Biometrika (1965), **52**, 1 *and* 2, *p.* 139
Printed in Great Britain

An approximate degrees of freedom solution to the multivariate Behrens–Fisher problem*

By YING YAO
Bell Telephone Laboratories, Inc., Holmdel, New Jersey

1. Introduction

The comparison of the means of two populations on the basis of two independent samples is one of the oldest problems in statistics. Indeed, it has been a testing ground for many methods of inference as well as for a variety of analytic approaches to practical problems. The univariate problem was first studied by Behrens (1929) and his solution was presented by Fisher (1935) in terms of the fiducial theory. Welch studied it in the confidence theory framework and provided an 'approximate degrees of freedom' solution as well as an asymptotic series solution (1936, 1947). Many others have investigated this topic and various methods of approach were also suggested by Jeffreys (1940), Scheffé (1943), McCullough, Gurland & Rosenberg (1960), Banerjee (1961), and Savage (1961). In the multivariate extension of the Behrens–Fisher problem, Bennett (1951) has extended the Scheffé solution, and James (1954) the Welch series solution. The present paper studies an extension of the Welch 'approximate degrees of freedom' (APDF) solution provided by Tukey (1959), and discusses the results of a Monte Carlo sampling study on this new APDF solution and its comparison with the James series solution.

2. The multivariate problem

We wish to compare the means $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$, of two normal p-dimensional populations on the basis of two random samples. Let the samples be denoted by

$$\{\mathbf{x}_{ij}, j = 1, 2, \ldots, f_i + 1\} \quad (i = 1, 2).$$

Define the sample means and their covariance matrices as

$$\bar{\mathbf{x}}_i = \frac{1}{f_i+1} \sum_{j=1}^{f_i+1} \mathbf{x}_{ij},$$

$$S_i = \frac{1}{f_i(f_i+1)} \sum_{j=1}^{f_i+1} (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)'.$$

Let Σ_1, Σ_2 be the true covariance matrices of the sample means, and '$\sim$' denote 'is distributed as'; then

$$\bar{\mathbf{x}}_i \sim N(\boldsymbol{\mu}_i, \Sigma_i),$$

$$f_i S_i \sim \text{Wishart}(f_i, \Sigma_i) \quad (i = 1, 2).$$

The set $(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, S_1, S_2)$ forms a set of joint minimal sufficient statistics for the entire sample.

Let the difference between the means be

$$\boldsymbol{\mu} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2.$$

* Part of a Ph.D Thesis at the University of Chicago.

We shall devote our attention to the problem of testing hypotheses of the form

$$H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0,$$

where $\boldsymbol{\mu}_0$ is a given fixed vector.

The difference between the sample means is an estimate for $\boldsymbol{\mu}$; in fact

$$\mathbf{x} = \bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 \sim N(\boldsymbol{\mu}, \Sigma),$$

where $$\Sigma = \Sigma_1 + \Sigma_2.$$

Writing $$\mathbf{y} = \mathbf{x} - \boldsymbol{\mu}_0,$$

so that $$\mathbf{y} \sim N(\boldsymbol{\mu} - \boldsymbol{\mu}_0, \Sigma),$$

then if Σ_1, Σ_2 were known, a test of H_0 with critical region

$$\mathbf{y}'\Sigma^{-1}\mathbf{y} \geqslant \chi^2_\alpha(p)$$

would have exact significance level α, or

$$\Pr\{\mathbf{y}'\Sigma^{-1}\mathbf{y} \geqslant \chi^2_\alpha(p)\} \equiv \alpha,$$

provided that H_0 is true.

If Σ_1, Σ_2 are unknown but assumed to be proportional ($\Sigma_1 = k\Sigma_2$, k known) a pooled estimate for Σ is

$$S_E = \left(\frac{f_1 S_1 + f_2 k S_2}{f}\right)\left(\frac{k+1}{k}\right),$$

where $$f = f_1 + f_2, \quad fS_E \sim \text{Wishart}\,(f, \Sigma),$$

and we may construct a Hotelling T^2 statistic

$$\mathbf{y}'S_E^{-1}\mathbf{y} \sim T^2(p, f),$$

where the T^2 distribution is related to an F distribution by

$$T^2_\alpha(p, f) = \left(\frac{fp}{f-p+1}\right) F_\alpha(p, f-p+1).$$

A T^2-test with critical region $$\mathbf{y}'S_E^{-1}\mathbf{y} \geqslant T^2_\alpha(p, f)$$

has exact level of significance α and is uniformly most powerful of all tests invariant under non-singular transformations of the sample space.

When there is no knowledge of Σ_1, Σ_2, if they are independently estimated by S_1, S_2, then

$$S = S_1 + S_2$$

is an estimate for Σ, and the quadratic form $\mathbf{y}'S^{-1}\mathbf{y}$ is in large samples asymptotically distributed as a chi-square with p degrees of freedom. The test with critical region

$$\mathbf{y}'S^{-1}\mathbf{y} \geqslant \chi^2_\alpha(p)$$

has type I error $$\Pr\{\mathbf{y}'S^{-1}\mathbf{y} \geqslant \chi^2_\alpha(p)\} = \alpha + O(f_i^{-1}). \quad (1)$$

The form of the large sample solution suggests that an improvement may be obtained by choosing a critical region of the form

$$\mathbf{y}'S^{-1}\mathbf{y} \geqslant c_\alpha,$$

where c_α is also a sample statistic; we hope that for some well chosen function c_α,

$$\Pr\{\mathbf{y}'S^{-1}\mathbf{y} \geqslant c_\alpha\} = \alpha + O(f_i^{-k}), \tag{2}$$

with $k > 1$.

Here we shall study a new solution which is a multivariate extension of the Welch 'approximate degrees of freedom' solution. We define an Approximate Degrees of Freedom (APDF) f_T by

$$\frac{1}{f_T} = \frac{1}{f_1}\left(\frac{\mathbf{y}'S^{-1}S_1S^{-1}\mathbf{y}}{\mathbf{y}'S^{-1}\mathbf{y}}\right)^2 + \frac{1}{f_2}\left(\frac{\mathbf{y}'S^{-1}S_2S^{-1}\mathbf{y}}{\mathbf{y}'S^{-1}\mathbf{y}}\right)^2 \tag{3}$$

and propose the critical region $\quad \mathbf{y}'S^{-1}\mathbf{y} \geqslant T^2_\alpha(p, f_T)$.

Thus once the sample value f_T is determined, the standard T^2-test may be carried out with test statistic $\mathbf{y}'S^{-1}\mathbf{y}$. The function f_T was suggested by Tukey (1959); a heuristic derivation is given below in § 3.

The APDF f_T solution may be compared with James's first-order asymptotic series test (1954). In general he suggests the use of a critical region

$$\mathbf{y}'S^{-1}\mathbf{y} \geqslant h_j(S_1, S_2; \alpha),$$

where the h_j are to be derived so that

$$\Pr\{\mathbf{y}'S^{-1}\mathbf{y} \geqslant h_j(S_1, S_2; \alpha)\} = \alpha + O(f_i^{-j-1}) \tag{4}$$

for $j = 0, 1, 2, \ldots$. In the notation adopted here he finds

$$h_0(S_1, S_2; \alpha) = \chi^2_\alpha(p), \tag{5}$$

$$h_1(S_1, S_2; \alpha) = \chi^2_\alpha(p)\left[1 + \frac{1}{2}\left(\frac{k_1}{p} + \frac{k_2\chi^2_\alpha(p)}{p(p+2)}\right)\right], \tag{6}$$

with

$$k_1 = \sum_{i=1}^{2} (\operatorname{tr} S^{-1}S_i)^2/f_i,$$

$$k_2 = \sum_{i=1}^{2} [(\operatorname{tr} S^{-1}S_i)^2 + 2\operatorname{tr}(S^{-1}S_iS^{-1}S_i)]/f_i.$$

A comparative Monte Carlo study for the type I errors of the tests at selected sets of $(f_1, f_2, \Sigma_1, \Sigma_2)$ as given in Table 1 suggests a slight superiority for the APDF f_T test.

We may use the numerical example (*a*) given by James (1954) to illustrate the APDF method and compare it with the series solution h_1.

The sample means and their covariances are

$$\bar{\mathbf{x}}_1 = \begin{pmatrix} 9{\cdot}82 \\ 15{\cdot}06 \end{pmatrix}, \quad \bar{\mathbf{x}}_2 = \begin{pmatrix} 13{\cdot}05 \\ 22{\cdot}57 \end{pmatrix},$$

$$S_1 = \begin{pmatrix} 7{\cdot}500 & -1{\cdot}019 \\ -1{\cdot}019 & 1{\cdot}112 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 7{\cdot}436 & 2{\cdot}918 \\ 2{\cdot}918 & 4{\cdot}891 \end{pmatrix},$$

$$f_1 = 15, \quad f_2 = 10.$$

The difference between the means is

$$\mathbf{x} = \bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 = \begin{pmatrix} -3{\cdot}23 \\ -7{\cdot}51 \end{pmatrix},$$

and the sample estimate for the covariance matrix is

$$S = S_1 + S_2 = \begin{pmatrix} 14{\cdot}936 & 1{\cdot}899 \\ 1{\cdot}899 & 6{\cdot}003 \end{pmatrix},$$

while

$$S^{-1} = \begin{pmatrix} 0{\cdot}06976 & -0{\cdot}02207 \\ -0{\cdot}02207 & 0{\cdot}17357 \end{pmatrix}.$$

The test statistic is

$$\mathbf{x}'S^{-1}\mathbf{x} = 9{\cdot}4471$$

and the APDF f_T is given by

$$1/f_T = \tfrac{1}{15}(0{\cdot}1657)^2 + \tfrac{1}{10}(0{\cdot}8343)^2, \quad f_T = 14{\cdot}00.$$

Using

$$T^2_\alpha(2, f_T) = \frac{2f_T}{f_T - 1} F_\alpha(2, f_T - 1)$$

the critical values $T^2_\alpha(2, f_T)$ may be compared with the values of the James series solution $h_1(S_1, S_2; \alpha)$ and the large sample solution $\chi^2_\alpha(2)$.

α	$\chi^2_\alpha(2)$	$h_1(S_1, S_2; \alpha)$	$T^2_\alpha(2, 14{\cdot}0)$
0·05	5·991	7·23	8·21
0·025	7·378	9·18	10·70
0·01	9·210	11·90	14·43

It is seen that here the critical values for the APDF test are higher than the h_1 and χ^2 values.

3. Motivation of the APDF solution

In constructing a test with approximate significance level α, we shall consider the test statistic $\mathbf{y}'S^{-1}\mathbf{y}$ and use the univariate Welch APDF method to suggest a multivariate generalization based on the T^2-distribution. We have

$$\mathbf{y} \sim N(\mathbf{0}, \Sigma).$$

If S were a Wishart matrix,

$$fS \sim \text{Wishart}(f, \Sigma),$$

then for an arbitrary constant vector $\mathbf{b}$ we should have,

$$\mathbf{b}'\mathbf{y} \sim N(\mathbf{0}, \mathbf{b}'\Sigma\mathbf{b}),$$

$$f(\mathbf{b}'S\mathbf{b}) \sim (\mathbf{b}'\Sigma\mathbf{b})\chi^2(f),$$

and

$$w_b = \frac{(\mathbf{b}'\mathbf{y})^2}{(\mathbf{b}'S\mathbf{b})} \sim t^2(f).$$

It can be shown (Bush & Olkin, 1959) that

$$\sup_{\mathbf{b}}(w_b) = w_{b*} = \frac{(\mathbf{b}^{*\prime}\mathbf{y})^2}{(\mathbf{b}^{*\prime}S\mathbf{b}^*)} = \mathbf{y}'S^{-1}\mathbf{y},$$

where the maximizing $\mathbf{b}^*$ is

$$\mathbf{b}^* = S^{-1}\mathbf{y};$$

also

$$\mathbf{y}'S^{-1}\mathbf{y} \sim T^2(p, f).$$

Thus for every fixed $\mathbf{b}$, we would have

$$\Pr\{w_b \geqslant t^2_\alpha(f)\} \equiv \alpha,$$

while for the supremum over all $\mathbf{b}$,

$$\Pr\{w_{b*} \geqslant T^2_\alpha(p,f)\} \equiv \alpha.$$

In actual fact, however, for every fixed $\mathbf{b}$, $\mathbf{b}'S\mathbf{b}$ is a linear combination of two chi-square variates, with f_1 and f_2 degrees of freedom,

$$\mathbf{b}'S\mathbf{b} = \mathbf{b}'S_1\mathbf{b} + \mathbf{b}'S_2\mathbf{b}.$$

Applying the Welch APDF method, we get

$$\Pr\{w_b \geqslant t^2_\alpha(f_b)\} = \alpha + O(f_i^{-2}),$$

where the APDF f_b is

$$\frac{1}{f_b} = \frac{1}{f_1}\left(\frac{\mathbf{b}'S_1\mathbf{b}}{\mathbf{b}'S\mathbf{b}}\right)^2 + \frac{1}{f_2}\left(\frac{\mathbf{b}'S_2\mathbf{b}}{\mathbf{b}'S\mathbf{b}}\right)^2.$$

Extending this to the supremum over all $\mathbf{b}$, we hope a similar condition exists so that

$$\Pr\{w_b^* \geqslant T^2_\alpha(p,f_b^*)\} = \alpha + O(f_i^{-2}).\dagger$$

Writing f_T instead of f_b^*, the expression for f_T becomes,

$$\frac{1}{f_T} = \frac{1}{f_1}\left(\frac{\mathbf{y}'S^{-1}S_1S^{-1}\mathbf{y}}{\mathbf{y}'S^{-1}\mathbf{y}}\right)^2 + \frac{1}{f_2}\left(\frac{\mathbf{y}'S^{-1}S_2S^{-1}\mathbf{y}}{\mathbf{y}'S^{-1}\mathbf{y}}\right)^2.$$

We are thus led to explore the test procedure with critical region

$$\mathbf{y}'S^{-1}\mathbf{y} \geqslant T^2_\alpha(p,f_T). \tag{7}$$

4. Computational methods

4·1. *Reduction of the probability integral*

For a given pair of positive definite symmetric matrices (Σ_1, Σ_2) there exists a non-singular matrix C and a diagonal matrix Λ, with diagonal elements $0 \leqslant \lambda_1 \leqslant \lambda_2 \ldots \leqslant \lambda_p \leqslant 1$ such that

$$\left.\begin{aligned} C\Sigma_1C' &= \Lambda = \Lambda_1, \\ C\Sigma_2C' &= I - \Lambda = \Lambda_2, \\ C\Sigma C' &= I, \end{aligned}\right\} \tag{8}$$

where I is the $(p \times p)$ identity matrix.

The type I error of the APDF test for a given set of (Σ_1, Σ_2) is a function of the diagonal matrix Λ only and may be written as

$$P_T(\Lambda) = \Pr\{\mathbf{z}'M^{-1}\mathbf{z} \geqslant T^2_\alpha(p,f_T)\}, \tag{9}$$

where

$$\mathbf{z} \sim N(\mathbf{0}, I),$$

$$f_iM_i \sim \text{Wishart}\,(f_i, \Lambda_i) \quad (i = 1, 2),$$

$$M = M_1 + M_2,$$

and

$$\frac{1}{f_T} = \frac{1}{f_1}\left(\frac{\mathbf{z}'M^{-1}M_1M^{-1}\mathbf{z}}{\mathbf{z}'M^{-1}\mathbf{z}}\right)^2 + \frac{1}{f_2}\left(\frac{\mathbf{z}'M^{-1}M_2M^{-1}\mathbf{z}}{\mathbf{z}'M^{-1}\mathbf{z}}\right)^2.$$

† It is not known whether this result holds for $p > 1$.

Similarly, the type I error of the James h_1 test is

$$P_J(\Lambda) = \Pr\{\mathbf{z}'M^{-1}\mathbf{z} \geqslant h_1(M_1, M_2; \alpha)\}. \tag{10}$$

For ease in numerical calculations, expressions (9) and (10) may be written as mathematical expectations,

$$P_T(\Lambda) = E[Q(R_T)], \tag{11}$$

$$P_J(\Lambda) = E[Q(R_J)], \tag{12}$$

where

$$R_T = \frac{fT^2_\alpha(p, f_T)}{(\mathbf{d}'M^{-1}\mathbf{d})\,w}, \tag{13}$$

$$R_J = \frac{fh_1(M_1, M_2; \alpha)}{(\mathbf{d}'M^{-1}\mathbf{d})\,w}, \tag{14}$$

with

$$\mathbf{d} = \mathbf{z}(\mathbf{z}'\mathbf{z})^{-\frac{1}{2}}, \quad w = \sum_{i=1}^{2} f_i \operatorname{tr}(\Lambda_i^{-1} M_i),$$

and $Q(a)$ is the complementary cumulative F-distribution with (p, fp) degrees of freedom.

Expression (11) is obtained from (9) by writing

$$\begin{aligned} P_T(\Lambda) &= \Pr\{\mathbf{z}'M^{-1}\mathbf{z} \geqslant T^2_\alpha(p, f_T)\} \\ &= \Pr\{(\mathbf{z}'\mathbf{z})(\mathbf{d}'M^{-1}\mathbf{d}) \geqslant T^2_\alpha(p, f_T)\} \\ &= \Pr\left\{\frac{f(\mathbf{z}'\mathbf{z})}{w} \geqslant \frac{fT^2_\alpha(p, f_T)}{(\mathbf{d}'M^{-1}\mathbf{d})\,w}\right\} \\ &= \Pr\{x \geqslant R_T\} \\ &= E[\Pr\{x \geqslant R_T \mid R_T\}] \\ &= E[Q(R_T)], \end{aligned}$$

where

$$x = \frac{f(\mathbf{z}'\mathbf{z})}{w} \sim F(p, fp),$$

and x is independent of R_T. A parallel analysis on (10) gives (12) immediately.

We shall present one more expression useful in the Monte Carlo estimation of $P_T(\Lambda)$ and $P_J(\Lambda)$. Let

$$\begin{aligned} P_E(\Lambda) &= \Pr\{\mathbf{z}'M^{-1}\mathbf{z} \geqslant T^2_\alpha(p, f)\} \\ &= E[Q(R_E)] \end{aligned} \tag{15}$$

where

$$R_E = \frac{fT^2_\alpha(p, f)}{w(\mathbf{d}'M^{-1}\mathbf{d})}.$$

For the special case

$$\Lambda = \Lambda_E = (f_1/f)\,I$$

we know that $\mathbf{z}'M^{-1}\mathbf{z}$ is a T^2-variable, so

$$P_E(\Lambda_E) \equiv \alpha.$$

4·2. *The Monte Carlo estimation*

The values $P_T(\Lambda)$ and $P_J(\Lambda)$ are mathematical expectations of the variables $Q(R_T)$ and $Q(R_J)$. For any set of parameters (f_1, f_2, Λ), R_T and R_J may be generated from a set of standard variables $(\mathbf{z}, W_1, W_2)$,

$$\mathbf{z} \sim N(\mathbf{0}, I) \quad (p \times 1),$$

$$W_i \sim \text{Wishart}\,(f_i, I) \quad (p \times p),$$

$$M_i = \frac{1}{f_i}(\Lambda_i^{\frac{1}{2}} W_i \Lambda_i^{\frac{1}{2}}) \quad (i = 1, 2).$$

These samples were generated from the Rand (1955) table of normal deviates, using Bartlett (1933) decomposition for the Wishart matrices.

Once a set of $(\mathbf{z}, W_1, W_2)$ is available we may use it to define R_T as well as R_J. This means that the resulting estimates P_T and P_J are correlated, thus providing a closer comparison between the two solutions. Re-use of the samples reduces computation both in the generating process and in the parallel computations for P_T and P_J. The matrices W_1, W_2 are added and further used as samples from Wishart $(f_1 + f_2, I)$.

Since the M_i's are obtained from the W_i's by transformation using Λ, the W_i's are re-used for various values of Λ, so that the effect of different Λ on P_T and P_J is reflected more readily in the samples.

If N independent sets of $(\mathbf{z}, W_1, W_2)$ are available we would then have N sets of (R_T, R_J) for each Λ, and the sample means

$$\bar{P}_T(\Lambda) = \frac{1}{N}\sum_{i=1}^{N} Q(R_{Ti}),$$

$$\bar{P}_J(\Lambda) = \frac{1}{N}\sum_{i=1}^{N} Q(R_{Ji})$$

may be used as estimates for P_T, P_J. However, we also know that an exact solution exists for

$$\Lambda_E = \frac{f_1}{f} I,$$

namely,

$$P_E(\Lambda_E) = E[Q(R_E)] \equiv \alpha,$$

as already shown in § 4·1. Thus if we obtain a set of R_E from the samples, the sample mean

$$\bar{P}_E(\Lambda_E) = \frac{1}{N}\sum_{i=1}^{N} Q(R_{Ei})$$

is an estimate for α. A regression method can then be applied to provide estimates and confidence intervals for P_T and P_J.

5. Results of Monte Carlo estimation

Estimates of P_T and P_J, with their 95% confidence intervals were computed for different sets of (f_1, f_2, Λ), where Λ is defined by expression (8) in § 4. The nominal significance levels for bivariate APDF f_T and James h_1 tests were $\alpha = 0{\cdot}01$, $0{\cdot}05$. For $(f_1, f_2) = (6, 12)$, 500 samples were used, while for $(f_1, f_2) = (12, 12), (6, 18)$, only 240 samples were used. The methods followed have been described in § 4·2.

In Table 1 below (λ_1, λ_2) are the diagonal elements of Λ; e.g. at $(f_1, f_2) = (6, 12)$, $\lambda_1 = \lambda_2 = 0{\cdot}1$, $\alpha = 0{\cdot}05$, the 95% confidence limits for P_T are $0{\cdot}0509 \pm 0{\cdot}0025$, centred at the regression estimate 0·0509. Correspondingly, the confidence interval for P_J is $0{\cdot}0575 \pm 0{\cdot}0025$. It may be seen from the table that P_T is in general smaller than P_J, and their standard errors are very similar. In terms of the original problem, with covariance matrices Σ_1, Σ_2 for the two sample means, if $\lambda_1 = \lambda_2 = \lambda$, so that $\Sigma_1 = [\lambda/(1-\lambda)]\Sigma_2$, we can see the effect of changing λ on P_T, P_J. When $\lambda_1 \neq \lambda_2$, P_T and P_J seems to decrease as λ_1 becomes much smaller than λ_2.

Table 1. *Monte Carlo estimates of type I errors of the two bivariate tests (with 95 % confidence limits)*

Degrees of freedom, etc.				Nominal level $\alpha = 0{\cdot}05$				Nominal level $\alpha = 0{\cdot}01$			
				APDF f_T		James series h_1		*APDF* f_T		James series h_1	
f_1	f_2	λ_1	λ_2								
6	12	1/10	1/10	0·0509	± 0·0025	0·0575	± 0·0025	0·0110	± 0·0004	0·0144	± 0·0004
		1/3	1/3	·0471	·0004	·0530	·0002	·0088	·0002	·0119	·0002
		1/2	1/2	·0468	·0016	·0532	·0014	·0086	·0014	·0117	·0002
		1/4	3/4	·0458	·0026	·0526	·0026	·0084	·0010	·0114	·0011
		1/10	9/10	·0435	·0035	·0528	·0038	·0080	·0013	·0117	·0018
		2/3	2/3	·0496	·0032	·0561	·0029	·0099	·0013	·0129	·0012
		9/10	9/10	·0574	·0074	·0698	·0079	·0156	·0032	·0206	·0036
		27/28	27/28	·0555	·0078	·0767	·0092	·0158	·0048	·0258	·0050
12	12	1/28	1/28	0·0502	0·0073	·00582	0·0073	0·0124	0·0038	0·0161	0·0040
		1/10	1/10	·0512	·0060	·0563	·0063	·0122	·0040	·0146	·0030
		1/4	1/4	·0490	·0034	·0531	·0031	·0104	·0001	·0122	·0002
		1/2	1/2	·0479	·0004	·0516	·0001	·0092	·0001	·0108	·0001
		1/4	3/4	·0485	·0025	·0523	·0025	·0094	·0001	·0112	·0001
		1/10	9/10	·0476	·0039	·0525	·0039	·0090	·0001	·0112	·0001
6	18	1/10	1/10	0·0510	0·0047	0·0545	0·0049	0·0106	0·0010	0·0124	0·0010
		1/4	1/4	·0478	·0003	·0514	·0003	·0091	·0002	·0109	·0001
		3/4	3/4	·0573	·0065	·0617	·0078	·0156	·0039	·0172	·0038
		9/10	9/10	·0646	·0134	·0749	·0129	·0225	·0178	·0262	·0075

The author wishes to thank Professor D. L. Wallace for his invaluable help during the course of the research work and Professor J. W. Tukey for his suggestion of the APDT method. Comments from the referee were also much appreciated.

REFERENCES

BANERJEE, S. (1961). On confidence interval for two-mean problem based on separate estimates of variances and tabulated values of *t*-table. *Sankhyā*, **10**, 359–78.

BARTLETT, M. S. (1933). On the theory of statistical regression. *Proc. Roy. Soc. Edin.* **53**, 260–83.

BEHRENS, W. V. (1929). Ein bertrag zur fehlerberechnung bei wenigen beobachtungen. *Landw. Jb.* **68**, 807–37.

BENNETT, B. M. (1951). Note on a solution of the generalized Behrens–Fisher problem. *Ann. Inst. Stat. Math.* **2**, 87–90.

BUSH, K. A. & OLKIN, I. (1959). Extrema of quadratic forms with applications to statistics. *Biometrika*, **46**, 483–6.

FISHER, R. A. (1935). The fiducial argument in statistical inference. *Ann. Eugen., Lond.*, **6**, 391–8.

JAMES, G. S. (1954). Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown. *Biometrika*, **41**, 19–43.

JEFFREYS, H. (1940). Notes on the Behrens–Fisher formula. *Ann. Eugen., Lond.*, **10**, 48–51.

McCullough, R. S., Gurland, J. & Rosenberg, L. (1960). Small sample behaviour of certain tests of the hypothesis of equal means under variance heterogeneity. *Biometrika*, **47**, 345–53.

Rand Corporation (1955). *One Million Random Digits and 100,000 Normal Deviates.* Glencoe, Illionois: Free Press.

Savage, L. J. (1961). The foundations of statistics reconsidered. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1, University of California Press, Berkeley and Los Angeles.

Scheffé, H. (1943). On solutions of the Behrens–Fisher problem based on *t*-distribution. *Ann. Math. Statist.* **10**, 129–37.

Tukey, J. (1959). Suggestion of the APDF f_T approach. Personal communication.

Welch, B. L. (1936). Specification of rules for rejecting too variable a product, with particular reference to an electric lamp problem. *J.R. Statist. Soc. Supp.*, **3**, 29–48.

Welch, B. L. (1947). The generalization of Student's problem when several populations are involved. *Biometrika*, **34**, 28–35.