"OPTIMUM" NONPARAMETRIC TESTS

WASSILY HOEFFDING
UNIVERSITY OF NORTH CAROLINA

1. Introduction

The problem of "optimum" tests has two aspects: (1) the choice of a definition of "optimum," and (2) the mathematical problem of constructing the test. The second problem may be difficult, but at least it is definite once an "optimum" test has been defined. But the definition itself involves a considerable amount of arbitrariness. Clearly, the definition should be "reasonable" from the point of view of the statistician (which is a very vague requirement) and it should be realizable, that is, an "optimum" test must exist, at least under certain conditions (which is trivial). Furthermore, even a theoretically "best" test is of no use if it cannot be brought into a form suitable for applications. When deciding which of two tests is "better" one ought to take into account not only their power functions but also the labor required for carrying out the tests.

The problem of "optimum" tests was first stated and partially solved by Neyman and E. S. Pearson. They, and most later writers, considered the parametric case, where the distributions are of known functional form which depends on a finite number of unknown parameters. A survey of the present status of the theory of testing hypotheses in the parametric case, with several extensions, will be found in a recent paper of Lehmann [3]. For the nonparametric case, where the functional form of the distributions is not specified, the problem has been attacked only recently. Wald's general theory of decision functions (see, for example, [8]) covers both the parametric and the nonparametric case, but its application to specific problems is often far from being trivial. The first (and at this writing only) publication which explicitly solves the problem of constructing tests of certain nonparametric hypotheses which are optimum in a specified sense is the paper of Lehmann and Stein [4] which appeared in 1949.

Many of the definitions formulated in parametric terms can easily be extended to the nonparametric case. I shall here mention some of these extensions which will be used in this paper.

Let Ω be a set of probability functions $P(A) = Pr\{X \in A\}$ of a random variable (usually a vector) X. Let ω be a subset of Ω, and let H be the hypothesis that P is in ω. A test is determined by a function $\phi(x)$, $0 \leqq \phi(x) \leqq 1$, measurable with respect to P, which is interpreted as the probability of rejecting H when $X = x$. If $\phi(x)$ can take only the values 0 and 1, it is the characteristic function of a set which is commonly known as the critical region. The probability that the test ϕ rejects H when P is the true distribution equals

$$E_P(\phi) = \int \phi(x)\, dP(x)$$

This work was done under the sponsorship of the Office of Naval Research, Project NR 042 031 at Chapel Hill, North Carolina.