ON A TEST FOR THE MULTIVARIATE BEHRENS-FISHER PROBLEM

by

Koichi Ito

Work at Chapel Hill sponsored by the Office of Naval Research under Contract NR 042 031. Reproduction in whole or in part is permitted for any purpose of the United States government.

Institute of Statistics
Mimeograph Series No. 132
June 17, 1955

ON A TEST FOR THE MULTIVARIATE BEHRENS-FISHER PROBLEM[1]

by Koichi Ito
University of North Carolina

1. Introduction. As a partial solution for the multivariate Behrens-Fisher problem, Bennett [1] suggested an extension of Scheffé's procedure [6] to the multivariate case. The purpose of this note is to show that Bennett's test is the most powerful among all tests obtainable under the procedure, and also to compare its power with that of the ordinary generalized Student T test, when both tests are used to test the same hypothesis.

2. Extension of Scheffé's procedure and a class D of tests for the problem. Suppose that (y_{ir}), $i = 1, \ldots, p$; $r = 1, \ldots, N_1$ and (z_{is}), $i = 1, \ldots, p$; $s = 1, \ldots, N_2$, where it is assumed that $N_1 \leq N_2$, represent two independent random samples from the p-variate normal populations $N(\eta, \Sigma_1)$ and $N(\zeta, \Sigma_2)$, respectively. η and ζ are the p x 1 column vectors whose elements are the means of the p characteristics of the two populations, respectively, and $\Sigma_1 = (\sigma_{ij}^{(1)})$ and $\Sigma_2 = (\sigma_{ij}^{(2)})$, $i, j = 1, \ldots, p$, are unknown covariance matrices of the two populations, respectively, both being positive definite. Suppose moreover that $\Sigma_1 \neq \Sigma_2$.

The problem of testing a null hypothesis H_0: $\eta = \zeta$ against any alternative H: not H_0, is a Behrens-Fisher problem in the multivariate case. We shall apply Scheffé's procedure to this case in a general way to obtain a class D of tests for H_0 which make use of the generalized Student T statistic, and choose the most powerful test in D.

Consider a p x N matrix (x_{it}), where N is to be determined later on, and whose elements are defined as linear combinations of the elements of (y_{ir}) and (z_{is}) as follows:

[1] Work at Chapel Hill sponsored by the Office of Naval Research under Contract NR 042 031. Reproduction in whole or in part is permitted for any purpose of the United States government.

(2.1) $$x_{it} = \sum_{r=1}^{N_1} a_{rt}^{(i)} y_{ir} - \sum_{s=1}^{N_2} b_{st}^{(i)} z_{is}$$

where $i = 1, \ldots, p$; $t = 1, \ldots, N$, and $(a_{rt}^{(i)})$ and $(b_{st}^{(i)})$ are $N_1 \times N$ and $N_2 \times N$ matrices, respectively, whose elements are non-stochastic constants independent of population parameters. They are determined subject to the following conditions:

(2.2) $$E(x_{it}) = \eta_i - \zeta_i = \xi_i$$

and

(2.3) $$E(x_{it} - \xi_i)(x_{ju} - \xi_j) = \delta_{tu}\, \sigma_{ij}$$

for $i, j = 1, \ldots, p$; $t, u = 1, \ldots, N$; where $\delta_{tu} = 1$ if $t = u$ and $\delta_{tu} = 0$ if $t \neq u$. Then the null hypothesis H_0: $\underline{\eta} = \underline{\zeta}$ is equivalent to H_0' : $\underline{\xi} = 0$, and can be tested by the generalized Student T [2]. If we set

(2.4) $$x_{i.} = \frac{1}{N} \sum_{t=1}^{N} x_{it},$$

(2.5) $$s_{ij} = \frac{1}{n} \sum_{t=1}^{N} (x_{it} - x_{i.})(x_{jt} - x_{j.})$$

where $n = N - 1$, then the test consists of the following critical region:

(2.6) $$T^2 = N \sum_{i,j=1}^{p} s^{ij}\, x_{i.}\, x_{j.} \geq T_\alpha^2 \;,$$

where $(s^{ij}) = (s_{ij})^{-1}$, and

$$T_\alpha^2 = \frac{np}{n-p+1}\, F_\alpha(p, n-p+1),$$

$F_\alpha(p, n-p+1)$ being the upper $100\,\alpha$ per cent point of the analysis of variance F distribution with degrees of freedom p and n-p+1. The power of this test is given by [3],

$$(2.7)\qquad P(\lambda|p,\ n\text{-}p\text{+}1,\ \alpha) = 1 - e^{-\lambda} \sum_{h=0}^{\infty} \frac{\lambda^h}{h!}\ I_x\left(\tfrac{1}{2}p + h,\ \tfrac{1}{2}(n - p + 1)\right),$$

where $x = \dfrac{T_\alpha^2}{n + T_\alpha^2}$, $\lambda = \dfrac{N}{2} \sum_{i,j=1}^{p} \sigma^{ij}\, \xi_i\, \xi_j$ and $(\sigma^{ij}) = (\sigma_{ij})^{-1}$

Now, taking expectation and variance of (2.1) we see that conditions (2.2) and (2.3) are equivalent to

$$(2.8)\qquad \sum_{r=1}^{N_1} a_{rt}^{(i)} = 1, \qquad \sum_{s=1}^{N_2} b_{st}^{(i)} = 1,$$

$$(2.9)\qquad \sum_{r=1}^{N_1} a_{rt}^{(i)}\ a_{ru}^{(j)} = \delta_{tu}\ c_{ij}^{(1)}, \quad \sum_{s=1}^{N_2} b_{st}^{(i)}\ b_{su}^{(j)} = \delta_{tu} c_{ij}^{(2)},$$

and

$$(2.10)\qquad \sigma_{ij} = c_{ij}^{(1)}\ \sigma_{ij}^{(1)} + c_{ij}^{(2)}\ \sigma_{ij}^{(2)},$$

where $i, j = 1, \ldots, p$; $t, u = 1, \ldots, N$, and $c_{ij}^{(1)}$ and $c_{ij}^{(2)}$ are independent of t. Following arguments similar to Scheffé's [6], it is easily shown that in order for $a_{it}^{(i)}$ and $b_{st}^{(i)}$ to satisfy (2.8) and (2.9), N, $C_{ij}^{(1)}$ and $c_{ij}^{(2)}$ must satisfy

$$(2.11)\qquad N \leq N_1 \leq N_2$$

and

$$(2.12)\qquad c_{ij}^{(1)} = c^{(1)} \geq \frac{N}{N_1}, \quad c_{ij}^{(2)} = c^{(2)} \geq \frac{N}{N_2},$$

where $c^{(1)}$ and $c^{(2)}$ are constants independent of i and j. Hence, we have obtained a class D of tests of the type (2.6), each of which is given by a solution of (2.8) and (2.9) for $a_{rt}^{(i)}$ and $b_{st}^{(i)}$ when N, $c^{(1)}$ and $c^{(2)}$ satisfying (2.11) and (2.12), respectively, are specified. We note here that the power of a test in D against any specific alternative $\underline{\xi} = \underline{\eta} - \underline{\zeta} \neq 0$ is determined by the values of N, $c^{(1)}$ and $c^{(2)}$, and so our next problem is to choose N, $c^{(1)}$ and $c^{(2)}$

so as to obtain the most powerful test in D.

3. Determination of the most powerful test in D. We shall first observe the following two properties of the power function given by (2.7).

(i) $P(\lambda \mid a, b, \alpha)$ is a non-decreasing function of b for fixed values of λ, a and α. This is easily proved from Hsu's theorem [4] concerning the power function of the analysis of variance F test. As a matter of fact, tables of $P(\lambda \mid a, b, \alpha)$ (e.g., Tang [7]) indicate that $P(\lambda|a,b,\alpha)$ is actually a monotonically increasing function of b for fixed λ and a when α is a usual level of significance, i.e., .05 or .01.

(ii) $P(\lambda \mid a, b, \alpha)$ is a monotonically increasing function of λ for fixed values of a, b and α [5].

We shall next show that the parameter λ in $P(\lambda \mid a, b, \alpha)$ is a monotonically decreasing function of $c^{(1)}$ for fixed $c^{(2)}$, and also a monotonically decreasing function of $c^{(2)}$ for fixed $c^{(1)}$. We write:

$$\lambda = \frac{N}{2} \sum_{i,j=1}^{p} \sigma^{ij} \xi_i \xi_j = \frac{N}{2} \underline{\xi}'(c^{(1)} \Sigma_1 + c^{(2)} \Sigma_2)^{-1} \underline{\xi}.$$

where $\underline{\xi}'$ denotes the transpose of $\underline{\xi}$.

Now, let $\Sigma_1 = M \cdot \mathrm{Diag}(e_1, e_2, \ldots, e_p) \cdot M'$ and $\Sigma_2 = MM'$, where M is a p x p non-singular matrix, and $\mathrm{Diag}(e_1, e_2, \ldots, e_p)$ is a diagonal matrix whose p elements are the characteristic roots (all positive) of $\Sigma_1 \Sigma_2^{-1}$. Then, we get

$$\lambda = \frac{N}{2} \underline{\xi}' \left(c^{(1)} M \,\mathrm{Diag}\,(e_1, e_2, \ldots, e_p) M' + c^{(2)} M M'\right)^{-1} \underline{\xi}$$

$$= \frac{N}{2} \underline{\xi}' \left\{M \,\mathrm{Diag}\,(c^{(1)}e_1 + c^{(2)}, c^{(1)}e_2 + c^{(2)}, \ldots, c^{(1)}e_p + c^{(2)}) M'\right\}^{-1} \underline{\xi}$$

Again, let $\underline{\theta}' = \underline{\xi}' M'^{-1}$. Then

$$\lambda = \frac{N}{2} \underline{\theta}' \operatorname{Diag} (c^{(1)} e_1 + c^{(2)}, \ldots, c^{(1)} e_p + c^{(2)})^{-1} \underline{\theta}$$

$$= \frac{N}{2} \sum_{i=1}^{p} \frac{\theta_i^2}{c^{(1)} e_i + c^{(2)}} .$$

Hence, since e_i and θ_i are independent of $c^{(1)}$ and $c^{(2)}$, our proposition follows.

It is now clear that the most powerful test in D against any specific alternative is obtained, when we ascribe to N its maximum value satisfying (2.11), i.e., N_1 and then to $c^{(1)}$ and $c^{(2)}$ their respective minimum values satisfying (2.12), i.e., 1 and N_1/N_2. Hence the most powerful test in D is given by solving,

$$(3.1) \qquad \sum_{r=1}^{N_1} a_{rt}^{(i)} = 1, \qquad \sum_{s=1}^{N_2} b_{st}^{(i)} = 1,$$

and

$$(3.2) \qquad \sum_{r=1}^{N_1} a_{rt}^{(i)} a_{ru}^{(j)} = \delta_{tu}, \qquad \sum_{s=1}^{N_2} b_{st}^{(i)} b_{su}^{(j)} = \delta_{tu} \frac{N_1}{N_2} ,$$

where i, j = 1, ..., p; t, u = 1, ..., N_1. Very simple solutions of (3.1) and (3.2) are an immediate application of Scheffé's solution which was proposed by Bennett [1]. They are as follows.

$$(3.3) \qquad a_{rt}^{(i)} = \delta_{rt},$$

$$b_{st}^{(i)} = \begin{cases} \delta_{st} \left(\frac{N_1}{N_2}\right)^{1/2} - \frac{1}{(N_1 N_2)^{1/2}} + \frac{1}{N_2} & \text{for } s \leq N_1 \\ \frac{1}{N_2} & \text{for } s > N_1 . \end{cases}$$

By substituting (3.3) into (2.1) we obtain

$$(3.4) \qquad x_{it} = y_{it} - \left(\frac{N_1}{N_2}\right)^{1/2} z_{it} + \frac{1}{(N_1 N_2)^{1/2}} \sum_{s=1}^{N_1} z_{is} - \frac{1}{N_2} \sum_{s=1}^{N_2} z_{is},$$

where $i = 1, \ldots, p$; $t = 1, \ldots, N_1$, thus the critical region of the most powerful test in D being completely specified.

4. Comparison of power. In this section we shall compare the power of the test given by (2.6) and (3.4) with that of an optimum test for the case where $\Sigma_1 = \Sigma_2$. The latter test is an ordinary generalized Student T test whose critical region consists of

$$T^2 = \frac{N_1N_2}{N_1+N_2} \sum_{i,j=1}^{p} v^{ij} (y_{i.}-z_{i.})(y_{j.}-z_{j.}) \geq T_\alpha^2 , \tag{4.1}$$

where

$$y_{i.} = \frac{1}{N_1} \sum_{r=1}^{N_1} y_{ir}, \quad z_{i.} = \frac{1}{N_2} \sum_{s=1}^{N_2} z_{is} ,$$

$$v_{ij} = \frac{1}{m} \left\{ \sum_{r=1}^{N_1} (y_{ir}-y_{i.})(y_{jr}-y_{j.}) + \sum_{s=1}^{N_2} (z_{is}-z_{i.})(z_{js}-z_{j.}) \right\} ,$$

$$(v^{ij}) = (v_{ij})^{-1},$$

$$m = N_1 + N_2 - 2$$

and

$$T_\alpha^2 = \frac{mp}{m-p+1} F_\alpha(p, m-p+1).$$

The power of this test is given by $P(\lambda|p, m-p+1, \alpha)$ where

$$\lambda = \frac{N_1N_2}{2(N_1+N_2)} \sum_{i,j=1}^{p} \sigma_{(1)}^{ij} (\eta_i-\zeta_i)(\eta_j-\zeta_j)$$

and

$$(\sigma_{(1)}^{ij}) = (\sigma_{ij}^{(1)})^{-1} = (\sigma_{ij}^{(2)})^{-1}.$$

The test (2.6) with $N = N_1$ and $\sigma_{ij} = \sigma_{ij}^{(1)} + \frac{N_1}{N_2} \sigma_{ij}^{(2)}$ which is specified by (3.4) is the most powerful in class D of tests for testing H_0: $\underline{\eta} = \underline{\zeta}$ under the assumption that $\Sigma_1 \neq \Sigma_2$, but since we deal with a less favorable situation where nothing is known of Σ_1 and Σ_2, this test is presumed to be less powerful than (4.1). For the sake of comparison we suppose that both tests are used to test H_0 assuming that $\Sigma_1 = \Sigma_2$. Then it is easily seen that in the power functions of the two tests, the parameter λ takes on the same value for any specified alternative of $\underline{\eta}$ and $\underline{\zeta}$. Hence, given λ and α, we may compare the two powers for various values of N_1 and N_2. The following tables give the ratios of powers of Bennett's test to that of (4.1) for several values of λ, α, p, N_1 and N_2. They are calculated by using Tang's table [7].

Table 1. Ratio of powers for α = .05 (above) and for α = .01 (below), p = 2, λ = 1.5

N_1 \ N_2	6	10	15	20	∞
6	.730 .566	.669 .473	.633 .422	.615 .394	.547 .307
10		.847 .710	.818 .664	.801 .628	.724 .507
15			.904 .805	.889 .784	.817 .650
20				.936 .864	.863 .729
∞					1.000 1.000

Table 2. Ratio of powers for α = .05 (above) and for α = .01 (below), p = 3, λ = 2.

N_1 \ N_2	7	10	15	20	∞
7	.667 .500	.620 .436	.576 .376	.554 .345	.475 .250
10		.778 .615	.739 .557	.716 .520	.626 .390
15			.859 .740	.843 .705	.751 .555
20				.904 .817	.813 .652
∞					1.000 1.000

Table 3. Ratio of powers for α = .05 (above) and α = .01 (below), p = 4 and λ = 2.5

N_1 \ N_2	8	10	15	20	∞
8	.616 .449	.584 .404	.537 .342	.509 .310	.422 .211
10		.701 .533	.655 .467	.630 .429	.528 .300
15			.816 .669	.792 .636	.684 .468
20				.876 .764	.765 .597
∞					1.000 1.000

5. Summary. As is clear from the present status of the theory of multivariate analysis, the existence and construction of an "optimum" test for the multivariate Behrens-Fisher problem is a formidable affair, but it is of interest to know that Bennett's test provides an exact test and is the most powerful among all tests obtainable under Scheffé's procedure. Also its power is fairly high when $N_1 = N_2$, as compared with that of the optimum test (4.1) for the case of $\Sigma_1 = \Sigma_2$.

6. Acknowledgment. The author is grateful to Professor Harold Hotelling under whose kind supervision this work was done.

REFERENCES

[1] B. M. Bennett, "Note on a solution of the generalized Behrens-Fisher problem," Annals of the Institute of Statistical Mathematics, Vol. 2, No. 2 (1951), pp. 87-90.

[2] H. Hotelling, "The generalization of Student's ratio," Annals of Mathematical Statistics, Vol. 2 (1931), pp. 360-378.

[3] P. L. Hsu, "Notes on Hotelling's generalized T," Annals of Mathematical Statistics, Vol. 9 (1938), pp. 231-243.

[4] P. L. Hsu, "Analysis of variance from the power function point of view," Biometrika, Vol. 32 (1941), pp. 62-69.

[5] S. N. Roy, "A report on some aspects of multivariate analysis," Institute of Statistics, University of North Carolina, Mimeograph Series No. 121, pp. 61-62.

[6] H. Scheffé, "On solutions of the Behrens-Fisher problem based on the t-distribution," Annals of Mathematical Statistics, Vol. 14 (1943), pp. 35-44.

[7] P. C. Tang, "The power function of the analysis of variance tests with tables and illustrations of their use," Statistical Research Memoir, Vol. 2(1938), pp. 126-157.