ON THE POWER FUNCTION OF THE "BEST" t-TEST SOLUTION OF THE BEHRENS-FISHER PROBLEM

By John E. Walsh

The Rand Corporation

1. Introduction. The Behrens-Fisher problem is concerned with significance tests for the difference of the means of two normal populations when the ratio of the variances of the populations is unknown. Denote one population by $N(a_1, \sigma_1^2)$ and the other by $N(a_2, \sigma_2^2)$, where the notation $N(a, \sigma^2)$ represents a normal population with mean a and variance σ^2. Let m sample values be drawn from $N(a_1, \sigma_1^2)$ and n sample values from $N(a_2, \sigma_2^2)$ where $m \leq n$. Then Scheffé [1] has shown that certain optimum properties are possessed by a t-test solution he proposed for the Behrens-Fisher problem, in which the numerator of t is based on the difference of the means of the samples while the denominator is based on the square root of a function of the sample values which has a χ^2-distribution with $m - 1$ degrees of freedom. The purpose of this note is to compare the power function of this t-test with the power function of the corresponding most powerful test for the case in which the ratio of variances σ_1^2/σ_2^2 is also known (only one-sided and symmetrical tests are considered). This comparison is made by computing the power efficiency (see section 2 for definition) of Scheffé's test.

It is sufficient to limit power efficiency investigations to one-sided tests. As shown in [2], a symmetrical t-test with significance level 2α has the same power efficiency as the corresponding one-sided t-test with significance level α. Equation (2) of section 2 furnishes an explicit formula whereby approximate power efficiencies can be computed for a wide range of values of α, m, n. Table 1 contains values of (2) for $\alpha = .05, .01$ and several values of m and n.

For the situation considered here, a power efficiency of $100r\%$ has the quantitative interpretation that the given test based on samples of size m and n has approximately the same power function as the corresponding most powerful test based on samples of size rm and rn. Intuitively the power efficiency of a test measures the percentage of available information per observation which is utilized by that test.

2. Power efficiency derivations. The basic notion of the power efficiency of a significance test is given in [2]. For the present case the problem is to determine the value r such that a most powerful test of the same hypothesis (same significance level) based on rm and rn sample values will have approximately the same power function as the given t-test based on m and n sample values (from $N(a_1, \sigma_1^2)$ and $N(a_2, \sigma_2^2)$ respectively). Here the value of σ_1^2/σ_2^2 is assumed to be known. Then the power efficiency of the given t-test equals $100r\%$.

If the ratio of variances σ_1^2/σ_2^2 is known, the most powerful significance test (one-sided and symmetrical) for the difference of means of the two normal populations is a t-test where the numerator of t is based on the difference of the

TABLE 1

Percentage Power Efficiencies for Certain Values of m and n

$\alpha = .05$

m \ n	4	6	10	15	20	30	50	100	∞
4	79.6	73.5	67.2	63.4	61.4	59.3	57.6	56.2	54.9
6		86.9	82.9	80.2	78.7	77.0	75.5	74.2	72.9
10			92.6	90.9	89.8	88.6	87.3	86.2	85.0
15				95.2	94.4	93.5	92.5	91.5	90.3
20					96.4	95.7	94.9	94.0	92.9
30						97.7	97.1	96.4	95.3
50							98.6	98.1	97.2
100								99.3	98.6
∞									100.0

$\alpha = .01$

m \ n	6	8	10	15	20	30	50	100	∞
6	74.9	70.2	66.7	61.2	57.9	54.3	51.1	48.6	45.9
8		81.3	78.8	74.7	72.1	69.1	66.3	63.9	61.4
10			85.3	81.9	79.8	77.2	74.7	72.5	69.9
15				90.4	88.9	87.0	85.0	83.1	80.7
20					92.9	91.4	89.8	88.1	85.8
30						95.3	94.1	92.8	90.7
50							97.2	96.3	94.5
100								98.6	97.3
∞									100.0

two sample means while the denominator is based on the square root of a function of the sample values and σ_1^2/σ_2^2 which has a χ^2-distribution with $m + n - 2$ degrees of freedom [1, p. 43]. Thus the problem is that of comparing the power functions of two t-tests.

As stated in section 1, it is sufficient to consider one-sided tests. We find, using a modification of the normal approximation to the power function of a one-sided t-test given in [3], that Scheffé's one-sided t-test for the Behrens-Fisher problem and the corresponding most powerful one-sided test (σ_1^2/σ_2^2 known) have approximately the same power function when r is chosen so that

$$K_\alpha - \delta\sqrt{r}\{1 - K_\alpha^2/2[(m + n)r - 2]\}^{1/2} = K_\alpha - \delta[1 - K_\alpha^2/2(m - 1)]^{1/2},$$

where α is the significance level of the tests, K_α is the value of the standardized normalized deviate exceeded with probability α, and δ is a function of m, n, a_1, a_2, σ_1^2, σ_2^2 and the given hypothetical value of $a_1 - a_2$ being tested. This condition for the approximate equality of the power functions is reasonably accurate for the following cases: $\alpha = .05$, $m \geq 4$; $\alpha = .025$, $m \geq 5$; $\alpha = .01$, $m \geq 6$; $\alpha = .005$, $m \geq 7$. The accuracy of the approximation increases as m increases.

Hence a value of r such that the two power functions are approximately equal is determined by the equation

$$r\{1 - K_\alpha^2/2[(m + n)r - 2]\} = 1 - K_\alpha^2/2(m - 1). \tag{1}$$

Let

$$A = A(m, \alpha) = 1 - K_\alpha^2/2(m - 1).$$

Then solving (1) for the appropriate root yields

$$r = \frac{1}{2(m + n)} \{2 + (m + n)A + K_\alpha^2/2 + \sqrt{[2 + (m + n)A + K_\alpha^2/2]^2 - 8(m + n)A}\}.$$

Thus the power efficiency of Scheffé's one-sided t-test solution to the Behrens-Fisher problem, for the case in which the ratio of the variances is also known, is approximately equal to

$$\frac{50}{(m + n)} \{2 + (m + n)A + K_\alpha^2/2 + \sqrt{[2 + (m + n)A + K_\alpha^2/2]^2 - 8(m + n)A}\}\%$$

for suitable values of α and m.

REFERENCES

[1] HENRY SCHEFFÉ, "On solutions of the Behrens-Fisher problem based on the t-distribution," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 35–44.

[2] 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.

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