$z = \frac{n_1}{2} \frac{v_1 + u_1}{v_1 - u_1} - \frac{n_2(v_2 + u_2)}{2(v_2 - u_2)}$ can be used to test the hypothesis of equality of means of any two rectangular populations, and has in the limit the distribution (10), if the means of the populations are equal.

6. The one-parameter rectangular distribution. If $f(x) = 1/\lambda$, $(0 \leq x \leq \lambda)$, then $f(x_1, \cdots, x_n \mid v) = v^{1-n}$. Thus v is a sufficient statistic and is evidently the maximum likelihood estimate of λ. Here $F(v) = (v/\lambda)^n$; $f(v) = nv^{n-1}\lambda^{-n}$; and $\mu_k(v) = \lambda^k n/(n + k)$. The normalized error $y = n(\lambda - v)/\lambda$ has the probability density function $f(y) = (1 - y/n)^{n-1}$, which tends to e^{-y} as n increases.

REFERENCE

[1] R. A. Fisher, "On the mathematical foundations of theoretical statistics," *Phil. Trans. Roy. Soc. London,* Series A, Vol. 222 (1921), pp. 309–368.

ON THE POWER FUNCTION OF THE SIGN TEST FOR SLIPPAGE OF MEANS

By John E. Walsh

Princeton University

1. Summary. This note compares the power functions of the sign test for slippage with the power functions of the most powerful test for the case of normal populations. The sign test is found to be approximately 95% efficient for small samples.

2. Introduction. Let us consider a univariate population whose mean equals its median and whose cumulative distribution function is continuous at the mean. A sampling method of testing the supposition that the mean of this population exceeds a given constant value μ_0 (slippage to the right) is furnished by considering how many values of the sample are less than μ_0. An analogous method applies for testing whether the mean is less than μ_0 (slippage to the left). A particular class of populations for which the sign test is valid are the normal populations. This note compares the power functions of the sign test with the power functions of the most powerful test for slippage for the case in which the population is normal (Table I). It is shown that the sign test is approximately 95% as efficient as the most powerful test (the Student t-test) for samples of size 4, 5 and 6, and that although the relative efficiency of the sign test decreases as the sample size increases, its efficiency is approximately 75% for samples of size 13. This supports the idea that for normal populations little efficiency is lost by using attributes instead of continuous variables if the sample size is small.

In choosing between the sign and Student t-tests for slippage the following considerations may be of interest:

(a) The sign test is valid for a more general class of populations than the t-test.

(b) The sign test is almost as efficient as the t-test for small samples from normal populations.

(c) The sign test is much more easily computed than the t-test.

(d) The sign test has a very limited choice of significance levels for small samples while the t-test can have any desired significance level for any size sample.

The considerations (a) to (d) also apply in choosing between the sign test and the Daly test based on $(\bar{x} - \mu_0)/R$, where $\bar{x}$ is the mean and R the range of the sample used for the test (see [1]).

In section 5, Table II shows that for small size samples the significance levels of the sign test do not change greatly if the mean is only approximately equal to the median.

3. Statement of sign test. Let $x_1, \cdots, x_n$ be a sample of size n from a univariate population whose mean equals its median and whose cumulative distribution function is continuous at the mean, that is, which has the property that

$$Pr(x < \mu) = Pr(x > \mu) = \tfrac{1}{2}, \tag{1}$$

where μ is the population mean.

The significance test to decide whether μ exceeds a given constant value μ_0 is defined by

(2) *If m or less of the sample values $x_1, \cdots, x_n$ are less than μ_0, accept $\mu > \mu_0$.*

The significance test to decide whether $\mu < \mu_0$ is given by

(3) *If m or less of $x_1, \cdots, x_r$ are greater than μ_0, accept $\mu < \mu_0$.*

It is to be observed that in both (2) and (3) the null hypothesis tested is that $\mu = \mu_0$. In (2) the alternative is $\mu > \mu_0$ and in (3) the alternative is $\mu < \mu_0$.

From (1) it follows immediately that (2) and (3) both have the same significance level $\alpha(m, n)$, where

$$\alpha(m, n) = (\tfrac{1}{2})^n \sum_{j=0}^{m} \frac{n!}{j!(n-j)!}.$$

Appropriate choices of m and n will result in values of $\alpha\,(m, n)$ suitable for significance tests. For example

$\alpha(0, 4) = .0624,$	$\alpha(1, 8) = .0352$
$\alpha(0, 5) = .0312,$	$\alpha(1, 9) = .0195$
$\alpha(0, 6) = .0156,$	$\alpha(1, 10) = .0107$
$\alpha(1, 7) = .0625,$	$\alpha(2, 13) = .0112.$

If the population has a continuous distribution function, $Pr(x_i = x_j\,; i \neq j) = 0$. In this case let $x_{(i)}$ be the ith largest of $x_1, \cdots, x_n$. Then (2) can be restated as

$$\text{If } x_{(m+1)} > \mu_0, \text{ accept } \mu > \mu_0. \tag{4}$$

Test (3) is seen to be equivalent to

(5) $$\textit{If } x_{(n-m)} < \mu_0 \text{, } \textit{accept } \mu < \mu_0 .$$

Thus for the case of populations with continuous distribution functions it is only necessary to determine one order statistic and compare it with μ_0 in order to apply a test.

It is to be observed that a particular class of populations which satisfy (1) are those which have distribution functions which are symmetrical and continuous. Thus the normal populations represent a particular class for which (4) and (5) are valid.

4. Comparison with Student *t*-test. Consider the case in which the population is normal with mean μ and variance σ^2. Then the power function for (4) is given by

$$\begin{aligned}\text{Power Function} &= Pr(x_{(m+1)} > \mu_0) \\ &= Pr\left(\frac{x_{(m+1)} - \mu}{\sigma} > \frac{\mu_0 - \mu}{\sigma}\right) \\ &= \frac{n!}{m!(n-m-1)!}\int_{\delta}^{\infty}\left(\int_{-\infty}^{x} f(y)\,dy\right)^{m}\left(\int_{x}^{\infty} f(y)\,dy\right)^{n-m-1} f(x)\,dx,\end{aligned}$$

where

$$f(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2} \qquad \text{and} \qquad \delta = \frac{\mu_0 - \mu}{\sigma}.$$

For a normal population, however, it is well known that the most powerful Studentized test of the one-sided alternative $\mu > \mu_0$ is the appropriate Student *t*-test. Values of the power function for the *t*-test are found for given values of δ by using the normal approximation given in [2].

The method of measuring the relative efficiencies of the two types of tests will be different from the common method of measuring the relative efficiencies of estimates, which consists in taking the ratio of the variances of the two estimates as the measure of their relative efficiency. The principle followed here will be to consider a sign test based on a given sample size and vary the degrees of freedom of the *t*-test having the same significance level until the power functions of the sign test and *t*-test agree in the sense that in the half-plane $\delta \leqq 0$ the area between the two power curves for which the sign test power function exceeds the *t*-test power function is equal to the analogous area for which the sign test power function is less than the *t*-test power function. The considerations are limited to the half-plane $\delta \leqq 0$ because the test is one-sided. The size of the *t*-test sample having this property divided by the size of the sign test sample is called the relative efficiency of that sign test. Intuitively this relative efficiency measures how much more data must be added if the sign test is to furnish an amount of information equivalent to that supplied by the t-test. In obtaining the relative efficiencies in the manner described above, the degrees of freedom of the t-test are allowed to assume fractional values and the values of the power function are computed using the normal approximation as if it were valid for fractional degrees of freedom. The number of degrees of freedom, of course, can only be integral. This method, however, gives an interpolated

TABLE I

A comparison of the power functions of the sign and t tests

Test	m	n	Approximate Relative Efficiency	Significance Level	Values of Power Function			
					$\delta=-\frac{1}{2}$	$\delta=-1$	$\delta=-1\frac{1}{2}$	$\delta=-2$
t		3.8		.0624	.219	.484	.755	.920
sign	0	4	95%	.0624	.229	.500	.755	.908
t		4.8		.0312	.150	.402	.709	.909
sign	0	5	96%	.0312	.159	.420	.703	.888
t		5.7		.0156	.098	.330	.660	.899
sign	0	6	95%	.0156	.110	.355	.655	.863
t		5.6		.0625	.306	.695	.932	.995
sign	1	7	80%	.0625	.311	.711	.920	.988
t		6.4		.0352	.225	.619	.908	.989
sign	1	8	80%	.0352	.239	.630	.869	.978
t		7.4		.0195	.171	.565	.893	.988
sign	1	9	82%	.0195	.182	.573	.879	.974
t		8		.0107	.117	.468	.848	.983
sign	1	10	80%	.0107	.137	.515	.853	.964
t		9.75		.0112	.162	.631	.950	.998
sign	2	13	75%	.0112	.165	.661	.949	.998

measure of the size sample of the t-test having the properties outlined above. Table I supplies a comparison of the relative efficiencies and the powers of the sign test and the t-test obtained in the manner just described. Thus for samples of size 4, 5 and 6 the sign test is approximately 95% as efficient as the Student t-test. The relative efficiency decreases as the size of the sample increases but even for samples as large as 13 is approximately 75%.

For normal populations it is also well known that the most powerful Studentized test of the alternative $\mu < \mu_0$ is given by the appropriate Student t-test. It is clear that Table I can also be considered as a comparison of test (5) with the corresponding Student t-test if δ is replaced by $-\delta$ and m by $n - m$.

5. Approximate cases. Suppose that (1) is only approximately satisfied by the population in question.

Let $Pr(x < \mu) = \frac{1}{2} + r$. Then the significance level of (2) is

$$\sum_{j=0}^{m} \frac{n!}{j!(n-j)!} \left(\tfrac{1}{2} + r\right)^{j}\left(\tfrac{1}{2} - r\right)^{n-j}. \tag{6}$$

Significance levels of (2) for small size samples are given in Table II as a function of r.

TABLE II

A comparison of the significance levels of the sign test when the mean differs from the median

m	n	Significance Level				
		$r=0$	$r=-.02$	$r=-.05$	$r=.02$	$r=.05$
0	4	.0624	.073	.091	.053	.041
0	5	.0312	.038	.050	.026	.019
0	6	.0156	.020	.028	.012	.008

Table II shows that for small samples the significance level of (2) does not change greatly from $\alpha(m, n)$ if (1) is only approximately satisfied. Expression (6) shows, however, that for large size samples even a small value of r can cause a large change in the significance level of (2).

For $Pr(x < \mu) = \frac{1}{2} + r$ it is apparent that the significance level of (3) is (6) with r replaced by $-r$ so that Table II applies to tests (3) if this replacement is made.

REFERENCES

[1] J. F. Daly, "On the use of the sample range in an analogue of Student's t-test", *Annals of Math. Stat.*, Vol. 17 (1946), pp. 71–74.

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