SOME SIGNIFICANCE TESTS FOR THE MEDIAN WHICH ARE VALID UNDER VERY GENERAL CONDITIONS[1]

By John E. Walsh

The Rand Corporation

1. Summary. Order statistics are used to derive significance tests for the population median which are valid under very general conditions. These tests are approximately as powerful as the Student t-test for small samples from a normal population. Also the application of a test requires very little computation. Thus the tests derived compare very favorably with the t-test for small sets of observations. Applications of these order statistic tests to certain well known statistical problems are given in another paper [1].

PART I. RESULTS AND DEFINITIONS

2. Introduction. Consider n independent observations drawn from n populations satisfying the conditions (A):

1) Each population is continuous (i.e. its cdf is continuous).
2) Each population is symmetrical.
3) The median of each population has the same value ϕ. (If the 50% point of a continuous symmetrical population is not unique, the median ϕ of the population is *defined* to be the midpoint of the segment of 50% values.)

It is to be emphasized that no two of the observations are necessarily drawn from the same population. Significance tests are derived to compare ϕ with a given constant value ϕ_0.

A general method of obtaining one-sided and symmetrical tests is given in section 8. This general method furnishes tests which have significance levels of the form $r/2^n$, $(r = 1, \cdots, 2^n - 1)$. Each value of r can be attained for some one-sided test. Unfortunately tests obtained by the general method are very difficult to apply from a computational viewpoint. If $n \geq 10$, the number of computations required for the application of a test is prohibitive.

To overcome the computational difficulty involved in using the general method, easily applied tests using order statistics are derived. These tests are based on order statistics of certain combinations of order statistics of the n observations, each combination being either a single order statistic of the n observations or one-half the sum of two order statistics. The tests are invariant under permutation of the n observations and have significance levels of the form $r/2^n$, $(r = 1, \cdots, 2^n - 1)$. Table 1 contains a list of some one-sided and symmetrical tests for $n \leq 15$ ($x_1, \cdots, x_n$ represent the n observations arranged in increasing order of magnitude). Additional significance tests can be obtained by use of Theorem 4 of section 6.

[1] The results presented in this paper were obtained in the course of research conducted under the sponsorship of the Office of Naval Research. This research was performed while the author was at Princeton University.

If a symmetrical population has a mean, the mean has the same value as the median. Thus if each population from which an observation is drawn satisfies the additional condition that its mean exists, the median tests derived in this paper are also tests of the mean.

Although it is unlikely that conditions (A) are ever exactly satisfied in practice, these conditions appear to be approximately satisfied in many practical situations. Moreover conditions (A) are of such a simple form that approximate verification can frequently be obtained without an extensive investigation.

Certain of the order statistic tests are very efficient if the n observations are a sample from a normal population. Efficiencies are listed for some of the tests in Table 1. These tests are approximately as efficient as the Student t-test. (The efficiency of a test, more precisely the power efficiency, is defined in section 3.)

The order statistic tests are competitive with the Student t-test. In choosing between the two types of tests the following considerations may be of interest:

(a) The order statistic tests are valid under much more general conditions than the t-test.

(b) The order statistic tests are almost as efficient as the t-test for small samples from a normal population.

(c) The order statistic tests are more easily computed than the t-test.

(d) For the case of a sample from a normal population and near significance the t-test gives more information than the order statistic tests.

In some cases a set of n independent observations satisfying only 1) and 3) of conditions (A) can be transformed into observations approximately satisfying all of conditions (A) by an appropriate continuous monotonic change of variable. For example, replacing each observation by the logarithm of the value of the observation sometimes results in a set of observations having approximately symmetrical distributions. Since the transformation, say $g(x)$, is continuous and monotonic, the resulting observations will have median $g(\phi)$ if the original observations have median ϕ. Confidence intervals can be found for ϕ by first obtaining confidence intervals for $g(\phi)$ on the basis of conditions (A) and then inverting. Significance tests can be obtained from these confidence intervals.

The tests of Part I can be applied to furnish generalized solutions for several well known statistical problems. Some of these applications are given in another paper [1].

One application occurs in cases where there is reason to believe that conditions (A) are satisfied but there is no reason to assume that the populations from which the observations were drawn are even approximately the same. Perhaps the most common situation of this type is that in which the value of a certain quantity is experimentally determined by several different methods, all of which should theoretically yield the same result. Then there is no reason to believe that all the experimental values have the same precision. It may be permissible, however, to assume that each value is an observation from a continuous symmetrical population and that all the populations have the same median. Then the order statistic tests can be used to test the true value of the quantity investigated. For example, consider the determination of a specified physical constant.

TABLE 1

Some one-sided and symmetrical significance tests for $n \leq 15$

n	Significance Level of Tests: One-sided	Significance Level of Tests: Symmetrical	Tests (Symmetrical: Accept $\phi \neq \phi_0$ if either) One-sided: Accept $\phi < \phi_0$ if	Tests (Symmetrical: Accept $\phi \neq \phi_0$ if either) One-sided: Accept $\phi > \phi_0$ if	Approx. Efficiency for Normality
	%	%			%
4	6.2	12.5	$x_4 < \phi_0$	$x_1 > \phi_0$	95
5	6.2	12.5	$\frac{1}{2}(x_4 + x_5) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$	98
	3.1	6.2	$x_5 < \phi_0$	$x_1 > \phi_0$	96
6	4.7	9.4	$\max[x_5, \frac{1}{2}(x_4 + x_6)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_3)] > \phi_0$	97
	3.1	6.2	$\frac{1}{2}(x_5 + x_6) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$	98
	1.6	3.1	$x_6 < \phi_0$	$x_1 > \phi_0$	95
7	5.5	10.9	$\max[x_5, \frac{1}{2}(x_4 + x_7)] < \phi_0$	$\min[x_3, \frac{1}{2}(x_1 + x_4)] > \phi_0$	95
	2.3	4.7	$\max[x_6, \frac{1}{2}(x_5 + x_7)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_3)] > \phi_0$	98
	1.6	3.1	$\frac{1}{2}(x_6 + x_7) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$	98
	0.8	1.6	$x_7 < \phi_0$	$x_1 > \phi_0$	95
8	4.3	8.6	$\max[x_6, \frac{1}{2}(x_4 + x_8)] < \phi_0$	$\min[x_3, \frac{1}{2}(x_1 + x_5)] > \phi_0$	94.5
	2.7	5.5	$\max[x_6, \frac{1}{2}(x_5 + x_8)] < \phi_0$	$\min[x_3, \frac{1}{2}(x_1 + x_4)] > \phi_0$	96
	1.2	2.3	$\max[x_7, \frac{1}{2}(x_6 + x_8)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_3)] > \phi_0$	98
	0.8	1.6	$\frac{1}{2}(x_7 + x_8) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$	98
	0.4	0.8	$x_8 < \phi_0$	$x_1 > \phi_0$	95
9	5.1	10.2	$\max[x_6, \frac{1}{2}(x_4 + x_9)] < \phi_0$	$\min[x_4, \frac{1}{2}(x_1 + x_6)] > \phi_0$	91
	2.2	4.3	$\max[x_7, \frac{1}{2}(x_5 + x_9)] < \phi_0$	$\min[x_3, \frac{1}{2}(x_1 + x_5)] > \phi_0$	96
	1.0	2.0	$\max[x_8, \frac{1}{2}(x_5 + x_9)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_5)] > \phi_0$	95.5
	0.6	1.2	$\max[x_8, \frac{1}{2}(x_7 + x_9)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_3)] > \phi_0$	99
	0.4	0.8	$\frac{1}{2}(x_8 + x_9) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$	98

10	5.6	11.1	$\max[x_6, \frac{1}{2}(x_4 + x_{10})] < \phi_0$	$\min[x_5, \frac{1}{2}(x_1 + x_7)] > \phi_0$	87.5
	2.5	5.1	$\max[x_7, \frac{1}{2}(x_5 + x_{10})] < \phi_0$	$\min[x_4, \frac{1}{2}(x_1 + x_6)] > \phi_0$	93
	1.1	2.1	$\max[x_8, \frac{1}{2}(x_6 + x_{10})] < \phi_0$	$\min[x_3, \frac{1}{2}(x_1 + x_5)] > \phi_0$	96.5
	0.5	1.0	$\max[x_9, \frac{1}{2}(x_6 + x_{10})] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_5)] > \phi_0$	96.5
11	4.8	9.7	$\max[x_7, \frac{1}{2}(x_4 + x_{11})] < \phi_0$	$\min[x_5, \frac{1}{2}(x_1 + x_8)] > \phi_0$	
	2.8	5.6	$\max[x_7, \frac{1}{2}(x_5 + x_{11})] < \phi_0$	$\min[x_5, \frac{1}{2}(x_1 + x_7)] > \phi_0$	89
	1.1	2.1	$\max[\frac{1}{2}(x_6 + x_{11}), \frac{1}{2}(x_8 + x_9)] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_6), \frac{1}{2}(x_3 + x_4)] > \phi_0$	
	0.5	1.1	$\max[x_9, \frac{1}{2}(x_7 + x_{11})] < \phi_0$	$\min[x_3, \frac{1}{2}(x_1 + x_5)] > \phi_0$	97
12	4.7	9.4	$\max[\frac{1}{2}(x_4 + x_{12}), \frac{1}{2}(x_5 + x_{11})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_9), \frac{1}{2}(x_2 + x_8)] > \phi_0$	
	2.4	4.8	$\max[x_8, \frac{1}{2}(x_5 + x_{12})] < \phi_0$	$\min[x_5, \frac{1}{2}(x_1 + x_3)] > \phi_0$	
	1.0	2.0	$\max[x_9, \frac{1}{2}(x_6 + x_{12})] < \phi_0$	$\min[x_4, \frac{1}{2}(x_1 + x_7)] > \phi_0$	93.5
	0.5	1.1	$\max[\frac{1}{2}(x_7 + x_{12}), \frac{1}{2}(x_9 + x_{10})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_6), \frac{1}{2}(x_3 + x_4)] > \phi_0$	
13	4.7	9.4	$\max[\frac{1}{2}(x_4 + x_{13}), \frac{1}{2}(x_5 + x_{12})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_{10}), \frac{1}{2}(x_2 + x_9)] > \phi_0$	
	2.3	4.7	$\max[\frac{1}{2}(x_5 + x_{13}), \frac{1}{2}(x_6 + x_{12})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_9), \frac{1}{2}(x_2 + x_8)] > \phi_0$	
	1.0	2.0	$\max[\frac{1}{2}(x_6 + x_{13}), \frac{1}{2}(x_9 + x_{10})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_8), \frac{1}{2}(x_4 + x_5)] > \phi_0$	
	0.5	1.0	$\max[x_{10}, \frac{1}{2}(x_7 + x_{13})] < \phi_0$	$\min[x_4, \frac{1}{2}(x_1 + x_7)] > \phi_0$	94.5
14	4.7	9.4	$\max[\frac{1}{2}(x_4 + x_{14}), \frac{1}{2}(x_5 + x_{13})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_{11}), \frac{1}{2}(x_2 + x_{10})] > \phi_0$	
	2.3	4.7	$\max[\frac{1}{2}(x_5 + x_{14}), \frac{1}{2}(x_6 + x_{13})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_{10}), \frac{1}{2}(x_2 + x_9)] > \phi_0$	
	1.0	2.0	$\max[x_{10}, \frac{1}{2}(x_6 + x_{14})] < \phi_0$	$\min[x_5, \frac{1}{2}(x_1 + x_9)] > \phi_0$	90.5
	0.5	1.0	$\max[\frac{1}{2}(x_7 + x_{14}), \frac{1}{2}(x_{10} + x_{11})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_8), \frac{1}{2}(x_4 + x_5)] > \phi_0$	
15	4.7	9.4	$\max[\frac{1}{2}(x_4 + x_{15}), \frac{1}{2}(x_5 + x_{14})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_{12}), \frac{1}{2}(x_2 + x_{11})] > \phi_0$	
	2.3	4.7	$\max[\frac{1}{2}(x_5 + x_{15}), \frac{1}{2}(x_6 + x_{14})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_{11}), \frac{1}{2}(x_2 + x_{10})] > \phi_0$	
	1.0	2.0	$\max[\frac{1}{2}(x_6 + x_{15}), \frac{1}{2}(x_{10} + x_{11})] < \phi_0$	$\min[\frac{1}{2}(x_1 + x_{10}), \frac{1}{2}(x_5 + x_6)] > \phi_0$	
	0.5	1.0	$\max[x_{11}, \frac{1}{2}(x_7 + x_{15})] < \phi_0$	$\min[x_5, \frac{1}{2}(x_1 + x_9)] > \phi_0$	92

Various scientists obtained experimental values for this constant by several different methods. If it can be assumed that each value is an observation from a continuous symmetrical population and that all the populations have the same median, the true value of the physical constant can be tested by applying the order statistic tests to the totality of the experimental values.

3. Power efficiency of tests. A problem which arises throughout the paper is that of determining how much information is lost by using some other test in place of the most powerful test of a given hypothesis. The quantitative measure of the amount of available information which is used by a test will be given as a percentage and is called the power efficiency of the test considered.

In all cases investigated the underlying population is normal with unknown variance and the hypotheses tested concern the population median (mean). Then the most powerful test (one-sided or symmetrical) is the appropriate Student t-test.

The procedure used to measure the power efficiency of a test is different from the common method of measuring the efficiency of an estimate. The efficiency of an estimate is obtained by taking the ratio of the variance of an efficient estimate with respect to the variance of the given estimate (expressed as a percentage). The method of determining the power efficiency of a test, however, consists in continuously varying the sample size of the appropriate most powerful test (same significance level) until the power functions of the given test and the most powerful test are equivalent in the following sense: The area between the two power curves for which the power function of the most powerful test exceeds the power function of the given test is equal to the analogous area for which the power function of the most powerful test is less than that of the given test. (It is assumed that the power functions of the tests can be made to depend on the values of a single parameter.) *The sample size (not necessarily integral) of the most powerful test with equivalent power function divided by the sample size of the given test is called the power efficiency of the given test (expressed as a percentage).*

In obtaining power efficiencies in the manner defined above, the sample size of the most powerful test is allowed to assume non-integral values. This furnishes an interpolated measure of the same size of the most powerful test which is power function equivalent to the given test. As pointed out above, the t-test is a most powerful test for the situations considered in this paper. A method of computing power function values for t-tests having non-integral sample sizes is given below.

The definition of power efficiency selected is very convenient from a computational point of view. Power function values for the t-test can be easily computed through use of the normal approximation given in [2]. For the significance levels considered in this paper, the normal approximation is reasonably accurate if the sample size is not too small. In the remaining cases the approximation underestimates some power function values and overestimates others. For the situations investigated, however, the error introduced by this combination of

TABLE 2

Efficiencies and power function values for certain order statistic tests

Significance Test	Sample Size	Approx. Efficiency	Significance Level	Values of Power Function $\delta = .6$	$\delta = 1.2$	$\delta = 1.8$
		%				
t	4.9		.0625	.337	.755	.964
$\frac{1}{2}(x_4 + x_5) < \phi_0$	5	98	.0625	.343	.755	.958
t	5.82		.0469	.327	.779	.980
$\max[x_5, \frac{1}{2}(x_4 + x_6)] < \phi_0$	6	97	.0469	.334	.779	.972
t	5.88		.0312	.244	.682	.951
$\frac{1}{2}(x_5 + x_6) < \phi_0$	6	98	.0312	.254	.687	.942
t	6.65		.0547	.406	.869	.994
$\max[x_5, \frac{1}{2}(x_4 + x_7)] < \phi_0$	7	95	.0547	.413	.867	.991
t	6.85		.0234	.239	.716	.969
$\max[x_6, \frac{1}{2}(x_5 + x_7)] < \phi_0$	7	98	.0234	.249	.717	.962
t	7.55		.0430	.395	.882	.996
$\max[x_6, \frac{1}{2}(x_4 + x_8)] < \phi_0$	8	94.5	.0430	.404	.879	.993
t	7.85		.0117	.174	.650	.956
$\max[x_7, \frac{1}{2}(x_6 + x_8)] < \phi_0$	8	98	.0117	.185	.656	.949
t	8.64		.0215	.302	.839	.994
$\max[x_7, \frac{1}{2}(x_5 + x_9)] < \phi_0$	9	96	.0215	.311	.834	.990
t	8.9		.0059	.127	.597	.947
$\max[x_8, \frac{1}{2}(x_7 + x_9)] < \phi_0$	9	99	.0059	.137	.599	.935
t	7.5		.0547	.450	.910	.998
$x_8 < \phi_0$	10	75	.0547	.454	.901	.995
t	9.65		.0107	.227	.790	.991
$\max[x_8, \frac{1}{2}(x_6 + x_{10})] < \phi_0$	10	96.5	.0107	.237	.786	.986
t	8.2		.0098	.176	.668	.964
$\max[x_9, \frac{1}{2}(x_1 + x_{10})] < \phi_0$	10	82	.0098	.191	.677	.952
t	8.9		.0059	.141	.621	.954
$x_{10} < \phi_0$	11	81	.0059	.152	.634	.942
t	11.22		.0102	.277	.870	.998
$\max[x_9, \frac{1}{2}(x_6 + x_{12})] < \phi_0$	12	93.5	.0102	.288	.862	.995

underestimation and overestimation tends to cancel out in the determination of power efficiencies if the above area definition of equality of power functions is used. Thus application of the normal approximation yields reasonably accurate power efficiencies for the cases considered in this paper. Use of the normal approximation furnishes an easily applied method of obtaining power function values for t-tests having non-integral sample sizes.

Table 2 contains examples of the above described method of determining power efficiencies. Here the power function values for the t-test were computed using the normal approximation. Examination of Table 2 shows that the maximum difference between corresponding power function values for the two types of tests is small for all the cases considered there. This holds in the determination of all the power efficiencies listed in Table 1.

Investigation indicates that the definition of power efficiency given here is for all practical purposes the same as that given in [3].

For the situations considered in this paper, it is sufficient to restrict power efficiency investigations to one-sided tests. Every symmetric test investigated can be considered as a combination of two non-overlapping one-sided tests, each having a significance level equal to half that of the symmetric test. Also, from symmetry, these one-sided tests (each considered as a separate test) have the same power efficiency. Thus it is an immediate consequence of the definition of power efficiency that the symmetric test has the same efficiency as each of the corresponding one-sided tests at half the significance level.

PART II. DERIVATIONS

4. Introduction. The purpose of the remainder of the paper is to present derivations of the significance test results stated in sections 1 and 2. The first derivations consist in obtaining confidence intervals for ϕ on the basis of conditions (A). Then properties of these confidence intervals are analyzed. Application of the confidence intervals and their properties to significance tests furnishes many of the results stated in sections 1 and 2. The remaining derivations are concerned with efficiencies and the general method mentioned in section 2.

5. Derivation of confidence intervals. Let us consider n independent observations, each observation being drawn from a possibly different population. Denote these observations by $y_1, \cdots, y_n$ and let the cdf of y_i be given by F_i, $(i = 1, \cdots, n)$. Furthermore let the n populations from which these n observations were drawn satisfy conditions (A). Then 1) of conditions (A) requires that each F_i is continuous, while 2) and 3) stipulate that

$$\int_{-\infty}^{-c} dF_i(y_i - \phi) = \int_{c}^{\infty} dF_i(y_i - \phi), \qquad (i = 1, \cdots, n),$$

for all values of c in the interval $-\infty < c < \infty$.

Let $x_1, \cdots, x_n$ represent $y_1, \cdots, y_n$ arranged in increasing order of magnitude. Since the cdf's are continuous, $Pr(x_i = x_j; i \neq j) = 0$. For the situations treated in this paper, it is sufficient to consider one-sided confidence intervals for ϕ. All one-sided confidence intervals derived have one of the forms

$$(1)\qquad \begin{aligned} g(x_1, \cdots, x_n) &< \phi, \\ h(x_1, \cdots, x_n) &> \phi, \end{aligned}$$

where g and h are Borel measurable functions of $x_1, \cdots, x_n$ such that

$$Pr[g(x_1, \cdots, x_n) < \phi] = Pr[g(x_1 - \phi, \cdots, x_n - \phi) < 0],$$

$$Pr[h(x_1, \cdots, x_n) > \phi] = Pr[h(x_1 - \phi, \cdots, x_n - \phi) > 0].$$

Consider the additional condition

(B) All populations are the same.

In terms of cumulative distribution functions, condition (B) requires that all the cdf's F_i are equal to some cdf F. A theorem will be proved which shows that all confidence intervals of the forms (1) derived on the basis of both conditions (A) and (B) are also valid if only conditions (A) necessarily hold; i.e. if

$$Pr[g(x_1, \cdots, x_n) < \phi] = p$$

whenever $x_1, \cdots, x_n$ are order statistics of observations from populations satisfying conditions (A) and (B), then this probability expression also has the value p if $x_1, \cdots, x_n$ are from populations necessarily satisfying only conditions (A). Similarly for $Pr[h(x_1, \cdots, x_n) > \phi]$.

THEOREM 1. *Let* $Q(x_1 - \phi, \cdots, x_n - \phi)$ *be a probability statement involving* $x_1 - \phi, \cdots, x_n - \phi$, *which defines a Borel measurable region* $R(x_1 - \phi, \cdots, x_n - \phi)$ *of the n-dimensional order statistic space. If*

$$(2)\qquad Q(x_1 - \phi, \cdots, x_n - \phi) = p$$

whenever $x_1, \cdots, x_n$ *are order statistics of n independent observations from populations satisfying conditions* (A) *and* (B), *then* (2) *also holds when* $x_1, \cdots, x_n$ *are order statistics of n independent observations from populations necessarily satisfying only conditions* (A).

PROOF. It is sufficient to consider the case in which $\phi = 0$. Then, if conditions (A) are satisfied, the joint probability element of $x_1, \cdots, x_n$ is

$$dF(x_1, \cdots, x_n) = \sum_{\pi} dF_1(x_{\pi(1)}) \cdots dF_n(x_{\pi(n)}),$$

where the summation is taken over all permutations π of the integers $1, \cdots, n$, and F's are cdf's of symmetrical populations with zero median. Let $R = R(x_1, \cdots, x_n)$ be the region of the n-dimensional order statistic space defined by the probability statement $Q(x_1, \cdots, x_n)$. Then Theorem 1 stipulates that

$$(3)\qquad \int_R dF(x_1, \cdots, x_n) = p$$

whenever $y_1, \cdots, y_n$ are from populations satisfying conditions (A) and (B) with zero median. In this case, however, each $F_i = F$ and (3) becomes

$$n! \int_R \prod_{i=1}^{n} dF(x_i) = p, \tag{4}$$

where F is the cdf of a population satisfying conditions (A) and (B) with zero median. Let

$$P = \prod_{i=1}^{n} \left(\sum_{j=1}^{n} dF_j(x_i) \right)$$

and define S_α^β to be the sum of all terms in the expansion of P which contain a specified α of $dF_1, \cdots . dF_n$ and no others; the particular set chosen is denoted by β, where $\beta = 1, \cdots, \binom{n}{\alpha}$. Then

$$P = F(x_1, \cdots, x_n) + \sum_\beta S_{n-1}^\beta + \cdots + \sum_\beta S_1^\beta .$$

Now consider any given S_α^β (i.e. α, β given). Define dH to be the sum of the α of $dF_1, \cdots, dF_n$ pertaining to β plus any set of zero or more of the remaining dF's. Then no matter which of the remaining dF's are chosen for dH, the sum of those terms in the expansion of $\prod_{i=1}^{n} dH(x_1)$ which contain the particular set of α of $dF_1, \cdots, dF_n$ is always equal to S_α^β. Let

$$P_\alpha = \sum_\beta \left(\prod_{i=1}^{n} dG_\alpha^\beta(x_i) \right),$$

where dG_α^β equals the sum of the α of $dF_1, \cdots, dF_n$ pertaining to β. Then from the above and the symmetrical fashion in which the dF's are treated,

$$P_\alpha = \sum_\beta S_\alpha^\beta + K_{\alpha-1}^{(\alpha)} \sum_\beta S_{\alpha-1}^\beta + \cdots + K_1^{(\alpha)} \sum_\beta S_1^\beta ,$$

where the $K_u^{(\alpha)}$ $(u = 1, \cdots, \alpha - 1)$, are constants.

Consider the case in which $\alpha = n - 1$. Using the above expression for P_α,

$$P = dF(x_1, \cdots, x_n) + P_{n-1} + (1 - K_{(n-2)}^{(n-1)}) \sum_\beta S_{n-2}^\beta + \cdots + (1 - K_1^{(n-1)}) \sum S_1^\beta .$$

Repeating this procedure successively for $\alpha = n - 2, n - 3, \cdots, 1$ shows that

$$dF(x_1, \cdots, x_n) = P + C_{n-1}P_{n-1} + \cdots + C_1 P_1 ,$$

where the C_v, $(v = 1, \cdots, n - 1)$, are constants.

Since each F_i is the cdf of a symmetrical population with zero median,

$$G_\alpha^\beta/\alpha = \frac{1}{\alpha} \text{ (sum of the } \alpha \text{ of } F_1, \cdots, F_n \text{ pertaining to } \beta)$$

is also the cdf for a continuous symmetrical population with zero median. But

$$P_\alpha = \alpha^n \left(\frac{P_\alpha}{\alpha^n}\right) = \alpha^n \sum_\beta \left(\prod_{i=1}^n dG_\alpha^\beta(x_i)/\alpha\right).$$

Hence $dF(x_1, \cdots, x_n)$ is equal to a sum of terms (multiplied by certain constants) of the form

$$n! \prod_{i=1}^n dF(x_i),$$

where F is the cdf of a continuous symmetrical population with zero median. Thus from (4) and the linear properties of the integral,

$$\int_R dF(x_1, \cdots, x_n) = p$$

if $y_1, \cdots, y_n$ are from populations necessarily satisfying only conditions (A). Q.e.d.

Next confidence intervals of the forms (1) will be derived for ϕ on the basis of conditions (A) and (B). Before stating the theorem on which these confidence intervals are based consider the following definition of notation: For each permissible selection of i and j, the symbol

$$\{i, j\} \qquad (1 \le i \le j \le n)$$

denotes an *arbitrary* but *fixed* selection of one or both of the inequality signs $<, >$. The selection of both inequality signs, denoted by $\lessgtr$, has the interpretation

$$x_i \lessgtr \phi \equiv -\infty < x_i < \infty$$

$$(x_i + x_j)/2 \lessgtr \phi \equiv -\infty < (x_i + x_j)/2 < \infty.$$

It is to be noted that $\{r, s\}$ is not necessarily equal to $\{i, j\}$ unless $r = i$ and $s = j$.

THEOREM 2. *Consider the probability statement*

$$(5) \qquad Pr[(x_i + x_j)/2 \; \{i, j\} \; \phi; 1 \le i \le j \le n].$$

Let this statement have the value q if $x_1, \cdots, x_n$ *are order statistics of a sample of size n drawn from the uniform population with range* $-\frac{1}{2}$ *to* $\frac{1}{2}$ *(then* $\phi = 0$*). Then* (5) *also has the value q if* $x_1, \cdots, x_n$ *are order statistics of a sample size n drawn from any population satisfying conditions* (A) *and* (B).

PROOF. Let $y_1, \cdots, y_n$ be a sample of n values from a population satisfying conditions (A) and (B) while $x_1, \cdots, x_n$ are the y's arranged in increasing order of magnitude. Then there is a monotone function π (see [4]) such that $\pi(z)$ will have the same cdf as $y_i - \phi$ if z is from a uniform population with range $-\frac{1}{2}$ to $\frac{1}{2}$. Since the y's are from a symmetrical population, $-\pi(z) = \pi(-z)$. Let $x_i - \phi = \pi(z_i)$, $(i = 1, \cdots, n)$, define the z_i. Then

$$Pr[(x_i + x_j)/2\{i, j\}\phi] = Pr[(\pi(z_i) + \pi(z_j))\{i, j\}0]$$
$$= Pr[\pi(z_i)\{i, j\} - \pi(z_j)].$$

From the monotone and symmetrical properties of the function π,

$$Pr[\pi(z_i)\{i, j\} - \pi(z_j)] = Pr[\pi(z_i)\{i, j\}\pi(-z_j)]$$
$$= Pr[z_i\{i, j\} - z_j].$$

By hypothesis this last expression has the value q, thus completing the proof.

Many of the probability statements of the form (5) have zero probability. For example, $Pr[x_1 > \phi, x_2 < \phi, \cdots] = 0$. Also many selections of the symbols $\{i, j\}$ result in equivalent probability statements. For example

$$Pr(x_1 \lessgtr \phi, x_2 < \phi) \equiv Pr(x_1 < \phi, x_2 < \phi).$$

An immediate consequence of Theorem 2 is that one-sided confidence intervals can be obtained for ϕ by choosing any specified subset of $(x_i + x_j)/2$, $(1 \leq i \leq j \leq n)$, and considering an arbitrary but fixed order statistic of the values of this subset. For example, consider the subset consisting of x_{n-1} and $(x_{n-2} + x_n)/2$. Then

$$Pr\{\max[x_{n-1}, (x_{n-2} + x_n)/2] < \phi\} = Pr[(x_i + x_j)/2\{i, j\}\phi],$$

where

$$\{i, j\} = \begin{cases} < \text{ if either } i = j = n - 1; \quad \text{or } i = n - 2, \quad j = n; \\ \lessgtr \text{ otherwise.} \end{cases}$$

In general, the confidence coefficient of any one-sided confidence interval formed by considering a certain order statistic of a specified subset of $(x_i + x_j)/2$, $(1 \leq i \leq j \leq n)$, can be expressed as a sum of probabilities of the form (5), where $\{i, j\} = \lessgtr$ if $(x_i + x_j)/2$ is not included in the specified subset, $(i \leq j)$.

It is usually preferable to select the subset of $(x_i + x_j)/2$, $(1 \leq i \leq j \leq n)$, in such a way that no two of the elements chosen necessarily have an order relation.

Satisfactory two-sided confidence intervals can usually be obtained as combinations of one-sided confidence intervals.

6. Confidence coefficients. The purpose of this section is to show that all the confidence coefficients for one-sided confidence intervals derived on the basis of Theorem 2 are of the form $r/2^n$, $(r = 1, \cdots, 2^n - 1)$. Also a method of determining confidence coefficient values for one-sided confidence intervals is developed.

First a theorem will be presented which shows that each of the one-sided confidence intervals derived in the preceding section has a confidence coefficient of the form $r/2^n$, $(r = 1, \cdots, 2^n - 1)$. On the basis of Theorem 2 it is sufficient to prove:

THEOREM 3. *Let* $x_1, \cdots, x_n$ *be the ordered values of a sample from the uniform population with range* $-\frac{1}{2}$ *to* $\frac{1}{2}$. *Then*

$$Pr[(x_i + x_j)/2\{i, j\}\ 0; 1 \leq i \leq j \leq n] = r/2^n$$

where r *has one of the values* $0, 1, \cdots, 2^n$. (*The symbol* $\{i, j\}$ *is defined in section* 5).

SKETCH OF PROOF. This theorem is proved by investigating how the hyperplanes

$$\tfrac{1}{2}(x_i + x_j) = 0 \qquad (1 \leq i \leq j \leq n),$$

intersect the n-dimensional order statistic space for the particular population considered. It is found that each relation of the form

$$\tfrac{1}{2}(x_i + x_j)\ \{i, j\}\ 0, \qquad (1 \leq i \leq j \geq n)$$

defines a region of the n-dimensional order statistic space which consists of a certain number r of n-dimensional "basic" cells each of which has an n-dimensional, "volume" equal to $(\frac{1}{2})^n$. A detailed proof of this theorem is given in [5].

Next a method will be developed whereby confidence coefficient values can be determined for any *one-sided* confidence interval of the form

$$\tfrac{1}{2}(x_i + x_j)\ \{i, j\}\ \phi, \qquad (1 \leq i \leq j \leq n). \tag{6}$$

For this purpose it is sufficient to derive a procedure for determining the confidence coefficient of any confidence interval of the form

$$\max\ [\text{certain subset of } \tfrac{1}{2}(x_i + x_j); 1 \leq i \leq j \leq n] < \phi. \tag{7}$$

The confidence coefficient of any one-sided confidence interval of the form $\min\ [\,] > \phi$ can be obtained by symmetry. The confidence coefficient of any other one-sided confidence interval of the form (6) can be found by expressing the value of

$$Pr\ [\tfrac{1}{2}(x_i + x_j)\ \{i, j\}\ \phi]$$

as a sum of terms of the form $Pr\{\max\ [\,] < \phi\}$ or as a sum of terms of the form $Pr\{\min\ [\,] > \phi\}$. That this is always possible for one-sided confidence intervals of the form (6) is shown by direct application of the results of page 17 of [6].

It is not difficult to show that any one-sided confidence interval of the form (7) can be expressed in the form

$$\max\ \{x(n - k), \tfrac{1}{2}[x(n - k + 1) + x(n - m_k - k + 1)], \cdots, \tfrac{1}{2}[x(n) + x(n - m_1)]\} < \phi,$$

where

$$x(i) = x_i, \qquad (i = 1, \cdots, n),$$

and $m_1, \cdots, m_k$ are k integers such that

$$n \geq m_1 > m_2 > \cdots > m_k > 0.$$

This is done by choosing $k, m_1, \cdots, m_k$ so that the two confidence intervals are equivalent.

Thus it is sufficient to prove the following theorem:

THEOREM 4. *Let $x(1), \cdots, x(n)$ represent the ordered values of n independent observations drawn from populations satisfying conditions* (A). *Choose a set of k integers $m_1, \cdots, m_k$ such that*

$$n \geq m_1 > m_2 > \cdots > m_k > 0.$$

Then the one-sided confidence interval

$$\max \{x(n-k), \tfrac{1}{2}[x(n-k+1) + x(n-m_k-k+1)], \cdots, \tfrac{1}{2}[x(n) + x(n-m_1)]\} < \phi, \tag{8}$$

where a term of the form $\frac{1}{2}[x(n-h+1) + x(n-m_h-h+1)]$, $(h = 1, \cdots, k)$, is to be deleted if[2] $n - m_h - h + 1 = 0$, *has the confidence coefficient*

$$2^{-n}\left[1 + m_1 + \sum_{i_1=1}^{m_2}(m_1 - i_1) + \sum_{i_2=1}^{m_3}\sum_{i_1=1}^{m_2-i_2}(m_1 - i_1 - i_2) + \cdots + \sum_{i_{k-1}=1}^{m_k}\sum_{i_{k-2}=1}^{m_{k-1}-i_{k-1}}\cdots\sum_{i_1=1}^{m_2-i_2-\cdots-i_{k-1}}(m_1 - i_1 - \cdots - i_{k-1})\right]. \tag{9}$$

SKETCH OF PROOF. It is sufficient to consider the case in which the n observations are a sample from the uniform population with range $-\frac{1}{2}$ to $\frac{1}{2}$ (then $\phi = 0$).

Let us consider the region of the n-dimensional order statistic space defined by (8). This region can be considered as an intersection of n-dimensional regions each of which is completely defined by a certain region in an x_j, x_i plane $(1 \leq i < j \leq n)$. Also the n-dimensional "volume" of this region equals the value of the confidence coefficient of the confidence interval (8).

By Theorem 3, the intersection region of (8) consists of a certain number of "basic" cells, each of n-dimensional "volume" $(\frac{1}{2})^n$. Theorem 4 is proved by developing a method for finding the number of "basic" cells in this intersection region on the basis of the corresponding regions in the x_i, x_j planes. It is found that the intersection region consists of

$$1 + m_1 + \cdots + \sum_{i_{k-1}=1}^{m_k}\cdots\sum_{i_1=1}^{m_2-i_2-\cdots-i_{k-1}}(m_1 - i_1 - \cdots - i_{k-1})$$

"basic" cells. A detailed derivation of this expression is given in [5].

Now consider some examples of the application of Theorem 4. Let $n = 11$, $m_1 = 11$, $m_2 = 5$, $m_3 = 2$. Then, by Theorem 4, the one-sided confidence interval

$$\max [x_8, \tfrac{1}{2}(x_9 + x_7), \tfrac{1}{2}(x_{10} + x_5)] < \phi$$

[2] For the trivial case in which $k = n$ the value of (9) is unity.

has a confidence coefficient equal to $103/2^{11}$. If $n = 12$ instead of 11, the confidence coefficient would be $103/2^{12}$ while the confidence interval becomes

$$\max [x_9, \tfrac{1}{2}(x_{10} + x_8), \tfrac{1}{2}(x_{11} + x_6), \tfrac{1}{2}(x_{12} + x_1)] < \phi.$$

As another example, let $n = 11$ and consider the confidence interval.

$$\text{Max} [x_3, \tfrac{1}{2}(x_9 + x_7), \tfrac{1}{2}(x_{10} + x_5), \tfrac{1}{2}(x_{11} + x_4)] < \phi.$$

Here $k = 3$ and comparison with (8) shows that this confidence interval satisfied Theorem 4 with $m_1 = 7$, $m_2 = 5$, $m_3 = 2$. Thus it has a confidence coefficient equal to $51/2^{11}$.

Theorem 3 shows that each one-sided confidence interval developed on the basis of Theorem 2 has a confidence coefficient of the form $r/2^n$, $(0 \leq r \leq 2^n)$. The question arises as to whether the one-sided confidence intervals defined by Theorem 4 have confidence coefficients which attain each of the values $1/2^n$, $2/2^n, \cdots, (2^n - 1)/2^n$. That this is not the case is proved as follows: The totality of different confidence intervals of the form (8) is equal to $2^n - 1$. This is shown by counting how many ways the integers $m_1, \cdots, m_k$ can be selected subject to the conditions $n \geq m_1 > m_2 > \cdots > m_k > 0$. It is easily seen that there are $\binom{n}{k}$ possible ways. Summing over the possible values of k yields $2^n - 1$. This figure is increased to 2^n if the confidence interval $x_n < \phi$ is also included. Examination of (9) shows, however, that two different selections of m_1, m_2, etc., will result in the same value of (9) for more than one case. Thus the one-sided confidence intervals of Theorem 4 do not have confidence coefficients which attain each of the values $1/2^n, \cdots, (2^n - 1)/2^n$.

Although the class of one-sided confidence intervals defined by Theorem 4 do not have confidence coefficients which attain each of the values $1/2^n, 2/2^n, \cdots,$ $(2^n - 1)/2^n$, they do have another property which is important from a practical point of view: If a certain confidence coefficient can be obtained for a particular value of n, then this confidence coefficient can also be obtained for all greater values of n. This result is a consequence of the following theorem:

THEOREM 5. *Let $x(1), \cdots, x(n)$ be the ordered values of n independent observations drawn from populations satisfying conditions* (A). *Then if a confidence interval of the form* (8) *has the confidence coefficient ϵ for a certain value n_0 of n, it is always possible to obtain another confidence interval of the form* (8), *which has the confidence coefficient ϵ for the value $n_0 + 1$.*

PROOF. Let $m_1, \cdots, m_k$ be the integers corresponding to the given confidence interval of form (8). These integers satisfy the condition

$$n_0 \geq m_1 > m_2 > \cdots m_k > 0.$$

Let n_0 be replaced by $n_0 + 1$ and consider the new set of integers $(m_1 + 1)$, $(m_2 + 1), \cdots, (m_k + 1)$, 1. Evidently

$$n_0 + 1 \geq m_1 + 1 > \cdots > m_k + 1 > 1 > 0.$$

Hence these integers can be used to define a confidence interval of the form (8). Also it is easily verified that

$$1 + (m_1 + 1) + \sum_{i_1=1}^{m_2+1} (m_1 + 1 - i_1)$$
$$+ \cdots + \sum_{i_{k-1}=1}^{m_k+1} \cdots \sum_{i_1=1}^{m_2+1-i_2-\cdots-i_{k-1}} (m_1 + 1 - i_1 - \cdots - i_{k-1})$$
$$+ \sum_{i_k=1}^{1} \sum_{i_{k-1}=1}^{m_k-1-i_k} \cdots \sum_{i_1=1}^{m_2+1-i_2-\cdots-i_k} (m_1 - 1 - i_1 - \cdots - i_k)$$
$$= 2\left[1 + m_1 + \sum_{i_1=1}^{m_2} (m_1 - i_1) + \cdots + \sum_{i_{k-1}=1}^{m_k} \cdots \sum_{i_1=1}^{m_2-i_2-\cdots-i_{k-1}} (m_1 - i_1 - \cdots - i_{k-1})\right].$$

Thus the new confidence interval has the same confidence coefficient as the given confidence interval.

From symmetry considerations, the one-sided confidence interval

$$\min \{x(k + 1), \tfrac{1}{2}[x(k) + x(m_k + k)], \cdots, \tfrac{1}{2}[x(1) + x(m_1 + 1)]\} > \phi,$$

where a term of the form $\frac{1}{2}[x(h) + x(m_h + h)]$, $(h = 1, \cdots, k)$, is to be deleted if $m_h + h = n + 1$, has the same confidence coefficient as the one-sided confidence interval (8); i.e. its confidence coefficient is given by (9).

7. Efficiency of some tests based on conditions (A). Let us consider the case in which the n observations used for a test are a sample from a normal population with unknown variance. The purpose of this section is to investigate the efficiency of some tests based on conditions (A) for this special case.

The method used to obtain efficiencies is outlined in section 3. Only one-sided and symmetrical tests are considered. For this purpose it is sufficient to limit investigations to one-sided tests of $\phi < \phi_0$.

If the subset of $\frac{1}{2}(x_i + x_j)$, $(1 \le i \le j \le n)$, chosen for a test is not of one of the forms

(a) x_i

(b) $\frac{1}{2}(x_i + x_j)$, $(i < j)$;

(c) x_j, $\frac{1}{2}(x_i + x_k)$, $(i < j < k)$,

the determination of power function values requires a numerical double or higher order integration. Such numerical integrations are extremely lengthy. For this reason only one-sided significance tests based on subsets of the forms (a) – (c) will be investigated.

Let the normal population have variance σ^2 and consider one-sided tests of $\phi < \phi_0$ based on subsets of the form (a). Then

Power Function $= Pr\ (x_i < \phi_0)$

$$= Pr\left(\frac{x_i - \phi}{\sigma} < \frac{\phi_0 - \phi}{\sigma}\right) = \sum_{s=1}^{n} \frac{n}{s!\,(n - s)!} (N(\delta))^s (1 - N(\delta))^{n-}$$

where

$$\delta = (\phi_0 - \phi)/\sigma, \qquad N(\delta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\delta} e^{-\frac{1}{2}y^2}\, dy.$$

The power function values listed for the test $x_i < \phi_0$ in Table 2 were computed from the above expression. The corresponding values for the t-test were computed from the normal approximation given in [2].

For subsets of forms (b) and (c) the expression for the power function is more complicated and will not be either derived or stated here. For any particular case, however, a simple analysis will yield an expression for the power function which requires only a first order numerical integration. General expressions for the power functions when the subsets are of the forms (b) and (c) are stated and derived in [5].

Table 2 contains power function values and efficiencies for several tests based on subsets of the forms (b) and (c). The power function values were computed by approximate integration (Simpson's rule, etc.). The t-test power function values were obtained by using the normal approximation. The power efficiencies listed in Table 1 for tests which do not appear in Table 2 were computed in [5], where a table of power function values is also given.

Examination of Table 2 shows that many of the tests formed from subsets of types (b) and (c) are very efficient for small values of n. The efficiency appears to decrease as n increases. Also the efficiency of a test depends strongly on the subset of $\frac{1}{2}(x_i + x_j)$, $(1 \leq i \leq j \leq n)$, used to form the test. For example, et $n = 10$. The test

$$Accept\ \phi < \phi_0\ if\ \max\,[x_9\,,\, \tfrac{1}{2}(x_1 + x_{10})] < \phi_0$$

l
has a significance level of approximately .01 but an efficiency of only 82%. However the test

$$Accept\ \phi < \phi_0\ if\ \max[x_8\,,\, \tfrac{1}{2}(x_6 + x_{10})] < \phi_0$$

also has a significance level of approximately .01 but an efficiency of 96.5%

An approximate set of rules for picking subsets which result in efficient tests of $\phi < \phi_0$ is suggested by the results of Table 2. Let $x(i_1), \cdots, x(i_r)$ be the order statistics which make up the elements of the particular subset of $\frac{1}{2}(x_i + x_j)$, $(1 \leq i \leq j \leq n)$, to be used for the test. The approximate rules are

1. Use the maximum of the values of the elements of the subset.
2. Choose $i_1, \cdots, i_r$ so that $\max(i_1, \cdots, i_r) = n$ and $\min(i_1, \cdots, i_r)$ is as large as possible subject to the restriction that the test is to have a significance level of a specified order of magnitude.

Symmetry considerations furnish the corresponding set of rules for obtaining efficient tests of $\phi < \phi_0$.

Other tests at approximately the same significance levels but not based on subsets of the forms (a)–(c) are undoubtedly more efficient than many of the tests considered in Tables 1 and 2 (particularly for the larger values of n). Computational difficulties, however, prevent consideration of more general situations.

8. A general solution.[3] A general method of obtaining one-sided tests of $\phi < \phi_0$ and $\phi > \phi_0$, also symmetrical tests of $\phi \neq \phi_0$, on the basis of conditions (A) is the following:

Let $y_1, \cdots, y_n$ be n independent observations drawn from populations satisfying conditions (A). Let

$$z_i = y_i - \phi_0 \qquad (i = 1, \cdots, n).$$

If the null hypothesis of $\phi = \phi_0$ is satisfied, each z_i is an observation from a population satisfying conditions (A) with zero median. Consider the 2^n sets of values obtained by the transformations

$$z_i \rightarrow \epsilon(i) z_i, \qquad (i = 1, \cdots, n).$$

where $\epsilon(i)$ is one of the signs $+$ or $-$. Form the mean of each of the 2^n sets of values. Then it is readily seen, from conditions (A), that the probability that $\bar{z}(= \Sigma z_i/n)$ is less than the $(r + 1)$th largest of the 2^n means has the value $r/2^n$ when the null hypothesis is true. Similarly the probability that $\bar{z}$ is greater than the $(2^n - r)$th largest of the 2^n means is equal to $r/2^n$ if the null hypothesis of $\phi = \phi_0$ is satisfied. Thus the test

Accept $\phi < \phi_0$ *if* $\bar{z}$ *is less than the* $(r + 1)$*th largest of the* 2^n *means.*

is a one-sided test of $\phi < \phi_0$ with significance level equal to $r/2^n$. Likewise the one-sided test

Accept $\phi > \phi_0$ *if* $\bar{z}$ *is greater than the* $(2^n - r)$*th largest of the* 2^n *means.*

has the significance level $r/2^n$. Consequently the symmetrical test

Accept $\phi \neq \phi_0$ *if* $\bar{z}$ *is either less than the* $(r + 1)$*th largest or greater than the* $(2^n - r)$*th largest of the* 2^n *means.*

has a significance level equal to $2r/2^n$.

The application of any of the above tests requires the computation of the 2^n means and a determination of where $\bar{z}$ falls in the ordering of these means. If $n = 5$, only 32 means need be computed. If $n = 10$, however, 1024 means must be computed. Evidently this test is too cumbersome to apply except for very small values of n.

9. Acknowledgements. The author would like to express his appreciation to Professors S. S. Wilks and John W. Tukey for valuable advice and assistance in

[3] This solution was derived independently by E. J. G. Pitman and the author. The fundamental idea on which the solution is based was presented by R. A. Fisher in [7].

the preparation of this paper, also to Mrs. Ruth S. Shafer for computational assistance.

REFERENCES

[1] JOHN E. WALSH, "Applications of some significance tests for the median which are valid under very general conditions," submitted to *Am. Stat. Assn. Jour.*

[2] N. L. JOHNSON AND B. L. WELCH, "Applications of the non-central t-distribution," *Biometrika*, Vol. 31 (1940), p. 376.

[3] JOHN E. WALSH, "On the power function of the sign test for slippage of means," *Annals of Math. Stat.*, Vol. 17 (1946), pp. 360–361.

[4] H. SCHEFFÉ AND J. W. TUKEY, "Non-parametric estimation. I. Validation of order statistics," *Annals of Math. Stat.*, Vol. 16 (1945), pp. 187–192.

[5] JOHN E. WALSH, "Some significance tests for the median which are valid under very general conditions," unpublished thesis, Princeton University.

[6] G. UDNY YULE AND M. G. KENDALL, *An Introduction to the Theory of Statistics*, Griffin and Co., 1947.

[7] R. A. FISHER, *The Design of Experiments*, Oliver and Boyd, 1942.