ON THE THEORY OF SOME NON-PARAMETRIC HYPOTHESES

By E. L. Lehmann and C. Stein

University of California, Berkeley

Summary. For two types of non-parametric hypotheses optimum tests are derived against certain classes of alternatives. The two kinds of hypotheses are related and may be illustrated by the following example: (1) The joint distribution of the variables $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ is invariant under all permutations of the variables; (2) the variables are independently and identically distributed. It is shown that the theory of optimum tests for hypotheses of the first kind is the same as that of optimum similar tests for hypotheses of the second kind. Most powerful tests are obtained against arbitrary simple alternatives, and in a number of important cases most stringent tests are derived against certain composite alternatives. For the example (1), if the distributions are restricted to probability densities, Pitman's test based on $\bar{y} - \bar{x}$ is most powerful against the alternatives that the X's and Y's are independently normally distributed with common variance, and that $E(X_i) = \xi$, $E(Y_i) = \eta$ where $\eta > \xi$. If $\eta - \xi$ may be positive or negative the test based on $| \bar{y} - \bar{x} |$ is most stringent. The definitions are sufficiently general that the theory applies to both continuous and discrete problems, and that tied observations present no difficulties. It is shown that continuous and discrete problems may be combined. Pitman's test for example, when applied to certain discrete problems, coincides with Fisher's exact test, and when $m = n$ the test based on $| \bar{y} - \bar{x} |$ is most stringent for hypothesis (1) against a broad class of alternatives which includes both discrete and absolutely continuous distributions.

1. Generalities. In the present paper we study the problem of determining optimum tests for certain non-parametric hypotheses. It is important in this connection to make some distinctions which are of lesser significance when the problem is approached from the intuitive point of view which has been customary in this field. Consider for example the hypothesis H that $Z_1, \cdots, Z_N$ are independently and identically distributed according to an unknown probability density function. All tests which have been suggested for testing H are valid also for testing the hypothesis H' that the unknown joint probability density function of the Z's is symmetric in its N arguments. On the other hand, tests which have optimum properties for testing H' against a certain class of alternatives will in general not possess the same properties when H' is replaced by H. From the present point of view the two hypotheses mentioned are essentially different. We shall be concerned in this paper primarily with generalizations of H', and we shall show that many of the tests suggested in the literature have optimum properties for testing hypotheses of this kind against certain classes of alternatives.

The corresponding general theory for hypotheses related to H is quite different.

However the two theories do coincide, provided tests of these latter hypotheses are restricted to similar regions. More specifically, all results on optimum tests of H' are equivalent to the corresponding results on optimum similar tests of H, and this equivalence holds also for many of the more general hypotheses considered in this paper.

It should be observed that in many experimental situations, the hypothesis H' that the joint distribution of the Z's is invariant under all permutations is more realistic than the hypothesis H that the Z's are independently and identically distributed. For example, suppose there is a block of land divided into $m + n$ plots, and the experimenter wants to test whether one of two fertilizers (used in fixed amounts) is more effective than the other in increasing the yield of a certain plant. Of the plots, m are chosen at random; fertilizer I is applied to these, and fertilizer II to the other n. If X_i denotes the yield from the ith plot to which fertilizer I has been applied and Y_j denotes the yield from the jth plot to which fertilizer II has been applied, where the plots are numbered at random, then the hypothesis that the two fertilizers are completely equivalent implies that the application of any permutation to $X_1, \cdots, X_m, Y_1, \cdots Y_n$ does not change their joint distribution. But it is not reasonable to suppose the X_i, Y_j are independently and identically distributed, since there may be intrinsic differences among the plots. For discussions of these and related points, see Fisher [1], Neyman [2], Pitman [3]. It may be that in many particular cases some hypothesis between the two is really appropriate but the hypothesis H is the only one that is evidently appropriate from a cursory inspection of the setup.

Many of the alternative hypotheses considered below, for example those involving normality, are dictated more by tradition and ease of treatment than by appropriateness in actual experiments. Thus this paper should not be considered as providing absolute justification for tests such as Pitman's but rather as suggesting a method of obtaining optimum non-parametric tests when the class of alternatives is fairly well specified.

Another possibility, first raised by Neyman [2], which has been ignored in this paper is the equality on the average of the two fertilizers but with fertilizer I having a larger dispersion than fertilizer II, or a distribution differing in some other characteristic. It would be reasonable to consider this as part of the hypothesis tested, but tests based on randomization may give a probability of rejection of the hypothesis of equivalence in this case which is much higher than the stated level of significance. We hope to return to problems of this type in later papers.

Let us make the following basic assumptions. $\mathcal{Z}$ is a space of points z and $\mathcal{C}$ is an additive class of subsets A of $\mathcal{Z}$. Any member of $\mathcal{C}$ will be said to be measurable. By a probability distribution we mean a measure F, defined over $\mathcal{C}$ for which $F(\mathcal{Z}) = 1$. We shall be concerned with two classes of probability distributions: One, the class of all distributions, and two, the class of distributions which are absolutely continuous with respect to a given measure μ, that is, the class of distributions F for which there exists a function f such that

$$(1.1) \qquad\qquad F(A) = \int_A f(z) \, d\mu(z).$$

We shall call f a generalized probability density function with respect to μ. By Z we denote a random variable such that for any A in $\mathcal{C}$,

$$(1.2) \qquad\qquad P\{Z \,\epsilon\, A\} = F(A).$$

For most of the applications we shall take $\mathcal{Z}$ to be a Euclidean space, and $\mathcal{C}$ to be the class of all Borel sets. Then if μ is Lebesgue measure, (1.1) states that f is a probability density function in the usual sense. However, we shall have occasion to consider also some measures other than Lebesgue measure. By a hypothesis H we mean a class of probability distributions. Next we describe the hypotheses with which we shall be concerned. Let Π be a partition of $\mathcal{Z}$, that is, let Π be a class of mutually exclusive subsets S of $\mathcal{Z}$ such that every point z of $\mathcal{Z}$ lies in one of the sets S. If two points z_1 and z_2 lie in the same set S, we shall say that z_1 is equivalent to z_2 with respect to Π: $z_1 \sim z_2 \pmod{\Pi}$. The set of all points which are equivalent to z will be denoted by $T(z)$, the number of points of $T(z)$ by $n(z)$. Concerning Π we make the following assumptions:

(i) All sets in Π are finite, so that $n(z)$ is finite for all z.

(ii) If we define S_n as the union of all those sets S of Π which contain exactly n points, there exist mutually exclusive sets $S_n^{(1)}, \cdots, S_n^{(n)}$ which are measurable and such that every element S of Π containing exactly n points has one and only one point in common with each $S_n^{(i)}$.

We shall say that a measure μ is invariant under Π if the following condition holds: For all n and $i, j \leq n$, if S is any set contained in $S_n^{(i)}$ and if S' denotes the set of equivalent points in $S_n^{(j)}$, then $\mu(S) = \mu(S')$.

Given a partition Π satisfying (i) and (ii), we formulate the hypothesis H that the distribution F of Z is invariant under Π. We shall refer to H as the hypothesis of invariance under Π. We shall also consider the hypothesis of invariance under a partition for a class of generalized densities f. In this case we assume that the measure μ of (1.1) is given, and that Π, in addition to (i) and (ii) satisfies the condition:

(iii) The measure μ is invariant under Π. The hypothesis H in this case states that $z_1 \sim z_2 \pmod{\Pi}$ implies $f(z_1) = f(z_2)$.

By a test of a hypothesis H we mean (see [4]) a measurable function φ on $\mathcal{Z}$ to the interval $[0, 1]$ which with every point $z \,\epsilon\, \mathcal{Z}$ associates a probability $\varphi(z)$ of rejection. This definition, slightly more general than the usual one, is particularly useful in non-parametric work. Among other advantages it automatically takes care of the problem of tied observations. It also disposes of the difficulties encountered by Scheffé [5] in his treatment of the problem of similar regions, as will be shown in Lemma 1.

The size of a test φ is defined to be

$$(1.3) \qquad\qquad \epsilon(\varphi) = \sup_{F \,\epsilon\, H} \int \varphi(z) \, dF(z).$$

If in particular

$$(1.4) \qquad \int \varphi \, dF = \epsilon(\varphi)$$

for all F in H, φ is said to be similar for testing H. Extending the terminology of Scheffé, we say that φ has structure $S(\epsilon)$ if for all z in S_n

$$(1.5) \qquad \sum_{z' \epsilon T(z)} \varphi(z') = n\epsilon.$$

The following lemma extends a result of Scheffé.

LEMMA 1. *For testing a hypothesis of invariance, any test of structure $S(\epsilon)$ is similar and of size ϵ.*

PROOF. For any F in H and any φ

$$(1.6) \qquad \int \varphi \, dF = \sum_{n=1}^{\infty} \sum_{i=1}^{n} \int_{S_n^{(i)}} \varphi \, dF = \sum_{n=1}^{\infty} \int_{S_n^{(1)}} \Big[\sum_{z' \epsilon T(z)} \varphi(z') \Big] \, dF(z).$$

But φ has structure $S(\epsilon)$ and hence (1.5) holds for all z. Therefore

$$(1.7) \qquad \int \varphi \, dF = \sum_{n=1}^{\infty} n\epsilon \int_{S_n^{(1)}} dF = \epsilon.$$

We shall show next that for testing a hypothesis of invariance at level of significance ϵ, only tests of structure $S(\epsilon)$ need be considered. In order to make this result applicable both to hypotheses referring to the class of all distributions and to those referring to a class of generalized densities, we shall state it in an asymmetric form which when taken together with lemma 1 indicates the essential equivalence of the two types of hypotheses.

LEMMA 2. *If φ is any test of a hypothesis of invariance for the class of generalized densities with respect to a fixed measure μ, and if the size of φ is less than or equal to ϵ, then there exists a test φ_1 of structure $S(\epsilon)$ such that*

$$(1.8) \qquad \int \varphi_1 \, dF \geq \int \varphi \, dF$$

for all probability distributions F.

PROOF. First we shall show that

$$(1.9) \qquad \frac{1}{n(z)} \sum_{z' \epsilon T(z)} \varphi(z') \leq \epsilon$$

almost everywhere μ. For let A be the set of points z such that

$$(1.10) \qquad \frac{1}{n(z)} \sum_{z' \epsilon T(z)} \varphi(z') > \epsilon$$

and suppose that $\mu(A)$ is positive. Let

$$(1.11) \qquad f(z) = \begin{cases} \dfrac{1}{\mu(A)} & \text{if } z \epsilon A; \\ 0 & \text{elsewhere.} \end{cases}$$

Then f is in H since by definition of A, whenever z is in A, $T(z)$ is contained in A. But

$$(1.12) \qquad \int \varphi f \, d\mu > \epsilon \int_A f \, d\mu = \epsilon,$$

in contradiction to the assumption that φ has size ϵ.

From (1.9) it follows easily that there exists a test φ_1 of structure $S(\epsilon)$ and such that for all z

$$(1.13) \qquad \varphi_1(z) \geq \varphi(z).$$

Since condition (1.8) is then satisfied, this completes the proof.

Lemma 2 raises the question whether it is possible to reduce the problem of testing a hypothesis of invariance still further, or whether the tests of structure $S(\epsilon)$ form, what Wald [6] has called an essentially complete class of admissible tests. This question is answered by

THEOREM 1. *Let μ be a measure defined over $\mathcal{C}$. Let Π_0 and Π_1 be two partitions of $\mathcal{Z}$ satisfying conditions* (i), (ii) *and* (iii), *and such that $z \sim z'$ (mod Π_1) implies $z \sim z'$ (mod Π_0). For the class of generalized densities with respect to μ denote by H_i $(i = 0, 1)$ the hypothesis of invariance relative to Π_i. Then for testing H_0 against H_1 at level of significance ϵ, the totality of tests which* (a) *have structure $S(\epsilon)$, and for which* (b) $z \sim z'$ (mod Π_1) *implies $\varphi(z) = \varphi(z')$, form an essentially complete class of admissible tests.*

PROOF. It is easily seen that we can restrict ourselves to that subclass of tests of structure $S(\epsilon)$ which possess property (b). For if φ is any test of structure $S(\epsilon)$ relative to Π_0, let

$$(1.14) \qquad \varphi^*(z) = \frac{1}{n(z)} \sum_{z' \epsilon T(z)} \varphi(z).$$

Then clearly φ^* possesses property (b) and has structure $S(\epsilon)$. Furthermore if f is any probability density function of H_1, then

$$(1.15) \qquad \int \varphi^* f \, d\mu = \int \varphi f \, d\mu,$$

so that φ and φ^* have identical power against H_1.

In order to complete the proof, we must show that if φ_1 and φ_2 are any two tests satisfying (a) and (b), and if φ_1 and φ_2 differ on a set of positive measure, there exists a probability density function f of H_1 for which

$$(1.16) \qquad \int \varphi_1 f \, d\mu > \int \varphi_2 f \, d\mu.$$

Since both φ_1 and φ_2 have structure $S(\epsilon)$, the set A of points z for which

$$(1.17) \qquad \varphi_1(z) > \varphi_2(z)$$

has positive measure. Also, because of (b), if two points are equivalent relative to Π_1, they are either both in A or both not in A. If $f(z)$ is defined as $1/\mu(A)$ for z in A and as zero elsewhere, then f is in H_1 and satisfies (1.16).

The theorem obtained from theorem 1 by letting the hypotheses H_0 and H_1 refer to the class of all probability distributions rather than to a particular class of generalized densities, is clearly also true, and cases between these two theorems could also be formulated.

Since the most powerful test φ for testing a hypothesis of invariance H_0 referring to a class of generalized densities against an alternative f from this class of densities has the correct size also for testing the wider hypothesis $\bar{H}_0$ referring to the class of all distributions, φ is also most powerful for testing $\bar{H}_0$ against f. The corresponding remark holds for most stringent tests. Therefore all optimum tests that will be derived in the sequel, through the use of theorems of this section, may be considered as tests of hypotheses referring to the class of all distributions: they are valid against these hypotheses, and no power is gained by restricting the hypothesis to the appropriate class of generalized densities.

2. Most powerful tests and most stringent tests. One of the main problems to be considered in this paper is the determination of a most powerful test of a hypothesis of invariance against a simple alternative. If we restrict our considerations to the class of generalized densities with respect to μ, a complete solution of this problem is given by the following

THEOREM 2. *Let H be the hypothesis of invariance under the partition Π, and let g be a probability density function not in H. For any z in S_n denote by $z^{(1)}, \cdots, z^{(n)}$ the n points of $T(z)$ arranged so that $g(z^{(1)}) \geq g(z^{(2)}) \geq \cdots \geq g(z^{(n)})$. For testing H against g a most powerful test of size ϵ is given by*

$$(2.1) \qquad \varphi(z) = \begin{cases} 1 & if \quad g(z) > g(z^{(1+[\epsilon n])}) \\ a & if \quad g(z) = g(z^{(1+[\epsilon n])}) \\ 0 & if \quad g(z) < g(z^{(1+[\epsilon n])}) \end{cases} \quad for\ z\ in\ S_n,$$

where $\sum_{i=1}^{n} \varphi(z^{(i)}) = n\epsilon$, $0 \leq a \leq 1$ and where a may depend on z through $T(z)$.

PROOF. First we observe that the number of $z^{(i)}$ for which $g(z^{(i)}) \geq g(z^{(1+[\epsilon n])})$ is greater than or equal to $1+[\epsilon n] \geq \epsilon n$ and that the number of $z^{(i)}$ for which $g(z^{(i)}) > g(z^{(1+[\epsilon n])})$ is less than or equal to $[\epsilon n] \leq \epsilon n$, so that there exists an a between 0 and 1 for which $\Sigma \varphi(z^{(i)}) = n\epsilon$. Since φ has structure $S(\epsilon)$, it follows from lemma 1 that it is similar and of size ϵ.

Let

$$(2.2) \qquad g^*(z) = g(z^{(1+[\epsilon n])}) \quad for \quad z \epsilon S_n.$$

To complete the proof consider first the special case that

$$(2.3) \qquad \int g^*(z)\ d\mu(z)$$

vanishes. Then

$$(2.4) \qquad \int \varphi g\ d\mu = \int g\ d\mu = 1$$

that is, the test φ has power 1, and therefore is clearly most powerful. Assume next that the integral (2.3) is positive. Then g^* is proportional to a probability density function of H. For it is measurable and satisfies the symmetry condition required of a member of H, and the integral (2.3) is finite since

$$
(2.5) \quad
\begin{aligned}
\sum_n \int_{S_n} g^*(z)\, d\mu(z) &\leq \sum_n \frac{1}{\epsilon n} \int_{S_n} \sum_{i=1}^{[\epsilon n]+1} g(z^{(i)})\, d\mu(z) \\
&\leq \frac{1}{\epsilon} \sum_n \int_{S_n} \frac{1}{n} \sum_{i=1}^{n} g(z^{(i)})\, d\mu(z) = \frac{1}{\epsilon}.
\end{aligned}
$$

The test φ therefore has the form of a probability ratio test. Since it is also similar, it follows from theorem 1 of [4] that φ is most powerful.

In practice one is usually interested in composite rather than simple alternatives. We shall therefore consider next the problem of deriving most stringent tests of hypotheses of invariance against certain classes of alternatives. This problem may be reduced to that of finding tests which maximize the minimum power over a class of alternatives by the following simple theorem of Hunt and Stein [7].

THEOREM 3. *Given a hypothesis H and a class of alternatives $\{g_\theta\}$, $\theta \in \Omega$, denote by $\beta^*(\theta)$ the envelope power function corresponding to the level of significance ϵ, that is, let*

$$
(2.6) \qquad \beta^*(\theta) = \sup_\varphi \beta(\varphi, \theta)
$$

where $\beta(\varphi, \theta)$ stands for the power of the test φ against the alternative g_θ and where the least upper bound is taken over all tests φ of size ϵ. Let $\{\Omega_\delta\}$ be a class of mutually exclusive subsets of Ω such that $\cup \Omega_\delta = \Omega$ and such that $\beta^*(\theta)$ is constant on each Ω_δ. Denote by φ_δ a test which maximizes the minimum power over Ω_δ. If $\varphi_\delta = \varphi$ is independent of δ, then φ is most stringent[1] for testing H against Ω at level of significance ϵ.

For obtaining tests which maximize the minimum power over a class of alternatives to a hypothesis of invariance, we can state the following simple extension of theorem 2.

THEOREM 4. *Let H be a hypothesis of invariance, and let H_1 be the class of alternatives $\{g_\theta\}$, $\theta \in \Omega$. Suppose there exists a subset Ω' of Ω and a probability measure λ over Ω' such that for the test φ of size ϵ defined as in theorem 2 with*

$$
(2.7) \qquad g(z) = \int_{\Omega'} g_\theta(z)\, d\lambda(\theta),
$$

the integral $\int \varphi g_\theta\, d\mu$ is constant for θ in Ω', and

$$
(2.8) \qquad \int \varphi g_\theta\, d\mu \geq \int \varphi g_{\theta'}\, d\mu \quad \text{for all} \quad \theta \in \Omega,\ \theta' \in \Omega'.
$$

Then φ maximizes the minimum power over Ω at level of significance ϵ.

---

[1] A test is said to be most stringent [16] if it minimizes the maximum difference between envelope power and power, that is, if it minimizes Sup $[\beta^*(\theta) - \beta(\varphi, \theta)]$.

PROOF. By theorem 2, φ is a most powerful test for testing H against g, that is, for any φ' of size ϵ

$$(2.9) \qquad \int \varphi'(z) \int_{\Omega'} g_\theta(z) \, d\lambda(\theta) \, d\mu(z) \leq \int \varphi(z) \int_{\Omega'} g_\theta(z) \, d\lambda(\theta) \, d\mu(z).$$

Consequently

$$\inf_{\theta \epsilon \Omega} \int \varphi'(z) g_\theta(z) \, d\mu(z) \leq \int_{\Omega'} d\lambda(\theta) \int \varphi'(z) g_\theta(z) \, d\mu(z)$$

$$(2.10) \qquad = \int \varphi'(z) \, d\mu(z) \int_{\Omega'} g_\theta(z) \, d\lambda(\theta) \leqq \int \varphi(z) \, d\mu(z) \int_{\Omega'} g_\theta(z) \, d\lambda(\theta)$$

$$= \int_{\Omega'} d\lambda(\theta) \int \varphi(z) g_\theta(z) \, d\mu(z) = \inf_{\theta \epsilon \Omega} \int \varphi(z) g_\theta(z) \, d\mu(z).$$

3. Normal alternatives. Let H be the hypothesis of invariance under Π, let $T(z)$ be the set of points equivalent to z (mod Π), and let f and g be two functions defined over $\mathfrak{Z}$. We shall write $f \sim g$ if there exists a function F such that

$$(3.1) \qquad f(z) = F[g(z), T(z)],$$

where for any fixed $T(z)$, F is a strictly increasing function of g. We note that $f \sim g$ in the following two special cases:

(i) $f(z) = F[g(z)]$ where F is strictly increasing;

(ii) $f(z) = a(z)g(z) + b(z)$ where $a(z) > 0$ for all z, and where $z_1 \sim z_2$ (mod Π) implies $a(z_1) = a(z_2)$, $b(z_1) = b(z_2)$.

The usefulness of this notation stems from the following remark. Let g^* and φ be defined as in (2.2) and (2.1) respectively and let $f \sim g$. If the test ψ is obtained from φ by substituting f and f^* for g and g^* respectively, then $\psi = \varphi$.

The purpose of the present section is to obtain most powerful and most stringent tests of some hypotheses of invariance against certain classes of normal alternatives. In particular, problems will be exhibited for which various non-parametric tests suggested in the literature possess these optimum properties.

PROBLEM 1. Suppose that the random variables Z_{ij} ($j = 1, \cdots, s_i$; $i = 1, \cdots, m$) have a joint probability density function, and denote by H the hypothesis that this probability density is invariant under all permutations of the s_i arguments within the ith group for $i = 1, \cdots, m$. Consider the alternative H_1 that all variables are independently distributed with common variance σ^2, and that

$$(3.2) \qquad E(Z_{ij}) = ax_{ij} + b_i,$$

where a, the b's and the x's are assumed known and where, without essential loss of generality, we assume $a > 0$. Assume further that

$$(3.3) \qquad \sum_{j=1}^{s_i} x_{ij} = 0.$$

In order to obtain the most powerful test of H against H_1, we apply theorem 2 with

$$(3.4) \qquad g(z) = c \exp\left[-\frac{1}{2\sigma^2}\Sigma\Sigma(z_{ij} - ax_{ij} - b_i)^2\right]$$

$$\sim \Sigma\Sigma(ax_{ij} + b_i)z_{ij} \sim \Sigma\Sigma x_{ij} z_{ij} .$$

The most powerful test is therefore given by (2.1), if we replace $g(z)$ by $\Sigma\Sigma x_{ij} z_{ij}$. This test being independent of σ^2, the b's and $a > 0$, it is uniformly most powerful against the class of alternatives obtained from H_1 by not specifying the values of these parameters but restricting a to be positive.

If we drop the restriction $a > 0$, a uniformly most powerful test no longer exists; we shall instead obtain the most stringent test against this extended class of alternatives, using theorems 3 and 4. Clearly the envelope power function is constant on the surfaces $|a|/\sigma^2 = $ constant. Take as the Ω of theorem 4, the set consisting of the two points $(a, b_1, \cdots b_m, \sigma)$ and $(-a, b_1, \cdots, b_m, \sigma)$. Let λ assign the probability $\frac{1}{2}$ to each of the two points. Then the function g of (2.7) becomes

$$\frac{1}{2}\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{\Sigma s_i} \exp\left\{\frac{1}{2\sigma^2}\Sigma\Sigma(z_{ij} - ax_{ij} - b_i)^2\right\} + \frac{1}{2}\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{\Sigma s_i}$$

$$(3.5) \qquad\qquad\qquad\qquad \exp\left\{-\frac{1}{2\sigma^2}\Sigma\Sigma(z_{ij} + ax_{ij} - b_i)^2\right\}$$

$$\sim \exp\{\Sigma\Sigma z_{ij}(ax_{ij} + b_i)\} + \exp\{\Sigma\Sigma z_{ij}(-ax_{ij} + b_i)\}$$

$$\sim \exp\{\Sigma\Sigma a x_{ij} z_{ij}\} + \exp\{-\Sigma\Sigma a x_{ij} z_{ij}\} \sim |\Sigma\Sigma x_{ij} z_{ij}|$$

The power of the test φ obtained by substituting this expression for g in (2.1) is the same at both points of Ω. For this test is most powerful for testing H against the simple alternatives H' that the density of the Z's is given by the first member of (3.5). But under the transformation $Z'_{ij} = -Z_{ij} + 2b_i$, H and H' and therefore the test φ are left invariant, while the two points of Ω are permuted.

Condition (2.8) of theorem 4 is therefore satisfied, and hence φ maximizes the minimum power over Ω. Since furthermore φ is independent of the particular set Ω chosen, it follows from theorem 3 that φ is most stringent for the problem under consideration. In case condition (3.3) is not satisfied, let $x'_{ij} = x_{ij} - x_i.$. Then $\Sigma\Sigma x'_{ij} = 0$ and $E(Z_{ij}) = ax'_{ij} + b'_i$.

Therefore the test criterion (3.5) becomes

$$(3.6) \qquad |\Sigma\Sigma z_{ij}(x_{ij} - x_i.)| = |\Sigma\Sigma(z_{ij} - z_i.)(x_{ij} - x_i.)| .$$

Some special cases of problem 1 are of particular interest.

a) Suppose that the variables of the ith group fall into two subgroups, and write for Z_{ij} : U_{ij} when $j = 1, \cdots, k_i$; V_{ij-k_i} when $j = k_i + 1, \cdots, k_i + l_i(k_i + l_i = s_i)$. Let

$$(3.7) \qquad x_{ij} = \begin{cases} 0 & \text{for} \quad j = 1, \cdots, k_i ; \\ 1 & \text{for} \quad j = k_i + 1, \cdots, k_i + l_i . \end{cases}$$

Then the alternatives ascribe to the variables normal distributions with common variance and such that

$$(3.8) \qquad E(U_{ij}) = b_i ; \qquad E(V_{ij}) = b_i + a.$$

The criterion becomes

$$(3.9) \qquad \Sigma\Sigma z_{ij}(x_{ij} - x_i) = \sum_{i=1}^{m} \left(\frac{k_i}{k_i + l_i} \Sigma v_{ij} - \frac{l_i}{k_i + l_i} \Sigma u_{ij} \right) = \Sigma \frac{1}{\frac{1}{k_i} + \frac{1}{l_i}} (v_i - u_i)$$

or

$$(3.10) \qquad \left| \Sigma \frac{1}{\frac{1}{k_i} + \frac{1}{l_i}} (v_i - u_i) \right|$$

according as a is restricted to positive values or not.

b) If we specialize still further and let $m = 1$, we are dealing with a problem which would coincide with the two sample problem if we added independence to the assumptions of the hypothesis. (3.10) becomes $| \bar{v} - \bar{u} |$, the criterion suggested by Pitman [3].

c) If instead of m we set $k_i = l_i = 1$ for $i = 1, \cdots, m$ we are testing interchangeability within each pair (u_i, v_i) against normal alternatives under which the means of U_i and V_i are different, the difference being independent of i. The criterion $| \Sigma (v_i - u_i) |$ to which (3.10) reduces was first suggested by R. A. Fisher [1].

d) As a last example set $m = 1$ in the original problem. Under the hypothesis the joint density of $Z_1, \cdots, Z_s$ is symmetric in its s arguments, while under the alternatives the Z's are normally distributed with common variance and mean $ax_i + b$. The criterion reduces to $| \Sigma (z_i - \bar{z})(x_i - \bar{x}) |$ which was proposed by Pitman [3].

We therefore see that several non-parametric tests which have been discussed in the literature are most powerful one-sided, or most stringent for testing a hypothesis of invariance against certain classes of normal alternatives. In a later section we shall indicate to what extent these results remain valid if to these hypotheses we add the assumption of independence.

The remaining problems will be considered somewhat more briefly since the proofs follow the same pattern as in problem 1.

PROBLEM 2. The conditions of problem 1d) are satisfied in particular if $x_1, \cdots, x_s$ are values taken on by random variables $X_1, \cdots, X_s$ and if under the alternatives the pairs (X_i, Z_i) have a common bivariate normal distribution with $\sigma_X^2 = \sigma_Z^2$. We are then concerned with a problem related to that of testing for absence of interclass correlation. For the corresponding intraclass problem, we consider random variables $X_1, \cdots, X_s, Z_1, \cdots, Z_s$, and test the hypothesis that the joint density of the $2s$ variables is symmetric in all its arguments, against the alternatives that the pairs (X_i, Z_i) have a common bivariate normal distribution, the means and variances of the X's and Z's being the same. We

shall only consider the case of positive correlation. Clearly, the criterion will be $\Sigma\, x_i z_i$ as in the one sided case of problem (d). However the tests differ, in that this expression must now be compared not only with the $s!$ expressions obtained by permuting the z's among themselves, but instead with the $(2s)!/2^s s!$ expressions obtained by considering all possible ways in which s pairs can be formed from the complete set of $2s$ observations.

PROBLEM 3. Consider once more the hypothesis that the joint density of $Z_1, \cdots, Z_n$ is symmetric in its n arguments, and consider the alternatives that the Z's are normally distributed with positive circular serial correlation. Then

$$(3.11) \quad g(z) = C \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} [(x_i - \xi) - \delta(x_{i+1} - \xi)]^2 \right\} \sim \sum z_i z_{i+1}$$

where $z_{n+1} = z_1$. The test based on this criterion, which was proposed by Wald and Wolfowitz [8], is therefore most powerful against the above class of alternatives.

PROBLEM 4. As a last problem, we shall test the hypothesis H that the joint density of $Z_1, \cdots, Z_n$ is symmetric in its n arguments and symmetric about each coordinate hyperplane, that is, invariant under the transformation $x_i' = -x_i$, $x_j' = x_j$ for all $j \neq i$, for $i = 1, \cdots, n$. This will be tested against the alternatives that the Z's are independently, identically distributed according to a normal distribution with non-zero mean. If we restrict this mean to positive values, we get

$$(3.12) \quad g(z) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left\{ -\frac{1}{2\sigma^2} \Sigma(z_i - \xi)^2 \right\} \sim \Sigma z_i .$$

If on the other hand both positive and negative values are allowed for the mean, the most stringent test is based on the statistic $|\Sigma z_i|$.

This test may be appropriate for some situations in which it is customary to use the sign test.

4. Binomial and other non-normal alternatives.

In the present section we shall be concerned mainly with generalisations of problems 1b) and 1c) of section 3. As described there, the hypotheses referred to the class of all probability densities in the usual sense. However, as was pointed out at the end of section 2, the same tests may be considered as referring to much wider hypotheses. If they are interpreted in this way, it is possible to greatly widen the class of alternatives without destroying the optimum properties of the tests.

Let $Z = (X_1, \cdots, X_n, Y_1, \cdots, Y_n)$ and denote by Π the partition under which two points z and z' are equivalent if they are obtainable from each other by a permutation of coordinates. Let H_0 be the hypothesis of invariance under Π. This is a generalization of the hypothesis of complete symmetry referring to a class of probability densities. Consider as alternative the class of distributions defined by

$$(4.1) \quad P\{Z \in A\} = \int_A C \exp\{\theta_1 \Sigma x_i + \theta_2 \Sigma y_i + \Sigma r(x_i) + \Sigma r(y_i)\}\, d\mu(z).$$

where the θ's are any real numbers, where μ is the $2n$th power of any one dimensional measure ν (and therefore invariant under Π), and where r is any ν-measurable function, subject only to the condition that the integral (4.1) converges when taken over the whole space.

We first consider the one-sided case $\theta_2 > \theta_1$. Using theorem 2 for a particular θ_1, θ_2, r and μ, we then have

$$g(z) = C \exp \{\theta_1 \Sigma x_i + \theta_2 \Sigma y_i + \Sigma r(x_i) + \Sigma r(y_i)\}$$

$$(4.2) \qquad \sim \theta_1 \Sigma x_i + \theta_2 \Sigma y_i \sim \theta_1 \Sigma x_i + \theta_2 \Sigma y_i - \tfrac{1}{2}(\theta_1 + \theta_2)\Sigma[x_i + y_i]$$

$$= \tfrac{1}{2}(\theta_1 - \theta_2)[\Sigma x_i - \Sigma y_i] \sim \Sigma y_i - \Sigma x_i .$$

Since this test does not depend on θ_1, θ_2, r or μ, it is uniformly most powerful against the one-sided class of alternatives $\theta_2 > \theta_1$.

Dropping the restriction $\theta_2 > \theta_1$, we apply theorem 4 with Ω the set consisting of the two points θ_1, θ_2, r, μ and θ_2, θ_1, r, μ. At these two points the envelope power function obviously takes on the same value. If for λ we select the distribution, which assigns equal probabilities to both points, then

$$g(z) \sim \exp \{\theta_1 \Sigma x_i + \theta_2 \Sigma y_i\} + \exp \{\theta_2 \Sigma x_i + \theta_1 \Sigma y_i\}$$

$$\sim \exp \{\tfrac{1}{2}(\theta_1 - \theta_2)[\Sigma x_i - \Sigma y_i]\} + \exp \{\tfrac{1}{2}(\theta_2 - \theta_1)[\Sigma x_i - \Sigma y_i]\}$$

$$\sim |\Sigma x_i - \Sigma y_i| \sim |\bar{y} - \bar{x}| .$$

The power of this test clearly is the same against both points of Ω. Since furthermore the test does not depend on the θ's, r, or μ, it is most stringent against H_1.

A univariate distribution such that

$$(4.4) \qquad P\{X \in A\} = \int_A C \exp \{\theta x + r(x)\} \, d\nu(x)$$

has been called Laplacian by Tweedie [9], who has studied these distributions in a different connection. Among others, the normal and χ^2, the binomial and Poisson distributions are Laplacian. To obtain, for example, the distribution of a characteristic variable, take for ν the measure ν^* which assigns to a set D the values 0. 1 or 2 according as D contains none, one or both of the points $x = 0$ and $x = 1$, and take as density the function

$$(4.5) \qquad p^x(1 - p)^{1-x} = (1 - p)e^{x \log(p/1-p)}$$

For comparison with tests which have been considered in the literature, one can specialize the problem just considered, so that the hypothesis $\bar{H}_0$ and the class of alternatives $\bar{H}_1$ consist only of those members of H_0 and H_1 which are generalized densities with respect to a fixed measure μ. One can specialize even further and take as alternative any subset of $\bar{H}_1$ provided with any point θ_1, θ_2, r, it also contains the point θ_2, θ_1, r. The test clearly will not change with these specializations, and the test based on (4.3) will therefore possess the same

optimum properties with respect to these special problems as with respect to the problem for which it was originally derived.

If in particular one selects for ν the measure ν^* mentioned above, one obtains the problem for which R. A. Fisher proposed the test based on (4.3). It follows that this test, Fisher's exact test, is most stringent in connection with the following problem: The random variables $X_1, \cdots, X_n, Y_1, \cdots, Y_n$ are characteristic variables, that is, they can take on only the values 0 and 1. If we let (4.6) $P\{X_1 = x_1, \cdots, Y_n = y_n\} = P(x_1, \cdots, y_n)$, the hypothesis states that the function P is invariant under all permutations of its arguments. An equivalent formulation is that the probability (4.6) depends only on $\Sigma\, x_i + \Sigma\, y_i$, the total number of "successes". Fisher's exact test is most stringent against the alternative that the X's and Y's are samples from two distinct populations of characteristic variables, that is, two populations corresponding to distinct probabilities of success.

Problem 1c) of section 3 can be extended quite analogously. Put again $Z = (X_1, \cdots, X_n, Y_1, \cdots, Y_n)$, and denote by Π the partition under which two points z and z' are equivalent provided they can be obtained from each other by a permutation of coordinates in which only the coordinates within pairs (X_i, Y_i) are interchanged. Consider the hypothesis of invariance under Π with reference to the class of all distributions and as alternative the class of distributions given by

$$(4.7) \qquad P\{Z \,\epsilon\, A\} = \int_A C \exp\left\{ \sum_{i=1}^{n} [\theta_1 x_i + \theta_2 y_i + r(x_i, y_i)] \right\} d\,\mu\,(z)\,.$$

The θ's here are any real numbers, μ is the $2n$th power of any one-dimensional measure ν, and r is any ν-measurable function such that (a) the integral (4.7) converges when A is the whole space, and such that (b) $r(x, y) = r(y, x)$.

Clearly in the one-sided case $\theta_2 > \theta_1$ we will again find $g(z) \sim \Sigma\, y_i - \Sigma\, x_i \sim \bar{y} - \bar{x}$, so that the associated test is uniformly most powerful against this one-sided class of alternatives, while the test based on $|\bar{y} - \bar{x}|$ is again most stringent against the full alternative H_1.

The class of distributions (4.7) contains the distributions (4.1) as a special case. If $(X_i, Y_i)\, i = 1, \cdots, n$ is a sample from a bivariate normal distribution with $\sigma_X^2 = \sigma_Y^2$, we get another case of (4.7).

As a last somewhat more special problem we mention a discrete analogue of problem 4 of section 3. Let $Z = (Z_1, \cdots, Z_n)$ and consider the class of generalized densities given by

$$(4.8) \qquad P\{Z \,\epsilon\, A\} = \int_A P(z_1, \cdots, z_n)\, d\mu(z)$$

where μ is the nth power of ν^*. Let H_0 be the hypothesis that P is invariant under permutations of the coordinates and under the group generated by the transformations $z_i' = 1 - z_i$, $z_j' = z_j\; j \neq i$ for $i = 1, \cdots, n$. This is an extension of the hypothesis that the probability of success in a binomial dis-

tribution equals $\frac{1}{2}$. The test of H_0 against the alternatives that $Z_1, \cdots Z_n$ is a sample of a characteristic variable is based on Σz_i or $|\Sigma z_i|$ as $P\{Z_i = 1\}$ is restricted to be greater than $\frac{1}{2}$ or is not so restricted. In the first case the test is most powerful, in the second most stringent.

5. Hypotheses of invariance for independent variables. To the results obtained so far, a different interpretation can be given, which throws some light on certain related problems. Theorem 2 gave sufficient conditions for a test to be most powerful against a simple alternative H_1 for the hypothesis H_0 of invariance under a partition Π. However, if taken in conjunction with section 1, the theorem can be intepreted as giving sufficient conditions for a test to be the most powerful test of structure $S(\epsilon)$ with respect to Π against H_1. That is, the theorem is really independent of the hypothesis, and depends solely on the alternative and on the class of tests admitted into competition, in our case the class of all tests having structure $S(\epsilon)$ with respect to Π. The same remark obviously also applies to most stringent tests.

Let us now consider a special class of partitions. Let Z stand for the m groups of random variables $(Z_{i1}, \cdots, Z_{is_i})$ $(i = 1, \cdots, m)$ and let Π denote the partition under which two points z and z' are equivalent provided they can be obtained from each other by a permutation of coordinates which however permutes only the coordinates within the m groups. Let μ be the power of a one-dimensional measure ν, and assume that the probability distribution of Z is absolutely continuous with respect to μ and that the Z's are independently distributed, so that

$$(5.1) \qquad P\{Z \epsilon A\} = \int_A \prod_{i,j} [f_{ij}(z_{ij}) \, d\nu(z_{ij})].$$

Under these assumptions consider the hypothesis H that f_{ij} is independent of j, that is, that the Z's are identically distributed within each group. It easily can be shown that not all admissible tests of H that have size ϵ, have structure $S(\epsilon)$. However a generalization of a result of Feller [10] and Scheffé [5] for the case $m = 1$ and $\mu =$ Lebesgue measure, states that the only tests which are of size ϵ and similar for H, are the tests of structure $S(\epsilon)$ with respect to Π [11]. It follows that any test which is most powerful or most stringent for testing the hypothesis H' of invariance under Π for the class of generalised densities with respect to μ, has the same property relative to the class of all tests which are similar for testing H.

As an example, take problem 1b) of section 3. Here μ is Lebesgue measure, m is 1, and we put

$$(5.2) \qquad Z_{ij} = \begin{cases} U_j & \text{for } j = 1, \cdots, k \\ V_{j-k} & \text{for } j = k+1, \cdots, k+l = s. \end{cases}$$

It was shown in section 3 that the test based on $|\bar{u} - \bar{v}|$, Pitman's test, is most stringent for testing the hypothesis that the joint density of the U's and V's is

symmetric in its $k + l$ arguments against the alternative that the variables are independently normally distributed with common variance and such that $E(U_i) = \xi$, $E(V_i) = \eta$ where ξ and η are any distinct real numbers. It follows now that the same test is most stringent similar for testing against the same class of alternatives the hypothesis that $U_1, \cdots, U_k, V_1, \cdots, V_l$ are independently distributed, all with the same probability density. This is the hypothesis for which Pitman proposed his test, and the result just stated is a partial solution of the problem recently raised by Wilks [12], to determine the class of alternatives for which Pitman's test is satisfactory.

If we modify the example by taking for μ instead of Lebesgue measure the $k + l$th power of the measure ν^* of section 4, we are dealing with characteristic variables $U_1, \cdots, U_k, V_1, \cdots, V_l$. We have shown earlier that if $k = l$ the test based on $|\bar{u} - \bar{v}|$ is most stringent for testing the hypothesis of complete permutability against the alternative that the U's and V's are samples from two distinct populations of characteristic variables. If we add to this hypothesis the assumption of independence of all variables, we obtain a parametric problem, namely essentially the problem of testing equality of probability of success in two binomial populations corresponding to the same number of trials. It now follows that the test based on $|\bar{u} - \bar{v}|$ is most stringent for this problem. As is well known, it is also the uniformly most powerful, unbiased similar test.

These two examples suffice to illustrate the type of result that can be obtained. It should perhaps be mentioned that the equivalence discussed at the beginning of this section, can be utilized also in the opposite direction. The fact, for example, that the test based on $|\bar{u} - \bar{v}|$ is known to be uniformly most powerful unbiased similar for testing equality of probability of success in two populations of characteristic variables from which the U's and V's are samples, proves that this test is uniformly most powerful unbiased for testing the hypothesis of complete symmetry for the joint generalized density of the U's and V's.

6. Extension to infinite equivalence classes. The definition of a hypothesis of invariance given in section 1—in spite of the restriction to finite equivalence classes—was sufficiently general to cover the non-parametric problems that we wanted to study. It is possible however to extend the definition so as to allow infinite equivalence classes. In this concluding section we shall briefly outline a theory based on such a broader definition. This will enable us to point out a relationship between the approach of the present paper and the standard parametric theory.

Let $\mathcal{Z}$ be a space of points z and $\mathcal{C}$ an additive class of subsets of $\mathcal{Z}$. We define a partition of $\mathcal{Z}$ into subsets $\{S_t\}$ as follows: Let $\mathcal{T}$ be some space, and for each $t \, \epsilon \, \mathcal{T}$ let S_t be a measurable subset of $\mathcal{Z}$ (i.e. an element of $\mathcal{C}$) such that the S_t are mutually exclusive and exhaustive. Let $\mathcal{C}_0$ be the class of all $C_0 \, \epsilon \, \mathcal{C}$ which can be expressed in the form

$$(6.1) \qquad\qquad C_0 = \bigcup_{t \, \epsilon \, D_0} S_t$$

and let $\mathfrak{D}_0$ be the class of all D_0 occurring in such relationships. For each $t \in \mathcal{T}$ let G_t be a specified probability measure over $\mathfrak{A}_t$, where $\mathfrak{A}_t$ is the class of A_t such that $A_t \in S_t$, $A_t \in \mathfrak{A}$. Let Z be a random variable distributed over $\mathfrak{Z}$ according to an unknown probability measure F. Let $\psi(z)$ be that $t \in \mathcal{T}$ for which $z \in S_t$, and let $T = \psi(Z)$. Let H_0 be the hypothesis that for each $t \in \mathcal{T}$ the conditional distribution of Z given $Z \in S_t$ is G_t, i.e. that there exists a probability measure Q_0 over $\mathfrak{C}_0$ such that for all $A \in \mathfrak{A}$

$$(6.2) \qquad F(A) = \int G_t(A \cap S_t) \, dQ_0(t).$$

It is seen that we have essentially the situation described in section 1, except that there we assumed further that each S_t was finite and for all t, G_t assigned equal probabilities to all points of S_t.

We say that a test φ of H_0 has structure $S(\epsilon)$ if the conditional expectation $E_t[\varphi(Z)]$ of $\varphi(Z)$ given $Z \in S_t$ satisfies

$$(6.3) \qquad E_t[\varphi(z)] = \int_{S_t} \varphi \, dG_t = \epsilon \quad \text{for all} \quad t.$$

The lemmas and theorems stated below are straight-forward generalizations of those in section 1 so that no proof will be given.

LEMMA 1'. *Any test φ of structure $S(\epsilon)$ with respect to H_0 is similar and of size ϵ for H_0.*

LEMMA 2'. *If φ is any test of H_0 of size $\leq \epsilon$, there exists a test φ_1 of H_0 having structure $S(\epsilon)$ and such that*

$$(6.4) \qquad \int \varphi_1 \, dF \geq \int \varphi \, dF$$

for all probability measures F, for which the conditional distribution of Z given $Z \in S_t$ is absolutely continuous with respect to G_t for all t.

Suppose next there is defined another partition of $\mathfrak{Z}$ into sets $\{S_u'\}$ by means of a space $\mathfrak{U}$, and let $\mathfrak{C}_1$, $\mathfrak{D}_1$ and $\mathfrak{A}_u$ refer to this second partition. We shall assume that for every $t \in \mathcal{T}$, $u \in \mathfrak{U}$ either $S_u' \subset S_t$ or $S_u' \cap S_t$ is empty. Let G_u' be a specified probability measure over $\mathfrak{A}_u$ and suppose that for each $t \in \mathcal{T}$ here exists a probability measure Q_t such that for all $A_t \in \mathfrak{A}_t$

$$(6.5) \qquad G_t(A_t) = \int G_u'(A_t \cap S_u') dQ_t(u).$$

If H_1 denotes the hypothesis that for each $u \in \mathfrak{U}$ the conditional distribution of Z given $Z \in S_u'$ is G_u', we can state

THEOREM 1'. *For testing H_0 against H_1 at level of significance ϵ, the totality of tests φ which have structure $S(\epsilon)$ and for which z, $z' \in S_u'$ implies $\varphi(z) = \varphi(z')$ form an essentially complete class of admissible tests.*

Let F_1 be a distribution not in H_0, and for each $t \in \mathcal{T}$ let G_{1t} be the conditional distribution of Z given $Z \in S_t$. We suppose that for each $t \in \mathcal{T}$, G_{1t} is chosen to be a true probability measure, which is possible in most cases of practical

interest (see Doob [13] for a discussion of this point). Then we have the equivalent of theorem 2:

THEOREM 2′. *Let*

(6.6)
$$G_{1t}(A_t) = \int_{A_t - H_t} g_t \, dG_t + G_{1t}(A_t \cap H_t)$$

for all $A_t \subset S_t$, *where in accordance with the Radon-Nikodym Theorem* [14], g_t *is a non-negative function integrable over* S_t, *and* $H_t \subset S_t$ *has* G_t *measure 0 and does not depend on* A_t. *For testing* H_0 *against* H_1, *a most powerful test of size* ϵ *is given by* $\varphi(z) = \varphi_t(z)$ *for* $z \, \epsilon \, S_t$ *where*

(6.7)
$$\varphi_t(z) = \begin{cases} 1 & if \quad z \, \epsilon \, H_t \\ 1 & if \quad g_t(z) > c_t \\ a_t & if \quad g_t(z) = c_t \\ 0 & if \quad g_t(z) < c_t \end{cases}$$

where c_t *and* a_t *are so chosen that* φ *has structure* $S(\epsilon)$.

Theorems 3 and 4 require no modification.

As in the case of finite equivalence classes the results just outlined can be interpreted differently. Again the theorems are really independent of the hypotheses, but depend only on the alternatives and on the class of tests admitted into competition. This class of tests φ is in the present case defined by condition (6.4), that the conditiona expectation of φ given $Z \, \epsilon \, S_t$ equals ϵ. But this is just the condition which in the standard approach to the problem of testing a composite parametric hypothesis for which T is a sufficient statistic, by means of similar regions is frequently found to be the necessary and sufficient condition for φ to be similar. (See for example [15]). For these cases therefore the hypotheses of the present section represent non-parametric analogues to which the same tests apply with the same optimum properties but without the a priori restriction to similar regions.

As a simple illustration of this remark, let $\hat{Z} = (Z_1, \cdots, Z_n)$, and let $T = \sum_{i=1}^{n} Z_i^2$. For the conditional distribution of Z given $T = t$ take the uniform distribution over the sphere $T = t$, and for μ take Lebesgue measure. Then the hypothesis H states merely that the joint probability density of the Z's is a function only of $\sum_{i=1}^{n} Z_i^2$. If we add to this the assumption of independence of the Z's, we obtain the new hypothesis H' that the Z's are a sample from a normal distribution with zero mean. The tests φ for which the conditional expectation over each sphere is ϵ, constitute the only admissible tests of H and the only admissible similar tests of H'. If as alternatives we consider that the Z's are a sample from a normal distribution with mean $\xi > 0$, the test

(6.8)
$$\frac{\bar{x}}{\sqrt{\Sigma(x_i - \bar{x})^2}} \geqslant C$$

is uniformly most powerful for H and uniformly most powerful similar for H'. If we do not restrict ξ to positive values, the test

$$(6.9) \qquad \left| \frac{\bar{x}}{\sqrt{\Sigma(x_i - \bar{x})^2}} \right| \geq C',$$

Student's test, is uniformly most powerful unbiased and most stringent for testing H, uniformly most powerful unbiased similar and most stringent similar for testing H'.

REFERENCES

[1] R. A. FISHER, *Design of Experiments*, Oliver and Boyd, Edinburgh, 1935.

[2] J. NEYMAN, K. IWASKIEWICZ AND ST. KOLODZIECZYK, "Statistical problems in agricultural experimentation," *Roy. Stat. Soc. Jour. Suppl.*, Vol. 2 (1935), p. 107.

[3] E. J. G. PITMAN, "Significance tests which may be applied to samples from any proportion, *Roy. Stat. Soc. Jour. Suppl.*, Vol. 4 (1937), p. 119; II. The correlation coefficient test, *Roy. Stat. Soc. Jour. Suppl.*, Vol. 4 (1937), p. 225; III. The analysis of variance test, *Biometrika*, Vol. 29 (1938), p. 322.

[4] E. L. LEHMANN AND C. STEIN, "Most powerful tests of composite hypotheses. I. Normal distributions," *Annals of Math. Stat.*, Vol. 19 (1948).

[5] H. SCHEFFÉ, "On a measure problem arising in the theory of non-parametric tests," *Annals of Math. Stat.*, Vol. 14 (1943), p. 227.

[6] A. WALD, "An essentially complete class of admissible decision functions," *Annals of Math. Stat.*, Vol. 18 (1947), p. 549.

[7] G. HUNT AND C. STEIN, "Most stringent tests of statistical hypotheses," *unpublished*.

[8] A. WALD AND J. WOLFOWITZ, "An exact test for randomness in the non-parametric case, based on serial correlation," *Annals of Math. Stat.*, Vol. 14 (1943), p. 378.

[9] M. C. K. TWEEDIE, "Functions of a statistical variate with given means, with special reference to Laplacian distributions," *Cam. Phil. Soc. Proc.*, Vol. 43 (1947), p. 41.

[10] W. FELLER, "Note on regions similar to the sample space," *Stat. Res. Memoirs*, Vol. 2 (1938), p. 117.

[11] E. L. LEHMANN AND H. SCHEFFÉ, "Completeness, similar regions and unbiased estimation," *unpublished*.

[12] S. S. WILKS, "Order Statistics," *Am. Math. Soc. Bull.*, Vol. 54 (1948), p. 6.

[13] J. L. DOOB, "Asymptotic properties of Markoff transition probabilities," *Trans. Amer. Math. Soc.*, Vol. 63 (1948), footnote p. 399.

[14] S. SAKS, *Theory of the Integral*, Stechert, 1937.

[15] J. NEYMAN AND E. S. PEARSON, "On the problem of the most efficient tests of statistical hypotheses," *Roy. Soc. London Phil. Trans.*, Ser. A., Vol. 231 (1933), p. 289.

[16] A. WALD, *On the Principles of Statistical Inference*, Notre Dame Mathematical Lectures, Number 1.