"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.

and is called the power function of the test ϕ (defined over Ω). The least upper bound of $E_P(\phi)$ for all P in ω is called the size of the test ϕ (with respect to ω).

Let P_1 be a probability function in $\Omega - \omega$ and H_1 the hypothesis that $P = P_1$. If ϕ is of size α and

$$(1) \qquad E_{P_1}(\phi) \geqq E_{P_1}(\phi') \text{ for all tests } \phi' \text{ of size } \alpha ,$$

ϕ is called a most powerful test of size α for testing H against H_1. If (1) holds for all P_1 in a set ω_1, then ϕ is said to be uniformly most powerful with respect to ω_1.

The function

$$\beta(P) = \sup_{\phi} E_P(\phi) ,$$

where the supremum is taken with respect to all ϕ of size α, is called the envelope power function. If ω_1 is a subset of $\Omega - \omega$, then a test which minimizes

$$\sup_{P \in \omega_1} \{ \beta(P) - E_P(\phi) \}$$

with respect to all tests of size α is called most stringent against the set of alternatives ω_1.

A test ϕ is called similar of size α for testing H if $E_P(\phi) = \alpha$ for all P in ω. The definitions of a most powerful similar and a most stringent similar test are obvious.

Restrictions other than similarity are sometimes imposed on tests (see, for example, section 3). Other types of tests will be considered in section 4.

If Ω is a set of absolutely continuous functions $P(A)$ and $f(x)$ denotes the density of $P(A)$, it will sometimes be convenient to refer to Ω as a set of densities f and to write $E_f(\phi)$ for $E_P(\phi)$. This will cause no confusion if we agree to regard two densities as equal when they are densities of the same function $P(A)$.

This paper presents a survey of known and some new results on "optimum" nonparametric tests. Section 2 on most powerful and most stringent tests is based on the paper of Lehmann and Stein [4]. In section 3 most powerful rank order tests are discussed. The final section deals with some tests which have optimum properties with respect to an extensive class of alternatives.

2. Most powerful and most stringent tests

I shall here outline some results of Lehmann and Stein [4]. For the sake of brevity I shall confine myself to a special but typical case.

Let $\mathfrak{X}$ be a Euclidean space of $N = n_1 + \ldots + n_k$ dimensions. The points of $\mathfrak{X}$ will be denoted by

$$x = (x_{11}, \ldots, x_{1n_1}, x_{21}, \ldots, x_{2n_2}, \ldots, x_{k1}, \ldots, x_{kn_k}) .$$

Let ω' be the set of probability densities $f(x)$ (in the ordinary sense) over $\mathfrak{X}$ which are invariant under any permutation within $(x_{i1}, \ldots, x_{in_i})$, $(i = 1, \ldots, k)$. Let H' be the hypothesis that $f(x)$ is in ω'.

Many common statistical hypotheses *imply* invariance of this type. For instance, the hypothesis that two independent random samples came from the same population implies invariance of the distribution under all permutations of the variables (in this case $k = 1$).

Let W be the subset of $\mathfrak{X}$ where $x_{ig} \neq x_{ih}$ if $g \neq h$, $(i = 1, \ldots, k)$. Since we are dealing with probability densities, we need consider only points in W. For any x in W denote by $T(x)$ the set of points obtained by permuting $(x_{i1}, \ldots, x_{in_i})$ in all possible ways, for $i = 1, \ldots, k$. Each set $T(x)$ consists of exactly $M = n_1! \ldots n_k!$ points.

A test $\phi(x)$ which satisfies the condition

$$\sum_{x' \in T(x)} \phi(x') = M\alpha$$

for all x in W has been called a test of structure $S(\alpha)$ by Scheffé [7] and Lehmann and Stein [4].

It is easily shown that a test of structure $S(\alpha)$ is similar with respect to ω' and of size α.

Let H_1 be the hypothesis that $f(x) = g(x)$, a probability density over $\mathfrak{X}$ not in ω'. For each x in W denote the points of $T(x)$ by $x^{(1)}, \ldots, x^{(M)}$ in such a way that

$$g(x^{(1)}) \geqq g(x^{(2)}) \geqq \ldots \geqq g(x^{(M)}).$$

For every x in W let

$$\phi(x) = \begin{cases} 1 & \text{if } g(x) > g(x^{(1+[M\alpha])}) \\ a(x) & \text{if } g(x) = g(x^{(1+[M\alpha])}) \\ 0 & \text{if } g(x) < g(x^{(1+[M\alpha])}), \end{cases} \tag{2}$$

where $[M\alpha]$ is the largest integer $\leqq M\alpha$ and $a(x)$ is uniquely determined by the condition

$$\sum_{i=1}^{M} \phi(x^{(i)}) = M\alpha.$$

It is easily seen that $0 \leqq a(x) < 1$, and that $\phi(x)$ is measurable. Hence $\phi(x)$ is a test of structure $S(\alpha)$.

Lehmann and Stein [4] show that $\phi(x)$ is a most powerful test of H' against H_1.

For certain rational values of α we have $a(x) = 0$ almost everywhere, and $\phi(x)$ reduces to the characteristic function of a critical region.

Now let H be the hypothesis that $f(x)$ is of the form

$$f(x) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} f_i(x_{ij}).$$

This type of hypothesis is frequent in applications. H implies, but is not equivalent to H'.

It follows from a result of Feller [1] (see also Scheffé [7]) that any similar test of H is of structure $S(\alpha)$. (In fact, this holds even for much more restricted hypotheses than H, for instance a parametric set of distributions each of which differs arbitrarily little from a normal distribution.) On the other hand, since $E_g(\phi) \geqq E_g(\phi')$ for any test ϕ' of size α with respect to ω', this inequality is *a fortiori* true for all ϕ' of structure $S(\alpha)$. It follows that ϕ is a most powerful similar test of H against H_1.

For example, let $k = 1$, $n_1 = n$, and

$$(3) \qquad g(x) = g(x, \mu, \nu, \sigma) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2\sigma^2)(x_i - a_i\mu - \nu)^2},$$

where $a_1, \ldots, a_n$ are given numbers, not all equal. If $\mu > 0$, the most powerful test ϕ of H can be expressed by (2) with $g(x)$ replaced by $L = \sum (a_i - \bar{a})(x_i - \bar{x})$, where $\bar{a} = \sum a_i/n$, $\bar{x} = \sum x_i/n$. Hence the test is uniformly most powerful (uniformly most powerful similar) for testing $H'(H)$ against the set of alternatives $g(x)$ with $\mu > 0$, ν and σ arbitrary.

Lehmann and Stein also give a method which in certain cases enables one to find a most stringent (most stringent similar) test of $H'(H)$ against a set of alternatives. The method will be described in section 4 in connection with a related problem. In the last example the most stringent test against $g(x)$ with $\mu \neq 0$, ν and σ arbitrary is found to be the test (2) with $g(x)$ replaced by $|L|$.

It is remarkable that essentially the same test was earlier proposed by Pitman [6] on intuitive grounds.

To apply this test one requires the distribution of

$$L = \sum (a_i - \bar{a})(x_{r_i} - \bar{x})$$

in the population of equally probable permutations $(x_{r_1}, \ldots, x_{r_n})$ of the fixed sample values $x_1, \ldots, x_n$. The exact distribution is very cumbersome to determine unless n is small. Wald and Wolfowitz [9] and Noether [5] have shown that under certain conditions on the a_i and x_i, L is asymptotically normal. Pitman [6] has suggested an approximation which amounts to applying the standard t-test. Exact conditions under which either approximation is valid are unknown. Practically nothing is known about the power function of the test. Similar remarks apply to other tests of this type.

The results of Lehmann and Stein also apply to generalized densities with respect to a measure which need not be Lebesgue measure (so that discrete distributions are covered) and to types of invariance other than invariance under permutations.

3. Most powerful rank order tests

In some statistical problems the numerical values of the observations are not given and nothing but the relative size, or rank order, of the observations in certain parts of the sample is known. In other problems one may want to have a test which is invariant under any order preserving transformation in certain parts of the sample. In either case one is confined to a class of tests which can be called rank order tests.

Let, as before, $\mathfrak{X}$ be the Euclidean space of points

$$x = (x_{11}, \ldots, x_{1n_1}, \ldots, x_{k1}, \ldots, x_{kn_k})$$

and W the subset of $\mathfrak{X}$ where $x_{ig} \neq x_{ih}$ if $g \neq h$. Denote by $P(A)$ the probability function of a random vector X whose values are in $\mathfrak{X}$. We shall assume that $P(W) = 1$.

Let ω be any set of probability functions $P(A)$ of X which are invariant under all permutations within $(X_{i1}, \ldots, X_{in_i})$, $(i = 1, \ldots, k)$. Let H be the hypothesis that $P(A)$ is in ω.

For any point x in W, the coordinate x_{ij} is said to have rank r_{ij} with respect to $(x_{i1}, \ldots, x_{in_i})$ if exactly $r_{ij} - 1$ coordinates x_{ih} (with i fixed) are less than x_{ij}. Let R be the permutation

$$R = (r_{11}, \ldots, r_{1n_1}, \ldots, r_{k1}, \ldots, r_{kn_k}) ,$$

and let $S(R)$ be the set of all x in W for which x_{ij} has rank r_{ij} $(j = 1, \ldots, n_i;$ $i = 1, \ldots, k)$.

The set W is the union of the $M = n_1! \ldots n_k!$ disjoint sets $S(R)$. Let $P[S(R)]$ be briefly denoted by $P(R)$. If P is in ω, we have $P(R) = 1/M$ for all R. Hence any union of m sets $S(R)$ is the critical region of a similar test of size m/M with respect to ω. A test of this type will be referred to as a rank order test.

Let H_1 be the hypothesis that $P(A) = P_1(A)$, a probability function of X not in ω. Denote the M permutations R by $R_1, \ldots, R_M$ in such a way that

$$P_1(R_i) \geqq P_1(R_j) \qquad \text{if } i = 1, \ldots, m ; \quad j = m+1, \ldots, M .$$

Clearly, $(R_1, \ldots, R_m)$ determines a rank order test which is most powerful for testing H against H_1.

Suppose that $P_1(A)$ is absolutely continuous with probability density $g(x)$. Then

$$P_1(R) = \int_{S(R)} g(x)\, dx = \int_{S(I)} g(x_R)\, dx ,$$

where $S(I)$ is the set

$$x_{11} < \ldots < x_{1n_1}, \ldots, x_{k1} < \ldots < x_{kn_k}$$

and x_R is obtained from x by applying the permutation R to the second subscripts of the coordinates.

Suppose, moreover, that ω contains an absolutely continuous distribution with density

$$f(x) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} f_i(x_{ij}) .$$

Let

$$X_{i1}^{(n_i)}, \ldots, X_{in_i}^{(n_i)}$$

have the probability density

$$n_i! \prod_{j=1}^{n_i} f_i(x_j) \text{ if } x_1 < \ldots < x_{n_i}$$

and zero elsewhere; that is, the $X_{ij}^{(n_i)}$ are the order statistics associated with the distribution

$$\prod_{j=1}^{n_i} f_i(x_j) .$$

Let

$$X^0 = (X_{11}^{(n_1)}, \ldots, X_{1n_1}^{(n_1)}, \ldots, X_{k1}^{(n_k)}, \ldots, X_{kn_k}^{(n_k)}) .$$

Then we can write

$$P_1(R) = \int_{S(I)} \frac{g(x_R)}{f(x_R)} f(x)\, dx$$

$$= \frac{1}{M} \int_{x_{11}<\dots<x_{1n_1}} \int \dots \int_{x_{k1}<\dots<x_{kn_k}} \int \frac{g(x_R)}{f(x_R)} \prod_{i=1}^{k} \left\{ n_i! \prod_{j=1}^{n_i} f_i(x_{ij})\, dx_{ij} \right\}$$

or

$$MP_1(R) = E\left\{ \frac{g(X_R^0)}{f(X_R^0)} \right\}.$$

Thus we can express $MP_1(R)$ as the expected value of a function of the order statistics, permuted according to R, the function being the ratio of the two probability densities involved.

As an example, let $k = 1$, let $g(x) = g(x, \mu, \nu, \sigma)$ be defined as in (3), and let $f(x) = g(x, 0, \nu - \bar{a}\mu, \sigma)$. After some simplification we obtain

$$n!\, P_1(R) = e^{-(\delta^2/2)\Sigma(a_i-\bar{a})^2} E\{ e^{\delta \Sigma(a_i-\bar{a}) X_{r_i}^{(n)}} \},$$

where $\delta = \mu/\sigma$ and $X_1^{(n)} < \dots < X_n^{(n)}$ are the standard normal order statistics.

Whereas the most powerful (similar) test depends only on the sign of μ, the most powerful rank order test depends on the value of δ. Also the evaluation of $P_1(R)$ presents considerable difficulties. We have

$$n!P_1(R) = 1 + \delta c_1(R) + 0(\delta^2),$$

where

$$c_1(R) = \sum_{i=1}^{N} (a_i - \bar{a}) E X_{r_i}^{(n)}.$$

Hence the rank order test which is most powerful against δ positive and "small" is based on the statistic $c_1(R)$, large values being significant. This kind of rank order test is essentially equivalent to one described by Fisher and Yates [2], where a table of $EX_i^{(n)}$ is given.

The distribution of

$$\frac{\sum (a_i - \bar{a}) E X_{r_i}^{(n)}}{\sqrt{\sum (a_i - \bar{a})^2 \sum (E X_j^{(n)})^2}}$$

when H is true can be approximated by the probability density

$$B\left(\frac{1}{2}, \frac{n-2}{2}\right)^{-1} (1 - x^2)^{(n-4)/2}, \qquad -1 \leqq x \leqq 1.$$

This approximation was suggested by Pitman [6] for statistics of a similar type and is apparently implied by Fisher and Yates [2]. The approximation appears to be satisfactory even for moderate values of n provided the a_i satisfy certain conditions.

It would be desirable to determine the most stringent rank order tests against $\delta > 0$ and $\delta \neq 0$, provided the test criteria are computable. It would also be interesting to compare the power of the c_1-test with that of the most powerful rank order test for an arbitrary positive value of δ.

4. Tests with optimum properties with respect to nonparametric alternatives

So far we have considered nonparametric tests which have optimum properties with respect to simple or parametric alternatives. On the other hand, in a problem where the functional form of the distribution is not specified by the null hypothesis one often may want to have a test which is "optimum" with respect to a set of alternatives which itself is nonparametric. Only in rare cases do there exist tests which are uniformly most powerful against a nonparametric set, and most stringent tests do not seem to be readily applicable to many types of nonparametric alternatives. (An example of a uniformly most powerful and a most stringent test against a special type of nonparametric alternative was given by Lehmann and Stein [4, p. 38].)

The tests to be considered in this section are based on the minimax principle which underlies Wald's theory of decision functions.

Let Ω be a set of probability functions P, ω a subset of Ω, H the hypothesis that P is in ω. The set of alternatives is $\Omega - \omega$.

We still confine ourselves to tests which are of size α for testing H, and by a test we shall always mean a test of size α for testing H.

Suppose there is given a nonnegative weight function $W(P)$, defined for all P in $\Omega - \omega$; it may be interpreted as expressing the loss caused by accepting H when the true distribution is P. The expected loss when P (in $\Omega - \omega$) is the distribution and test ϕ is used is $W(P)[1 - E_P(\phi)]$ and is called the risk. The maximum risk associated with test ϕ is

$$R(\phi) = \sup_{P \in \Omega - \omega} W(P)[1 - E_P(\phi)] .$$

We assume that there exists a ϕ for which $R(\phi)$ is finite.

A test ϕ which minimizes the maximum risk, so that $R(\phi) \leqq R(\phi')$ for all ϕ' of size α, will be called a test of minimax risk with respect to the weight function $W(P)$. This is the immediate extension to the nonparametric case of a test considered by Lehmann [3].

In statistical practice it is often impossible without undue arbitrariness to assign a numerical weight $W(P)$ to any alternative. But in many cases one will be able to decide of any two alternatives P_1, P_2 whether they should be assigned equal weights or which of them should have greater weight. This leads to a partition of $\Omega - \omega$ into disjoint sets $\Omega(d)$, where d is a real parameter such that $d < d'$ implies that the P in $\Omega(d)$ have smaller weight than the P in $\Omega(d')$, and all P in $\Omega(d)$ have equal weight.

If a weight function $W(P)$ is given and $\{\Omega(d)\}$ is a partition of $\Omega - \omega$ such that $W(P)$ is constant over each set $\Omega(d)$, we shall say that $\{\Omega(d)\}$ is a partition induced by $W(P)$. [$\Omega(d)$ need not consist of all P for which $W(P)$ equals a constant value.] The same weight function induces more than one partition, and the same partition is induced by more than one function. In particular, a partition induced by $W(P)$ is also induced by any strictly monotone function of $W(P)$.

Let $\Omega(d)$ be a partition of $\Omega - \omega$, and suppose there exists a test ϕ such that

$$\inf_{P \in \Omega(d)} E_P(\phi) \geqq \inf_{P \in \Omega(d)} E_P(\phi')$$

for all ϕ' and all $\Omega(d)$. Then ϕ will be referred to as a test[1] which maximizes the minimum power (uniformly) with respect to the partition $\{\Omega(d)\}$. Such a test may be regarded as satisfactory provided its minimum power over $\Omega(d)$ increases with d.

A test which uniformly maximizes the minimum power does not always exist. If it exists, it is "optimum" in a stronger sense than a test of minimax risk, as is shown by the following theorem.

THEOREM 1. *If a test maximizes the minimum power with respect to the partition* $\{\Omega(d)\}$, *it has minimax risk with respect to any weight function which induces this partition.*

The proof follows from the relation

$$R(\phi) = \sup_d \sup_{P \in \Omega(d)} W(P)[1 - E_P(\phi)] = \sup_d W_d [1 - \inf_{P \in \Omega(d)} E_P(\phi)],$$

where W_d denotes the constant value of $W(P)$ on $\Omega(d)$.

Theorem 1 is analogous to a theorem of Hunt and Stein (quoted by Lehmann [3, theorem 8.2] from an unpublished paper) which can be stated as follows: If a test maximizes the minimum power with respect to a partition induced by the envelope power function, then the test is most stringent. It follows incidentally that the test has minimax risk with respect to any weight function which is a nondecreasing function of the envelope power function.

If ϕ_d denotes a test which maximizes the minimum power over $\Omega(d)$ and $\phi_d = \phi$ is independent of d, then ϕ clearly maximizes the minimum power uniformly with respect to $\{\Omega(d)\}$.

A method for finding a test which maximizes the minimum power over a given set when the distributions are absolutely continuous was given by Lehmann [3, theorem 8.3]. It is stated in parametric terms but can be immediately extended to the nonparametric case as follows.

THEOREM 2. *Let ω_0 and ω_1 be two sets of probability densities $f(x)$ with respect to a fixed measure μ. Suppose additive classes of sets have been defined over ω_0 and ω_1, and let λ_i be a probability measure over ω_i which assigns measure* 1 *to a parametric subset of ω_i $(i = 0, 1)$; then*

$$h_i(x) = \int_{\omega_i} f(x)\, d\lambda_i(f), \qquad i = 0, 1,$$

are probability densities with respect to μ. Let ϕ be a most powerful test of size α for testing the simple hypothesis h_0 against the simple alternative h_1. Let β be the power of ϕ against h_1. Then if

$$E_f(\phi) \leqq \alpha \qquad \text{for all } f \in \omega_0$$

and

$$E_f(\phi) \geqq \beta \qquad \text{for all } f \in \omega_1,$$

[1] Since tests of this type frequently occur, it would be desirable to have a shorter name for them. One might call them tests of maximin power with respect to $\{\Omega(d)\}$.

we have for all ϕ' of size $\leqq a$

$$\inf_{f\in\omega_1} E_f(\phi) \geqq \inf_{f\in\omega_1} E_f(\phi') .$$

As an example, let $g(y)$ denote any density with respect to a fixed measure ν on the real line such that $\nu\{(y \leqq 0)\} > 0$ and $\nu\{(y > 0)\} > 0$, let μ be the n-th power of ν, and let Ω be the set of all densities with respect to μ of the form

$$(x) = g(x_1) \dots g(x_n) , \tag{4}$$

where

$$G(0) = \int_{-\infty}^{0} g(y)\, d\nu(y) \geqq q ,$$

q a fixed number, $0 < q < 1$. Let ω be the subset of Ω where $G(0) = q$, and let $\Omega(d)$ be the subset of $\Omega - \omega$ where $G(0) = q + d$, $(0 < d \leqq 1 - q)$.

The problem can also be interpreted as one of testing whether the q-quantile of $g(y)$ is zero against the alternative that it is less than zero.

Let b be a positive constant, I_1 and I_2 the intervals $-b < y \leqq 0$ and $0 < y \leqq b$, where b is so chosen that $\nu(I_1) > 0$, $\nu(I_2) > 0$, and let

$$f_d(x) = \prod_{i=1}^{n} \left(\frac{q+d}{\nu(I_1)}\right)^{c_1(x_i)} \left(\frac{1-q-d}{\nu(I_2)}\right)^{c_2(x_i)} , \qquad 0 \leqq d \leqq 1 - q ,$$

where $c_j(y)$ is the characteristic function of I_j, $(j = 1, 2)$.

It is easily verified that f_d is in $\Omega(d)$, $(d > 0)$ and f_0 is in ω.

Let λ_1 be the measure over $\Omega(d)$ which assigns measure 1 to the one point set $\{f_d\}$ and measure 0 to $\Omega(d) - \{f_d\}$, and λ_0 the corresponding measure over ω. Applying theorem 2 with $\omega_0 = \omega$, $\omega_1 = \Omega(d)$ we have $h_0 = f_0$, $h_1 = f_d$. By Neyman-Pearson's lemma, a most powerful test of size a for testing f_0 against f_d is

$$\phi(x) = \begin{cases} 1 \\ a \\ 0 \end{cases} \text{if } \sum_{i=1}^{n} c(x_i) \begin{cases} > k \\ = k \\ < k \end{cases} \tag{5}$$

where $c(y) = 1$ or 0 according as $y \leqq 0$ or $y > 0$, and the constants k, a (k an integer, $0 \leqq a < 1$) are determined by the condition that $\phi(x)$ be of size a. This is the well known sign test.

The power of ϕ for any f in Ω depends only on $G(0)$ and hence is constant over $\Omega(d)$. The conditions of theorem 2 are satisfied. Since the test is independent of d, it maximizes the minimum power with respect to $\{\Omega(d)\}$. The power of ϕ increases with d.

If the size a is so chosen that $a = 0$, the test ϕ is essentially unique. This follows from the fact that for any test ϕ' which differs from ϕ on a set of positive μ-measure we can choose b so large that

$$E_{f_d}(\phi') < E_{f_d}(\phi) .$$

Now let Ω be the set of all densities of the form (4) with no restriction on $G(0)$, let ω be the subset with $G(0) = \frac{1}{2}$, and let $\Omega(d)$ be the set where $|G(0) - \frac{1}{2}| = d$,

$(0 < d \leqq \frac{1}{2})$. Let $f_d(x)$ be defined as before, with $q = \frac{1}{2}$, $-\frac{1}{2} \leqq d \leqq \frac{1}{2}$. Let λ_1 assign measure $\frac{1}{2}$ to f_d and to f_{-d}, measure 0 to $\Omega(d) - \{f_d, f_{-d}\}$. Let λ_0 assign measure 1 to f_0, measure 0 to $\omega - \{f_0\}$. Then $h_0 = f_0$, $h_1 = (f_d + f_{-d})/2$. The test which maximizes the minimum power with respect to $\{\Omega(d)\}$ is found to be of the form (5) with $\sum c(x_i)$ replaced by $\left| \sum c(x_i) - n/2 \right|$.

The test can be interpreted as a test of whether the median of g is zero against the alternative that it is different from 0. If g is normal with mean μ and standard deviation σ, $|G(0) - \frac{1}{2}|$ is an increasing function of $|\mu/\sigma|$, so that the power of the test can be directly compared with that of Student's test.

It has yet to be investigated whether the tests here discussed can be applied to more complicated nonparametric problems, or whether different types of tests will be required.

REFERENCES

[1] W. Feller, "Note on regions similar to the sample space," *Stat. Res. Memoirs*, Vol. 2 (1938), pp. 117–125.

[2] R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, 3rd ed., Hafner, New York, 1949.

[3] E. L. Lehmann, "Some principles of the theory of testing hypotheses," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 1–26.

[4] E. L. Lehmann and C. Stein, "On the theory of some nonparametric hypotheses," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 28–45.

[5] G. Noether, "On a theorem by Wald and Wolfowitz," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 455–458.

[6] E. J. G. Pitman, "Significance tests which may be applied to samples from any population," *Jour. Roy. Stat. Soc., Suppl.*, Vol. 4 (1937), pp. 119–130 and 225–232.

[7] H. Scheffé, "On a measure problem arising in the theory of non-parametric tests," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 227–233.

[8] A. Wald, "Statistical decision functions," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 165–205.

[9] A. Wald and J. Wolfowitz, "Statistical tests based on permutations of the observations," *Annals of Math. Stat.*, Vol. 15 (1944), pp. 358–372.