ON THE POWER EFFICIENCY OF A t-TEST FORMED BY PAIRING SAMPLE VALUES

BY JOHN E. WALSH

Princeton University

1. Introduction. Consider two equal sized samples, one from a normal population with mean μ and the other from a normal population with mean ν. Let $x_1, \cdots, x_n$ be the sample values from the population with mean μ and $y_1, \cdots, y_n$ the values from the population with mean ν. If the two populations have the same variance and the two samples are independent, the most powerful tests for comparing μ and ν using these samples (one-sided and symmetrical two-sided) are based on the statistic

$$t_2 = \frac{[\bar{x} - \bar{y} - (\mu - \nu)]\sqrt{n(n-1)}}{\sqrt{\sum_1^n (x_i - \bar{x})^2 + \sum_1^n (y_i - \bar{y})^2}},$$

which has a Student t-distribution with $2n - 2$ degrees of freedom. Tests based on t_2 also have the desirable property of being invariant under permutation of the data in each sample.

Sometimes, however, it is useful to combine the sample values in the form

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

Examples:

(a). When the samples are independent but it is not known that the two populations have the same variance (Behrens-Fisher problem).

(b). When there may be correlation between x_i and y_i, $(i = 1, \cdots, n)$, this correlation being the same for each value of i (i.e. x_i is independent of y_j if $i \neq j$ while each pair x_i, y_i, $(i = 1, \cdots, n)$, has the same normal bivariate distribution).

In both (a) and (b) the z_i are independently normally distributed with the same variance and mean $\mu - \nu$.

The Student t-test for comparing μ and ν using the z_i is based on the statistic

$$t_1 = \frac{[\bar{z} - (\mu - \nu)]\sqrt{n(n-1)}}{\sqrt{\sum_1^n (z_i - \bar{z})^2}} = \frac{[\bar{x} - \bar{y} - (\mu - \nu)]\sqrt{n(n-1)}}{\sqrt{\sum_1^n [x_i - y_i - (\bar{x} - \bar{y})]^2}},$$

which has a Student t-distribution with $n - 1$ degrees of freedom. These tests are not invariant under permutation of the data in each sample.

If it is true that all the sample values are independently distributed with the same variance σ^2, efficiency will be lost by using the test based on t_1 instead of the most powerful test based on t_2. The purpose of this note is to determine the power efficiency of the tests based on t_1 as compared with the corresponding tests based on t_2 for this case.

TABLE I

Power Function Values for the t_1 and t_2 Tests

Test	n	Approx. Efficiency	α	Approx. Values of Power Function			
				$\delta = \frac{1}{2}$	$\delta = 1$	$\delta = 1\frac{1}{2}$	$\delta = 2$
t_1	6	87%	.05	.276	.674	.933	.994
t_2	5.2		.05	.275	.672	.932	.994
t_1	6	82.5%	.025	.159	.486	.822	.970
t_2	4.95		.025	.160	.488	.823	.970
t_1	8	90%	.05	.355	.812	.985	
t_2	7.2		.05	.354	.813	.985	
t_1	8	86.5%	.025	.226	.674	.952	.998
t_2	6.9		.025	.225	.675	.951	.998
t_1	8	82%	.01	.112	.458	.843	.983
t_2	6.55		.01	.112	.457	.842	.983
t_1	10	92%	.05	.425	.898	.997	
t_2	9.2		.05	.425	.897	.997	
t_1	10	90%	.025	.289	.802	.988	
t_2	9		.025	.290	.803	.988	
t_1	10	85.5%	.01	.159	.626	.950	.999
t_2	8.55		.01	.159	.627	.950	.999
t_1	15	95.5%	.05	.579	.980		
t_2	14.3		.05	.579	.980		
t_1	15	93%	.025	.437	.950	1.000	
t_2	13.95		.025	.437	.949	1.000	
t_1	15	90%	.01	.278	.876	.998	
t_2	13.5		.01	.278	.876	.998	
t_1	25	98%	.05	.784	.999		
t_2	24.5		.05	.784	.999		
t_1	25	96%	.025	.670	.998		
t_2	24		.025	.670	.998		
t_1	25	94.5%	.01	.514	.992		
t_2	23.7		.01	.514	.992		

Consideration is limited to one-sided tests, which is not a serious limitation since any two-sided test can be considered as a combination of two one-sided tests. Table II contains approximate power efficiencies of one-sided tests for $n \geq 4$ at the significance levels $\alpha = .05, .025, .01$.

It is found that the efficiency of the t_1 test increases with the sample size but is high even for small size samples.

2. Outline of computations. The method of obtaining power efficiencies used here will be that outlined in [1]. Essentially this consists in computing the power function for the test based on t_1 and then adjusting the sample size for the corresponding test based on t_2 until its power function is approximately the same as for the t_1 test. The ratio of the sample size (perhaps fractional) of the adjusted t_2 test to that of the t_1 test is called the power efficiency of the t_1 test. Intuitively this efficiency measures the fraction of the total available information which is being used when the t_1 test is applied (since the t_2 test is most powerful).

TABLE II

Approximate Power Efficiencies for Given n and α

α \ n	4	5	6	7	8	9	10	15	25	∞
.05	82.5%	85%	87%	88.5%	90%	91%	92%	95.5%	98%	100%
.025	77%*	80%*	82.5%	84.5%	86.5%	88.5%	90%	93%	96%	100%
.01	73%	75.5%	78%	80%	82%	84%	85.5%	90%	94.5%	100%

* These values were obtained by comparison with the corresponding values for $\alpha = .05$ and $.01$.

It is easily seen from symmetry that a one-sided t_1 test of $\mu < \nu$ has the same power efficiency as the corresponding one-sided t_1 test of $\mu > \nu$. Thus it is sufficient to consider the one-sided tests of $\mu > \nu$.

The power function is found as a function of the parameter δ, where

$$\delta = \frac{\mu - \nu}{\sigma \sqrt{2}}.$$

Most of the approximate power efficiencies were determined by using the normal approximation given in [2] to compute the power function values. This approximation was used for fractional values of n. Table I contains the results of these computations for one-sided tests of $\mu > \nu$.

Exact values of the power function for integral values of n and $\alpha = .05, .01$ can be found from the tables in [3]. A comparison of the power function values obtained from the normal approximation with these exact values shows that, for $n \leq 6$, $\alpha = .01$ and $n \leq 4$, $\alpha = .05, .025$, the approximation underestimates the true values for small δ and overestimates for large δ. Although this combination of underestimation and overestimation tends to cancel out in the determination of power efficiencies, so that little error in power efficiencies would be expected if the approximation were used for $n = 6$, $\alpha = .01$ or $n = 4$, $\alpha = .05$, the efficiencies given in Table II for $n = 4$, $\alpha = .05$ and $n = 4, 6$, $\alpha = .01$ were obtained from the exact values by graphical interpolation and cross-interpolation.

Power efficiencies were not considered for $n < 4$ because of the difficulties of interpolation and the inexactness of the normal approximation in this range.

For $n = \infty$, t_1 and t_2 both have a normal distribution with zero mean and unit variance. Thus the power efficiency is 100% at all significance levels for this case.

These computations furnish approximate power efficiencies for $n = 6, 8, 10, 15, 25, \infty$ at $\alpha = .05, .025, .01$, and for $n = 4$ at $\alpha = .05$ and $.01$. The other approximate power efficiencies listed in Table II were obtained by graphical interpolation from these values.

The results of this note can be roughly summarized for $n \leq 15$ by stating that of the $2n$ sample values

(i). approximately 1.6 values are lost at the 5% significance level,

(ii). approximately 2.1 values are lost at the 2.5% significance level,

(iii). approximately 2.8 values are lost at the 1% significance level, if the tests based on t_1 are used instead of the corresponding tests based on t_2. Examination of Table I shows that the number of sample values lost decreases as n increases for $n > 15$.

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. 360, 361.

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

[3] J. NEYMAN, "Statistical problems in agricultural experimentation", *Roy. Stat. Soc. Suppl.*, Vol. 2 (1935), pp. 131, 132.

NOTE ON THE LIAPOUNOFF INEQUALITY FOR ABSOLUTE MOMENTS

BY MAURICE H. BELZ

The University of Melbourne

For a variate x measured from the mean of the population, the absolute moment of order r is defined by

$$\nu_r = \int_{-\infty}^{\infty} | x |^r \, dF(x),$$

where $F(x)$ is the cumulative distribution function. Treating r as continuous, we have

$$\frac{d\nu_r}{dr} = \int_{-\infty}^{\infty} | x |^r \log_e | x | \, dF(x),$$

the integral on the right existing if ν_{r+1} exists.