

CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION CFSTI
DOCUMENT MANAGEMENT BRANCH 410.11

LIMITATIONS IN REPRODUCTION QUALITY

ACCESSION #

- 1. WE REGRET THAT LEGIBILITY OF THIS DOCUMENT IS IN PART UNSATISFACTORY. REPRODUCTION HAS BEEN MADE FROM BEST AVAILABLE COPY.
- 2. A PORTION OF THE ORIGINAL DOCUMENT CONTAINS FINE DETAIL WHICH MAY MAKE READING OF PHOTOCOPY DIFFICULT.
- 3. THE ORIGINAL DOCUMENT CONTAINS COLOR, BUT DISTRIBUTION COPIES ARE AVAILABLE IN BLACK-AND-WHITE REPRODUCTION ONLY.
- 4. THE INITIAL DISTRIBUTION COPIES CONTAIN COLOR WHICH WILL BE SHOWN IN BLACK-AND-WHITE WHEN IT IS NECESSARY TO REPRINT.
- 5. LIMITED SUPPLY ON HAND: WHEN EXHAUSTED, DOCUMENT WILL BE AVAILABLE IN MICROFICHE ONLY.
- 6. LIMITED SUPPLY ON HAND: WHEN EXHAUSTED DOCUMENT WILL NOT BE AVAILABLE.
- 7. DOCUMENT IS AVAILABLE IN MICROFICHE ONLY.
- 8. DOCUMENT AVAILABLE ON LOAN FROM CFSTI (TT DOCUMENTS ONLY).
- 9.

NBS 9 64

PROCESSOR:

SOME BOUNDED SIGNIFICANCE LEVEL TESTS FOR
THE MEDIAN

John Walsh

P-156

16 June 1950

The **RAND** *Corporation*

1500 FOURTH ST. • SANTA MONICA • CALIFORNIA

1

SOME BOUNDED SIGNIFICANCE LEVEL TESTS FOR THE MEDIAN

John E. Walsh
The RAND Corporation

✓ In practice it is often permissible to assume that the observations of a set are statistically independent and from continuous populations with a common median. Then the population median can be investigated by using the sign test. For small numbers of observations, however, the sign test does not furnish very many suitable significance levels. Also, some of the sign tests with suitable significance levels are not very efficient. This note presents some tests whose significance levels are only approximate but cover a wide range of suitable values. The significance levels of these tests are exactly determined if the populations are symmetrical; they are bounded otherwise. Some of these bounded significance level tests have high efficiencies. ←

INTRODUCTION AND STATEMENT OF RESULTS

For practical situations involving a small number n of observations, it is frequently permissible to assume that

- (i). The observations are statistically independent.
- (ii). The observations are from continuous populations.
- (iii). The populations have a common median ϕ .

This is the case, for example, if the observations are a sample from a continuous population. Then the sign test can be used to compare ϕ with a given hypothetical value ϕ_0 .

Let x_1, \dots, x_n denote the values of the n observations arranged in increasing order of magnitude. Then x_1 = value of smallest observation, etc. The one-sided sign test of $\phi < \phi_0$ can be expressed in the form

$$(1) \quad \text{Accept } \phi < \phi_0 \text{ if } x_1 < \phi_0.$$

The significance level of this test has the value

$$(2) \quad \left(\frac{1}{2}\right)^n \sum_{s=1}^n \frac{n!}{s!(n-s)!}$$

if assumptions (i)-(iii) hold. The one-sided test of $\phi > \phi_0$ can be defined as

$$\text{accept } \phi > \phi_0 \text{ if } x_{n+1-i} > \phi_0.$$

The significance level of this test also equals (2) when (i)-(iii) hold. The equal-tail sign test of $\phi \neq \phi_0$ is defined by

$$\underline{\text{accept}} \phi \neq \phi_0 \underline{\text{if either}} x_i < \phi_0 \underline{\text{or}} x_{n+1-i} > \phi_0, .$$

where $i > (n+1)/2$. The significance level of this test equals twice the value of (2) when (i)-(iii) are satisfied.

The sign tests are easy to apply and valid under very general conditions. However, for small values of n (say, $n \leq 10$) they have the drawbacks

(I). For given n , the number of tests with suitable significance levels is quite limited.

(II). If $i < n$, the efficiency of the sign test is not very great. First consider (I). For $n = 10$, there are only three suitable significance levels (i.e., in the range .001-.07 for one-sided tests or in the range .002-.014 for equal-tail tests). Moreover, these values are not very near each other. For $n < 10$, there are at most two satisfactory significance levels for given n , and these are not very close together.

Now consider (II). If $i = n$, the efficiency is approximately 95% for the special case of a sample from a normal population (unknown standard deviation). If $i = n - 1$, this efficiency is only about 80% and drops further to about 75% if $i = n - 2$ (see [1]). Accepting the efficiency for normality as an indication of that for most situations of a practical nature, the sign test is not very efficient if $i < n$.

The one-sided test of $\phi < \phi_0$ presented in this note is

$$(3) \quad \underline{\text{accept}} \phi < \phi_0 \underline{\text{if}} \max \left[x_{n-1}, \frac{1}{2}(x_n + x_{n+1-j}) \right] < \phi_0.$$

If the condition

(iv). The populations are symmetrical.

is satisfied in addition to conditions (i)-(iii), the significance level for test (3) is $j(\frac{1}{2})^n$. The one-sided test of $\phi > \phi_0$ is

$$\underline{\text{accept}} \phi > \phi_0 \underline{\text{if}} \min \left[x_2, \frac{1}{2}(x_1 + x_j) \right] > \phi_0.$$

The significance level of this test also equals $j(\frac{1}{2})^n$ if (i)-(iv) hold. The "symmetrical" test of $\phi \neq \phi_0$ is

$$\underline{\text{Accept}} \phi \neq \phi_0 \underline{\text{if either}} \max \left[x_{n-1}, \frac{1}{2}(x_n + x_{n+1-j}) \right] < \phi_0 \\ \underline{\text{or}} \min \left[x_2, \frac{1}{2}(x_1 + x_j) \right] > \phi_0.$$

The significance level of this test has the value $j\left(\frac{1}{2}\right)^{n-1}$ when conditions (i)-(iv) are satisfied.

If only conditions (i)-(iii) are necessarily satisfied, the significance levels of the tests presented are bounded by those of sign tests. For example, the significance levels of (1) with $i = n$ and $i = n - 1$ are lower and upper bounds, respectively, for test (3) when only (i)-(iii) need hold.

The bounds obtained for a test are not very close together. Usually the value of the upper bound differs from that of the lower bound by a factor somewhere between 5 and 11. It should be remembered, however, that the upper and lower limits represent the worst that can happen; these bounds are attained only for the most extreme types of situations. For the usual run of practical cases, the populations tend to be unimodal with moderate skewness. For these situations, the significance levels should not differ appreciably from the values for symmetry (for the bounded significance level tests of this paper). It seems unlikely that the values of the limits will be approached for many situations of practical interest unless the significance level for symmetry is already near a bound.

Table 1 contains a list of the tests presented. The significance level bounds were computed under the assumption that only (i)-(iii) necessarily hold. In obtaining the significance levels listed, it was assumed that (iv) is also satisfied.

No explicit efficiency investigation was carried out for the tests of Table 1. However, the results of [2] indicate that the efficiency of a test is reasonably high if its significance level (for symmetry) is not near the upper significance level bound. As an example, consider the case where $n = 7$. The one-sided tests with significance levels .039, .031, .023, .016, .008 likely have fairly high efficiencies. On the other hand, the one-sided test with significance level .054 is probably not very efficient. As a general rule, if the significance level of a test is near or below the middle of the range of possible values (i.e., the point halfway between the bounds), the efficiency of the test can be assumed to be rather high.

Examination of Table 1 shows that, for given n , several of the tests may have the same upper and lower significance level bounds, even though their significance levels are different. For example, let $n = 7$. The

one-sided tests with listed significance levels (for symmetry) .054, .047, .039, .031, .023, .016 all have upper bound .062 and lower bound .008. The tests of Table 1 with common bounds (n fixed) have two useful properties. First, the significance levels (for symmetry) of these tests rather thoroughly cover the range between the upper limit and the lower limit. Second, the actual significance level values always have the same ordering as that for the case of symmetry. As an illustration, consider a case where $n = 8$. The significance level of the test

$$\text{accept } \phi < \phi_0 \text{ if } \max\left[x_7, \frac{1}{2}(x_3 + x_8)\right] < \phi_0 .$$

is always at least as great as the significance level for the test

$$\text{accept } \phi < \phi_0 \text{ if } \max\left[x_7, \frac{1}{2}(x_4 + x_8)\right] < \phi_0 .$$

This is a direct consequence of the relation

$$\max\left[x_7, \frac{1}{2}(x_3 + x_8)\right] \leq \max\left[x_7, \frac{1}{2}(x_4 + x_8)\right] .$$

In fact, all the significance level orderings follow from relations of this type.

The two properties are helpful in deciding which test of Table 1 to apply and in intuitively justifying its use. For example, let $n = 7$ and consider one-sided tests of $\phi < \phi_0$, the desired significance level being .03. Then the test with significance level .031 would appear suitable for use. The existence of several tests with significance levels necessarily above and necessarily below that for the chosen test would seem to furnish a significance level "cushion." That is, if the significance level for the chosen test approaches one of the bounds, then the significance levels of all the "intermediate" tests must be "squeezed" into a very short interval. Examination of the statistics used for the tests would seem to indicate that such a "squeezing" will occur only for extreme situations. This leads to the hope that the significance levels for the tests of Table 1 will usually be much nearer the values for symmetry than indicated by the bounds.

If $n \leq 6$, even use of Table 1 does not yield very many tests with suitable significance levels. However, the tests of [3, Table 2] can be used to partially fill this gap. These tests have bounded significance levels on the basis of (i)-(iii) alone. Table 2 lists the tests of [3, Table 2] along with their significance level bounds. All of these tests can be assumed to have high efficiencies.

Use of the tests of Table 1 with high efficiencies is indicated even when there is reason to believe that condition (iv) is satisfied. This procedure furnishes a safety factor with respect to violation of condition (iv), the only one of the four conditions not readily justifiable on intuitive grounds alone (ordinarily).

The significance level limits for the tests of Tables 1 and 2 are obtained by finding sign tests whose significance levels bound those for the given tests. Conditions (i)-(iii) are sufficient for determining the exact significance level of a sign test for the median. Let us consider a "symmetrical" test of Table 1 or Table 2. The upper bound of the significance level for this test is given by the significance level of an equal-tail sign test. Similarly for the lower bound (if a lower bound is listed for the test). Thus the significance level limits for the "symmetrical" tests of Tables 1 and 2 tend to remain fixed even when conditions (ii) and (iii) are violated. This follows from the results of [4], where it is shown that the significance level of an equal-tail sign test tends to remain about the same under extremely general conditions.

MATHEMATICAL ANALYSIS

This section contains an outline of the methods used in obtaining the results stated in the preceding sections.

In Table 1, the significance levels for the case where all of (i)-(iv) hold were determined by use of Theorem 4 of [2].

In Tables 1 and 2, the test criteria for the alternative $\phi < \phi_0$ (i.e., tests of the form: Accept $\phi < \phi_0$ if ...) are of the types

$$(a) \quad x_n < \phi_0 ,$$

$$(b) \quad A^2 x_n + B^2 x_{n-1} < \phi_0 , \quad (A^2 + B^2 = 1)$$

$$(c) \quad C^2 x_n - D^2 x_1 < \phi_0 , \quad (C^2 - D^2 = 1)$$

$$(d) \quad \max[x_{n-1}, h(x_1, \dots, x_n)] < \phi_0 , \quad h(x_1, \dots, x_n) \leq x_n .$$

Let us examine each of these four cases.

Form (a) represents a sign test itself (i.e., accept $\phi < \phi_0$ if $x_n < \phi_0$ is a sign test). The significance level of this sign test is $(\frac{1}{2})^n$.

Now consider case (b). Since

$$x_n \geq A^2 x_n + B^2 x_{n-1} \geq x_{n-1} ,$$

it follows that

$$\Pr(x_n < \phi_0) \leq \Pr(A^2 x_n + B^2 x_{n-1} < \phi_0) \leq \Pr(x_{n-1} < \phi_0).$$

This relation gives upper and lower significance level limits in terms of sign tests. The significance level for the sign test based on $x_{n-1} < \phi_0$ is $(n+1)\left(\frac{1}{2}\right)^n$.

Case (c) was analyzed in [5] and the reasoning will not be repeated here. The relation obtained is

$$\Pr(C^2 x_n - D^2 x_1 < \phi_0) \leq \Pr(x_n < \phi_0).$$

Finally, let us consider form (d). Since

$$\max[x_{n-1}, h(x_1, \dots, x_n)] \leq x_n,$$

it follows that

$$\Pr(x_n < \phi_0) \leq \Pr\{\max[x_{n-1}, h(x_1, \dots, x_n)] < \phi_0\}.$$

Also,

$$\begin{aligned} \Pr\{\max[x_{n-1}, h(x_1, \dots, x_n)] < \phi_0\} &= \Pr[x_{n-1} < \phi_0, h(x_1, \dots, x_n) < \phi_0] \\ &\leq \Pr(x_{n-1} < \phi_0). \end{aligned}$$

This establishes upper and lower significance level limits in terms of the significance levels of sign tests.

The results for the corresponding one-sided tests of $\phi > \phi_0$ are obtained by obvious modifications of the analysis given above. The results for "symmetrical" tests of $\phi \neq \phi_0$ are a direct consequence of the non-overlapping property of the two one-sided tests combined to form a "symmetrical" test.

REFERENCES

- [1] John E. Walsh, "On the power function of the sign test for slippage of means," Annals of Math. Stat., Vol. 17 (1946), pp. 358-62.
- [2] _____, "Some significance tests for the median which are valid under very general conditions," Annals of Math. Stat., Vol. 20 (1949), pp. 64-81.
- [3] _____, "Applications of some significance tests for the median which are valid under very general conditions," Jour. Amer. Stat. Assoc., Vol. 44 (1949), pp. 342-55.

- [4] _____, "Some bounded significance level properties of the equal-tail sign test." Abstracted in Annals of Math. Stat., Vol. 19 (1948), p. 601.
- [5] _____, "On the range-midrange test and some tests with bounded significance levels," Annals of Math. Stat., Vol. 20 (1949), pp. 257-67.

TABLE 1

Some Bounded Significance Level Median Tests for $n \leq 10$

n	Significance Level (i) - (iv)		TESTS "SYMMETRICAL": Accept $\phi \neq \phi_0$ if either		Significance Level Bounds for Conditions (i)-(iii)			
	One-sided	"Symmetrical"	ONE-SIDED:	ONE-SIDED:	One-sided		"Symmetrical"	
			Accept $\phi < \phi_0$ if	Accept $\phi > \phi_0$ if	Upper	Lower	Upper	Lower
4	.062	.125	$x_4 < \phi_0$	$x_1 > \phi_0$.062	.062	.125	.125
5	.062	.125	$\frac{1}{2}(x_4 + x_5) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$.187	.031	.374	.062
	.031	.062	$x_5 < \phi_0$	$x_1 > \phi_0$.031	.031	.062	.062
6	.062	.125	$\max[x_5, \frac{1}{2}(x_3 + x_6)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_4)] > \phi_0$.109	.016	.218	.032
	.047	.094	$\max[x_5, \frac{1}{2}(x_4 + x_6)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_3)] > \phi_0$				
	.031	.062	$\frac{1}{2}(x_5 + x_6) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$				
	.016	.032	$x_6 < \phi_0$	$x_1 > \phi_0$				
7	.054	.108	$\max[x_6, \frac{1}{2}(x_1 + x_7)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_7)] > \phi_0$.062	.008	.125	.016
	.047	.094	$\max[x_6, \frac{1}{2}(x_2 + x_7)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_6)] > \phi_0$				
	.039	.078	$\max[x_6, \frac{1}{2}(x_3 + x_7)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_5)] > \phi_0$				
	.031	.062	$\max[x_6, \frac{1}{2}(x_4 + x_7)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_4)] > \phi_0$				
	.023	.047	$\max[x_6, \frac{1}{2}(x_5 + x_7)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_3)] > \phi_0$				
	.016	.032	$\frac{1}{2}(x_6 + x_7) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$				
	.008	.016	$x_7 < \phi_0$	$x_1 > \phi_0$				
8	.031	.062	$\max[x_7, \frac{1}{2}(x_1 + x_8)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_8)] > \phi_0$.035	.004	.070	.008
	.027	.054	$\max[x_7, \frac{1}{2}(x_2 + x_8)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_7)] > \phi_0$				
	.023	.047	$\max[x_7, \frac{1}{2}(x_3 + x_8)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_6)] > \phi_0$				
	.020	.040	$\max[x_7, \frac{1}{2}(x_4 + x_8)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_5)] > \phi_0$				
	.016	.032	$\max[x_7, \frac{1}{2}(x_5 + x_8)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_4)] > \phi_0$				
	.012	.023	$\max[x_7, \frac{1}{2}(x_6 + x_8)] < \phi_0$	$\min[x_2, \frac{1}{2}(x_1 + x_3)] > \phi_0$				
	.008	.016	$\frac{1}{2}(x_7 + x_8) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$				
	.004	.008	$x_8 < \phi_0$	$x_1 > \phi_0$				

Some Bounded Significance Level Median Tests for $n \leq 10$

n	Significance Level (i) - (iv)		TESTS		Significance Level Bound for Conditions (i) - (iii)			
	One-sided	"Symmetrical"	"SYMMETRICAL": Accept $\phi \neq \phi_0$ if either		One-sided		"Symmetrical"	
			ONE-SIDED: Accept $\phi < \phi_0$ if	ONE-SIDED: Accept $\phi > \phi_0$ if	Upper	Lower	Upper	Lower
9	.018	.035	$\max [x_9, \frac{1}{2}(x_1+x_9)] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_9)] > \phi_0$.020	.002	.040	.004
	.016	.032	$\max [x_8, \frac{1}{2}(x_2+x_9)] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_8)] > \phi_0$				
	.014	.027	$\max [x_8, \frac{1}{2}(x_3+x_9)] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_7)] > \phi_0$				
	.012	.023	$\max [x_8, \frac{1}{2}(x_4+x_9)] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_6)] > \phi_0$				
	.010	.020	$\max [x_8, \frac{1}{2}(x_5+x_9)] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_5)] > \phi_0$				
	.008	.016	$\max [x_8, \frac{1}{2}(x_6+x_9)] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_4)] > \phi_0$				
	.006	.012	$\max [x_8, \frac{1}{2}(x_7+x_9)] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_3)] > \phi_0$				
	.004	.008	$\frac{1}{2}(x_8 + x_9) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$				
	.002	.004	$x_9 < \phi_0$	$x_1 > \phi_0$.002	.002	.004	.004
10	.010	.020	$\max [x_9, \frac{1}{2}(x_1+x_{10})] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_{10})] > \phi_0$.011	.001	.022	.002
	.009	.018	$\max [x_9, \frac{1}{2}(x_2+x_{10})] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_9)] > \phi_0$				
	.008	.016	$\max [x_9, \frac{1}{2}(x_3+x_{10})] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_8)] > \phi_0$				
	.007	.014	$\max [x_9, \frac{1}{2}(x_4+x_{10})] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_7)] > \phi_0$				
	.006	.012	$\max [x_9, \frac{1}{2}(x_5+x_{10})] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_6)] > \phi_0$				
	.005	.010	$\max [x_9, \frac{1}{2}(x_6+x_{10})] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_5)] > \phi_0$				
	.004	.008	$\max [x_9, \frac{1}{2}(x_7+x_{10})] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_4)] > \phi_0$				
	.003	.006	$\max [x_9, \frac{1}{2}(x_8+x_{10})] < \phi_0$	$\min [x_2, \frac{1}{2}(x_1+x_3)] > \phi_0$				
	.002	.004	$\frac{1}{2}(x_9 + x_{10}) < \phi_0$	$\frac{1}{2}(x_1 + x_2) > \phi_0$				
.001	.002	$x_{10} < \phi_0$	$x_1 > \phi_0$.001	.001	.002	.002	

TABLE 2

Significance Level Bounds for the Tests of [1, Table 2]

TESTS n	"SYMMETRICAL": Accept $\phi \neq \phi_0$ if either		Significance Level Bounds for Conditions (i) - (iv)				Significance Level Bounds for Conditions (i) - (iii)			
	ONE-SIDED		One-sided Tests		"Symmetrical" Tests		One-sided Tests		"Symmetrical" Tests	
	Accept $\phi < \phi_0$ if	Accept $\phi > \phi_0$ if	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
4	$1.055x_4 - .055x_1 < \phi_0$	$1.055x_1 - .055x_4 > \phi_0$.062		.125		.062		.062	
5	$.63x_5 + .37x_4 < \phi_0$	$.63x_1 + .37x_2 > \phi_0$.062	.031	.125	.062	.187	.031	.374	.062
	$1.02x_5 - .02x_1 < \phi_0$	$1.02x_1 - .02x_5 > \phi_0$.031		.062		.031		.062	
6	$.63x_6 + .37x_5 < \phi_0$	$.63x_1 + .37x_2 > \phi_0$.031	.016	.062	.031	.109	.016	.218	.031
	$1.06x_6 - .06x_1 < \phi_0$	$1.06x_1 - .06x_6 > \phi_0$.016		.031		.016		.031	
7	$.785x_7 + .215x_6 < \phi_0$	$.785x_1 + .215x_2 > \phi_0$.016	.008	.031	.016	.070	.008	.140	.016
	$1.05x_7 - .05x_1 < \phi_0$	$1.05x_1 - .05x_7 > \phi_0$.008		.016		.008		.016	
8	$\max [x_7, (.5x_8 + .28x_6 + .22x_7)] < \phi_0$	$\min [x_2, (.5x_1 + .28x_3 + .22x_2)] > \phi_0$.012	.008	.023	.016	.036	.004	.072	.008
	$.785x_8 + .215x_7 < \phi_0$	$.785x_1 + .215x_2 > \phi_0$.008	.004	.016	.008	.036	.004	.072	.008
9	$\max [x_8, (.5x_9 + .28x_7 + .22x_8)] < \phi_0$	$\min [x_2, (.5x_1 + .28x_3 + .22x_2)] > \phi_0$.006	.004	.012	.008	.020	.002	.040	.004