

An Alternative Solution to Hotelling T square under the Heteroscedasticity of the dispersion Matrix

P.O. Adebayo* and G.M. Oyeyemi

Department of Statistics, University of Ilorin, Ilorin, Nigeria

Research Article

Received date: 12/11/2018

Accepted date: 23/11/2018

Published date: 28/11/2018

*For Correspondence

Department of Statistics, University of Ilorin, Ilorin, Nigeria.

E-mail: bayooni3@gmail.com

Keywords: Variance co-variance matrix, linear combination, Type I error rate, power the test, Heteroscedasticity, R statistical package

ABSTRACT

This work focus on developing an alternative procedure to the multivariate Behrens-Fisher problem by using an approximate degree of freedom test which was adopted from Satterthwaite univariate procedure. The proposed procedure was compared via R package simulation and real-life data used by Tim (1975) with six (6) existing procedures namely: Johanson, Yao, Krishnamoorthy, Hotelling T square, Nel and Van der Merwe and Yanagihara. And it was discovered that proposed procedure performed better in term of power of the test than all existing procedures considered in all the scenarios that are at different: (i) random variables (p), (ii) variance-covariance matrix, (iii) sample size and (iv) significant level (α). And compete favorably well in term of type I error rate with Johanson, Yao, Krishnamoorthy, Nel, and Van der Merwe.

INTRODUCTION

Suppose we have a random sample of size n_1 , $x_{11}, x_{12}, x_{13}, \dots, x_{1n_1}$ for $N(\mu_1, \sigma_1^2)$ and a second random sample of size n_2 , $x_{21}, x_{22}, x_{23}, \dots, x_{2n_2}$ for $N(\mu_2, \sigma_2^2)$. It is desired to test $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$. If σ_1 and σ_2 are both known a normal test is used. If $\sigma_1 = \sigma_2$ but both are unknown a t-test is commonly used with the test statistics.

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\text{Where } S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$\bar{x}_i = \frac{1}{n_i} \sum x_{ij} \text{ and } S_i^2 = \frac{\sum (x_{ij} - \bar{x}_i)^2}{n_i - 1}, i=1,2$$

A t-test is a type of statistical test that is used to compare the means of two groups, such as men vs. women, athletes vs. non-athletes, young vs. elderly, or you may want to compare means measured on a single group under two different experimental conditions or at two different times. T-tests are a type of parametric method; they can be used when the samples satisfy the conditions of normality, equal variance, and independence. T-tests can be divided into two types. There is the independent t-test, which can be used when the two groups under comparison are independent of each other, and the paired t-test, which can be used when the two groups under comparison are dependent on each other. If $\sigma_1 \neq \sigma_2$ and both are unknown then we are confronted with the Behrens-Fisher problem. There is no universally accepted testing procedure for this problem although arrays of tests have been developed and will be discussed in the Review of Literature. Behrens^[1] proposed the statistics.

In the literature, there is a modified test statistic (t-test) when the assumption of equal variances is violated has been known as the Behrens-Fisher problem ^[1,2]. Early investigations showed that the problem can be overcome by substituting separate-variances tests, such as the ones introduced by Welch ^[3,4], and Satterthwaite ^[5], for the Student t-test. These modified significance tests, unlike the usual two-sample Student t-test, do not pool variances in the computation of an error term. Moreover, they alter the degrees of freedom by a function that depends on sample data. It has been found that these procedures in many cases restore Type I error probabilities to the nominal significance level and also counteract increase or decrease of Type II error ^[6-9].

Student's t-test is univariate and analog to Hotelling T square which is the multivariate version of T-test and this Hotelling's T² has three basic assumptions that are fundamental to the statistical theory: independent, multivariate normality and equality of variance-covariance matrices. A statistical test procedure is said to be robust or insensitive if departures from these assumptions do not greatly affect the significance level or power of the test.

To use Hotelling's T² one must assume that the two samples are independent and that their variance-covariance matrices are equal ($\Sigma_1 = \Sigma_2 = \Sigma$). When variance-covariance matrices are not homogeneous and unknown, the test statistic will not be distributed as a T². This predicament is known as the multivariate Behrens-Fisher problem.

The Behrens-Fisher Problem addresses interval estimation and hypothesis testing concerning the differences between the means of two normally distributed populations when the variances of the two populations are not equal. While Multivariate Behrens-Fisher problem deal with testing the equality of two normal mean vectors under heteroscedasticity of dispersion matrices. These are the some of the existing Multivariate Behrens-Fisher problem: Yao ^[10], Johansen ^[11], Nel et al ^[12], Kim ^[13], Krishnamoorthy and Yu ^[14], Gamage et al ^[15] Yanagihara and Yuan ^[16], and Kawasaki and Seo ^[17] and so on. But with all these procedures there is no one with a hundred percent (100%) satisfactory in term of power of the test and type I error rate. And each of these scholars works on the degree of freedom using a different method which is classified into four (4): Approximate degree of freedom tests, Series expansion-based tests, Simulation-based tests, and Transformation-based tests.

Scheffe ^[18], Lauer and Han ^[19], Lee and Gurland ^[20], Murphy ^[21], Yao ^[10], Algina and Tang ^[22], Kim ^[13], De la Rey and Nel's ^[23], Christensen and Rencher ^[24], Oyeyemi and Adebayo ^[25]. All these authors mentioned and many more have work on the comparison of some of the Multivariate Behrens-Fisher problem procedures.

The purpose of this work is to develop an alternative procedure for multivariate data that will be robust compared to other procedures and the work will begin with an introduction to the statistical notation that will be helpful in understanding the concepts. This is followed by a discussion of procedures that can be used to test the hypothesis of multivariate mean equality when statistical assumptions are and are not satisfied. We will then show how to obtain a test that is robust to the covariance heterogeneity.

MULTIVARIATE BEHRENS-FISHER PROBLEM (EXISTING PROCEDURE)

Consider two p-variate normal populations $N(\underline{\mu}_1, \Sigma_1)$ and $N(\underline{\mu}_2, \Sigma_2)$ where $\underline{\mu}_1$ and $\underline{\mu}_2$ are unknown $p \times 1$ vector and Σ_1 and Σ_2 are unknown $p \times p$ positive definite matrices.

Let $X_{\alpha 1} \sim N(\underline{\mu}_1, \Sigma_1), \alpha = 1, 2, \dots, n_1$, and

$X_{\alpha 2} \sim N(\underline{\mu}_2, \Sigma_2), \alpha = 1, 2, \dots, n_2$, denote random samples from these two populations, respectively. We are interested

in the testing problem

$$H_0 : \underline{\mu}_1 = \underline{\mu}_2 \quad \text{vs} \quad H_1 : \underline{\mu}_1 \neq \underline{\mu}_2$$

For $i=1,2$, let

$$\bar{X}_i = \frac{1}{n_i} \sum_{\alpha=1}^{n_i} X_{\alpha i}$$

$$A_i = \sum_{\alpha=1}^{n_i} (X_{\alpha i} - \bar{X}_i)(X_{\alpha i} - \bar{X}_i)'$$

$$S_i = A_i / (n_i - 1), \quad i = 1, 2.$$

Then $\bar{X}_1, \bar{X}_2, A_1$ and A_2 , which are sufficient for the mean vectors and dispersion matrices, are independent random variables having the distributions:

$$\bar{X}_i \sim N\left(\mu_i, \frac{\Sigma_i}{n_i}\right), \text{ and } A_i \sim W_p(n_i - 1, \Sigma_i), i = 1, 2$$

Where $W_p(r, \Sigma)$ denotes the p -dimensional wishart distribution with $df=r$ and scale matrix Σ .

$$\tilde{S}_i = S_i / n_i, \quad i = 1, 2$$

$$\tilde{S} = \tilde{S}_1 + \tilde{S}_2$$

$$T^2 = (\bar{X}_1 - \bar{X}_2)' \tilde{S}^{-1} (\bar{X}_1 - \bar{X}_2)$$

The following are the existing procedures or solutions to Multivariate Behrens-Fisher problem considered in this study

Yao Procedure

Yao ^[10], invariant test. This is a multivariate extension of the Welch ‘approximate degree of freedom’ solution provided by Turkey and the test statistic is based on a transformation of Hotelling T^2 . And is based on $T^2 \sim (vp / (v - p + 1)) F_{p, v-p+1}$ with the degrees of freedom v given by

$$\frac{1}{v} = \frac{1}{(T^2)^2} \sum_{i=1}^2 \frac{1}{n_i - 1} \left[(\bar{x}_1 - \bar{x}_2)' \tilde{S}^{-1} \tilde{S}_i \tilde{S}^{-1} (\bar{x}_1 - \bar{x}_2) \right]^2$$

$$T_{Yao} = \frac{(v - p + 1) T^2}{pv}$$

Statistical significance is then assessed by comparing the T_{Yao} statistic to its critical value $F_{\alpha}(p, v-p+1)$, that is, a critical value from the F distribution with p and $v-p+1$ degrees of freedom (df)

Johansen Procedure

Johansen ^[11], invariant test, Yanagihara and Yuan ^[16], Kawasaki and Takashi ^[17]. They used $T^2 \sim qF_{pv}$ where

$$q = p + 2D - 6D / [p(p - 1) + 2],$$

$$v_{Joh} = p(p + 2) / 3D$$

$$D = \frac{1}{2} \sum_{i=1}^2 \left\{ tr \left[\left(I - (\tilde{S}_1^{-1} + \tilde{S}_2^{-1})^{-1} \tilde{S}_i^{-1} \right)^2 \right] + tr \left[\left(I - (\tilde{S}_1^{-1} + \tilde{S}_2^{-1})^{-1} \tilde{S}_i^{-1} \right) \right]^2 \right\} / n_i$$

And his proposed test statistic $T_{Johan} = \frac{T^2}{q}$

Statistical significance is then assessed by comparing the T_{Johan} statistic to its critical value $F_{\alpha}(p, v_{Joh})$, that is, a critical value from the F distribution with p and v_{Joh} degrees of freedom (df).

Nel and Van der Merwe (1986) noninvariant solution

Here we use

$$T^2 \sim (v_{Nv} p / (v_{Nv} - p + 1)) F_{p, v_{Nv} - p + 1}$$

except that v is defined by

$$v_{Nv} = \frac{tr(\tilde{S}^2) + [tr(\tilde{S})]^2}{\left(\frac{1}{n_1}\right)\left\{tr(\tilde{S}_1^2) + [tr(\tilde{S}_1)]^2\right\} + \left(\frac{1}{n_2}\right)\left\{tr(\tilde{S}_2^2) + [tr(\tilde{S}_2)]^2\right\}}$$

Krishnamoorthy and Yu (2004)'s, Lin and Wang (2009), Modified Nel/ Van der Merwe Invariant Solution

We use the idea as before, namely,

$$T^2 \sim (v_{kY}p / (v_{kY} - p + 1)) F_{p, v_{kY}-p+1}$$
 with the d.f.v defined by

$$v_{kY} = (p + p^2) / C$$

$$C = \frac{1}{n_1} \left\{ tr \left[(\tilde{S}_1 \tilde{S}_1^{-1})^2 \right] + \left[tr \left(\tilde{S}_1 \tilde{S}_1^{-1} \right) \right]^2 \right\} + \frac{1}{n_2} \left\{ tr \left[(\tilde{S}_2 \tilde{S}_2^{-1})^2 \right] + \left[tr \left(\tilde{S}_2 \tilde{S}_2^{-1} \right) \right]^2 \right\}$$

$$T_{krish} = \frac{(v_{kY} - p + 1) T^2}{pv_{kY}}$$

Statistical significance is then assessed by comparing the T_{krish} statistic to its critical value $F_{\alpha}(p, v_{kY}-p+1)$, that is, a critical value from the F distribution with p and $v_{kY}-p+1$ degrees of freedom (df)

Yanagihara and Yuan Procedure used the Series Expansion Based Test to Developed an Alternative Procedure to Multivariate Behrens-Fisher problem

$$T_{Yan} = \frac{n-2-\hat{\theta}_1}{(n-2)p} T \sim F_{p, \hat{v}_{yan}}$$

$$\hat{v}_{yan} = \frac{(n-2-\hat{\theta}_1)^2}{(n-2)\hat{\theta}_2 - \hat{\theta}_1}$$

$$\hat{\theta}_1 = \frac{p\hat{\psi}_1 + (p-2)\hat{\psi}_2}{p(p+2)} \quad \text{and} \quad \hat{\theta}_2 = \frac{\hat{\psi}_1 + 2\hat{\psi}_2}{p(p+2)}$$

$$\hat{\psi}_1 = \frac{n_2^2(n-2)}{n^2(n_1-1)} \left\{ tr(S_1 \bar{S}^{-1}) \right\}^2 + \frac{n_1^2(n-2)}{n^2(n_2-1)} \left\{ tr(S_2 \bar{S}^{-1}) \right\}^2$$

$$\hat{\psi}_2 = \frac{n_2^2(n-2)}{n^2(n_1-1)} tr(S_1 \bar{S}^{-1} S_1 \bar{S}^{-1}) + \frac{n_1^2(n-2)}{n^2(n_2-1)} tr(S_2 \bar{S}^{-1} S_2 \bar{S}^{-1})$$

$$\bar{S} = \frac{n_2}{n} S_1 + \frac{n_1}{n} S_2$$

Hotelling's T²

$$T_1 = \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_1 - \bar{x}_2)' S_{pl}^{-1} (\bar{x}_1 - \bar{x}_2)$$

Where

$$S_{pl} = \frac{1}{n_1 + n_2 - 2} \left[(n_1 - 1) S_1 + (n_2 - 1) S_2 \right]$$

The test statistic can also be converted to an F statistic,

$$v_{hotel} = \frac{N - p - 1}{p(N - 2)} T_1$$

Where $N = n_1 + n_2$. Statistical significance is then assessed by comparing the v_{hotel} statistic to its critical value $F_{\alpha}(p, N - p - 1)$, that is, a critical value from the F distribution with p and $N - p - 1$ degrees of freedom (df).

THE MEAN AND VARIANCE OF THE CHI-SQUARE DISTRIBUTION WITH N DEGREES OF FREEDOM

The chi-square distribution is defined with n degrees of freedom by

$$\chi_n^2 = Z_1^2 + Z_2^2 + \dots + Z_n^2, \text{ where } Z_1, Z_2, \dots, Z_n \text{ are independent random variables, each with distribution } N(0,1).$$

Find the expected value and variance of both sides, then we have

$$E(\chi_n^2) = E(Z_1^2) + \dots + E(Z_n^2) \text{ and}$$

$$Var(\chi_n^2) = Var(Z_1^2) + \dots + Var(Z_n^2)$$

And all the instances of Z_i have identical distributions, then

$$E(\chi_n^2) = nE(Z^2) \text{ and}$$

$$Var(\chi_n^2) = nVar(Z^2)$$

Where Z is the random variable with distribution $N(0, 1)$. Then

$$E(Z^2) = E[(Z - 0)^2] = E[(Z - \mu_z)^2] = Var(Z) = 1$$

Therefore

$$E(\chi_n^2) = n \cdot 1 = n \tag{1}$$

For $Var(Z^2)$

$$Var(Z^2) = E[(Z - \mu_{z^2})^2] = E[(Z^2 - 1)^2] = E(Z^4 - 2Z^2 + 1)$$

$$= E(Z^4) - 2E(Z^2) + 1 = E(Z^4) - 2 \cdot 1 + 1$$

Now

$$Var(Z^2) = E(Z^4) - 1 \tag{2}$$

To find $E(Z^4)$, we will use the fact that for any continuous random variable X with probability density function f , and any exponent k ,

$$E(X^k) = \int_{-\infty}^{\infty} x^k f(x) dx$$

And that the probability density function f of the $N(0,1)$ random variable is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

$$\text{Then, } E(Z^4) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^4 e^{-\frac{z^2}{2}} dz$$

By integration by parts, we have, $E(Z^4) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^4 e^{-\frac{z^2}{2}} dz$

$$= \frac{1}{\sqrt{2\pi}} \left\{ \left[z^3 \cdot \left(-e^{-\frac{z^2}{2}} \right) \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} 3z^2 \left(-e^{-\frac{z^2}{2}} \right) dz \right\}$$

$$= \frac{1}{\sqrt{2\pi}} \left\{ 0 + \int_{-\infty}^{\infty} 3z^2 e^{-\frac{z^2}{2}} dz \right\}$$

$$= 3 \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} dz$$

$$= 3E(z^2) = 3 \cdot 1 = 3 \tag{3}$$

Therefore substitute equation (3) into equation (2) then we have

$$Var(Z^2) = 3 - 1 = 2, \text{ that is } Var(\chi_n^2) = n \cdot 2 = 2n \tag{4}$$

For two sample t-test, we will limit this work to the version of the test where we do not assume that the two populations have equal variances. Let random sample x_1, \dots, x_{n_1} from a random variable X with distribution $N(\mu_1, \sigma_1^2)$ and a random sample y_1, \dots, y_{n_2} from a random variable Y with distribution (μ_2, σ_2^2) . We have

$$t = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Strictly speaking, this statistic does not follow t-distribution, therefore;

The variance of $\bar{X} - \bar{Y}$ is $\sigma_B^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ and as an estimator for σ_B^2 we have $s_B^2 = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$

For t to bet-distribution, there would have to be some multiple of s_B^2 that is chi-square distribution and this is not the case. However, remember that in the one-sample case, $\frac{(n-1)s^2}{\sigma^2}$ had a chi-square distribution with n-1 degrees of freedom. By analogy, we have $\frac{rS_B^2}{\sigma_B^2}$ has a chi-square distribution with r degrees of freedom. Satterthwaite found the true distribution of s_B^2 and showed that if r is chosen so that the variance of the chi-square distribution with r degrees of freedom is equal to the true variance of $\frac{rS_B^2}{\sigma_B^2}$, then, under certain conditions, this chi-square distribution with r degrees of freedom is a good approximation to be the true distribution of $\frac{rS_B^2}{\sigma_B^2}$. so from this point, we are assuming that $\frac{rS_B^2}{\sigma_B^2}$ has distribution χ_r^2 . So from equation

$$Var\left(\frac{rS_B^2}{\sigma_B^2}\right) = 2r \tag{5}$$

$$Var\left(\frac{rS_B^2}{\sigma_B^2}\right) = \frac{r^2}{\sigma_B^4} Var(s_B^2) \tag{6}$$

Equating the equation (5) and (6)

$$2r = \frac{r^2}{\sigma_B^4} \text{Var}(s_B^2)$$

$$\frac{2}{r} = \frac{1}{\sigma_B^4} \text{Var}(s_B^2) \tag{7}$$

Now $s_B^2 = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$, and s_1 and s_2 are independent so

$$\text{Var}(s_B^2) = \frac{1}{n_1^2} \text{Var}(s_1^2) + \frac{1}{n_2^2} \text{Var}(s_2^2) \tag{8}$$

We know that $\frac{(n_1 - 1)S_1^2}{\sigma_1^2}$ has a chi-square distribution with $n_1 - 1$ degrees of freedom, from equation (3)

$$\text{Var}\left[\frac{(n_1 - 1)S_1^2}{\sigma_1^2}\right] = 2(n_1 - 1)$$

$$\frac{(n_1 - 1)^2}{\sigma_1^2} \text{Var}(s_1^2) = 2(n_1 - 1)$$

$$\text{Var}(s_1^2) = \frac{2\sigma_1^4}{n_1 - 1} \text{ and similarly } \text{Var}(s_2^2) = \frac{2\sigma_2^4}{n_2 - 1} \tag{9}$$

Substitute equation (9) into equation (8)

$$\text{Var}(s_B^2) = \frac{1}{n_1^2} \cdot \frac{2\sigma_1^4}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{2\sigma_2^4}{n_2 - 1} \tag{10}$$

Substitute equation (10) into equation (7)

$$\frac{2}{r} = \frac{1}{\sigma_B^4} \left(\frac{1}{n_1^2} \cdot \frac{2\sigma_1^4}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{2\sigma_2^4}{n_2 - 1} \right)$$

$$r = \frac{\sigma_B^2}{\frac{1}{n_1^2} \cdot \frac{2\sigma_1^4}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{2\sigma_2^4}{n_2 - 1}} \tag{11}$$

In practice, the values of the population variances, σ_1^2 and σ_2^2 , are unknown, and so we replace σ_1^2 , σ_2^2 , and σ_B^2 by their estimators s_1^2 , s_2^2 and s_B^2 also $s_B^2 = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$ from equation (11)

$$r = \frac{s_B^2}{\frac{1}{n_1^2} \cdot \frac{2s_1^4}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{2s_2^4}{n_2 - 1}} = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1} \cdot \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1} \cdot \left(\frac{s_2^2}{n_2}\right)^2}$$

MULTIVARIATE APPROACH

We shall consider the test statistic $y'S^{-1}y$ and use Univariate Satterthwaite approximation of degrees of freedom method to suggest multivariate generalization based on the T^2 -distribution. We have

$$S = \hat{S}_1 + \hat{S}_2, \hat{S}_i = \frac{S_i}{n_i} \quad \text{and} \quad y = \bar{X}_1 - \bar{X}_2$$

$$y \sim N(0, \Sigma)$$

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

Where b is an arbitrary constant vector

$$v(b'Sb) \sim (b'\Sigma b) \chi_{(v)}^2$$

$$\frac{v(b'Sb)}{(b'\Sigma b)} \sim \chi_{(v)}^2 \quad \text{Multivariate of a version of equation (5)}$$

$$Var\left(\frac{v(b'Sb)}{(b'\Sigma b)}\right) = 2v \tag{12}$$

Multivariate of a version of equation (5) is equation (12)

$$\frac{v^2}{(b'\Sigma b)^2} Var(b'Sb) = 2v$$

$$Var(b'Sb) = \frac{2v(b'\Sigma b)^2}{v^2}$$

$$\frac{2(b'\Sigma b)^2}{v} = Var(b'Sb)$$

$$\frac{2}{v} = \frac{1}{(b'\Sigma b)^2} Var(b'Sb)$$

Note $S = \frac{S_1}{n_1} + \frac{S_2}{n_2}$ that is $(b'Sb) = \frac{(b'S_1b)}{n_1} + \frac{(b'S_2b)}{n_2}$ and $(b'\Sigma b) = \frac{(b'\Sigma_1 b)}{n_1} + \frac{(b'\Sigma_2 b)}{n_2}$ (14)

Put equation (14) into equation (13)

$$\frac{2}{v} = \frac{1}{(b'\Sigma b)^2} Var\left(\frac{(b'S_1b)}{n_1} + \frac{(b'S_2b)}{n_2}\right)$$

$$\frac{2}{v} = \frac{1}{(b'\Sigma b)^2} Var\left(\frac{1}{n_1^2} Var(b'S_1b) + \frac{1}{n_2^2} Var(b'S_2b)\right) \tag{15}$$

Multivariate of a version of equation (9) is

$$Var(b'S_1b) = \frac{2((b'\Sigma_1b))^2}{n_1 - 1} \text{ and } Var(b'S_2b) = \frac{2((b'\Sigma_2b))^2}{n_2 - 1} \tag{16}$$

Put equation (16) into equation (15)

$$\begin{aligned} \frac{2}{v} &= \frac{1}{(b'\Sigma b)^2} \left[\frac{1}{n_1^2} \cdot \frac{2((b'\Sigma_1b))^2}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{2((b'\Sigma_2b))^2}{n_2 - 1} \right] \\ \frac{2}{v} &= \frac{2}{(b'\Sigma b)^2} \left[\frac{1}{n_1^2} \cdot \frac{((b'\Sigma_1b))^2}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{((b'\Sigma_2b))^2}{n_2 - 1} \right] \\ \frac{1}{v} &= \frac{1}{(b'\Sigma b)^2} \left[\frac{1}{n_1^2} \cdot \frac{((b'\Sigma_1b))^2}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{((b'\Sigma_2b))^2}{n_2 - 1} \right] \end{aligned} \tag{17}$$

Equation (17) becomes

$$v = \frac{(b'\Sigma b)^2}{\frac{1}{n_1^2} \cdot \frac{((b'\Sigma_1b))^2}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{((b'\Sigma_2b))^2}{n_2 - 1}}$$

The values of the population variances Σ_i are unknown, and so we replace Σ_i and $b'\Sigma b$ by their estimators S_i and $b'Sb$

$$v = \frac{(b'Sb)^2}{\frac{1}{n_1^2} \cdot \frac{(b'\Sigma_1b)^2}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{(b'\Sigma_2b)^2}{n_2 - 1}} \tag{18}$$

Put equation (14) into equation (18) to have

$$v = \frac{\frac{(b'S_1b)}{n_1} + \frac{(b'S_2b)}{n_2}}{\frac{1}{n_1^2} \cdot \frac{(b'S_1b)^2}{n_1 - 1} + \frac{1}{n_2^2} \cdot \frac{(b'S_2b)^2}{n_2 - 1}} \tag{19}$$

Set $b=S^{-1}y$ then equation (19) becomes

$$v = \frac{\left(\frac{y'S^{-1}S_1S^{-1}y}{n_1} + \frac{y'S^{-1}S_2S^{-1}y}{n_2} \right)^2}{\frac{(y'S^{-1}S_1S^{-1}y)^2}{n_1^2(n_1 - 1)} + \frac{(y'S^{-1}S_2S^{-1}y)^2}{n_2^2(n_2 - 1)}} \tag{20}$$

Equation (20) can be in this form

$$f = \frac{\left(\sum \frac{1}{n_i} (yS^{-1}S_iS^{-1}y) \right)^2}{\sum \frac{1}{n_i^2(n_i - 1)} (yS^{-1}S_iS^{-1}y)^2} \tag{21}$$

Let $y = \bar{X}_1 - \bar{X}_2$ then equation (22) become

$$T^2 = \frac{\left(\sum \frac{1}{n_i} ((\bar{X}_1 - \bar{X}_2) S^{-1} S_i S^{-1} (\bar{X}_1 - \bar{X}_2)) \right)^2}{\sum \frac{1}{n_i^2 (n_i - 1)} ((\bar{X}_1 - \bar{X}_2) S^{-1} S_i S^{-1} (\bar{X}_1 - \bar{X}_2))}$$

$$T^2 \sim \left(\frac{f_{prop} \times p}{(f_{prop} - p + 1)} \right) F_{p, f_{prop} - p + 1} \tag{22}$$

$$T_{prop} = \frac{(f_{prop} - p + 1) T^2}{f_{prop} \times p} \tag{23}$$

Then equation (23) is the test statistic of the proposed procedure. Statistical significance is then assessed by comparing the T_{prop} statistic to its critical value $F_{\alpha}(p, f_{prop} - p + 1)$, that is, a critical value from the F distribution with p and $f_{prop} - p + 1$ degrees of freedom (df)

SIMULATION STUDY

A simulation study using R package was conducted in order to estimate and compare the Type I error rate and power for each of the previously discussed approximate solution [10,11,14], Proposed procedure, Hotelling’s T square, [14,16]. The simulations are carried out when the null hypothesis is true and not true, for Multivariate normal distribution, when there are unequal variance-covariance matrix. Five (5) factors were varied in the simulation: the sample size, the number of variables p, variance covariance matrices, mean vectors, and significant levels. For each of the above combinations, an $n_i \times p$ data matrix $X_i (i=1 \text{ and } 2)$ were replicated 1,000. The comparison criteria; type I error rate and power of the test were therefore obtained and the results were presented in both tabular.

The following are the levels used for each of the three factors.

Table 1. Levels used for each of the three factors.

Multivariate Distribution	P	α	Sample size
Normal	2, 3, 4	0.01	20, 10
	2, 3, 4	0.025	50, 30
	2, 3, 4	0.05	100, 60

These levels provide 36-factor combinations the values for sample size are shown in **Table 1**.

RESULT

Table 2. Power of the test.

P=2	$\alpha=0.01$							
$X_1=(20 \ 30)$	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.3852	0.3872	0.3872	0.3861	0.2146	0.3915	0.2487
$X_2=(10 \ 30)$	50, 30	0.8332	0.8349	0.8346	0.8374	0.6672	0.8348	0.7691
	100, 60	0.991	0.9911	0.9911	0.9912	0.9497	0.9911	0.9865
$S_1 = \begin{pmatrix} 267 & 200 \\ 200 & 267 \end{pmatrix}$	$\alpha=0.025$							
	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.5037	0.5043	0.5045	0.5037	0.3158	0.508	0.3787
	50, 30	0.8903	0.8911	0.8909	0.8926	0.7609	0.891	0.8528

$S_2 = \begin{pmatrix} 35 & 25 \\ 25 & 35 \end{pmatrix}$	100, 60	0.9959	0.9959	0.9959	0.996	0.973	0.9959	0.9941
	$\alpha=0.05$							
	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
$\bar{x}_1 - \bar{x}_2 = (10 \ 0)$	20, 10	0.6106	0.6105	0.6104	0.6099	0.4201	0.6133	0.5092
	50, 30	0.9319	0.932	0.9319	0.9329	0.8348	0.932	0.911
	100, 60	0.9982	0.9982	0.9982	0.9983	0.9855	0.9982	0.9975

From the **Table 2** Nel and Van der Merwe have the highest power of the test when the sample sizes are small (20, 10) but at (50, 30) and (100, 60) the proposed procedure has the highest power than all other procedures.

Table 3. Type I error rate.

P=2		$\alpha=0.01$						
$X_1=(20 \ 30)$	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.011	0.011	0.011	0.01	0.002	0.011	0.002
$X_2=(20 \ 30)$	50, 30	0.007	0.007	0.007	0.007	0.001	0.006	0.004
	100, 60	0.008	0.008	0.008	0.008	0	0.008	0.008
$S_1 = \begin{pmatrix} 267 & 200 \\ 200 & 267 \end{pmatrix}$	$\alpha=0.025$							
	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.017	0.017	0.017	0.018	0.003	0.018	0.005
	50, 30	0.026	0.027	0.026	0.027	0.008	0.026	0.015
$S_2 = \begin{pmatrix} 35 & 25 \\ 25 & 35 \end{pmatrix}$	100, 60	0.015	0.015	0.015	0.015	0.002	0.015	0.009
	$\alpha=0.05$							
	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
$\bar{x}_1 - \bar{x}_2 = (0 \ 0)$	20, 10	0.048	0.049	0.049	0.048	0.006	0.049	0.023
	50, 30	0.045	0.044	0.045	0.046	0.011	0.045	0.035
	100, 60	0.057	0.057	0.057	0.057	0.016	0.057	0.051

From **Table 3**, when the sample size is (20,10) the proposed procedure is on nominal level exactly while Hotelling T square and Yanagihara ^[16] are below the nominal level, but at (50,30) and (100,60) all the procedures are below the nominal level, at significant level 0.01. At $\alpha=0.025$, all the procedures are inflated at (50,30) and deflated at (20,10) and (100,60).

Table 4. Power of the test.

P=3		$\alpha=0.01$						
$x_1=(30 \ 24 \ 50)$	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.2029	0.2198	0.2054	0.2147	0.1313	0.2278	0.0895
$x_2=(15 \ 14 \ 29)$	50, 30	0.6107	0.6182	0.6154	0.6217	0.4472	0.6203	0.5187
	100, 60	0.9295	0.9312	0.9307	0.9322	0.8107	0.9315	0.9073
$S_1 = \begin{pmatrix} 500 & 450 & 350 \\ 450 & 500 & 180 \\ 350 & 180 & 500 \end{pmatrix}$	$\alpha=0.025$							
	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.2954	0.312	0.2983	0.3062	0.206	0.3205	0.1684
	50, 30	0.7137	0.7195	0.7173	0.7222	0.5602	0.7209	0.6451
$S_2 = \begin{pmatrix} 90 & 30 & 20 \\ 30 & 90 & 10 \\ 20 & 10 & 90 \end{pmatrix}$	100, 60	0.9596	0.9606	0.9603	0.9611	0.877	0.9608	0.9478
	$\alpha=0.05$							
	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
$\bar{x}_1 - \bar{x}_2 = (15 \ 10 \ 21)$	20, 10	0.4095	0.4255	0.4122	0.4207	0.3002	0.4334	0.2743
	50, 30	0.7952	0.7992	0.7978	0.8012	0.6624	0.8003	0.7475
	100, 60	0.9754	0.976	0.9758	0.9763	0.9199	0.9761	0.9692

From **Table 4**, It is obvious that proposed procedure performed better than all other procedures at (50,30) and (100, 60) but Nel and Van der Merwe is better when the sample size is (20,10) in all the scenarios considered.

Table 5. Type I error rate.

P=3		α=0.01						
$x_1 = (30 \ 24 \ 50)$	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.007	0.008	0.008	0.006	0.014	0.013	0
$x_2 = (15 \ 14 \ 29)$	50, 30	0.011	0.01	0.011	0.09	0.013	0.011	0.006
	100, 60	0.015	0.015	0.015	0.015	0.018	0.015	0.013
$S_1 = \begin{pmatrix} 500 & 450 & 350 \\ 450 & 500 & 180 \\ 350 & 180 & 500 \end{pmatrix}$	α=0.025							
	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.035	0.033	0.035	0.027	0.033	0.041	0.003
	50, 30	0.025	0.027	0.026	0.024	0.032	0.028	0.011
$S_2 = \begin{pmatrix} 90 & 30 & 20 \\ 30 & 90 & 10 \\ 20 & 10 & 90 \end{pmatrix}$	100, 60	0.029	0.031	0.031	0.03	0.039	0.031	0.026
	α=0.05							
	$n_1 \neq n_2$	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.059	0.066	0.068	0.048	0.049	0.073	0.013
$\bar{x}_1 - \bar{x}_2 = (15 \ 10 \ 21)$	50, 30	0.053	0.059	0.055	0.056	0.052	0.057	0.042
	100, 60	0.061	0.061	0.062	0.061	0.056	0.063	0.051

Table 5, the type I error rate of all procedures considered are fluctuating, either inflated or deflated. At significant level 0.01, 0.02, 0.05 their inflation in type I error rate, when sample sizes are (50,30) and (100,60), but at (20, 10) all most all the procedures are deflated.

Table 6. Type I error rate

P=4		α=0.01						
$x_1 = (2.9 \ 1.8 \ 2.31 \ 3.21)$		Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.343	0.3589	0.3599	0.3508	0.1974	0.3737	0.078
$x_2 = (1.89 \ 0.8 \ 0.3 \ 0.2)$	50, 30	0.86	0.8712	0.8711	0.875	0.7027	0.8734	0.6825
	100, 60	0.9968	0.9971	0.9971	0.9973	0.9733	0.9972	0.9902
$S_1 = \begin{pmatrix} 5 & 3 & 2 & 1 \\ 3 & 5 & 2 & 1 \\ 2 & 2 & 5 & 0.9 \\ 1 & 1 & 0.9 & 5 \end{pmatrix}$	α=0.025							
		Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.4656	0.48	0.4817	0.4721	0.297	0.4946	0.1604
	50, 30	0.918	0.9249	0.9247	0.9271	0.799	0.926	0.8057
$S_2 = \begin{pmatrix} 0.9 & 0.7 & 0.5 & 0.2 \\ 0.7 & 0.9 & 0.3 & 0.1 \\ 0.5 & 0.3 & 0.9 & 0.1 \\ 0.2 & 0.1 & 0.1 & 0.9 \end{pmatrix}$	100, 60	0.9991	0.9992	0.9992	0.9992	0.9888	0.9992	0.9971
	α=0.05							
		Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.5662	0.5793	0.5804	0.5731	0.3905	0.5914	0.2585
$\bar{x}_1 - \bar{x}_2 = (1.01 \ 1 \ 2.10 \ 3.01)$	50, 30	0.9507	0.9547	0.9547	0.9561	0.8649	0.9555	0.886
	100, 60	0.9995	0.9996	0.9996	0.9996	0.9938	0.9996	0.9987

Table 6 shows when that sample size is (20,10) Nel and Van der Merwe performed better than other procedures, but when sample size increases to (50,30) proposed procedure is better. And there was a great competition among the procedures at (100,60).

Table 7. Power of the test.

P=4		α=0.01						
$x_1 = (2.9 \ 1.8 \ 2.31 \ 3.21)$		Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.008	0.013	0.011	0.011	0.002	0.013	0
$x_2 = (1.89 \ 0.8 \ 0.3 \ 0.2)$	50, 30	0.011	0.012	0.012	0.012	0.001	0.012	0.002
	100, 60	0.007	0.007	0.007	0.007	0	0.007	0.004

$S_1 = \begin{pmatrix} 5 & 3 & 2 & 1 \\ 3 & 5 & 2 & 1 \\ 2 & 2 & 5 & 0.9 \\ 1 & 1 & 0.9 & 5 \end{pmatrix}$	$\alpha=0.025$							
		Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.024	0.029	0.031	0.026	0.004	0.033	0
$S_2 = \begin{pmatrix} 0.9 & 0.7 & 0.5 & 0.2 \\ 0.7 & 0.9 & 0.3 & 0.1 \\ 0.5 & 0.3 & 0.9 & 0.1 \\ 0.2 & 0.1 & 0.1 & 0.9 \end{pmatrix}$	50, 30	0.02	0.026	0.026	0.026	0.003	0.026	0.006
	100, 60	0.021	0.022	0.022	0.022	0.002	0.022	0.01
	$\alpha=0.05$							
$\bar{x}_1 - \bar{x}_2 = (1.01 \quad 1 \quad 2.10 \quad 3.01)$		Johan	Yao	Krish	Propo	Hotel	Nel	Yana
	20, 10	0.05	0.055	0.055	0.051	0.009	0.06	0
	50, 30	0.048	0.051	0.051	0.051	0.012	0.051	0.022
	100, 60	0.052	0.053	0.052	0.053	0.008	0.053	0.034

Hotelling T ^[26] square and yanagihara ^[16] are below the nominal level in all the scenarios considered, while other procedures fluctuated (Inflated or deflated) round the nominal level as in **Table 7**.

DATA SET (FOR ILLUSTRATED EXAMPLE)

The data used here is an illustrated example used by Timm. The two sample sizes considered are ten and twenty respectively ($n_1=10$ and $n_2=20$) and two random variables ($p=2$) form each population.

The sample means and their covariances are

$$y_1 = \begin{pmatrix} 45.0 \\ 90.0 \end{pmatrix}, y_2 = \begin{pmatrix} 40.0 \\ 80.0 \end{pmatrix}$$

$$S_1 = \begin{pmatrix} 80.0 & 30.0 \\ 30.0 & 20.0 \end{pmatrix}, S_2 = \begin{pmatrix} 120.0 & -100.0 \\ -100.0 & 200.0 \end{pmatrix}$$

$$n_1=10, n_2=20$$

The difference between the means is

$$y = y_1 - y_2 = \begin{pmatrix} 5.0 \\ 10.0 \end{pmatrix}$$

And the test statistic is

$$T^2 = (y_1 - y_2)' \left(\frac{S_1}{n_1} + \frac{S_2}{n_2} \right)^{-1} (y_1 - y_2)$$

$$(5.0, 10.0)' \begin{pmatrix} 0.0731707 & 0.0121951 \\ 0.0121951 & 0.0853659 \end{pmatrix} \begin{pmatrix} 5.0 \\ 10.0 \end{pmatrix}$$

$$T^2=11.58542$$

Table 8. The result from the illustrated example.

	$\alpha=0.05$						
	Johan	Yao	Krish	Propo	Hotel	Nel	Yana
Critical value	6.978	7.2012	7.223	7.7396	6.9567	6.9601	10.0088
Power	0.4979	0.5109	0.5121	0.868	0.5068	0.4969	0.6244
	$\alpha=0.025$						
Critical value	8.852	9.1661	9.1987	9.9867	8.7984	8.8036	13.8708
Power	0.62	0.6334	0.6347	0.9325	0.6273	0.618	0.7527
	$\alpha=0.01$						
Critical value	11.4986	11.9613	12.0129	13.2753	11.3828	11.3908	20.3456
Power	0.7503	0.7625	0.7638	0.9732	0.7551	0.7456	0.868

From **Table 8**, the proposed procedure has the highest power followed by Yanagihara ^[16], Krishnamoorthy ^[14], Yao ^[10], Hotelling T square ^[26], Johansen ^[11] and Nel and Van der Merwe ^[14] at all the significant level α considered ($\alpha=0.05, 0.025$ and 0.01).

FINDINGS

From the simulation, it is obvious from **Tables 1, 3 and 5** that when the sample size is very small (20,10) proposed procedure is not at his best, but when sample size increases to (50,30) and (100,60), the proposed procedure performed better than the all procedures considered. Nel and Van der Merwe ^[14] performed better when the sample size is very small (20,10) followed by Yao ^[10], Krishnamoorthy ^[14] and Proposed procedures in term of power of the test in all the scenarios considered.

In term of Type I error rate, proposed procedure compete favorably well with the other procedures selected for this study. Yao ^[10], Krishnamoorthy ^[14], Johansen ^[11], Nel and Van der Merwe ^[14] and the proposed procedures are fluctuating (Inflated and Deflated) around the nominal level while Hotelling T ^[26] square and Yanagihara ^[16] are below the nominal level.

REFERENCES

1. Behrens WU. Ein Beitrag zur Fehlerberechnung bei wenigen Beobachtungen. Institut für Pflanzenbau. 1929;68:807-837.
2. Fisher RA. The Fiducial argument in statistical inference. Annals of Eugenics. 1935;6:391–398.
3. Welch BL. The significance of the difference between two means when the population variances are unequal. Biometrika. 1938;29:350-362.
4. Welch BL. The generalization of Student's problem when several different population variances are involved. Biometrika. 1947;34:29-35.
5. Satterthwaite FE. An approximate distribution of estimates of variance components. Biometrics Bulletin. 1946;2:110-114.
6. Overall JE, et al. Tests that are robust against variance heterogeneity in $k \times 2$ designs with unequal cell frequencies. Psychological Reports. 1995a;76:1011-1017.
7. Overall JE, et al. Power of a test that is robust against variance heterogeneity. Psychological Reports. 1995b;77:155-159.
8. Zimmerman DW. A note on preliminary tests of equality of variances. Br J Math Stat Psychol. 2004;57:173-181.
9. Zimmerman DW, et al. Rank transformations and the power of the Student t-test and Welch t' test for non-normal populations with unequal variances. Can J Exp Psychol. 1993;47:523-539.
10. Yao Y. An approximation degrees of freedom solution to the multivariate Behrens-Fisher problem, Biometrika. 1965;52:139-147.
11. Johansen S. The Welch-James approximation to the distribution of the residual sum of squares in a weighted linear regression. Biometrika. 1980;67:85-95.
12. Nel DG, et al. A solution to the multivariate Behrens-Fisher problem. Commun Stat-Simul C. 1986;15:3719-3735.
13. Kim S. A Practical Solution to the Multivariate Behrens-Fisher Problem. Biometrika. 1992;79:171-176.
14. Krishnamoorthy K, et al. Modified Nel and Van der Merwe test for the multivariate Behrens-Fisher problem. Stat Probab Lett. 2004;66:161-169.
15. Gamage J, et al. Generalized p-values and generalized confidence regions for the multivariate Behrens-Fisher problem and MANOVA. J Multivar Anal. 2004;88:177-189.
16. Yanagihara H, et al. Three approximate solutions to the multivariate behrens fisher problem. Comm Statist Simulation Comput. 2005;34:975-988.
17. Kawasaki T, et al. A two-sample test for mean vectors with unequal covariance matrices. American Journal of Mathematical and Management Sciences. 2012;31:117-135.
18. Scheffe' H. Practical solutions of the Behrens-Fisher problem. J Am Stat Assoc. 1970;65:1501-1508.
19. Lauer GN, et al. Power of Cochran's test in the Behrens-Fisher Problem. Technometrics. 1974;16:545–549.
20. Lee A, et al. Size and power of tests for equality of means of two normal populations with unequal variances. J Am Stat Assoc. 1975;70:933–947.
21. Murphy BP. Comparison of some two sample means tests by simulation. Commun Stat-Simul C. 1976;5:23-32.

22. Algina J, et al. Type I error rates for Yao's and James' tests of equality of mean vectors under variance-covariance heteroscedasticity. *Journal of Educational Statistics*. 1988;13:281–290.
23. De la Rey DG. A comparison of the significance levels and power functions of several solutions to the multivariate Behrens-Fisher problem. *S Afr Stat J*. 1993;27:129-148.
24. Christensen WF, et al. A comparison of Type I error rates and power levels for seven solutions to the multivariate Behrens-Fisher problem. *Commun Stat Simul Comput*. 1997;26:1251–11273.
25. Oyeyemi GM, et al. Comparison of the invariant solution to the multivariate Behrens-Fisher problem. *J Nigerian Assoc Math Phy*. 2016;34:199-208.
26. Hotelling H. The generalization of student's ratio. *Annals of Mathematical Statistics*. 1931;2:360-378.