RANDOM VARIABLES FROM THE POINT OF VIEW OF A GENERAL THEORY OF VARIABLES

KARL MENGER
ILLINOIS INSTITUTE OF TECHNOLOGY

1. Introduction

In his great book *Sequential Analysis*, Wald defines (see p. 5 in [1]) a random variable as a variable x such that "for any given number c a definite probability can be ascribed to the event that x will take a value less than c." As a first example of a random variable, Wald mentions the outcome x of the experiment of weighing an object selected at random from a lot of n known objects. He calls x a random variable "since a probability can be ascribed to the event that x will take a value less than c, for any given c." If n_c is the number of objects in the lot whose weight is less than c, that probability is n_c/n. On page 11, Wald says that "statistical problems arise when the distribution function of a random variable is not known and we want to draw some inference concerning the unknown distribution function on the basis of a limited number of observations." He then mentions, as an example, the random variable x assuming the value 0 if a unit selected from a completely unknown lot of products is nondefective, and the value 1 if the unit is defective.

In 1947, I submitted to Wald the following two observations: (1) the concept "variable" on which the notion of random variable is based (see p. 5 in [1]) does not appear to be that of a numerical variable, the only one then clearly defined; (2) the statement and example on page 11 seem to be at variance with the definition of random variables on page 5.

I believe that I carry out Wald's intentions by saying that he fully agreed with both remarks and expressed the hope to clarify the statistical concept of random variables at a later occasion. His untimely death in 1950, after the completion of his fundamental book on statistical decision functions (in which he essentially retained the treatment of random variables of *Sequential Analysis*) prevented him from carrying out this plan.

For the past few years I have tried to analyze the ideas behind the general term "variable"—a term that, in spite of its frequent and heretofore indiscriminate use, has never been introduced by a comprehensive definition (either explicitly, in terms of other concepts, or implicitly, by postulates). As a result of these studies [2], [3], [4], and especially [5], it appears that there is not one comprehensive concept of variable. The underlying material has been resolved into an extensive spectrum of concepts. That array begins in mathematical logic; it traverses algebra, analysis, the various types of geometry, and physical science; it touches social science, and it ends in statistics. Some of those concepts have only one common bond—the name variable. In content, they differ about as much as do the tangent of an angle in trigonometry and the tangent to a curve in geometry. But whereas no one has ever confused the latter two ideas because of a flimsy equivoca-