SOME BOUNDED SIGNIFICANCE LEVEL PROPERTIES OF THE EQUAL-TAIL SIGN TEST

By John E. Walsh

The Rand Corporation[1]

1. Summary. In addition to being easily applied and reasonably efficient for small samples, the equal-tail sign procedure for testing hypotheses about, or setting confidence intervals for, the population median is valid under very general conditions. (For brevity, the equal-tail sign procedure will be referred to as Procedure E.) Rarely, if ever, however, are these conditions exactly satisfied in practice. Thus the actual significance level or confidence coefficient for Procedure E is only an approximation to the standard value (which holds when the conditions are satisfied). Undoubtedly the equal-tail sign procedure is used in many cases when these conditions are only roughly approximated. The purpose of this paper is to investigate under what conditions Procedure E has significance levels and confidence coefficients which are satisfactory approximations to the standard values. It is found that the approximation is reasonably good for a wide variety of situations if the number of observations is not large. Thus, as far as errors of Type I are concerned, Procedure E is a sufficiently close approximation for many practical cases. This significance level stability, combined with its other favorable properties, suggests that the equal-tail sign procedure be seriously considered for application when an inference is to be made from a small number of observations to the population median.

2. Introduction and discussion. Let us consider testing whether the population median μ equals a given hypothetical value μ_0 for situations where alternative values of the median greater than μ_0 are to receive the same emphasis as those less than this value. The equal-tail sign test represents a solution to this problem which is of great practical utility. The computation required for the application of an equal-tail sign test is small. The efficiency of these tests is reasonably high for small samples from normal populations (see [1]). Also the equal-tail sign test is valid under very general conditions. Sufficient conditions are that the observations used for the test are *statistically independent* and from populations which satisfy

(i) the populations have a common median value μ, and

(ii) no population has a discrete amount of probability concentrated at μ_0; i.e., $Pr(x = \mu_0) = 0$ for each population.

Here it should be emphasized that μ is not necessarily unique; there may be an entire interval of points which satisfy (i).

Situations where μ is not unique but represents a set of points cause little difficulty if suitably interpreted. An equal-tail sign test of the null hypothesis

[1] The author is now with the U. S. Bureau of the Census, Washington 25, D. C.

$\mu = \mu_0$ is merely a method of deciding whether μ_0 is a point having the property that $Pr(x < \mu_0) = Pr(x > \mu_0) = \frac{1}{2}$ for each population. The location of μ_0 among the 50% points of a population is usually not of importance. Thus the null hypothesis $\mu = \mu_0$ has the interpretation μ_0 in μ.

Let the n independent observations on which a test is based be denoted by $y_1, \cdots, y_n$. Subtract μ_0 from each of these observations. Then n nonzero numbers will be obtained (the probability of the number zero occurring is zero). The equal-tail sign test for the median can be expressed in terms of the signs (+ or −) of these numbers. Let p be the number of positive signs (whence $n - p$ is the number of negative signs). The equal-tail sign test for comparing μ with the given hypothetical value μ_0 is *Accept* $\mu \neq \mu_0$ *if either* $p \geqq i$ *or* $p \leqq n - i$, *where* $i > (n + 1)/2$. An equivalent way to state this test is in terms of order statistics. Let $x_1, \cdots, x_n$ represent the values of $y_1, \cdots, y_n$ arranged in increasing order of magnitude. Using order statistics, the equal-tail sign test for the median is

TEST 1. *Accept* $\mu \neq \mu_0$ *if either* $x_i < \mu_0$ *or* $x_{n+1-i} > \mu_0$, *where* $i > (n + 1)/2$.

The significance level of Test 1 is a function of i and n which has the value

$$Pr(x_i < \mu) + Pr(x_{n+1-i} > \mu) = \left(\frac{1}{2}\right)^{n-1} \sum_{s=i}^{n} \frac{n!}{s!(n-s)!} \tag{1}$$

when conditions (i) and (ii) hold.

The statement of the equal-tail sign test in terms of order statistics is convenient because equal-tail confidence intervals for μ can also be derived. Since $i > (n + 1)/2$, it follows from (1) that

$$(x_{n+1-i}, x_i)$$

is an equal-tail confidence interval for μ with confidence coefficient

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

if (i) and (ii) hold.

When conditions (i) and (ii) are not necessarily satisfied, Test 1 is no longer exact. Its significance level may differ substantially from the value of (1). The null hypothesis may not be expressible in the form $\mu = \mu_0$. In many cases, however, the equal-tail sign test furnishes a reasonably close approximation to a fairly large class of tests. This approximation is close in the sense that each test of the class has a significance level which is near the value of (1) when conditions (i) and (ii) are even roughly satisfied. The principal purpose of this paper is to define this class of tests and investigate their significance level properties.

First, let us consider the form and properties of the null hypotheses for the class of tests to be investigated. Since condition (i) is not necessarily satisfied, the null hypothesis can no longer be expressed in the form $\mu = \mu_0$. Let μ_j represent the median value (or set of median values) for the population from which the observation y_j was drawn ($j = 1, \cdots, n$). For each test of the class, the null hypothesis is required to be some function of $\mu_1, \cdots, \mu_n$ which reduces to $\mu = \mu_0$ when condition (i) is satisfied. Since these null hypotheses represent generalizations of the null hypothesis for the sign test ($\mu = \mu_0$), they will be referred to as generalized null hypotheses. Hence the generalized null hypotheses considered will be of the form

$$\mu_0 \textit{ is contained in } h(\mu_1, \cdots, \mu_n),$$

where the set function h is restricted so that it is contained in the set of 50 % points common to all populations (denoted by μ) when condition (i) holds. If h is not unique, the generalized null hypothesis has the interpretation μ_0 in h.

The function h is also restricted so that it is nearly the same as μ when condition (i) is approximately satisfied. Stated in another way, the function chosen for h should not be sensitive to condition (i); i.e., a moderate deviation from the existence of a common median value should not have an appreciable effect on h. For example, let $\mu_1, \cdots, \mu_n$ be unique and large. Then the function

$$\frac{1}{n}\sum_1^n \mu_j + \mu_1 \cdots \mu_n \sum_1^{n-1} (\mu_{j+1} - \mu_j)^2$$

would not be suitable for use as h even though it reduces to μ when all the μ_j have the value μ.

Now let us define the class of tests which are investigated in this paper. All tests of the class reduce to the equal-tail sign test when conditions (i) and (ii) hold. Consequently, each test of the class will be referred to as a generalized test. A generalized test is defined by

TEST 2. *Accept that μ_0 is outside of h if either $x_i < \mu_0$ or $x_{n+1-i} > \mu_0$, where $i > (n + 1)/2$.*

The significance level of this test equals

$$Pr(x_i < \mu_0 \mid \mu_0 \text{ in } h) + Pr(x_{n+1-i} > \mu_0 \mid \mu_0 \text{ in } h). \tag{2}$$

The value of (2) is not completely determined by i and n. It also depends on many other factors such as the populations from which the observations were drawn and the value of μ_0. In spite of this inexactness, the value of (2) is usually rather closely fixed if h is a reasonable type of function and conditions (i) and (ii) are even roughly satisfied. The statement of Test 2 defines a class of tests rather than a single test because of the possible choices for the function h.

It should be pointed out that Test 2 does not necessarily have equal tails. That is, the value of $Pr(x_i < \mu_0 \mid \mu_0 \text{ in } h)$ is not necessarily equal to the value of $Pr(x_{n+1-i} > \mu_0 \mid \mu_0 \text{ in } h)$. In extreme cases, Test 2 might even be one-sided.

The main problem of the paper is to show that in practice the value of (1) is usually a close approximation to the value of (2). This, of course, is not always true. For example, consider the case where some or all of the populations from which the observations were drawn have a large proportion of their probability concentrated at or near the median. Then the value of (2) may differ greatly from that of (1) even though conditions (i) and (ii) are very nearly satisfied. For populations of the type ordinarily encountered in practice and a reasonable choice of h, however, the value of (1) is usually near that of (2) even when conditions (i) and (ii) are only roughly satisfied. This is proved by obtaining upper and lower bounds for (2) as functions of n, i and a quantity β. Here β is defined to be the greater of

$$\max_j \left| Pr(y_j < \mu_0 \mid \mu_0 \text{ in } h) - \frac{1}{2} \right|, \qquad \max_j \left| Pr(y_j > \mu_0 \mid \mu_0 \text{ in } h) - \frac{1}{2} \right|.$$

If $\beta = 0$, the significance level of Test 2 equals that of Test 1. If β is small, the value of (2) is very near that of (1). Table 1 contains upper and lower bounds for the significance level of Test 2 for $\beta = .02, .05, .08, .10, .15, .20$, and $n \leqq 15$. If the populations are continuous (or very nearly so) at μ_0, the value of the lower bound is noticeably increased (see Table 2). Thus, for $n \leqq 15$, the value of (1) does not differ greatly from that of (2) even for β moderately large. A value of β as large as .05 would seem unusual for the ordinary type of practical situation where there is reason to believe that conditions (i) and (ii) are approximately satisfied.

Let us consider the practical implications of the fact that the equal-tail sign test approximates Test 2 in the sense of significance level. Suppose the experimenter recognizes the possibility that conditions (i) and (ii) may not hold for his experiment. He then selects the function $h(\mu_1, \cdots, \mu_n)$ which is of principal interest to him and uses Test 2. In this manner he obtains an accurate test of the null hypothesis in which he is interested. On the other hand, suppose that the experimenter applies the equal-tail sign test without considering the possibility that conditions (i) and (ii) may be violated. The results of this paper show that he is protected if the appropriate function h (which he would have chosen) and the populations from which the observations were drawn are of a reasonable nature. Then he is testing the appropriate null hypothesis at approximately the specified significance level even though he may not think of the test in this light.

Since for the case of a sample from a normal population the efficiency of the equal-tail sign test decreases as n increases, much of the investigation is limited to tests based on 15 or fewer observations. Table 1 contains a list of the tests investigated along with their efficiency for normality. The efficiency of a significance test (more precisely, the power efficiency) is defined in [1]. Intuitively the efficiency of a test measures the percentage of available information per observation which is utilized by that test.

The equal-tail sign test for the median may be useful for situations where there is not much information available concerning properties of the populations from which the observations were taken. Due to the extremely general conditions under which its significance level is approximately determined, this test can be used in cases where more specialized tests are not necessarily applicable.

Approximate confidence intervals for $h(\mu_1, \cdots, \mu_n)$ can be obtained from Test

TABLE 1

Some properties of Test 1 and Test 2 for $n \leq 15$

n	Test 1: Accept $\mu \neq \mu_0$ if / Test 2: Accept μ_0 outside h if	Test 1 normality efficiency	Test 1 significance level	Significance level bounds for Test 2: $\beta = .02$ Upper	$\beta = .02$ Lower	$\beta = .05$ Upper	$\beta = .05$ Lower	$\beta = .08$ Upper	$\beta = .08$ Lower	$\beta = .10$ Upper	$\beta = .10$ Lower	$\beta = .15$ Upper	$\beta = .15$ Lower	$\beta = .20$ Upper	$\beta = .20$ Lower
4	$x_4 < \mu_0$ or $x_1 > \mu_0$	95%	.1250	.1262	.1062	.1325	.0820	.1443	.0622	.1552	.0512	.1924	.0300	.2482	.0162
5	$x_5 < \mu_0$ or $x_1 > \mu_0$	96%	.0625	.0635	.0510	.0688	.0369	.0787	.0261	.0880	.0205	.1213	.0105	.1705	.0049
6	$x_6 < \mu_0$ or $x_1 > \mu_0$	95%	.0312	.0320	.0245	.0360	.0166	.0436	.0110	.0508	.0082	.0773	.0037	.1184	.0015
7	$x_7 < \mu_0$ or $x_1 > \mu_0$	95%	.0156	.0162	.0117	.0190	.0075	.0244	.0046	.0296	.0033	.0497	.0013	.0826	.0004
	$x_6 < \mu_0$ or $x_2 > \mu_0$	80%	.1250	.1271	.1008	.1381	.0714	.1586	.0492	.1775	.0377	.2428	.0180	.3332	.0076
8	$x_8 < \mu_0$ or $x_1 > \mu_0$	95%	.0078	.0082	.0056	.0101	.0034	.0138	.0019	.0175	.0013	.0321	.0005	.0577	.0001
	$x_7 < \mu_0$ or $x_2 > \mu_0$	80%	.0704	.0721	.0545	.0813	.0362	.0987	.0233	.1149	.0170	.1727	.0071	.2566	.0026
9	$x_9 < \mu_0$ or $x_1 > \mu_0$	95%	.0039	.0041	.0027	.0054	.0015	.0078	.0008	.0103	.0005	.0208	.0002	.0404	.0000
	$x_8 < \mu_0$ or $x_2 > \mu_0$	82%	.0390	.0404	.0291	.0476	.0182	.0613	.0109	.0743	.0076	.1225	.0028	.1964	.0009
10	$x_9 < \mu_0$ or $x_2 > \mu_0$	80%	.0214	.0225	.0154	.0278	.0090	.0380	.0051	.0480	.0034	.0865	.0011	.1495	.0003
	$x_8 < \mu_0$ or $x_3 > \mu_0$	75%	.1094	.1122	.0839	.1270	.0548	.1543	.0344	.1796	.0246	.2664	.0096	.3844	.0032
11	$x_{10} < \mu_0$ or $x_2 > \mu_0$	81%	.0117	.0124	.0081	.0161	.0044	.0236	.0023	.0310	.0015	.0608	.0004	.1130	.0001
	$x_9 < \mu_0$ or $x_3 > \mu_0$	76%	.0654	.0678	.0483	.0800	.0296	.1032	.0174	.1248	.0118	.2022	.0041	.3133	.0012
12	$x_{10} < \mu_0$ or $x_3 > \mu_0$	75%	.0386	.0404	.0274	.0500	.0158	.0685	.0086	.0863	.0056	.1521	.0017	.2530	.0004
13	$x_{11} < \mu_0$ or $x_3 > \mu_0$	75%	.0224	.0238	.0153	.0310	.0083	.0453	.0042	.0592	.0026	.1135	.0007	.2026	.0001
	$x_{10} < \mu_0$ or $x_4 > \mu_0$	70%	.0924	.0956	.0675	.1133	.0407	.1461	.0233	.1764	.0156	.2808	.0050	.4213	.0013
14	$x_{12} < \mu_0$ or $x_3 > \mu_0$	78%	.0130	.0139	.0085	.0192	.0043	.0297	.0021	.0404	.0012	.0841	.0003	.1609	.0001
	$x_{11} < \mu_0$ or $x_4 > \mu_0$	73%	.0574	.0601	.0404	.0746	.0229	.1022	.0123	.1282	.0078	.2216	.0022	.3554	.0005
15	$x_{13} < \mu_0$ or $x_3 > \mu_0$	78%	.0074	.0081	.0046	.0118	.0022	.0194	.0010	.0274	.0006	.0618	.0001	.1268	.0000
	$x_{12} < \mu_0$ or $x_4 > \mu_0$	74%	.0350	.0373	.0238	.0487	.0127	.0709	.0063	.0924	.0039	.1732	.0010	.2970	.0002
	$x_{11} < \mu_0$ or $x_5 > \mu_0$	70%	.1186	.1229	.0860	.1459	.0509	.1882	.0285	.2266	.0187	.3548	.0057	.5162	.0013

2. For populations of the type usually encountered in practice and a reasonable function h,

$$(x_{n+1-i}, x_i)$$

is a confidence interval for h with confidence coefficient approximately equal to unity minus the value of (1).

The material presented in this paper is limited to investigation of Type I errors of the equal-tail sign test when the conditions on which it is based are generalized. Due to the extremely general situations considered, an investigation of Type II errors was not feasible. However, the results obtained for the particular case of a sample from a normal population indicate that the efficiency of the equal-tail sign test is reasonably high for most situations if the number of observations is small.

3. Outline of results. This section contains a statement of the main results of the paper. The proofs of these statements are given in Section 4.

The method followed in obtaining bounds for (2) consists in fixing n, i, β and then finding the largest and smallest values of (2) possible on the basis of these and any additional restrictions. Thus the bounds represent the worst possible situations for the given restrictions. For most situations, the value of (2) would likely be nowhere near the values of the bounds. Consequently, for most cases the value of (1) will be much nearer (2) than is indicated by the upper and lower limits in the tables.

Let us consider the general case where both conditions (i) and (ii) could be violated. Values of upper and lower bounds for the significance level of Test 2 as functions of n, i, and β are given by

$$\text{upper bound} = \sum_{s=i}^{n} \frac{n!}{s!(n-s)!} \cdot \left[\left(\frac{1}{2}+\beta\right)^{s}\left(\frac{1}{2}-\beta\right)^{n-s} + \left(\frac{1}{2}-\beta\right)^{s}\left(\frac{1}{2}+\beta\right)^{n-s}\right], \tag{3}$$

$$\text{lower bound} = 2\sum_{s=i}^{n} \frac{n!}{s!(n-s)!}\left(\frac{1}{2}-\beta\right)^{s}\left(\frac{1}{2}+\beta\right)^{n-s}.$$

Thus, if $\beta = 0$ the value of (2) equals (1) while if β is small the value of (2) is very nearly equal to (1). Table 1 contains values of these upper and lower bounds for the tests considered. A visual example of how the bounds given by (3) vary as functions of β for fixed n and i is given by Figure 1, which contains a plot of these bounds for the case $n = 9$, $i = 8$. If $\beta \to \frac{1}{2}$, the upper bound $\to 1$ and the lower bound $\to 0$.

A case of practical interest is that where condition (i) is not violated to any appreciable extent; i.e., none of the populations has a noticeable amount of probability concentrated at μ_0. Then the upper bound given in (3) still holds but the lower bound is greatly improved. Table 2 contains a list of some numerical values for this lower bound. These values are only slightly less than the value of

(1) except for large values of β. The dotted curve in Figure 1 represents a plot of this lower bound as a function of β for the case $n = 9$, $i = 8$.

In all the above results, the n observations on which tests are based were assumed to be independent. Although no analysis will be made for cases in

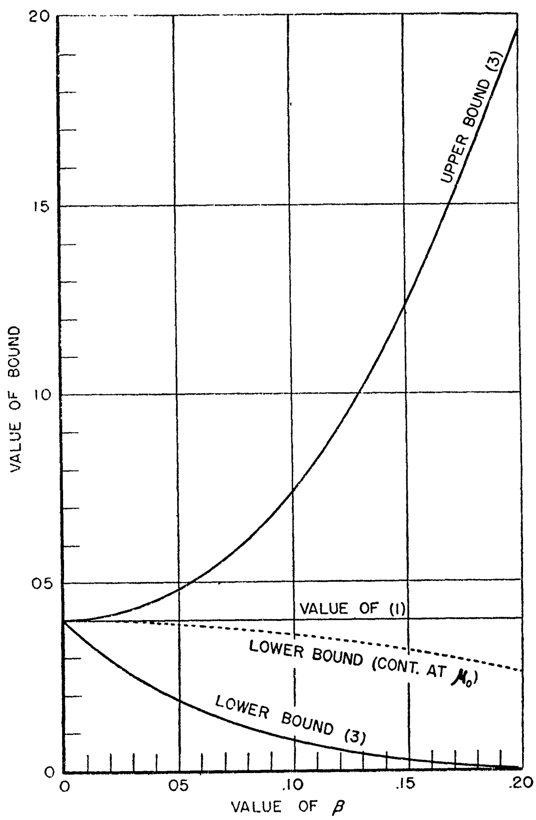
TABLE 2

Lower bounds for the significance level of Test 2 when populations are continuous at μ_0

n	Test: Accept μ_0 outside h if either	Value of (1) ($\beta = 0$)	Lower bound for significance level of Test 2: $\beta = .02$	$\beta = .05$	$\beta = .08$	$\beta = .10$	$\beta = .15$	$\beta = .20$	$\beta = .30$	$\beta = .40$
4	$x_4 < \mu_0$ or $x_1 > \mu_0$	.1250	.1246	.1225	.1187	.1152	.1035	.0882	.0512	.0162
5	$x_5 < \mu_0$ or $x_1 > \mu_0$	.0625	.0623	.0613	.0593	.0576	.0519	.0441	.0256	.0081
6	$x_6 < \mu_0$ or $x_1 > \mu_0$	.0312	.0311	.0303	.0289	.0276	.0235	.0185	.0082	.0015
7	$x_7 < \mu_0$ or $x_1 > \mu_0$	.0156	.0156	.0152	.0145	.0138	.0118	.0093	.0041	.0007
	$x_6 < \mu_0$ or $x_2 > \mu_0$	.1250	.1247	.1231	.1202	.1175	.1082	.0953	.0604	.0214
8	$x_8 < \mu_0$ or $x_1 > \mu_0$	.0078	.0078	.0075	.0070	.0066	.0043	.0039	.0013	.0001
	$x_7 < \mu_0$ or $x_2 > \mu_0$	.0704	.0701	.0688	.0663	.0641	.0567	.0469	.0236	.0049
9	$x_9 < \mu_0$ or $x_1 > \mu_0$	.0039	.0039	.0038	.0035	.0033	.0027	.0019	.0007	.0001
	$x_8 < \mu_0$ or $x_2 > \mu_0$	.0390	.0389	.0381	.0367	.0354	.0310	.0254	.0125	.0025
10	$x_9 < \mu_0$ or $x_2 > \mu_0$	.0214	.0214	.0208	.0198	.0188	.0158	.0121	.0047	.0005
	$x_8 < \mu_0$ or $x_3 > \mu_0$	.1094	.1091	.1075	.1044	.1016	.0919	.0785	.0436	.0104
11	$x_{10} < \mu_0$ or $x_2 > \mu_0$	.0117	.0117	.0113	.0107	.0102	.0085	.0065	.0024	.0003
	$x_9 < \mu_0$ or $x_3 > \mu_0$	.0654	.0652	.0641	.0621	.0602	.0538	.0453	.0241	.0055
12	$x_{10} < \mu_0$ or $x_3 > \mu_0$	.0386	.0384	.0376	.0360	.0346	.0298	.0237	.0102	.0014
13	$x_{11} < \mu_0$ or $x_3 > \mu_0$	.0224	.0224	.0218	.0208	.0200	.0170	.0133	.0055	.0007
	$x_{10} < \mu_0$ or $x_4 > \mu_0$	.0924	.0920	.0907	.0882	.0859	.0779	.0669	.0381	.0096
14	$x_{12} < \mu_0$ or $x_3 > \mu_0$	.0130	.0129	.0125	.0118	.0112	.0092	.0068	.0022	.0002
	$x_{11} < \mu_0$ or $x_4 > \mu_0$	.0574	.0572	.0561	.0540	.0522	.0459	.0375	.0176	.0027
15	$x_{13} < \mu_0$ or $x_3 > \mu_0$	.0074	.0073	.0071	.0067	.0063	.0051	.0037	.0012	.0001
	$x_{12} < \mu_0$ or $x_4 > \mu_0$	.0350	.0350	.0343	.0329	.0317	.0275	.0221	.0099	.0015

which the observations are not independent, examination of the significance level expression (2) for Test 2 indicates that the value of (2) will often be approximately equal to (1) when the observations are mildly dependent. This follows from the intuitive observation that in many cases dependence changes $Pr(x_i < \mu_0 \mid \mu_0 \text{ in } h)$ and $Pr(x_{n+1-i} > \mu_0 \mid \mu_0 \text{ in } h)$ in such a way that one probability expression is increased while the other is decreased; consequently the value of (2) tends to remain near that of (1).

4. Derivations. The purpose of this section is to present derivations of the results stated in the preceding sections.



BOUNDS OF (2) FOR n = 9, i = 8

FIG. 1.

The expressions for (1) and (2) follow from $i > (n + 1)/2$, conditions (i) and (ii), and elementary probability considerations. Consider relations (3). Let

$$Pr(y_j < \mu_0 \mid \mu_0 \text{ in } h) = \tfrac{1}{2} + \alpha_j, \qquad Pr(y_j = \mu_0 \mid \mu_0 \text{ in } h) = \epsilon_j,$$
$$Pr(y_j > \mu_0 \mid \mu_0 \text{ in } h) = \tfrac{1}{2} + \gamma_j \qquad (j = 1, \cdots, n).$$

Then

$$\gamma_j = -(\alpha_j + \epsilon_j)$$

and

$$Pr(x_i < \mu_0 \mid \mu_0 \text{ in } h) = \prod_{j=1}^{n}\left(\frac{1}{2} + \alpha_j\right) + \sum_{r=1}^{n-i} \sum_{j_1 > \cdots > j_r = 1}^{n} \left[\Pi'\left(\frac{1}{2} + \alpha_j\right)\right] \cdot \left[\prod_{k=1}^{r}\left(\frac{1}{2} - \alpha_{j_k}\right)\right],$$

$$(4) \quad Pr(x_{n+1-i} > \mu_0 \mid \mu_0 \text{ in } h) = \prod_{j=1}^{n}\left(\frac{1}{2} - \alpha_j - \epsilon_j\right) + \sum_{r=1}^{n-i} \sum_{j_1 > \cdots > j_r = 1}^{n} \left[\Pi'\left(\frac{1}{2} - \alpha_j - \epsilon_j\right)\right]\left[\prod_{k=1}^{r}\left(\frac{1}{2} + \alpha_{j_k} + \epsilon_{j_k}\right)\right],$$

where the notation Π' denotes the product over those values of j ($j = 1, \cdots, n$) which are different from $j_1, \cdots, j_r$. If $i = n$, each double summation in (4) is taken to be zero.

Examination of (4) shows that (2) can be written in the form

$$f(\alpha_1, \cdots, \alpha_n; \epsilon_1, \cdots, \epsilon_{j-1}, \epsilon_{j+1}, \cdots, \epsilon_n) - \epsilon_j\, g(\alpha_1, \cdots, \alpha_n; \epsilon_1, \cdots, \epsilon_{j-1}, \epsilon_{j+1}, \cdots, \epsilon_n),$$

where

$$g(\alpha_1, \cdots, \alpha_n; \epsilon_1, \cdots, \epsilon_{j-1}, \epsilon_{j+1}, \cdots, \epsilon_n) \geqq 0$$

for each value of j. Thus, since setting $\epsilon_j = 0$ places no additional restrictions on the possible values of the α's and the other ϵ's, to obtain the maximum value for (2) all the ϵ's should be zero. Now consider (2) with all the ϵ's equal to zero. It can be written in the form

$$(5) \quad u(\alpha_1, \cdots, \alpha_{j-1}, \alpha_{j+1}, \cdots, \alpha_n) + \alpha_j v(\alpha_1, \cdots, \alpha_{j-1}, \alpha_{j+1}, \cdots, \alpha_n)$$

for each value of j. Since $-\beta \leqq \alpha_j \leqq \beta$ for all j, the maximum value of (2) is obtained when the α_j are restricted to be of the form

$$(6) \quad \alpha_j = \eta_j \beta \qquad (j = 1, \cdots, n),$$

where each η_j equals either $+1$ or -1. Assume that an arbitrary but fixed choice has been made for the η_j. Then (4) shows that (2) is a polynomial in α which is an even function. Consider the coefficient of an arbitrary even power of α in this polynomial. Examination shows that this coefficient is maximum (algebraically) for the case where all the η_j are chosen to have the same value. Hence (2) is maximum when

$$(7) \quad \epsilon_j = 0, \qquad \alpha_j = \beta \qquad (j = 1, \cdots, n).$$

Thus the upper bound for the significance level of Test 2 is that given in (3).

Now consider the lower bound for (2). Examination of (4) shows that $Pr(x_i < \mu_0 \mid \mu_0 \text{ in } h)$ is minimum when $\alpha_j = -\beta$ ($j = 1, \cdots, n$). Similarly,

$Pr(x_{n+1-i} > \mu_0 \mid \mu_0 \text{ in } h)$ is minimum when $\alpha_j + \epsilon_j = \beta$ $(j = 1, \cdots, n)$. Thus (2) is minimum when

$$\epsilon_j = 2\beta, \qquad \alpha_j = -\beta \qquad\qquad (j = 1, \cdots, n).$$

Substitution of these values into (4) verifies the expression given in (3) for the lower bound.

If the populations for Test 2 all satisfy condition (ii), $\epsilon_j = 0$ $(j = 1, \cdots, n)$. From (7), the upper bound of (2) given in (3) is unchanged for this case. The lower bound, however, can be noticeably larger than the value stated in (3). Since for each value of j $(j = 1, \cdots, n)$, the value of (2) can be expressed in the form (5), the lower bound of the significance level of Test 2 is equal to the minimum value which can be obtained for (2) when the α_j are restricted to be of the form (6) and the ϵ_j have the value zero. As (2) is invariant with respect to permutations of $\alpha_1, \cdots, \alpha_n$, the problem of obtaining the lower bound of the significance level of Test 2 is reduced to that of determining the number m of the η_j which equal $+1$ when the resulting value of (2) is minimum. Since the lower bound for the significance level of Test 2 is only required for $n \leqq 15$ and $i \geqq n - 3$ (see Table 2), an analytical method of determining the value of m which minimizes (2) will not be developed; the values for the lower bounds listed in Table 2 were obtained by substituting numerical values for m and computing the resulting values of (2). For example, if $i = n$ and m of the $\alpha_j = +\beta$ while the remaining α_j equal $-\beta$, the value of (2) is

$$(\tfrac{1}{2} + \beta)^m(\tfrac{1}{2} - \beta)^{n-m} + (\tfrac{1}{2} - \beta)^m(\tfrac{1}{2} + \beta)^{n-m}.$$

If $i < n$, the expressions become much more complicated and will not be given here.

REFERENCE

[1] JOHN E. WALSH, "Some significance tests for the median which are valid under very general conditions," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 64–81.