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 oc-