ON SOME SOLUTIONS OF THE MULTIVARIATE BEHRENS FISHER PROBLEM

[1]Gulumbe SU, [2]Shehu B, [2]Dahiru S

Professor of Statistics
[1]Department of Mathematics
Usmanu Danfodiyo University Sokoto
Email: ugulumbe@gmail.com

[2]Department of Mathematics
Kebbi State University of Science and Technology, Aliero
Email: bsbshehu0@gmail.com

[3]Department of Mathematics
Kebbi State University of Science and Technology, Aliero
Email: shamsssdeeny@gmail.com

Abstract

Multivariate Behrens-Fisher Problem is a problem that deals with testing the equality of two means from multivariate normal distribution when the covariance matrices are unequal and unknown. However, there is no single procedure served as a better performing solution to this problem, Adebayo (2018). In this study effort is made in selecting five different existing procedures and examined their power and rate to which they control type I error using a different setting and conditions observed from previous studies. To overcome this problem a code was designed via R Statistical Software, to simulate random normal data and independently run 1000 times using MASS package in other to estimate the power and rate at which each procedure control type I error. The simulation result depicts that, in a setting when variance covariance matrices $S_1 > S_2$ associated with a sample sizes ($n_1 > n_2$) in Table 4.1, 4.2, 4.5, and 4.6, shows that, Adebayos' procedure performed better but at a sample sizes ($n_1 = n_2$ and $n_1 < n_2$) Hotelling T^2 is recommended in terms of power. For type I error rate where robustness and nominal level matters we found that under some settings none of the procedure maintained nominal level as revealed in Table 4.11 and 4.15. The results presented in Table 4.9 to 4.16 shows that when nominal level matters Krishnamoorthy came first, followed by Adebayos', Yaos', Johansons' then Hotelling T^2 were recommended in the sequentially under the settings used in this study.

Keywords: *Behrens-Fisher, Power, Type I error, Hotelling's T^2, Yoa, Johansen, Krishnamoorthy, Adebayo*

INTRODUCTION

A statistician by name Walter Behrens in (1929) propose the problem of testing the equality of two population means without assuming equal population variances (Yao 1965 and Wang 1971). Six years later another statistician by name Ronald Fisher (1935) noted that Behrens' solution could be derived using Fishers' concept of fiducial distributions. This makes the problem to be known as Behrens-Fisher problem. Many scholars have made efforts extending univariate form of this problem to its multivariate procedures. However, each procedure has its own good and weak part. One procedure may be good under a particular condition and become weak or moderately perform under another condition and this motivates the researcher to designed different conditions under which all selected procedures will be tested and judge according to their performances.

Adebayo and Oyeyemi (2018) developed an alternative procedure to multivariate Behrens–Fisher problem by using approximate degree of freedom test which was adopted from Satterthwaite univariate procedure. They discovered that in Table 2, 4 and 6 when sample size are very small (20, 10) proposed procedure is not the 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 performed better when sample size is very small (20, 10) followed by Yao, Krishnamoorthy and proposed procedure in term of power of the test in all the scenarios considered. In terms of type I error rate, proposed procedure competed favorably well with the other procedures selected for their study. Yao, Krishnamoorthy, Johanson, Nel and Van der Merwe and the proposed procedures are fluctuating (inflated and deflated) around the nominal level while Hotellling T square and Yanagihara are below the nominal level.

COMPUTATIONAL PROCEDURES

Consider two ρ-variate normal populations $N(\mu_1, \sum_1)$ and $N(\mu_2, \sum_2)$ where μ_1 and μ_2 are unknown p x 1 vectors and $\sum_1$ and $\sum_2$ are unknown $p \times p$ positive definite matrices.

Let $X\alpha_1 \sim N(\mu_1, \sum_1)$, $\alpha = 1, 2, \ldots, n_1$, and $X\alpha_2 \sim N(\mu_2, \sum_2)$, $\alpha = 1, 2, \ldots, n_2$ denote random samples from these two populations, respectively. We are interested in the testing problem.

$$Ho: \mu_1 = \mu_2 \text{ against } H_1 : \mu_1 \neq \mu_2 \quad (1)$$

For $i = 1, 2$ Let

$$\bar{X}_i = \frac{1}{n_i} \sum_{\alpha=1}^{ni} X_{\alpha i}, \quad (2)$$

$$A_i = \sum_{\alpha=1}^{ni} (X_{\alpha i} - \bar{X}_i)(X_{\alpha i} - \bar{X}_i)' \quad (3)$$

$$S_i = A_i/(ni - 1), \; i = 1,2 \quad (4)$$

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:

$$\bar{X}_i \sim N\left(\mu_i, \frac{\sum i}{n^i}\right). \text{ and } A_i \sim W_p(n_i - 1, \textstyle\sum_i), \; i = 1,2 \quad (5)$$

Where $Wp(r, \sum)$ denotes the p-dimensional Wishart distribution with df = r and scale matrix $\sum$. $\bar{X}_i$ *and* S_i *are* the sample mean vector and sample variance covariance of the i^{th} sample.

The Hotelling T-square (1930): The two sample statistic is given by

$T^2 = \frac{n_1 n_2}{n_1+n_2}(\bar{X}_1 - \bar{X}_2)`\tilde{S}_i^{-1}(\bar{X}_1 - \bar{X}_2)$
where

$$\tilde{S}_{pl} = \frac{1}{n_1 + n_2 - 2}[(n_1 - 1)S_1(n_2 - 1)S_2]$$

$$v_{hotel} = \frac{(N-p+1)}{p(N-2)}T^2 \quad (6)$$

Yao (1965) procedure:

The procedure is based on $T^2 \sim (vp/(v\text{-}p\text{+}1)\ F_{pv-p+1}$ with the degrees of freedom v given by:

$$v = \left[\frac{1}{n_1}\left(\frac{\bar{X}\prime_d \tilde{S}^{-1}\tilde{S}_1\tilde{S}^{-1}\bar{X}_d}{\bar{X}_d\tilde{S}^{-1}\bar{X}_d}\right) + \frac{1}{n_2}\left(\frac{\bar{X}\prime_d \tilde{S}^{-1}\tilde{S}_2\tilde{S}^{-1}\bar{X}_d}{\bar{X}_d\tilde{S}^{-1}\bar{X}_d}\right)\right]$$

$$T_{Yao} = \frac{(v-p+1)T^2}{vp} \quad (7)$$

Johansen (1980) procedure:

The procedure is based on $T^2 \sim qF_{p,v}$

where

$q = p + 2D - 6D/\,[p(p\text{-}1)\text{+}2]$,
and $v_{joh}=p(p+2)/3D$

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

$$T_{Joh} = \frac{T^2}{q} \quad (8)$$

Krishnamoorthy and Yu (2004) procedure:

The procedure is based on $T^2 \sim (v_{ky}p/(v\text{-}p\text{+}1)\ F_{pv-p+1}$ with the *d.f.v* defined by

$$v_{ky}p = p + p^2/C(\tilde{S}_1, \tilde{S}_2)$$

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

$$+ {}^1/_{n_2}\left\{tr\left[\left(\tilde{S}_2, \tilde{S}^{-1}\right)^2\right] + \left[tr\left(\tilde{S}_2, \tilde{S}^{-1}\right)\right]^2\right\}$$

$$T_{krish} = \frac{(v_{ky}p - p + 1)T^2}{v_{ky}p} \qquad (9)$$

Adebayo's (2019) procedure:

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

and $T^2 \sim \left(\frac{f_{Adeb} \times p}{(f_{Adeb} - p + 1)}\right) F_{p, f_{Adeb} - p + 1}$ approximately

$$T_{Adebayo} = \frac{(f - p + 1)T^2}{f \times p} \qquad (10)$$

Statistical significance is assessed by comparing the T_{Adeb} statistic to its critical value $F_{\propto}\,(p, f_{Adeb} - p + 1)$, that is, a critical value from the F distribution with p and degrees of freedom, $f_{Adeb} - p + 1$.

To compute the above procedures, each method is encoded in R software and the program designed in sequential order analysing either power or type I error rate depending on the mean setting. The codes were designed with the ability of generation multivariate normal data randomly from package called MASS. For each run the program will execute the process 1000 times out of which the average number null hypothesis is rejected will be considered as power or type I error rate depending on the mean setting.

NUMERICAL SIMULATION AND DISCUSSION OF RESULTS

The analysis is sectionalise in to two different categories; power of a test and level at which each procedure control type I error rate at a giving condition. Many factors were also considered such as number of variables (P), Sample sizes, Level of significance (alpha, α) and variance covariance matrices.

For P, the number of variables we consider two levels which are 2 and 3, for alpha we consider 0.01 and 0.05, for sample sizes we consider $(n_1 = n_2, n_1 < n_2, n_1 > n_2)$ and for covariance matrices we consider when $(S_1 > S_2 \; and \; S_1 < S_2)$ respectively.

4.1 Power of a test

4.1.1 Power of a test when S₁ > S₂ and P=2 at both level of significances 0.01 and 0.05

Table 4.1: Power of the Test

P=2	$S_1 > S_2$	Combination of different sample size					$\alpha = 0.01$
$\bar{X}_1 = (12 \quad 10)$	Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (3.7 \quad 2.4)$	TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 280 & 98 \\ 98 & 280 \end{pmatrix}$	Hotel	**0.1198**	**0.3566**	0.0984	0.6779	**0.1949**	**0.6120**
	Adebayo	0.1137	0.3511	**0.1899**	**0.8982**	0.0827	0.4243
$S_2\begin{pmatrix} 35 & 20 \\ 20 & 35 \end{pmatrix}$	Krish	0.1121	0.3505	0.1884	0.8977	0.0860	0.4259
	Yao	0.1137	0.3511	0.1887	0.8978	0.0886	0.4264
$(\bar{X}_1 - \bar{X}_2) = (8.3 \quad 7.1)$	Johan	0.1121	0.3490	0.1875	0.8974	0.0880	0.4244

Table 4.2: Power of the Test

P=2	$S_1 > S_2$	Combination of different sample size					$\alpha = 0.05$
$\bar{X}_1 = (12 \quad 10)$	Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (3.7 \quad 2.4)$	TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 280 & 98 \\ 98 & 280 \end{pmatrix}$	Hotel	**0.2739**	**0.5620**	0.2458	0.8441	**0.3535**	**0.7840**
	Adebayo	0.2711	0.5602	**0.3831**	**0.9629**	0.2192	0.6398
$S_2\begin{pmatrix} 35 & 20 \\ 20 & 35 \end{pmatrix}$	Krish	0.2697	0.5598	0.3815	0.9627	0.2231	0.6408
	Yao	0.2711	0.5602	0.3818	0.9628	0.2261	0.6410
$(\bar{X}_1 - \bar{X}_2) = (8.3 \quad 7.1$	Johan	0.2710	0.5594	0.3815	0.9627	0.2272	0.6404

Simulation result of power from Table 4.1 and 4.2 depicts that when $S_1 > S_2$ and P=2 at both level of significances 0.01 and 0.05, for a sample sizes ($n_1 = n_2$ and $n_1 < n_2$) under the following settings (15, 15), (50, 50), (10, 20) and (60, 100) Hotelling T^2 is recommended but when ($n_1 > n_2$) associated with (30, 15) and (200, 100) Adebayos' is recommended for practical use.

4.1.2 Power of a test when $S_1 < S_2$ and P=2 at both level of significances 0.01 and 0.05

Table 4.3: Power of the Test

P=2	$S_1 < S_2$	Combination of different sample size						$\alpha = 0.01$
$\bar{X}_1 = (3.3 \quad 2.5)$		Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (8.6 \quad 5.1)$		TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix}3.2 & 2.7\\2.7 & 3.2\end{pmatrix}$		Hotel	0.7558	0.9995	**0.9393**	1	0.6352	**1**
		Adebayo	**0.7564**	**0.9996**	0.7873	1	0.7920	**1**
$S_2\begin{pmatrix}9.5 & 4\\4 & 9.5\end{pmatrix}$		Krish	0.7476	0.9995	0.7980	1	**0.8082**	**1**
		Yao	**0.7564**	**0.9996**	0.8060	1	**0.8082**	**1**
$(\bar{X}_1 - \bar{X}_2) = (-5.3 \quad -2.6)$		Johan	0.7477	0.9995	0.7977	1	0.8093	**1**

Power of a test in Table 4.3, when $S_1 < S_2$ at significance level 0.01 revealed that Adebayos' and Yaos' performed better with equal power when sample size are equal (n_1=n_2). For a sample sizes (60, 100) and (200, 100) all procedures are recommended and produces equal power. Krishnamoorthy and Yao yield a better result at a sample size (10, 20) having equal power but for a sample size (30, 50) only Hotelling T^2 is recommended.

Table 4.4: Power of the Test

P=2	$S_1 < S_2$	Combination of different sample size						$\alpha = 0.05$
$\bar{X}_1 = (3.3 \quad 2.5)$		Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (8.6 \quad 5.1)$		TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix}3.2 & 2.7\\2.7 & 3.2\end{pmatrix}$		Hotel	0.8916	**0.9999**	**0.9811**	1	0.8121	0.8121
		Adebayo	**0.8960**	**0.9999**	0.9208	1	0.9115	0.9115
$S_2\begin{pmatrix}9.5 & 4\\4 & 9.5\end{pmatrix}$		Krish	0.8927	**0.9999**	0.9247	1	*0.9176*	*0.9176*
		Yao	**0.8960**	**0.9999**	0.9271	1	*0.9176*	*0.9176*
$(\bar{X}_1 - \bar{X}_2) = (-5.3 \quad -2.6)$		Johan	0.8940	**0.9999**	0.9257	1	**0.9190**	**0.9190**

Power of a test in Table 4.3 when $S_1 < S_2$ at significance level 0.05 revealed that Adebayos' and Yaos' performed better with equal power when sample size is (15, 15). For a sample sizes (60, 100) and (10, 20) Johanson produce a better power. For a sample sizes (50, 50) and (200, 100) all procedures have equal power and were good to use for practical perspectives.

4.1.3 Power of a test when $S_1 > S_2$ and P=3 at both level of significances 0.01 and 0.05

Table 4.5: Power of the Test

P=3 $S_1 > S_2$	Combination of different sample size						$\alpha = 0.01$
$\bar{X}_1 = (30\ \ 24\ \ 50\,)$	Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (15\ \ 14\ \ 29\,)$	TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1 = \begin{pmatrix} 800 & 500 & 300 \\ 500 & 800 & 250 \\ 300 & 250 & 800 \end{pmatrix}$	Hotel	**0.1432**	**0.5207**	0.1300	0.8884	**0.2185**	**0.8274**
	Adebayo	0.1301	0.5066	**0.2597**	**0.9894**	0.0812	0.6127
$S_2 \begin{pmatrix} 100 & 50 & 25 \\ 50 & 100 & 20 \\ 25 & 20 & 100 \end{pmatrix}$	Krish	0.1288	0.5080	0.2584	0.9893	0.0859	0.6175
	Yao	0.1301	0.5066	0.2573	0.9893	0.0928	0.6163
$(\bar{X}_1 - \bar{X}_2) = (15\ \ 10\ \ 2\,1)$	Johan	0.1235	0.4973	0.2495	0.9890	0.0863	0.6070

Power of a test presented in Table 4.5 depicts that, for a given sample sizes ($n_1=n_2$ and $n_1<n_2$) Hotelling T^2 has the highest power but when ($n_1 > n_2$) Adebayos' procedure has the highest power.

Table 4.6: Power of the Test

P=3 $S_1 > S_2$	Combination of different sample size						$\alpha = 0.05$
$\bar{X}_1 = (30\ \ 24\ \ 50\,)$	Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (15\ \ 14\ \ 29\,)$	TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1 = \begin{pmatrix} 800 & 500 & 300 \\ 500 & 800 & 250 \\ 300 & 250 & 800 \end{pmatrix}$	Hotel	**0.3197**	**0.7125**	0.3005	0.9622	**0.4011**	**0.9208**
	Adebayo	0.3078	0.7053	**0.4761**	**0.9978**	0.2231	0.7955
$S_2 \begin{pmatrix} 100 & 50 & 25 \\ 50 & 100 & 20 \\ 25 & 20 & 100 \end{pmatrix}$	Krish	0.3068	0.7065	0.4749	**0.9978**	0.2311	0.7980
	Yao	0.3078	0.7053	0.4736	**0.9978**	0.2383	0.7973
$(\bar{X}_1 - \bar{X}_2) = (15\ \ 10\ \ 2\,1)$	Johan	0.3001	0.6996	0.4667	0.9977	0.2315	0.7924

Power of a test presented in Table 4.6 depicts that, for a given sample sizes ($n_1=n_2$ and $n_1<n_2$) Hotelling T^2 has the highest power. When ($n_1 > n_2$) for a sample size (30, 15) Adebayos' perform better but when sample size increase to (200, 100) Adebayo, Krishnamoorthy and Yao perform good having equal power.

4.1.4 Power of a test when $S_1 < S_2$ and P=3 at both level of significances 0.01 and 0.05

Table 4.7: Power of the Test

P=3 $S_1 < S_2$	Combination of different sample size						α = 0.01
$\bar{X}_1 = (10\quad 15\quad 20)$	Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (5\quad 10\quad 15)$	TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 5.5 & 3.2 & 2.5 \\ 3.2 & 5.5 & 2.1 \\ 2.5 & 2.1 & 5.5 \end{pmatrix}$	Hotel	**0.0942**	**0.2900**	**0.2103**	**0.8680**	0.0514	0.3455
	Adebayo	0.0813	0.2754	0.0733	0.5694	**0.1054**	0.5638
$S_2\begin{pmatrix} 95 & 80 & 70 \\ 80 & 95 & 60 \\ 70 & 60 & 95 \end{pmatrix}$	Krish	0.0822	0.2787	0.0767	0.5723	0.1038	**0.5645**
	Yao	0.0813	0.2754	0.0766	0.5705	0.1019	0.5610
$(\bar{X}_1 - \bar{X}_2) = (5\quad 5\quad 5)$	Johan	0.0791	0.2712	0.0740	0.5653	0.0995	0.5585

Table 4.8: Power of the Test

P=3 $S_1 < S_2$	Combination of different sample size						α = 0.05
$\bar{X}_1 = (10\quad 15\quad 20)$	Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (5\quad 10\quad 15)$	TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 5.5 & 3.2 & 2.5 \\ 3.2 & 5.5 & 2.1 \\ 2.5 & 2.1 & 5.5 \end{pmatrix}$	Hotel	**0.2269**	**0.4953**	**0.4057**	**0.9442**	0.1563	0.5612
	Adebayo	0.2118	0.4846	0.2173	0.7631	**0.2559**	0.7539
$S_2\begin{pmatrix} 95 & 80 & 70 \\ 80 & 95 & 60 \\ 70 & 60 & 95 \end{pmatrix}$	Krish	0.2135	0.4875	0.2229	0.7648	0.2542	**0.7543**
	Yao	0.2118	0.4846	0.2222	0.7637	0.2515	0.7523
$(\bar{X}_1 - \bar{X}_2) = (5\quad 5\quad 5)$	Johan	0.2091	0.4810	0.2185	0.7607	0.2485	0.7508

Power of a test presented in Table 4.7 and 4.8 Hotelling T^2 has the highest power when sample sizes are equal ($n_1 = n_2$ and $n_1 > n_2$). But when ($n_1 < n_2$) associated with a small sample size (10, 20) Adebayo perform better and when the sample size increases to (60, 100) Krishnamoorthy is recommend.

4.2 Type I Error Rate

4.2.1 Type I Error when $S_1 > S_2$ and P=2 at both level of significance 0.01 and 0.05

Table 4.9: Type I error

P=2	$S_1 > S_2$	Combination of different sample size						$\alpha = 0.01$
$\bar{X}_1 = (10 \quad 10)$		Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (10 \quad 10)$		TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 280 & 98 \\ 98 & 280 \end{pmatrix}$		Hotel	**0.01**	0.004	0.001	0.00*	0.067*	0.057*
		Adebayo	0.008	0.004	0.012	0.011	0.009	0.016
$S_2\begin{pmatrix} 35 & 20 \\ 20 & 35 \end{pmatrix}$		Krish	0.008	0.004	0.011	0.011	**0.01**	0.016
		Yao	0.008	0.004	0.011	0.011	0.011	0.018*
$(\bar{X}_1 - \bar{X}_2) = (0 \quad 0)$		Johan	0.008	0.004	0.011	0.011	**0.01**	0.015

Table 4.9, depicts that, Hotelling T^2 (15, 15), Krishnamoorthy and Johanson (10, 20) have maintained the nominal level. Hotelling T^2 (200, 100), (10, 20), (60, 100) and Yao (60, 100) are not robust. There is inflation in type I error rate, when sample sizes are ($n_1 = n_2$) and when sample size is ($n_1 > n_2$) all procedures fluctuate within the range except Hotelling T^2 which deflated.

Table 4.10: Type I error

P=2	$S_1 > S_2$	Combination of different sample size						$\alpha = 0.05$
$\bar{X}_1 = (10 \quad 10)$		Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (10 \quad 10)$		TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 280 & 98 \\ 98 & 280 \end{pmatrix}$		Hotel	0.052	0.042	0.004*	0.003*	0.172*	0.127*
		Adebayo	0.053	0.042	0.043	0.043	0.047	0.049
$S_2\begin{pmatrix} 35 & 20 \\ 20 & 35 \end{pmatrix}$		Krish	0.052	0.042	0.043	0.043	0.051	**0.05**
		Yao	0.053	0.042	0.043	0.043	0.054	0.049
$(\bar{X}_1 - \bar{X}_2) = (0 \quad 0)$		Johan	0.053	0.042	0.043	0.043	0.054	0.049

Table 4.10, depicts that, Krishnamoorthy (60, 100) have maintained the nominal level. Hotelling T^2 (30, 15), (200, 100), (10, 20) and (60, 100) are not robust.

4.2.2 Type I Error when $S_1 < S_2$ and P=2 at both level of significance (0.01 and 0.05)

Table 4.11: Type I error

P=2	$S_1 < S_2$	Combination of different sample size					$\alpha = 0.01$
$\bar{X}_1 = (3.5 \quad 3.5)$		Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$
$\bar{X}_2 = (3.5 \quad 3.5)$		TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1 = \begin{pmatrix} 3.2 & 2.7 \\ 2.7 & 3.2 \end{pmatrix}$		Hotel	0.015	0.016	0.042*	0.064*	0.004	0.005
		Adebayo	0.018*	0.016	0.007	0.007	0.009	0.017*
$S_2 \begin{pmatrix} 9.5 & 4 \\ 4 & 9.5 \end{pmatrix}$		Krish	0.015	0.016	0.009	0.007	0.009	0.017*
		Yao	0.018*	0.016	0.009	0.007	0.009	0.017*
$(\bar{X}_1 - \bar{X}_2) = (0 \quad 0)$		Johan	0.015	0.016	0.009	0.007	0.009	0.017*

Table 4.11, depicts that, none of the procedures maintained nominal level also at a sample size (50, 50) all procedures inflated with equal rate. For a sample sizes (200, 100) and (10, 20) all procedures deflated. For a sample size (60, 100) all procedures are not robust except Hotelling T^2 which reverse is the case in the two mention scenarios. Adebayo and Yao at a sample size (15, 15) are not robust.

Table 4.12: Type I error

P=2	$S_1 < S_2$	Combination of different sample size					$\alpha = 0.05$
$\bar{X}_1 = (3.5 \quad 3.5)$		Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$
$\bar{X}_2 = (3.5 \quad 3.5)$		TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1 = \begin{pmatrix} 3.2 & 2.7 \\ 2.7 & 3.2 \end{pmatrix}$		Hotel	0.042	0.04	0.148*	0.14*	0.01*	0.022*
		Adebayo	0.041	0.04	0.058	0.039	0.051	0.052
$S_2 \begin{pmatrix} 9.5 & 4 \\ 4 & 9.5 \end{pmatrix}$		Krish	0.041	0.041	0.062	0.039	**0.05**	0.052
		Yao	0.041	0.04	0.066*	0.04	0.048	0.052
$(\bar{X}_1 - \bar{X}_2) = (0 \quad 0)$		Johan	0.043	0.041	0.062	0.039	**0.05**	0.052

Table 4.12, depicts that, Krishnamoorthy and Yao have maintained the nominal level. Hotelling T^2 at sample sizes (30, 15), (200, 100), (10, 20), 60, (100) and Yao (30, 15) are not robust. In the remaining settings all procedures either deflated or inflated.

4.2.3 Type I Error when $S_1 > S_2$ and P=3 at both level of significance (0.01 and 0.05)

Table 4.13: Type I error

P=3	$S_1 > S_2$	Combination of different sample size					α = 0.01
$\bar{X}_1 = (10\ \ 15\ \ 20)$	Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (10\ \ 15\ \ 20)$	TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 800 & 500 & 300 \\ 500 & 800 & 250 \\ 300 & 250 & 800 \end{pmatrix}$	Hotel	0.011	**0.01**	0*	0*	0.075*	0.052*
	Adebayo	**0.01**	**0.01**	0.009	0.005	0.011	0.012
$S_2\begin{pmatrix} 100 & 50 & 25 \\ 50 & 100 & 20 \\ 25 & 20 & 100 \end{pmatrix}$	Krish	**0.01**	**0.01**	0.009	0.005	**0.01**	0.012
	Yao	**0.01**	**0.01**	0.009	0.005	0.016	0.014
$(\bar{X}_1 - \bar{X}_2) = (0\ \ 0\ \ 0)$	Johan	**0.01**	0.008	0.008	0.005	0.009	0.012

Table 4.13, depicts that, all procedures maintained nominal level except Hotelling T^2 and Johanson in one condition each also at a sample size (15, 15) and (50, 50) also at a sample size (10, 20) Krisnamoorthy maintained nominal level. Hotelling T^2 at a sample sizes (30, 15), (200, 100), (10, 20) and (60, 100) is not robust.

Table 4.14: Type I error

P=3	$S_1 > S_2$	Combination of different sample size					α = 0.05
$\bar{X}_1 = (10\ \ 15\ \ 20)$	Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (10\ \ 15\ \ 20)$	TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 800 & 500 & 300 \\ 500 & 800 & 250 \\ 300 & 250 & 800 \end{pmatrix}$	Hotel	0.051	0.066*	0.003*	0.003*	0.191*	0.143*
	Adebayo	0.052	0.065	0.045	0.057	**0.05**	0.06
$S_2\begin{pmatrix} 100 & 50 & 25 \\ 50 & 100 & 20 \\ 25 & 20 & 100 \end{pmatrix}$	Krish	0.048	0.065	0.046	0.057	0.052	0.06
	Yao	0.052	0.065	0.046	0.057	0.066*	0.06
$(\bar{X}_1 - \bar{X}_2) = (0\ \ 0\ \ 0)$	Johan	0.047	0.059	0.046	0.057	0.052	0.058

Table 4.14, depicts that, at a sample size (10, 20) Adebayos' maintained nominal level. Hotelling T^2 at a sample sizes (50, 50), (30, 15), (200, 100), (10, 20), (60, 100) and Yao (10, 20) were not robust. In the remaining settings all procedures either deflated or inflated.

4.2.4 Type I Error when $S_1 < S_2$ and P=3 at both level of significance (0.01 and 0.05)

Table 4.14: Type I error

P=3	$S_1 < S_2$	Combination of different sample size						$\alpha = 0.01$
$\bar{X}_1 = (10\ \ 10\ \ 10)$		Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (10\ \ 10\ \ 10)$		TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 5.5 & 3.2 & 2.5 \\ 3.2 & 5.5 & 2.1 \\ 2.5 & 2.1 & 5.5 \end{pmatrix}$		Hotel	0.021*	0.004	0.092*	0.098*	0.003*	0.00*
		Adebayo	0.02*	0.004	0.008	0.008	0.011	0.013
$S_2\begin{pmatrix} 95 & 80 & 70 \\ 80 & 95 & 60 \\ 70 & 60 & 95 \end{pmatrix}$		Krish	0.013	0.003*	0.009	0.008	0.011	0.013
		Yao	0.02*	0.004	0.011	0.009	0.012	0.013
$(\bar{X}_1 - \bar{X}_2) = (0\ \ 0\ \ 0)$		Johan	0.011	0.003*	0.008	0.008	0.011	0.013

Table 4.14, depicts that, none of the procedure maintained nominal level, and only Johanson robust at a sample size (15, 15) but when sample size is (50, 50) Johansons and Krishnamoorthy were not robust. Hotelling T^2 is not robust in all scenarios except when sample size are (n_1=n_2).

Table 4.15: Type I error

P=3	$S_1 < S_2$	Combination of different sample size						$\alpha = 0.05$
$\bar{X}_1 = (10\ \ 10\ \ 10)$		Sample size	$n_1 = n_2$		$n_1 > n_2$		$n_1 < n_2$	
$\bar{X}_2 = (10\ \ 10\ \ 10)$		TEST	15, 15	50, 50	30, 15	200, 100	10, 20	60, 100
$S_1=\begin{pmatrix} 5.5 & 3.2 & 2.5 \\ 3.2 & 5.5 & 2.1 \\ 2.5 & 2.1 & 5.5 \end{pmatrix}$		Hotel	0.055	0.042	0.215*	0.204*	0.011*	0.011*
		Adebayo	**0.05**	0.041	0.052	**0.05**	0.043	**0.05**
$S_2\begin{pmatrix} 95 & 80 & 70 \\ 80 & 95 & 60 \\ 70 & 60 & 95 \end{pmatrix}$		Krish	0.048	0.041	0.055	**0.05**	0.041	0.049
		Yao	**0.05**	0.041	0.057	**0.05**	0.042	0.049
$(\bar{X}_1 - \bar{X}_2) = (0\ \ 0\ \ 0)$		Johan	0.044	0.041	0.052	0.049	0.041	0.047

Table 4.15, depicts that, at a sample sizes (15, 15) Adebayo, Yao and at a sample size (200, 100) Adebayo, Krishnamoorthy and Yao have maintained nominal level; also Adebayo (60, 100) maintained nominal level. Hotelling T^2 is not robust in all scenarios except when sample size are (n_1=n_2). In the remaining settings all procedures either deflated or inflated within the nominal range.

CONCLUSION

In a setting when variance covariance matrices $S_1 > S_2$ associated with a sample sizes ($n_1 > n_2$) in respective of the level of significance and P (number of variable) as depicts in Table 4.1, 4.2, 4.5, and 4.6, we concluded that, Adebayos' procedure performed better but at a sample sizes ($n_1 = n_2$ and $n_1 < n_2$) Hotelling T^2 is recommended. In a setting when variance covariance matrices $S_1 < S_2$, many procedures tend to have equal performance as depicts in Table 4.3 and 4.4; However, when $\alpha = 0.01$ under the following sample sizes (200, 100) and (60, 100) all procedure have equal performance, also when $\alpha = 0.05$ at sample sizes (50, 50) and (200, 100) all procedures have equal power; Furthermore, Adebayo and Yao at a sample size (15, 15) have equal power, Krishnamoorthy and Yao at a sample sizes ($n_1 < n_2$) have equally power. For a setting when p=3 associated with $S_1 < S_2$, at a sample sizes ($n_1 = n_2$ and $n_1 > n_2$) as depicts in 4.7 and 4.8 the use of Hotelling T^2 is recommended and when ($n_1 < n_2$) associated with a small sample size (10, 20) Adebayo performed better but sample size increases to (60, 100) Krishnamoorthy is recommend.

For type I error rate where robustness and nominal level matters we found that under some settings none of the procedure maintained the nominal level as revealed in Table 4.11 and 4.15. The results presented from Table 4.9 to 4.16 depicts the number of times each procedure maintained a nominal level: Krishnamoorthy (7), Adebayo (6), Yao (4), Johanson (3) and Hotelling T^2 (2) times respectively. Therefore, when nominal level matters Krishnamoorthy came first, followed by Adebayos', Yaos', Johansons' then Hotelling T^2 were recommended in the sequential order mention.

FURTHER WORK

We recommend that the subsequent researchers should design more complicated settings and conditions in which more procedures could be investigated. Also need for better extension of univariate form of this problem to multivariate form is seriously required, so that one may be able to achieve a better performing procedure under all settings.

Furthermore, a high value of ($P >= 4$) and different significant level other than the ones considered in this research may be considered.

References

1. Adebayo P.O and G.M. Oyeyemi (2018). "An Alternative Solution to Hotelling T Square under the Heteroscedasticity of the Dispersion Matrix". Department of Statistics, University of Ilorin, Ilorin, Nigeria. The Pacific Journal of Science and Technology. http://www.akamaiuniversity.us/PJST.htm.

2. Behrens W. H. V. (1929). "Ein Beitrang Zur Fehlerberchnung bei wenigen beobachtungen". Landwirtsch Jahrbucher 68, 807–837.

3. Fisher, R. A. (1935). "The fiducial argument in statistical inference". Annals of Human

Genetics 6, 391–398.

4. Hotelling, H. (1931). “The generalization of student’s ratio”. Annals of Mathematical Statistics. 2:360-378.

5. Johansen, S. (1980). “The welch-james approximation to the distribution of the residual sum of squares in a weighted linear regression”. Biometrika 67, 85–92.

6. Krishnamoorthy, K. and J. Yu. (2004). “Modified Nel and Van der Merwe Test for the Multivariate Behrens – Fisher Problem”. Statistica and Probability Letters. 66:161 – 169.

7. Yao, Y. (1965). “An Approximation Degrees of Freedom Solution to the Multivariate Behrens – Fisher Problem”. Biometrika. 52:139 – 147.