Heteroscedasticity in One Way Multivariate Analysis of Variance

Oyeyemi GM, Adebayo PO* and Adeleke BL
Department of Statistics, University of Ilorin, Ilorin, Nigeria

Abstract

This work aimed at developing an alternative procedure to MANOVA test when there is problem of heteroscedasticity of dispersion matrices and compared the procedure with the existing multivariate test for vector of means. The alternative procedure was developed by adopting Satterthwaite’s approach of univariate test for unequal variances. The approach made use of approximate degree of freedom method in one way MANOVA when the dispersion matrices are not equal and unknown but positive definite. The new procedure was compared by using simulated data when it is Multivariate normal, Multivariate Gamma and real life data. The new procedure performed better in terms of power of the test and type I error rate when compared with Johanson procedure.

Keywords: Multivariate analysis of variance; Type I error rate; Power the test; Heteroscedasticity equality of variance co-variance; Balance design; Unbalance design; Alternative hypothesis; R statistical package

Introduction

Multivariate Analysis of Variance (MANOVA) can be viewed as a direct extension of the univariate (ANOVA) general linear model that is most appropriate for examining differences between groups of means on several variables simultaneously [1,2]. In ANOVA, differences among various group means on a single-response variable are studied. In MANOVA, the number of response variables is increased to two or more variables. MANOVA has three basic assumptions that are fundamental to the statistical theory: (i) independent, (ii) multivariate normality and (iii) 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. The violations in assumptions of multivariate normality and homogeneity of covariances may affect the power of the test and type I error rate of multivariate analysis of variance test [3-6].

The problem of comparing the mean vectors that are more than two multivariate normal populations is called Multivariate Analysis of Variance (MANOVA). If the variance - covariance matrices of the populations are assumed to be equal, then there are some accepted tests available to test the equality of the normal mean vectors, which are: [7] largest root, the trace [8-10] likelihood ratio, and the [11,12]. Contrary to popular belief, they are not competing methods, but are complementary to one another. However when the assumption of equality of variance-covariance matrix failed or violated it means that none of the aforementioned test statistic is appropriate for the analysis otherwise the result will be prejudiced. This predicament is known as the multivariate Behrens - Fisher problem which deal with testing the equality of normal mean vector under heteroscedasticity of dispersion matrices. If the covariance matrices are unknown and arbitrary, then the problem of testing equality of the mean vectors is more complex, and only approximate solutions are available.

Johansen et al. [13-15] proposed multivariate tests for the situation in which the covariance matrices could be unequal. In this study, an approximate degree of freedom used [16] for comparing k normal mean vectors when the population variance - covariance matrices are unknown is proposed and compared with an existing procedure (by Johanson) when the groups (k) and random variables (p) are three respectively.

Abstract

This work aimed at developing an alternative procedure to MANOVA test when there is problem of heteroscedasticity of dispersion matrices and compared the procedure with the existing multivariate test for vector of means. The alternative procedure was developed by adopting Satterthwaite’s approach of univariate test for unequal variances. The approach made use of approximate degree of freedom method in one way MANOVA when the dispersion matrices are not equal and unknown but positive definite. The new procedure was compared by using simulated data when it is Multivariate normal, Multivariate Gamma and real life data. The new procedure performed better in terms of power of the test and type I error rate when compared with Johanson procedure.

Keywords: Multivariate analysis of variance; Type I error rate; Power the test; Heteroscedasticity equality of variance co-variance; Balance design; Unbalance design; Alternative hypothesis; R statistical package

Introduction

Multivariate Analysis of Variance (MANOVA) can be viewed as a direct extension of the univariate (ANOVA) general linear model that is most appropriate for examining differences between groups of means on several variables simultaneously [1,2]. In ANOVA, differences among various group means on a single-response variable are studied. In MANOVA, the number of response variables is increased to two or more variables. MANOVA has three basic assumptions that are fundamental to the statistical theory: (i) independent, (ii) multivariate normality and (iii) 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. The violations in assumptions of multivariate normality and homogeneity of covariances may affect the power of the test and type I error rate of multivariate analysis of variance test [3-6].

The problem of comparing the mean vectors that are more than two multivariate normal populations is called Multivariate Analysis of Variance (MANOVA). If the variance - covariance matrices of the populations are assumed to be equal, then there are some accepted tests available to test the equality of the normal mean vectors, which are: [7] largest root, the trace [8-10] likelihood ratio, and the [11,12]. Contrary to popular belief, they are not competing methods, but are complementary to one another. However when the assumption of equality of variance-covariance matrix failed or violated it means that none of the aforementioned test statistic is appropriate for the analysis otherwise the result will be prejudiced. This predicament is known as the multivariate Behrens - Fisher problem which deal with testing the equality of normal mean vector under heteroscedasticity of dispersion matrices. If the covariance matrices are unknown and arbitrary, then the problem of testing equality of the mean vectors is more complex, and only approximate solutions are available.

Johansen et al. [13-15] proposed multivariate tests for the situation in which the covariance matrices could be unequal. In this study, an approximate degree of freedom used [16] for comparing k normal mean vectors when the population variance - covariance matrices are unknown is proposed and compared with an existing procedure (by Johanson) when the groups (k) and random variables (p) are three respectively.
\[ T(x; \delta) = T(x, \ldots, x; \delta, \ldots, \delta) \quad (4) \]

Johanson’s test [17]:
\[ J_{OH} = T(x, \ldots, x; \delta, \ldots, \delta) \quad (5) \]

Where,
\[ c = p(k - 1) + 2A - \frac{6A}{p(k - 1) + 2} \quad (6) \]

And,
\[ A = \sum_{i=1}^{k} \{I - w_i^2\} \quad (7) \]

Johanson showed that, under \( H_0 \), \( J_{OH} \) is approximately distributed as a \( F_{f/1, f/2} \) random variable, where the \( f_1 = p(k - 1) \) and \( f_2 = \frac{p(k - 1)[p(k - 1) + 2]}{3A} \). Thus, the Johanson test rejects the null hypothesis in eqn. (3) whenever \( J_{OH} \geq F_{f/1, f/2, 1-\alpha} \).

**Proposed Method**

The entire aforementioned scholars worked on the degree of freedom by using various methods to get approximate degree of freedom to the test statistic, which the proposed procedure intended to, by extending Satterthwaite’s procedure (two moment solution to the behrens-fisher problem) in univariate to a multivariate Behrens-Fisher problem. In Satterthwaite [16] proposed a method to estimate the distribution of a linear combination of independent chi-square random variables with a chi-square distribution. Let \( \sum_{i=1}^{k} u_i \) where \( a \) are known constants, and \( U_i \) are independent random variables such that
\[ U_i = \frac{(n_i - 1)s_i^2}{\sigma_i^2} - X^2(n_i - 1) a_i = \frac{c_i^2a_i^2}{n_i(n_i - 1)}, f = 1, 2 \quad (8) \]

Since linear combination of random variable does not, in general, possess a chi-square distribution. Satterthwaite [16] suggested the use of a chi-square distribution, Say \( X^2(f) \) as an approximation to the distribution of \( \frac{f}{E[L]} \). This notion is compactly written as:
\[ \frac{f}{E[L]} \sim X^2(f) \quad (9) \]

Where “\( \sim \)” is taken to mean “is approximately distributed as.”

From an intuitive standpoint, the distribution of \( \frac{f}{E[L]} \) should have characteristics similar to some member of the chi-square family of densities [17-21]. But recall that if a chi-square distribution has degrees of freedom \( n_i - 1 \), then its mean is and variance is \( 2(n_i - 1) \).

Symbolically, this requires that, the first moment of the statistic is:
\[ E \left[ \frac{f}{E[L]} \right] = f \quad (10) \]

This implies that a chi-square with \( f \) degrees of freedom should be used.

Let consider the second moment. The variance of the statistic is:
\[ Var \left[ \frac{f}{E[L]} \right] = 2f \quad (11) \]

The first two central moments of \( L \) are obtained:

We shall consider the test statistic \( Y^tS^{-1}Y \) and use Univariate Satterthwaite approximation of degrees of freedom method to suggest multivariate generalization based on the \( T^2 \)- distribution [22-29]. Let
\[ S = \sum_{i=1}^{k} S_i \quad \text{and} \quad Y = x - m \quad \text{where} \quad i = 1, 2, \ldots, k. \]

\[ y \sim N(0, \sum) \]

If \( S \) were a Wishart matrix \((n_i - 1)S \sim \text{wishart}(n_i - 1, \sum)\) then for an arbitrary constant vector \( b \) we should have
\[ b'y \sim N(0, b'b \sum) \]

\[ (n_i - 1)(b'b \sum) \sim \chi^2(n_i - 1) \]

That is \( m_i = \frac{(n_i - 1)b'b \sum}{b'b \sum} \quad X^2(n_i - 1) \]

Eqn. (12) is the multivariate version of eqn. (8) given by Satterthwaite

A linear combination of \( p \) (random) variables
\[ h = r_{m_1} + r_{m_2} + \ldots + r_{m_p} \]

\[ E[h] = E[r_{m_1} + r_{m_2} + \ldots + r_{m_p}] \quad (13) \]

Substitute eqn. (12) into eqn. (13)
\[ E[h] = \frac{d_1 d_2 d_3 \sum_{i=1}^{k} b_i}{n_i} + \ldots + \frac{d_1 d_2 d_3 \sum_{i=1}^{k} b_i}{n_i} \]

Eqn. (14) is the multivariate version of eqn. (8) given by Bush and Olkin, [19] that
\[ \text{Var}[h] = \text{Var}[r_{m_1} + r_{m_2} + \ldots + r_{m_p}] \quad (15) \]

Substitute eqn. (12) into eqn. (15)
\[ \text{Var}[h] = \frac{d_1 d_2 d_3 \sum_{i=1}^{k} b_i}{n_i} + \ldots + \frac{d_1 d_2 d_3 \sum_{i=1}^{k} b_i}{n_i} \]

Note that
\[ \text{Var}[h] = \frac{(n_i - 1)b'b \sum}{b'b \sum} \sim \chi^2(n_i - 1) \]

\[ E[h] = \frac{d_1 d_2 d_3 \sum_{i=1}^{k} b_i}{n_i} + \ldots + \frac{d_1 d_2 d_3 \sum_{i=1}^{k} b_i}{n_i} \]

Substituting eqns. (14) and (16) into eqn. (11)
\[ 2f = \frac{f}{E[h]} \quad \text{Var}[h] = \frac{2d_1 d_2 d_3 \sum_{i=1}^{k} b_i}{n_i} + \ldots + \frac{2d_1 d_2 d_3 \sum_{i=1}^{k} b_i}{n_i} \]

\[ \text{Substituting eqns. (14) and (16) into eqn. (11)} \]

\[ 2f = \frac{f}{E[h]} \quad \text{Var}[h] = \frac{(d_1 d_2 d_3 \sum_{i=1}^{k} b_i)}{n_i} + \ldots + \frac{(d_1 d_2 d_3 \sum_{i=1}^{k} b_i)}{n_i} \]

\[ f = \frac{(d_1 d_2 d_3 \sum_{i=1}^{k} b_i)}{n_i} + \ldots + \frac{(d_1 d_2 d_3 \sum_{i=1}^{k} b_i)}{n_i} \]

Yao [30] showed that
\[ \text{Var}[h] = \frac{(b'y)^2}{b'b \sum} - f^2(n_i - 1) \]

And also it was shown by the authors that
\[ \text{sup}(wb) = wb^* = \frac{(b^* - y)^2}{(b^* - Sb^*)} = y' S^{-1} y \]

Where the maximizing \( b^* = S^{-1} y \) and \( d_i; d_i = 1 \), then eqn. (17) becomes

\[ f = \frac{\left( \frac{1}{n_i - 1} \sum_{y_i} x_i y \right)}{\left( \frac{1}{n - 1} \sum_{y} x y \right)} \]

When \( y = \overline{x} - \mu^*_i \) eqn. (18) becomes:

\[ f = \frac{\left( \frac{1}{n_i - 1} \sum_{y_i} (x_i - \mu_i)(x_i - \mu_i)^* \right)}{\left( \frac{1}{n - 1} \sum_{y} x y \right)} \]

Therefore \( T(\overline{x}, \mu) = \frac{\beta \rho}{f - p + 1} F_p f - p + 1 \)

where

\[ T(\overline{x}, \mu) = \frac{1}{n_i} \sum_{y_i} (x_i - \mu_i) \]

Data simulation

Data was simulated in R environment to estimate power of the test and Type I error rate when the alternative hypothesis is true (that is when the mean vectors are not equal).

Data analysis

Simulated and real life data sets from previous study [17] were used to compare the proposed alternative procedure with the existing one (Johanson). For the simulated data, three factors were varied namely: number of groups (k), the number of variables (p) and significant levels (α).

In each of the 1000 replications and for each of the factor combination, an \( n_i \times p \) (where \( i = 1, \ldots, 4 \)) data matrix \( X_i \) were generated using an R package for Multivariate Normal. The programme also performs the Box-M test for equality of covariance matrices using the test statistic:

\[ M = c \sum_{i=1}^{n_i} (n_i - 1) \log \left| S_i \right| \]

Where

\[ S_i = \sum_{y_i} (n_i - 1) X_i \]

\[ c = \frac{2 p^2 + 3 p - 1}{6(k - 1)(p + 1)} \left( \sum_{y_i} \frac{1}{n_i} - \frac{1}{n - k} \right) \]

\[ X_i^2 = (1 - C) M \]

And \( S_i \) and \( S_p \) are the \( i \)-th unbiased covariance estimator and the pooled covariance matrix respectively. Box’s M has an asymptotic chisquare distribution with \( \frac{1}{2}(p + k)(k - 1) \) degree of freedom. Box’s approximation seems to be good if each \( n_i \) exceeds 20 and if k and p do not exceed 5 [11]

\( H_i \) is rejected at the significance level \( \alpha \) if \( X_i^2 \geq X_{\alpha(i)} \) where

\[ v = \frac{1}{2}(p + 1)(k - 1) \]

Result

Table 1 shows that irrespective of the sample size and significant level \( \alpha \), the propose procedure has the higher power of the test and less Type I error rate compared to Johanson when the alternative hypothesis is true. The two only have the same type I error rate when the sample sizes are large (100’s and 200’s), but then the powers of the test are not the same throughout the sample sizes considered (5’s, 10’s, 50’s, 100’s and 200’s).

From Table 2, when the sample size are not equal and very small [(5,10,15) and (20,25,30)], Johanson procedure perceived to be better than the propose procedure in terms of power of the test but poor in type I error rate at significant level \( \alpha = 0.01 \), but when sample sizes increases to (50,70,90) and (100,150,200) the propose procedure performed better at the two significant level \( (a=0.01 \text{ and } 0.05) \).

Table 3, when the sample sizes are small [(5,5,5) and (10,10,10)] and equal in all the groups, Johanson performed better at significant level \( \alpha = 0.01 \) in terms of power of the test while propose procedure are better in terms of type I error rate in all the scenario, but when sample sizes are (100,100,100) and (200,200,200) they both perform the same.

From Table 4, when the stimulated data are multivariate gamma and unbalance, the propose procedure are better than Johanson procedure in the entire scenario both in terms of power of the test and type I error rate.

Illustrative example

The real life data used by Krishnamoorthy and Xia [27] was used

<table>
<thead>
<tr>
<th>Power of the test</th>
<th>Sample size</th>
<th>0.01</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johanson</td>
<td>Propose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P=2 and k=3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5, 5, 5</td>
<td>0.0381</td>
<td>0.0391</td>
<td>0.1386</td>
</tr>
<tr>
<td>10, 10, 10</td>
<td>0.0651</td>
<td>0.0803</td>
<td>0.1904</td>
</tr>
<tr>
<td>50, 50, 50</td>
<td>0.3732</td>
<td>0.4589</td>
<td>0.5895</td>
</tr>
<tr>
<td>100, 100, 100</td>
<td>0.7304</td>
<td>0.9109</td>
<td>0.8752</td>
</tr>
<tr>
<td>200, 200, 200</td>
<td>0.9744</td>
<td>0.9901</td>
<td>0.9933</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type I error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
</tr>
<tr>
<td>Johanson</td>
</tr>
<tr>
<td>P=2 and k=3</td>
</tr>
<tr>
<td>5, 5, 5</td>
</tr>
<tr>
<td>10, 10, 10</td>
</tr>
<tr>
<td>50, 50, 50</td>
</tr>
<tr>
<td>100, 100, 100</td>
</tr>
<tr>
<td>200, 200, 200</td>
</tr>
</tbody>
</table>

Table 1: Multivariate Normal Distribution (For balanced design).

<table>
<thead>
<tr>
<th>Power of the test</th>
<th>Sample size</th>
<th>0.01</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johanson</td>
<td>Propose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P=2 and k=3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5, 10, 15</td>
<td>0.0624</td>
<td>0.0469</td>
<td>0.1866</td>
</tr>
<tr>
<td>20, 25, 30</td>
<td>0.1546</td>
<td>0.0761</td>
<td>0.3366</td>
</tr>
<tr>
<td>50, 70, 90</td>
<td>0.4994</td>
<td>0.5626</td>
<td>0.7132</td>
</tr>
<tr>
<td>100, 150, 200</td>
<td>0.8848</td>
<td>0.962</td>
<td>0.9587</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type I error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
</tr>
<tr>
<td>Johanson</td>
</tr>
<tr>
<td>P=2 and k=3</td>
</tr>
<tr>
<td>5, 10, 15</td>
</tr>
<tr>
<td>20, 25, 30</td>
</tr>
<tr>
<td>50, 70, 90</td>
</tr>
<tr>
<td>100, 150, 200</td>
</tr>
</tbody>
</table>

Table 2: Multivariate normal distribution (For unbalanced design).
The summary statistics for the four groups are given below

\[
\begin{pmatrix}
X_{11} & X_{12} & X_{13} & X_{14} \\
X_{21} & X_{22} & X_{23} & X_{24} \\
X_{31} & X_{32} & X_{33} & X_{34} \\
X_{41} & X_{42} & X_{43} & X_{44}
\end{pmatrix}
\]

The matrices

\[
\begin{pmatrix}
0.862 & -0.173 & -0.210 & -1.174 \\
-0.604 & 0.138 & 0.308 & 0.076 \\
-0.493 & -3.080 & -2.717 \\
0.973 & 0.126 & 0.922 \\
0.925 & 0.091 & 0.070 & -0.015 \\
-0.953 & -0.146 & 0.227 & 0.158 \\
-0.625 & -1.409 & 0.085 & -0.430 \\
0.964 & 0.095 & -0.666 \\
-0.640 & 0.362 & 2.174 \\
-0.268 & 0.355 & 0.043 & 0.126 \\
\end{pmatrix}
\]

\[
S_i = w_i^{(-1)} \text{ where } w_i = S_i^{(-1)} \text{ and } S = \sum_{i=1}^{4} S_i
\]

\[
\begin{pmatrix}
2.380 & 0.494 & 0.407 & 0.750 \\
1.872 & 0.414 & 0.161 \\
2.356 & 0.062 \\
0.538 \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
2.051 & -0.325 & 0.308 & 0.071 \\
1.902 & 0.370 & 0.175 \\
0.655 & 0.347 \\
0.737 \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
1.110 & -0.176 & -0.136 & -0.051 \\
1.733 & 0.029 & 0.217 \\
0.693 \\
1.029 \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
0.758 & 0.032 & -0.053 & 0.151 \\
1.457 & 0.515 & 0.539 \\
1.912 & 0.466 \\
0.732 \\
\end{pmatrix}
\]
Remark

power of the test than Johanson. Than Johanson procedure because propose procedure has the higher illustrative example, it is observed that propose procedure performed level \( \alpha \) varies, when the design are balance and unbalance. Also from the in the entire scenario that is, when sample size differs, when significant procedure) power of the test are higher than that of Johanson procedure

\[ p \text{-value less than } \alpha. \]

null hypothesis since 7.6763 is greater than 3.1377 and 3.8459 with 2.7464 with \( p \)-value greater than \( \alpha \), while propose procedure rejected accepted the null hypothesis because 2.275 is less than 2.3451 and than 0.05, but when significant level \( \alpha \) are 0.025 and 0.01, Johanson is less than 0.05 and that of propose procedure is 0.0001 which is less 7.6763 is greater than 2.6138, also \( p \)-values of Johanson is 0.0294 which is greater than critical value that is 2.275 is greater than 2.0443 and 4.1377 and 3.8459 with \( p \)-value less than \( \alpha \).

Remark

From the simulated data, it is obvious that the propose procedure performed better than Johanson procedure because its (propose procedure) power of the test are higher than that of Johanson procedure in the entire scenario that is, when sample size differs, when significant level \( \alpha \) varies, when the design are balance and unbalance. Also from the illustrative example, it is observed that propose procedure performed than Johanson procedure because propose procedure has the higher power of the test than Johanson.

References

7. Roy SN (1945) The individual sampling distribution of the maximum, the minimum, and any intermediate of the p-statistics on the null hypothesis. Sankhya 7: 113-158