International Research Journal of Statistics and Mathematics
Volume 13 Issue 2, April-June 2025
ISSN: 2995- 4363
Impact Factor: 9.41
https://kloverjournals.org/index.php/sm

A MEDIAN-BASED APPROACH TO FISHER'S TEST: OVERCOMING THE INFLUENCE OF OUTLIERS IN STATISTICAL ANALYSIS

Chigozie Ugochukwu Okeke and Amarachi Nkechi Igwe
Mathematics/Statistics Department, Ignatius Ajuru University of Education, Rumuolumeni, Port Harcourt, Rivers State, Nigeria
DOI:https://doi.org/10.5281/zenodo.15482131

ABSTRACT: The study considered Median based Fisher's Test: An alternative to Fishers' Least Significant Difference in the presence of outlier for interval data. It started with running outlier test by using Grubbs method and analysis of variance for three population samples were carried out. In the general sense, in problems like this, the hypothesis is tested, if the null hypothesis is rejected, and the alternative hypothesis accepted, that is to say, there is significant difference in the means of the samples, one will usually conduct multiple comparison test, to know the pair that contributed to the significant difference. For the purpose of this study, only Fisher's test was considered, because it centered their operations on the difference in means of the data. This research, introduced difference in median for Fisher's test as a result, proposed M-Fisher's method of knowing if a pair of median differs or not. These method used median in place of mean in Sum of Squares, Mean Squares, F-value, t-test and Fisher's test. This is because of the fact that the mean is always affected by outliers and extreme values. So, method that are not sensitive to outliers is proposed. Three data sets were used. Data 1 without outlier was the data of advertisement on new product of orange juice focused on three cities 1, 2, 3. city 1 centered on convenience, city 2 centered on quality and city 3 centered on price for 20 sample size, data 2 has the introduction of outliers for 20 sample size and data 3 is 100 sample size and it is a simulated data. The results showed that the Fisher's method alongside M-Fisher's had similar results for data 1. For data 2, the M-Fisher's showed robustness, when outliers were introduced. While Fishers was heavily affected by the presence of the outliers.

Keywords: Mean, Median, Fisher's test, Median Fisher;s test, ANOVA, Sum of Squares.

1. Introduction

The classic Analysis of Variance is a general linear model that has been in use for over 100 years (Midway et al., 2020) and is always used when factor or categorical data need to be analyzed. Though, an ANOVA can only produce an F-statistic and its associated p-value for the whole model. That is to say, an ANOVA reports whether one or more significant differences among group levels exist, it does not provide any information about specific group means compared to each other. It is possible that group differences exist that ANOVA does not detect. Sometimes, interests are not always in comparison of two groups per experiment. From time to time though practically, very often there may be need to determine whether differences exist among the means of three or more groups. The most commonly used technique for such determinations is analysis of variance (ANOVA). By rejecting the null hypothesis (H_o) after ANOVA, that is, in the case of three groups, H_o: $\mu_A = \mu_B = \mu_C$, we do not know how one group differs from a certain group. The result of ANOVA does not provide detailed information regarding the differences among various combinations of groups, (Sangseok & Dong, 2018). Therefore, researchers usually perform additional analysis to clarify the differences between particular pairs of experimental groups. If the null hypothesis (H_o) is rejected in the ANOVA for the three groups, interest is shifted to the Multiple Comparisons Test (MCT). Obviously, the inability to specifically compare group means with ANOVA has long been famous and a sub-field of multiple comparisons tests (MCTs) began to be developed by the middle of the 20th century (Harter, 1980). Certainly, when the analysis only includes two groups for example as in a t-test, then a significant result from the model is consistent with a difference between groups. But the usefulness of this approach is obviously limited and as stated by Zar (2010) that: "employing a series of two-sample tests to address a multisample hypothesis is invalid." What has developed over the last several decades have been a bounty of statistical procedures that can be applied to the evaluation of multiple comparisons. On the surface, this long list of options for MCTs is a good thing for researchers and data analysts. However, all the tests are unique, and some are better suited to different data sets. In this research, the Fisher's Least Significance difference is applied on three data sets known as small and large sample sizes. The small sample size was obtained from Keller and Warrack (2003), this data was used to run analysis of ANOVA, and Fisher's was applied alongside the proposed median Fishers. Outlier is introduced to this data in order to see how the Fisher's and the Median based Fisher's tests will look like. The third data is that of a simulation of the original data, as a large sample size. It has been known that Fishers' test was based on mean of the data set, from which the sum of squares and their respective mean squares were obtained. In the phase of extreme value or outlier, any statistic that is mean dependent face the risk of false proposition, this may also affect the use of Fisher's test which depends on the sum of squares that are dependent on the mean value of the data set. In order to overcome such bottle neck, the median of the data set will be used in place of the means in calculating the sum of squares. This work tends to propose a median based Fishers' test, where the median of the data sets will be used instead of their means and their sum of squares and the mean squares based on the median value will be calculated. The proposed method will be compared with the already known Fishers' LSD to see which method will be better for a particular sample size. The work is aimed at proposing a median based Fishers' least significant difference as a better approach to be used instead of the Fishers' when the null hypothesis have been rejected. Ramsey and Ramsey (2008) did not consider using the Tukey–Kramer test as the omnibus test in the Hayter–Fisher procedure. However, based on their results as well as previous ones, making this modification should result in a procedure superior to both TK and HF for detecting pairwise differences. The focus of this investigation is to compare, for unequal sample sizes, the power of the modified Hayter–Fisher procedure using the Tukey–Kramer test of the largest observed studentized pairwise difference as the omnibus test, to that of the usual Hayter–Fisher procedure using the F-test as the omnibus test. Onu, et al. (2021) Studied discriminating between second-order model with or without interaction for central tendency estimation using ordinary least square for estimation of the model parameters. Two data sets were considered and they are known as small sample size which is data of unemployment rate as response, inflation rate and exchange rate as the predictors from 2007 to 2018 and large sample size which is data of flow rate on hydrate formation for Niger Delta deep offshore field. The R squared, AIC, SBC and SSE were applied for both data sets to test for the adequacy of the models. The small data was used as illustration 1 and the large data was used as illustration 2. It was revealed that, model centered on mean with interaction proved better than mean model without interaction. Median and mode values were found to be equal, as a result, the estimates of the median models were equal to the estimates of the mode models in all cases for both large and small data. The models centered on median and mode with interaction were better than those without interaction for both illustrations. Mean and mode models with or without interactions were found better than the mean models with or without interactions for both illustrations.

2. Materials and Methods

The research starts with running an outlier test on the data sets, to know if any of the data contains outlier or not. If there is no outlier in the data set, an artificial outlier will be created in order to study the Fisher's least significant difference that is based on the mean value of the data relative to the proposed median based fisher's test. The test statistic to be used for the outlier test is the Grubbs' test.

Grubbs' test

The Grubbs' test is given as seen in Sanchez et al. (2019) as

$$G = \frac{\bar{Y} - Y_{min}}{S} \quad (1)$$

Y_{min} is the minimum value of Y. The above formula is used to test whether the minimum value is an outlier. It can also be give as

$$G = \frac{Y_{max} - \bar{Y}}{S} \quad (2)$$

Where Y_{max} is the maximum value of Y. The above test is used to test if the maximum value of the data is an outlier.

Generally,

$$G = \frac{|Y_i - \bar{Y}|}{S} \quad (3)$$

$\bar{Y}$ is the sample mean and S is the sample standard deviation.

For two sided test, the null hypothesis of no outlier is rejected if

$$G > \frac{N-1}{\sqrt{N}} \sqrt{\frac{\left(t_{\propto/(2N),N-2}\right)^2}{N-2+\left(t_{\propto/(2N),N-2}\right)}} \quad (4)$$

($t_{\propto/2N,N-2}$) is the critical value of t-test and N-2 is the degree of freedom and $\propto/(2N)$ is the significant level.

Application of One Way Analysis of Variance

$H_0: \mu_1 = \mu_2 = . . . \mu_i$ (5)

While the alternative hypothesis is stated as

H_1:At least two means differ (6)

The analysis of variance can be facilitated by the notations in table 1 below.

Table 1 one-way ANOVA notation

Treatment

					J				
	1	2							k
	x_{11}	x_{12}	.	. .	x_{1j}	.	.	.	x_{1k}
	x_{12}	x_{22}	.	. .	x_{1j}	.	.	.	x_{1k}
	.	.							.
	.	.							.
	.	.							.
	x_{n1}	x_{n2}			x_{nj}				x_{nk}
Sample size	n_1	n_2			n_j				n_k
Sample mean	$\bar{x}_1$	$\bar{x}_2$			$\bar{x_j}$				$\bar{x}_k$

x_{ij} is the i^{th} observation of the j^{th} sample, n_j is the number of observations in the sample taken from the j^{th} population and

$$\bar{x}_i = \frac{\sum_{nj=i\,1} x_{ij}}{n_j} \quad (7)$$

is the mean of the j^{th} population, while

$$\bar{x} = \frac{\sum_{kj=1} \sum_{ni=i1} x_{ij}}{n} \quad (8)$$

is the grand mean or the mean of the means of all the observations, where$n = n_1 + n_2 + . . . + n_k$and k is the number of population.x is the response variable and the unit we measure in the experiment is called the experimental unit.

Test Statistic

This test statistic is always computed according to the following;

-if the null hypothesis is true, the population means would have equal value, which means that the sample means will be close to one another.
-if the alternative hypothesis is true, there would be large differences between some of the sample means.
The proximity of the sample means to each other is measured by a test statistic known as between treatment variation or sum of square treatment denoted by

$SS_t = \sum_{j=1}^{k} n_j (x_j - x)^2$ (9)

Note that if $x_1 = x_2 = . . . = x_k$, then
$SS_t = 0$
The small value of SS_t supports the null hypothesis, while large value supports the alternative hypothesis, but to know how large SS_t will be to indicate that the population means differ is measured by within treatment variation or sum of square error denoted by SS_E it is given as

$SS_E = \sum_{j=1}^{k} \sum_{i=1}^{n_i} (x_{ij} - x_j)^2$ (10)

$= \sum_{i=1}^{n_i} (x_{i1} - x_1)^2 + \sum_{i=1}^{n_i} (x_{i2} - x_2)^2 + . . . \sum_{i=1}^{n_i} (x_{ik} - x_k)^2$ (11)

$= (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 . . . (n_k - 1)s_k^2$ (12) where s_i^2 is the sample variances of the sample j.
Another test statistic to be applied is the sum of square total given as
$SSTotal$ (13)
The mean squares can then be obtained as follows
Mean Square Treatment which is given as
SS
$MSt = k - t1$
The other test statistic to be applied is the Mean Square error given as (14)
SS
$MSE = n - EK$
F-Statistic
The F-Statistic is given as (15)
MS
$F = t$
MSE
ANOVA TABLE
The Analysis of Variance table is as shown
Table 2: (Analysis of Variance)

SV	Df	SS	MS	F
				(16)
treatment	$k - 1$	SSE	SSt	MSt

International Research Journal of Statistics and Mathematics
Volume 13 Issue 2, April-June 2025
ISSN: 2995- 4363
Impact Factor: 9.41
https://kloverjournals.org/index.php/sm

t

$k-1$ MSE

Error $n-k$ SSE SS

$MSE = n - EK$

Total $n2-1$ $SSTotal$

Rejection Region and P-value

$MS =$ ____ $F =$ ____

The null hypothesis is rejected if $F_{cal} > F_{\propto, k-1, n-k}$ if not, the alternative hypothesis is accepted, as stated in equation (5) and (6), If the alternative hypothesis is accepted, that means that there is significant difference among the treatment means, then multiple comparison tests are applied to determine the very pair that contributed to the said difference.

Multiple Comparison Test (MCT)

The multiple comparison test that will be applied in this study is the Fisher's Least Significant Difference and the proposed median based Fisher's test known as M-Fisher's test as seen;

Fisher's Least Significant Difference (LSD)

The Fisher's Least Significant Difference is defined as seen in Keller & Warrack, (2003) for equal sample size as

$$LSD = t_{\frac{\propto}{2}}, N - k\sqrt{\frac{2MSE}{n}} \quad (17)$$

For equal sample sizes, the pairs of means are said to be significantly different if absolute value of their mean differences is greater than that in equation (16). For unequal sample sizes, a pair of means are said to be significantly different if the absolute value of their mean differences is greater than

$$= t_{\frac{\propto}{2}} = \sqrt{\quad (\frac{1}{ni} - \frac{1}{nj})}$$

$LSDMSE$ (18)

In order to determine whether differences exist between each pair of population means is to compare the absolute value of the difference between their two sample means and the LSD value; that is to say, to conclude that μ_i and μ_j differ is to compare if

$$|\bar{x}_i - \bar{x_j}| > LSD \quad (19)$$

Note that for equal sample sizes, the LSD value remains the same. Recall that the t-test statistic is given as

$$t = \frac{(x_1 - x_2) - (\mu_1 - \mu_2)}{\sqrt{S_P^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (20)$$

The confidence interval is given as

$$(x_1 - x_2) \pm t_{\frac{\propto}{2}}\sqrt{S_P^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad (21)$$

With $v = n_1 + n_2 - 2$ degrees of freedom, where S_P^2 is the pooled variance estimate which is an unbiased estimator of variance of the population, (Iwundu, 2017 and Keller & Warrack, 2003).

Proposed Multiple Comparison Tests (M-Fisher's Test)

This study will propose a new approach of multiple comparison test known as M-Fisher's test. This method will be derived from the ideas of Fisher's least significant difference. It modifies the t-test statistic by substituting median in the formula of the t-test. That means that the t-test will be based on the median marks of the data instead of the mean mark as was based upon by the researcher. Since median is not affected by outlier or extreme values, that is to say, when there is outlier in the data set, this method will be recommended. It is given in the sub-section below;

The proposed Median Fishers'least significant difference or simply M-Fisher's test for equal sample sizes is found from the t-test with median oriented given as

$$t_0 = \frac{\bar{\bar{y}}_i - \bar{\bar{y}}_j}{\sqrt{MSE\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}} \quad (22)$$

Where $\bar{\bar{y}}_i - \bar{\bar{y}}_j$ is difference in median, instead of difference in mean, for equality of sample sizes, $i = j$. For equal sample sizes, the pairs of median are said to be significantly different if absolute value of their median differences is greater than that in equation (17). For unequal sample sizes, a pair of median are said to be significantly different if the absolute value of their median differences is greater than that in equation (18).

The t-test can also be given as seen in Onu et al. (2021) and Kutner et al. (2005) as

$$t = \frac{\hat{m}_1}{S(\hat{m}_1)}$$

The Mean squares to be used will be obtained from median based sum of square treatment given as

$$SS_t = \sum_{j=1}^{k} n_j\left(x_j - \bar{\bar{x}}\right)^2 \quad (23)$$

The sum of square total given as

$$SS_{Total} = \sum_{j=1}^{k}\sum_{i=1}^{n_i}\left(x_{ij} - \bar{\bar{x}}\right)^2 \quad (24)$$

Where $\bar{\bar{x}}$ is the median of the data set used to replace the mean as seen in equations (9) and (12) above.

3. Results and Discussion

Summary of the results are presented in the following tables.

Table 3.1: comparisons of results of Fisher's least significant difference and the proposed Median Fisher's test for data without outlier for $\propto = 0.05$ and small sample size, n=20.

	Critical	$\lvert\bar{x}_1 - \bar{x}_2\rvert$	$\lvert\bar{x}_1 - \bar{x}_3\rvert$	$\lvert\bar{x}_2 - \bar{x}_3\rvert$
FISHER'S	59.71	**75.45**	31.10	44.35

	Critical	$\lvert\bar{\bar{x}}_1 - \bar{\bar{x}}_2\rvert$	$\lvert\bar{\bar{x}}_1 - \bar{\bar{x}}_3\rvert$	$\lvert\bar{\bar{x}}_2 - \bar{\bar{x}}_3\rvert$
M-FISHER'S	59.82	**75**	42	33

This table reveals that only the pair $\bar{x}_1$ and $\bar{x}_2$ contributed to the significant difference observed in the hypothesis testing. This is true for both Fisher's and M-Fisher's for small sample size of n=20 without outlier for $\propto = 0.05$.

Table 3.2: comparisons of results of Fisher's and the proposed Median Fisher's for data with outlier for $\propto = 0.05$ and small sample size, n=20.

	Critical	$\lvert\bar{x}_1 - \bar{x}_2\rvert$	$\lvert\bar{x}_1 - \bar{x}_3\rvert$	$\lvert\bar{x}_2 - \bar{x}_3\rvert$
FISHER'S	59.71	**Null hypothesis is accepted**		
	Critical	$\lvert\bar{\bar{x}}_1 - \bar{\bar{x}}_2\rvert$	$\lvert\bar{\bar{x}}_1 - \bar{\bar{x}}_3\rvert$	$\lvert\bar{\bar{x}}_2 - \bar{\bar{x}}_3\rvert$
M-FISHER'S	59.82	**75**	42	33

This table shows that as outliers are introduced in the small sample size of n=20, the results of the Fisher's changed, such that, instead of the rejection of the null hypothesis in the previous data, that led to employing Fisher's test, the null hypothesis was accepted. Hence needless of conducting Fisher's test, but the outliers introduced had no effect on the proposed method M-Fisher's.

Table 3.4: comparison of sum of squares, mean squares and F-values of Data with and without outlier for both mean and median based analysis.

	SST		SSE		MST		MSE		F-Value	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
NO OUTLIER	57512.23	**57060**	506983.5	**508930.2**	28756.12	**28530**	8894.45	**8928.6**	3.23 (3.15)	**3.20 (3.15)**

WITH OUTLIER	83703.34	**57060**	2430783.24	**508930.2**	41851.67	**28530**	42645.32	**8928.6**	1.0 (3.15)	**3.20 (3.15)**

The table reveals that outliers affects Sum of Square Total SST, Sum of Square Error SSE, Mean Squares and even the F-values, but when median was used in place of the mean in calculations of SST, SSE, MST, MSE and F-values, it was observed that the presence of outliers did not affect the results.

Discussion of Results

Fishers methods and the proposed median Fisher's test for data without outlier small sample size for $\propto = 0.05$ For small sample for data without outlier, if was revealed that the critical value for the Fishr's LSD was 59.71, while the proposed median method for fishers was 59. This result shows that the median based fisher's can compete favorably with the already known fisher's method, in the absence of outlier for small sample size. The result also revealed that in the already known fisher's method, the pair of x_1 and x_2 differs significant that is to say, the pair contributed to the significant difference observed in the hypothesis testing. This result appears to be the same for the proposed methods M-Fishers. This result is in line with what was obtained in Onu et al. (2021), except that they did not consider Fisher's test but the median proved better than the mean.

Fisher's method and their proposed median methods for data with outlier Small sample size. For $\propto = 0.05$

The results revealed that, in the presence of outlier, the already known Fisher's method was seriously altered, that is to say, instead of accepting the alternative hypothesis, the null hypothesis is accepted. Hence, needless of multiple comparison test. For the proposed methods, M-Fishers it was observed that, the presence of outliers has minimum or no effect on the M-Fisher's method, simply because of the use of the median of the data sets that are not affected by outliers and extreme values. This suggests that, the proposed method is more robust than the already known method.

Sum of square and mean squares for data with or without outliers.

The results show that the SST based on median for data with outlier was equal to that without outlier. The SST that was based on mean computation of the data without outlier is less than the mean with outlier. The mean square total based on median also remains unchanged while the mean square total based on mean for data without outlier changed significant as compared with that with outlier. Same result was observed for MSE and P-values. These differences are high to the ton of 261911.11 for SST,

Volume 13 Issue 2, April-June 2025
ISSN: 2995- 4363
Impact Factor: 9.41
https://kloverjournals.org/index.php/sm

1923799.2 for SSE and 2.23 for P-value. This simplify means that, with the presence of outlier, the result of ANOVA will be different with the data without outlier.

Conclusion

From the findings in this research, the study concluded that, as a result of outliers in a set of data, the results obtained by mean imputations, such as, variance, standard deviation, sum of squares, fishers LSD (Least Significant difference) may result in the loss of generality and cannot stand the test of time. The introduction of median in place of mean has shown to give equal or approximately equal results for data without outlier as compared to that of the mean. As evident in the decisions of the fisher's method relative to the M-Fishers.

Recommendations

The study recommends to statisticians and decision making organizations that;

1. The median should be used in place of the mean in analysis involving mean, especially when outlier is suspected. This is because of the effects of outliers and extreme values on the mean.
2. The M-Fishers can also be used in place of Fisher's LSD for data without outliers.
3. The variations in the significant levels affects the test of hypothesis, that is to say, it can reverse an already taken decision. It also affects the Fisher's result.
4. Increasing sample size of a set of data, completely affects the results of Fisher's, and MFisher's.

Contribution to Knowledge

The research have so far contributed to knowledge based in the following ways;

1. It has revealed to Statisticians and other researchers and cooperate bodies that, the mean is always affected by an outlier, while median is not easily affected by outliers.
2. The median can be used to calculate Variance, Standard deviation, sum of squares, and Fisher's test.
3. The proposed M-Fisher's is more robust than the Fisher's in the presence of outliers.

REFERENCES

Fisher, R. A.(1918). The correlation between relatives on the supposition of Mendelian inheritance. Transactions of the Royal Society of Edinburgh, 52, 399-433.

Harter, H. L.(1980). In: Vinod H, Rao C, eds. Hand Book of Statistics, Elsevier. 1, 617-622.

Iwundu, M. P. (2017). Principles of Experimental Design.CHADIK printing press, No 9 East West road, Port Harcourt, Nigeria.

Keller G. and Warrack B. (2003). Statistics for management and economics (6th ed): Thomson Brooks /Cole – 603-641.

Kutner M. H., Nschtsheim C. J., Neter J. and Li W. 2005. Applied linear statistical model, fifth edition, McGrawHilit:a Irwin, Boston Burr RIdge, IL Dubuque, IA MadIson, WI New York San FrancIsco St LoUIs Bangkok Bogota Caracas Kuala Lumpur LIsbon London

MadndMexIcoCItyMIlan Montreal New DeihlSantIago Seoul Smgapore Sydney TaIpeI Toronto

Midway, S., Robertson, M., Flinn, S., & Kally, M. (2020). Comparing multiple comparisons:

Practical guidance for choosing the best multiple comparisons test. Peer J8:e0387https:/doi.org/10.7717/Peer J.10387.

Onu, O. H., Owhondah, P. S., Nelson, M., (2021). Discriminating between second-order models with variation in model parameters based on central tendency estimation. International Journal of Science Academic Research, 2(5), 1478-1487.

Ramsey, P. H., & Ramsey, P. P. (2008). Power of pairwise comparisons in the equal variance and unequal sample size case. British Journal of Mathematical and Statistical

Sanchez, J., Caceres, J. R., & Rojas, J. P.(2019). Determination of expandability clay from San Josede Cucuta Metropolitan area by varying Calcium temperature and time. Journal of Physics conference series, 1388(1).

Sangseok, L., & Dong, K. L. (2018). What is the proper way to apply the multiple comparison test? Korean Journal Anesthesiol, 71(5), 353-360.

Zar, J. H.(2010). . Chennai, India: Pearson Education India.