean algebra can be imbedded in an implicative Boolean algebra and hence the introduction of the operator $\subset$ produces no essential restriction. In a paper as yet unpublished Professor Richard Büchi shows how to construct a natural model for every implicative Boolean algebra.

5. Probabilities

We shall introduce the following postulates for probabilities:

P5.1. $x \in B \Rightarrow p(x)$ is a nonnegative real number.

P5.2. $x \wedge y = 0 \Rightarrow p(x + y) = p(x) + p(y)$.

P5.3. $p(1) \neq 0$.

P5.4. $p(x \times y) = p(x)p(y)$.

The following theorems are easily proved.

T5.1. $p(1) = 1$.

T5.2. $0 \leqq p(x) \leqq 1$.

T5.3. $p(x \vee y) = p(x) + p(y) - p(x \wedge y)$.

T5.4. $p(x \wedge y) = p(x)\, p(y \subset x) = p(y)\, p(x \subset y)$.

For $p(x \wedge y) = p[x \times (y \subset x)] = p(x)\, p(y \subset x)$.

Independence is defined as follows.

D5.1. $x_1, x_2, \cdots, x_n$ *are independent provided*

$$p(x_1 \wedge x_2 \wedge \cdots \wedge x_n) = p(x_1 \times x_2 \times \cdots \times x_n) \tag{5.1}$$

and provided a similar condition holds for every subset of these elements.

6. Development of the formalization

We have seen that the verification of an element x is the mapping $B \rightarrow B \subset x$. The verification of x followed by the verification of y (where both x and y are regarded as elements of B) consists in the two mappings

$$B \rightarrow B \subset x \rightarrow (B \subset x) \subset (y \subset x) = B \subset (x \wedge y)\,. \tag{6.1}$$

The verification of y followed by the verification of x consists in

$$B \to B \subset y \to (B \subset y) \subset (x \subset y) = B \subset (x \wedge y) . \tag{6.2}$$

Hence verification is commutative. Clearly it is also associative.

The verification $B \to B \subset x$ maps y into $y \subset x$. If $y \subset x = 1$ and the verification of x is followed by the verification of y then the second verification is the identity mapping and hence the verification of x is equivalent to the verification of x followed by the verification of y. We then say that the verification of x produces the verification of y. For the verification of an element x to produce the verification of an element y it is necessary and sufficient that $x \wedge y = x$ and also necessary and sufficient that x strictly imply y. The verification of an element x is said to produce an observation on an element y if it produces the verification either of y or of $\sim y$.

Suppose that $B \to B \subset a$ produces an observation on x and an observation on y. Then this mapping produces a verification of $x \wedge y$ if it produces a verification of x and of y. Otherwise it produces a verification of $\sim(x \wedge y)$. Furthermore, the mapping produces a verification of $x \vee y$ if it produces a verification of x or of y or of both. Otherwise it produces a verification of $\sim(x \vee y)$. However, it is possible to produce a verification of $x \vee y$ without producing an observation either on x or on y and it is possible to produce a verification of $\sim(x \wedge y)$ without producing an observation either on x or on y. If a mapping $B \to B \subset a$ produces a verification of $x \vee y$ but does not produce observations on both x and y then we can follow the mapping by another mapping

$$B \subset a \to (B \subset a) \subset b = B \subset (a \times b) \tag{6.3}$$

which produces observations on both x and y. Since the first mapping carries $x \vee y$ into 1 the mapping $B \to B \subset (a \times b)$ must carry $x \vee y$ into 1 and hence must produce a verification of x or of y or of both. A verification of $\sim(x \wedge y)$ which does not produce observations on both x and y can be similarly treated.

An observation on $y \subset x$ shall mean a verification of x followed by an observation on y. The verification of x maps y into $y \subset x$ and the observation on y then maps $y \subset x$ into 1 or 0. If $y \subset x$ maps into 1 it is said to be verified and if it maps into 0 then $\sim(y \subset x)$ is verified. No mapping which produces a verification of $\sim x$ shall be regarded as producing an observation on $y \subset x$. If a mapping $B \to B \subset a$ produces a verification of x and an observation on y then it can be regarded as the resultant of the two mappings

$$B \to B \subset x \to (B \subset x) \subset (a \subset x) = B \subset (x \wedge a) = B \subset a . \tag{6.4}$$

The second mapping carries $y \subset x$ into 1 or 0.

Next consider a mapping $B \to B \subset a$ which produces observations on both x and y. Then x and y map into $x \subset a$ and $y \subset a$. Consider the element

$$(y \subset a) \subset (x \subset a) = y \subset (x \wedge a) . \tag{6.5}$$

If x is verified by the mapping then $y \subset (x \wedge a) = y \subset a$ and y or $\sim y$ is verified according as $y \subset (x \wedge a)$ is 1 or 0. If $\sim x$ is verified then $y \subset (x \wedge a)$ is undefined. Thus we obtain a verification of $y \subset x$ or verification of $\sim(y \subset x)$ or no observation on $y \subset x$ according as $(y \subset a) \subset (x \subset a)$ is 1 or 0 or undefined. It would have been much more satisfying if we could have characterized observations on $y \subset x$ in terms of the possible values of the element

$$(y \subset x) \subset a = y \subset (x \times a) , \tag{6.6}$$

but unfortunately such a characterization could not give us an accurate picture of the way in which observations on $y \subset x$ are actually made. It is not in general true that $y \subset (x \times a)$ is equal to $y \subset (x \wedge a)$.

We consider next the formalization of predictions. We start by assigning probabilities to the elements of B in accordance with P5.1 to P5.4. Then we choose a number λ such that $0 < \lambda < 1$. We predict every element x of B for which $\lambda \leqq p(x)$. Thus the assignment of probabilities and the selection of λ produce the formalization of predictions. The usual practice is to choose $\lambda = .95$.

It remains to decide what can be said of those elements x for which $p(x) < \lambda$. We concern ourselves with success ratios of finite sets of such elements rather than with the individual elements and hence we are led to the problem of formalizing the concept of success ratio. If we are given a set of elements $x_1, x_2, \cdots, x_n$ then we can form the element

$$(6.7) \qquad y_r = \bigcup_{0<k_1<k_2<\cdots<k_r<n} x_{k_1} \wedge x_{k_2} \wedge \cdots \wedge x_{k_r} \quad \text{if} \quad 0 < r \leqq n .$$

We also let $y_r = 1$ if $r < 1$ and $y_r = 0$ if $n < r$. If a verification $B \to B \subset a$ produces observations on all of the elements $x_1, x_2, \cdots, x_n$ then the success ratio for these elements is at least r/n if and only if y_r is verified. The success ratio is exactly r/n if and only if $y_r \wedge \sim y_{r+1}$ is verified. Next let I be any interval and let

$$(6.8) \qquad z_I = \bigcup_{r/n \in I} y_r \wedge \sim y_{r+1} .$$

Then z_I is the element which is verified if and only if the success ratio lies in I.

Suppose that $x_1, x_2, \cdots, x_n$ are independent, that each $p(x) = p$ and that p is interior to I. It is well known that a sufficiently large n will then produce a z_I for which $p(z_I) \geqq \lambda$. In this case we predict z_I, that is, we predict that the success ratio will lie in I. Now suppose that we make an observation on z_I. If z_I is verified we decide that the common probability is correct to within the degree of accuracy determined by the interval I. If $\sim z_I$ is verified we decide that p is not correct. This does not mean that the observation on z_I proves either the correctness or incorrectness of p. Rather this observation produces a criterion by which we come to a tentative decision. Further observations may cause us to reverse this decision.

Our procedure gives us a check on the assignment of a common probability to a finite set of elements assumed to be independent. The number of elements need not be large but if we wish to check the probability to within a high degree of accuracy a large number of elements is required. The probability of a single element x is checked only when this probability is such that we predict a success ratio of 1 or 0. Even an observation with respect to such a prediction does not produce much of a check on our particular choice of $p(x)$. In this connection it should be noted that the above procedure for checking the common probability p of a set of events is applicable even when $p > \lambda$.

The above criterion for deciding the correctness or incorrectness of probabilities gives to probabilities a meaning which is significant with regard to the predicted behavior of the physical phenomena to which the probabilities are attached. When we speak of the probability that a given event will occur, then we would seem to imply that this probability has something to do with occurrence or nonoccurrence. In fact, if there were no relation between probabilities and occurrences then the phrase "probability of occurrence" would be misleading. The phrase is however suggestive rather than misleading and the above criterion furnishes the relation between probabilities and occurrences.