ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)

FURTHER SOLUTIONS OF THE MULTIVARIATE BEHRENS FISHER PROBLEM

G. I. Onwuka[1], B. Shehu[2], and S. U. Gulumbe[3]

Dr. of Statistics
[1]Department of Mathematics
Kebbi State University of Science and Technology, Aliero

[2]Department of Mathematics
Kebbi State University of Science and Technology, Aliero
.
Professor of Statistics
[3]Department of Mathematics
Usmanu Danfodiyo University Sokoto

Corresponding Author Email: bsbshehu0@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. In this study efforts were made in selecting four different existing procedures and examined their power and rate to which they control type I error using different setting and conditions designed in the study. To overcome this problem a code was designed via R Statistical Software and simulate random normal data which independently run 1000 times, using MASS package in order to estimate the power and rate at which each procedure control type I error rate. In the simulation result we discovered that some of these existing procedures have equal and highest power in some certain settings like Yao and Adebayo, Johansen and Yao, Krishnamoorthy and Adebayo, Yao and Krishnamoorthy but when P-variables is increase we also found that these procedures with equal power varies significantly, where as some procedures' power decrease while some increases in 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 and some procedures lie outside the interval and considered non-robust. Yao and Adebayos were found good when P=2 and sample size $n_1 > n_2$, it is discovered that at a sample size (300, 200) all procedures attained the nominal level.

Keywords: *Multivariate Behrens-Fisher, Johansen, Yao, Krishnamoorthy and Adebayos' procedures, Power of Test, Error Rate.*

INTRODUCTION

1.0 Background of the Study

The well-known multivariate Behrens-Fisher problem is a problem which deals with testing the equality of two normal mean vectors under heteroscedasticity of dispersion matrices (Junyong Park, 2009). This problem is applicable when testing the equality of two means from multivariate normal distribution when the covariance matrices are unequal and unknown. This problem is a generalization of the univariate Behrens-Fisher problem; perhaps it inherits all of the difficulties that arise in the univariate problem.

ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)

In 1929 Walter Behrens instigated a problem that have driven the world of statistician into different researches, six years later in 1935, Ronald Fisher used a concept of his fiducial distribution and able to succeeded in some aspect of this problem, since then the problem was generally known as Behrens Fisher Problem. This problem is a problem that deals with testing the equality of two population means without assuming equal population variances (Yao 1965 and Wang 1971). In its univariate concepts there were enormous scholars who lay hands to overcome this problem such as Kim and Cohen (1998), Welch (1947) and Satterthwaite (1946) among others. Efforts were made by numbers of researchers extending univariate form of this problem to multivariate eversion, such as Yao 1965, Krishnamoorthy (2004), Johanson (1980), *Algina et al.* (1991), Adebayo and Oyeyemi (2018), more recently study Gulumbe *et al.* (2021) have designed some complicated settings and conditions under which the robustness of some existing procedures under Multivariate Behrens Fisher Problems were investigated. In their work they found that Hotelling T^2 may compete under some setting when power matters with some existing procedures under Multivariate Behrens Fisher Problems.

It is universally accepted that, Behrens Fisher Problem do not have one procedure or method that could provide a general solution to all problems in the area, 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 researchers to designed different conditions under which all procedures discuss in the methodology will be tested and judge according to their performances

The aim of this study is to propose new settings and conditions under which the power and rate at which each procedure control Type I error will be investigated. It also put into consideration both old and newly extended procedures in the field of Multivariate Behrens Fisher Problem. However, in the result finding we discovered that under the settings used in the study some procedures tend to have equal power where as in some settings the powers varies. We also observed that with the increase of P-variables some of these procedures with equal power tend to varies, one may increase while the other one decreases. Therefore, we have enough and cleared evidence to say P-variable has some effect to power and robustness. Under the settings used we also found that, there were some conditions where none of the procedure attained a nominal-level but in most of the conditions designed one procedure or the other will exactly be at nominal level of , $\alpha = 0.01$, $\alpha = 0.025$, $\alpha = 0.05$, respectively.

This study is limited to the use of multivariate normal data generated based on two random samples using a command mvrnom found in a package MASS in R statistical software. The study is also limited to the use of alpha at three different significance level, $\alpha = 0.01$, $\alpha = 0.025$, $\alpha = 0.05$, and when p=2 and p=3 respectively. The used of $\alpha = 0.025$ and other settings were made as suggested by Gulumbe *et al.* (2021)

COMPUTATIONAL PROCEDURES

Consider two ρ-variate normal populations $N(\mu_1, \sum_1)$ and $N(\mu_2, \sum_2)$ where $\mu_1$andμ_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.

ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)

$Ho: \mu_1 = \mu_2$ against $H_1: \mu_1 \neq \mu_2$ (1)

For i = 1, 2 Let

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

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

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

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.

Yao (1965) procedure:

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

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

$$T_{Yao} = \frac{(v-p+1)T^2}{vp} \quad (6)$$

Johansen (1980) procedure:

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

where

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

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

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

$$T_{Joh} = \frac{T^2}{q} \quad (7)$$

Krishnamoorthy and Yu (2004) procedure:

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

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

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

3

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

$$T_{krish} = \frac{(v_{ky}p - p + 1)T^2}{v_{ky}p} \qquad (8)$$

<u>Adebayo's (2019) procedure:</u>

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

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

$$T_{Adebayo} = \frac{(f - p + 1)T^2}{f \times p} \qquad (9)$$

Statistical significance is assessed by comparing theT_{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 analyzing 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.

RESULTS AND DISCUSSION

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 four discussed approximate solution (Johanson, Yao, Krishnamoorthy and Adebayos' procedures). 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 considered in the simulation: the sample size, the number of variables p, variance co-variance matrices, mean vectors and significant levels.

<u>Power of the test</u>

The settings and conditions considered are when $S_1 > S_2$, P=2, P=3and $\alpha = 0.01$, 0.025 and 0.05. The power was tested using both small and large sample sizes under the following settings " $n_1 = n_2$, $n_1 > n_2$ and $n_1 < n_2$" respectively.

ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)

Table 1: Power of the test

P=2 $S_1 > S_2$		Combination of different sample size					$\alpha = 0.01$
$\bar{X}_1 = (20 \quad 30)$ $\bar{X}_2 = (15 \quad 20)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$\bar{X}_1 - \bar{X}_2 = (5 \quad 10)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1=\begin{pmatrix} 800 & 200 \\ 200 & 800 \end{pmatrix}$	John	0.06085	0.6894	0.11488	0.8432	0.07196	0.7177
	Yao	**0.06202**	**0.6906**	0.11636	0.8438	**0.07286**	**0.7187**
$S_2=\begin{pmatrix} 90 & 60 \\ 60 & 90 \end{pmatrix}$	Krish	0.05966	0.6900	0.11535	0.8435	0.07025	0.7184
	Adebayo	**0.06202**	**0.6906**	**0.11814**	**0.8442**	0.06830	0.7179

Table 2: Power of the test

P=2 $S_1 > S_2$		Combination of different sample size					$\alpha = 0.025$
$\bar{X}_1 = (20 \quad 30)$ $\bar{X}_2 = (15 \quad 20)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$\bar{X}_1 - \bar{X}_2 = (5 \quad 10)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1=\begin{pmatrix} 800 & 200 \\ 200 & 800 \end{pmatrix}$	John	0.1150	0.7717	0.1746	0.9076	**0.1158**	0.8028
	Yao	**0.1154**	**0.7724**	0.1756	0.9079	**0.1158**	**0.8033**
$S_2=\begin{pmatrix} 90 & 60 \\ 60 & 90 \end{pmatrix}$	Krish	0.1130	0.7720	0.1349	0.1748	0.1133	0.8031

Table 3: Power of the test

P=2 $S_1 > S_2$		Combination of different sample size					$\alpha = 0.05$
$\bar{X}_1 = (20 \quad 30)$ $\bar{X}_2 = (15 \quad 20)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$\bar{X}_1 - \bar{X}_2 = (5 \quad 10)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1=\begin{pmatrix} 800 & 200 \\ 200 & 800 \end{pmatrix}$	John	**0.1839**	0.8448	0.2476	0.9350	**0.1158**	0.8621
	Yao	0.1833	**0.8452**	0.2485	0.9351	**0.1158**	**0.8623**
$S_2=\begin{pmatrix} 90 & 60 \\ 60 & 90 \end{pmatrix}$	Krish	0.1810	0.8450	0.2474	0.9350	0.1133	0.8622

ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)

From the Table 1 and 2 Yao and Adebayos' procedure tend to have equal and highest power when we have equal sample sizes at both small and large sizes. In Table 1, 2 and 3 when the sample size of $n_1 > n_2$ Adebayos' procedure has the highest power. At a sample size (200, 500) Yaos' procedure has the highest power in all the scenarios. Table 2 and 3 has shown that Johanson and Yao have the highest and equal power.

Table 4: Power of the test

P=3 $S_1 > S_2$		Combination of different sample size					$\alpha = 0.01$
$\bar{X}_1 = (10\ 30\ 20)$ $\bar{X}_2 = (8\ 27\ 18)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$(\bar{X}_1 - \bar{X}_2) = (2\ 3\ 2)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1 = \begin{pmatrix} 900 & 700 & 500 \\ 700 & 900 & 300 \\ 500 & 300 & 900 \end{pmatrix}$	John	0.02939	0.06052	0.03516	0.06959	0.03094	0.06467
	Yao	**0.03027**	0.06079	0.03595	0.06980	**0.03312**	0.06507
$S_2 = \begin{pmatrix} 200 & 90 & 50 \\ 90 & 200 & 20 \\ 50 & 20 & 200 \end{pmatrix}$	Krish	0.02979	**0.06084**	**0.03603**	**0.06984**	0.03135	**0.06512**
	Adebayo	**0.03027**	0.06079	0.03553	**0.06984**	0.02949	0.06473

Table 5: Power of the test

P=3 $S_1 > S_2$		Combination of different sample size					$\alpha = 0.025$
$\bar{X}_1 = (10\ 30\ 20)$ $\bar{X}_2 = (8\ 27\ 18)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$(\bar{X}_1 - \bar{X}_2) = (2\ 3\ 2)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1 = \begin{pmatrix} 900 & 700 & 500 \\ 700 & 900 & 300 \\ 500 & 300 & 900 \end{pmatrix}$	John	0.06439	0.1127	0.07309	0.1262	0.06713	0.1136
	Yao	**0.06552**	0.1130	0.07424	0.1265	0.07123	**0.1142**
$S_2 = \begin{pmatrix} 200 & 90 & 50 \\ 90 & 200 & 20 \\ 50 & 20 & 200 \end{pmatrix}$	Krish	0.06509	**0.1131**	**0.07447**	**0.1266**	0.06783	**0.1142**
	Adebayo	**0.06552**	0.1130	0.07374	**0.1266**	0.06525	0.1137

Table 6: Power of the test

P=3 S₁ > S₂		Combination of different sample size					α = 0.05
$\bar{X}_1 = (10\ 30\ 20)$, $\bar{X}_2 = (8\ 27\ 18)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$(\bar{X}_1 - \bar{X}_2) = (2\ 3\ 2)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1 = \begin{pmatrix} 900 & 700 & 500 \\ 700 & 900 & 300 \\ 500 & 300 & 900 \end{pmatrix}$	John	0.1133	0.1721	0.1201	0.1883	0.1152	0.1751
	Yao	**0.1152**	0.1725	0.1216	0.1886	**0.1199**	**0.1757**
$S_2 = \begin{pmatrix} 200 & 90 & 50 \\ 90 & 200 & 20 \\ 50 & 20 & 200 \end{pmatrix}$	Krish	0.114	**0.1726**	**0.1218**	0.1886	0.1161	**0.1757**

When P variable increases to P=3 the power of Yao decreases and unable to equalize with Adebayo under the sample size is large and equal (200, 200). Adebayo also decreases when sample size is small and unequal (25, 15). But the power of Krishnamoorthy increases as seen in Table 4 and 5 when sample size is large and unequal (200, 500).

Type I Error Rate

The upper limit was estimated using $\hat{\alpha} = \alpha + 2\sqrt{\frac{\alpha\,(1-\alpha)}{N}}$ and lower limit was obtained using $\hat{\alpha} = \alpha - 2\sqrt{\frac{\alpha\,(1-\alpha)}{N}}$, N is the number of time the process runs in R and N= 1000.

Thus, when α = 0.05 the interval ranges from 0.036 to 0.064, α = 0.01 the interval ranges from 0.004 to 0.016 and α = 0.025 the interval ranges from 0.015 to 0.035. That means all values for the estimated error which are lower than or higher than the respective values under their corresponding significance level (α) will be considered as non-robust since they lie outside the interval and will be marked with a star (*).

The settings and conditions considered are when $S_1 > S_2$, P=2, P=3 and α = 0.01, 0.025 and 0.05 respectively. The estimation of type I error was carryout with small and large sample sizes under the following settings " $n_1 = n_2$, $n_1 > n_2$ and $n_1 < n_2$".

ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)

Table 7: Type I error rate

P=2	$S_1 > S_2$	Combination of different sample size					$\alpha = 0.01$
$\bar{X}_1 = (2 \quad 3)$, $\bar{X}_2 = (2 \quad 3)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$\bar{X}_1 - \bar{X}_2 = (5 \quad 10)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1 = \begin{pmatrix} 800 & 200 \\ 200 & 800 \end{pmatrix}$	John	0.014	0.006	0.012	0.012	0.009	0.013
	Yao	0.015	0.006	0.012	0.012	0.011	0.013
$S_2 = \begin{pmatrix} 90 & 60 \\ 60 & 90 \end{pmatrix}$	Krish	0.014	0.006	0.012	0.012	0.008	0.013

From Table 7, all procedures deflated when sample size is (200, 200) and (10, 20) except Yao which is close to nominal level. When sample sizes are $n_1 > n_2$ all procedures are close to nominal level. All procedures inflated when sample sizes are(10, 10)

Table 8: Type I error rate

P=2	$S_1 > S_2$	Combination of different sample size					$\alpha = 0.025$
$\bar{X}_1 = (20 \quad 30)$, $\bar{X}_2 = (15 \quad 20)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$\bar{X}_1 - \bar{X}_2 = (5 \quad 10)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1 = \begin{pmatrix} 800 & 200 \\ 200 & 800 \end{pmatrix}$	John	0.026	**0.025**	0.012	0.021	**0.025**	**0.025**
	Yao	0.026	**0.025**	0.012	0.021	0.026	**0.025**
$S_2 = \begin{pmatrix} 90 & 60 \\ 60 & 90 \end{pmatrix}$	Krish	**0.025**	**0.025**	0.012	0.021	**0.025**	**0.025**

From Table 8; When sample sizes are (200, 200) and (200, 500) all procedures attained the nominal level while in the rest of the cases most of the procedures are close to nominal level.

Table 9: Type I error rate

P=2 S₁ > S₂	Combination of different sample size						α = 0.05
$\bar{X}_1 = (2 \quad 3)$ $\bar{X}_2 = (2 \quad 3)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$\bar{X}_1 - \bar{X}_2 = (5 \quad 10)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1=\begin{pmatrix} 800 & 200 \\ 200 & 800 \end{pmatrix}$	John	0.053	0.055	0.047	0.043	0.041	0.051
	Yao	0.057	0.055	0.046	0.043	0.041	0.051
$S_2=\begin{pmatrix} 90 & 60 \\ 60 & 90 \end{pmatrix}$	Krish	0.053	0.055	0.047	0.043	0.04	0.051

From Table 9; all procedures are closer to nominal level at sample sizes (200, 500) and (25, 15) while Adebayo is exactly at nominal level. There is deflation and inflation in the rest of the setting.

Table 10: Type I error rate

P=3 S₁ > S₂	Combination of different sample size						α = 0.01
$\bar{X}_1 = (10 \ 10 \ 10)$ $\bar{X}_2 = (10 \ 10 \ 10)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$(\bar{X}_1 - \bar{X}_2) = (0 \ 0 \ 0)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1=\begin{pmatrix} 900 & 700 & 500 \\ 700 & 900 & 300 \\ 500 & 300 & 900 \end{pmatrix}$	John	0.005	0.009	**0.01**	**0.01**	**0.01**	0.008
	Yao	0.005	**0.01**	0.011	**0.01**	0.014	0.009
$S_2=\begin{pmatrix} 200 & 90 & 50 \\ 90 & 200 & 20 \\ 50 & 20 & 200 \end{pmatrix}$	Krish	0.006	0.009	0.011	**0.01**	**0.01**	0.009

From Table 10, at sample sizes (300, 200) all procedures attained the nominal level. At a sample size (200, 200) Yao and Adebayo attained exactly nominal level and at a sample size (10, 20) Johanson and Krishnamoorthy attained exactly the nominal level. There is deflation when sample size is (10, 10).

Table 11: Type I error rate

P=3 $S_1 > S_2$	Combination of different sample size						$\alpha = 0.025$
$\bar{X}_1 = (10\ 10\ 10)$ $\bar{X}_2 = (10\ 10\ 10)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$(\bar{X}_1 - \bar{X}_2) = (0\ 0\ 0)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1 = \begin{pmatrix} 900 & 700 & 500 \\ 700 & 900 & 300 \\ 500 & 300 & 900 \end{pmatrix}$	John	0.035*	0.024	0.023	0.219	0.027	**0.025**
	Yao	0.037*	0.024	0.023	0.221	0.033	**0.025**
$S_2 = \begin{pmatrix} 200 & 90 & 50 \\ 90 & 200 & 20 \\ 50 & 20 & 200 \end{pmatrix}$	Krish	0.035*	0.024	0.025	0.221	0.028	**0.025**

From Table 11, all procedures lie outside the interval at a sample sizes (10, 10) and consider non-robust and when sample sizes are (200, 500) all procedures were exactly at nominal level.

Table 12: Type I error rate

P=3 $S_1 > S_2$	Combination of different sample size						$\alpha = 0.05$
$\bar{X}_1 = (10\ 10\ 10)$ $\bar{X}_2 = (10\ 10\ 10)$	Sample size	Equal sample size $n_1 = n_2$		Unequal sample size $n_1 > n_2$		Unequal sample size $n_1 < n_2$	
$(\bar{X}_1 - \bar{X}_2) = (0\ 0\ 0)$	TEST	10, 10	200, 200	25, 15	300, 200	10, 20	200, 500
$S_1 = \begin{pmatrix} 900 & 700 & 500 \\ 700 & 900 & 300 \\ 500 & 300 & 900 \end{pmatrix}$	John	0.049	0.048	0.043	0.051	0.053	0.052
	Yao	0.058	0.048	0.045	0.051	0.065*	0.052
$S_2 = \begin{pmatrix} 200 & 90 & 50 \\ 90 & 200 & 20 \\ 50 & 20 & 200 \end{pmatrix}$	Krish	**0.05**	0.048	0.045	0.051	0.057	0.052

From Table 12, Krishnamoorthy is exactly at nominal level when sample size is (10, 10) and Yao is non-robust at sample size (10, 20) while in the rest of the settings most of the procedures are close to nominal level and said to be good.

ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)

Summary of the Findings

Yao and Adebayos' procedures tend to have equal and highest power when we have equal sample sizes at both small and large sample sizes when P=2. In Table 1, 2 and 3 when the sample size $n_1 > n_2$, Adebayos' procedure has the highest power. At a sample size (200, 500) Yaos' procedure has the highest power in all the scenarios. Table 2 and 3 depict that Johanson and Yao have equal and highest power. When P variable increases to P=3 the power of Yao decreases and unable to equalize with Adebayo under the sample size (200, 200). Adebayo also decreases under a sample size (25, 15). But the power of Krishnamoorthy increases as seen in Table 4 and 5 under (200, 500) sample size.

From Table 7, all procedures deflated when sample size is (200, 200) and (10, 20) except Yao which is close to nominal level. When sample sizes are $n_1 > n_2$ all procedures are closer to nominal level. All procedures inflated when sample size is (10, 10). From Table 11, all procedures lie outside the interval and consider non-robust and when sample size is (200, 500) all procedures were at exactly nominal level. From Table 12, Krishnamoorthy is exactly at nominal level when sample size is (10, 10) and Yao is non-robust at sample size (10, 20) while in the rest of the settings most of the procedures are close to nominal level.

CONCLUSION

It is universally accepted that to date, there is no single procedure that can positively contest or compete in all conditions under Multivariate Behrens Fisher Problem. However, in this comparative study we are able to discovered and conclude that some of these existing procedures have equal and highest power in some certain settings like Yao and Adebayo, Johansen and Yao, Krishnamoorthy and Adebayo, Yao and Krishnamoorthy but when P-variables increases we also found that these procedures with equal power varies significantly, where as some procedures' power decrease while some increases in 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 and some procedures lie outside the interval and considered non-robust. Yao and Adebayos were found good when P=2 and sample size are $n_1 > n_2$, it is discovered that at a sample size (300, 200) all procedures attained the nominal level.

Further Research

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) using different settings and conditions other than the ones used in this research is recommended.

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. Algina, J. Oshima, T. C, and Tang, K. L. (1991). Robustness of Yao's, James', and Johansen's tests under variance-covariance heteroscedasticity and nonnormality. Journal of Educational Statistics, 16, 125-139.

3. Gulumbe, S.U, Basiru S and Dahiru S (2021). On Some Solution of the Multivariate Behrens Fisher Problem.ISSN 2688-8300 (Print) ISSN 2644-3368 (Online). JMSCM, Vol.2, No.2, January 2021. https://jmscm.smartsociety.org/vol2_no2.html

4. 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.

5. Junyong Park and Bimal Sinha (2009). Some aspects of multivariate behrens-fisher problem. Department of Mathematics/Statistics University of Maryland, Baltimore. DOI: 10.1177/0008068320090107.

6. Kim, S.-H.and Cohen, A. S. (1998). On the Behrens-fisher problem: a review. Journal of Educational and Behavioral Statistics 23, 356–377.

7. 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.

8. Satterthwaite, F.E. (1946). “An Approximate Distribution of Estimates of Variance Components”.Biometrics Bulletin. 2:110-114.

9. Wang, Y. Y. (1971). Probabilities of the Type I errors of the Welch tests for the BehrensFisher problem. Journal of the American Statistical Association, 66, 605-608.

10. Welch, B.L. (1947). “The Generalization of Student’s Problem when Several Different Population Variances are Involved”. Biometrika. 34:29-35.

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