ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

A STUDY OF THE EFFECT OF FOUR DIFFERENT MEDICINES ON THE PATIENTS WITH HIGH BLOOD PRESSURE USING ANALYSIS OF VARIANCE (ANOVA)

Aminu Suleiman*, Aliyu Ismail Ishaq, Abubakar Usman

* M. Sc. Statistics, Mathematics Department, Sharda University, Greater Noida-201306, India.
M. Sc. Statistics, Mathematics Department, Sharda University, Greater Noida-201306, India.
M. Sc. Statistics, Mathematics Department, Sharda University, Greater Noida-201306, India.

ABSTRACT

High blood pressure is a major determinant of risk for Coronary Heart Disease (CHD) and stroke, leading causes of death in anywhere of the world. However, the factors which increase blood pressure levels in human body are not well comprehended, but body fat is expected to be one of the major determinant of blood pressure level. This research carried out in Sharda University, to examine the effect of four different treatments taken by the patients with high blood pressure by using analysis of Variance (ANOVA). By using one way ANOVA in Minitab we have seen that the treatment are not equally effective. However, by using two factor ANOVA without replication in excel, it shows there is no statistically significant difference between effects of the treatments with respect to the patient age. It also affirm our one way ANOVA that there is a statistically significant difference between the treatments as treatment C happens to be less effective than others treatments. And we have seen that with respect to age the treatment are equally effective but regardless of age the treatments are not equally effective.

KEYWORDS:.

INTRODUCTION

The incidence, prevalence and consequences of elevated blood pressure has been extensively documented. The definition of elevated blood pressure (hypertension) has evolved over time; the present definition being a blood pressure greater than 140/90mmHg. Since the prevalence of hypertension in the western adult population exceeds 20%, this issue is very important. Elevated blood pressure greatly increases the risk of cardiovascular disease, the leading cause of death in most western countries. This topic has been explored extensively across many age groups and amongst people of different racial backgrounds for many years though the cause of hypertension remains unknown. Though body fat is thought to be a major casual factor of increased pressure, the true relationship between high blood pressure and body mass remains obscure.

Over time, the increased pressure can cause a wide range of problems. Small bulges, called aneurysms, may form in blood vessels. The heart can become enlarged, increasing the danger of heart failure. Damage to blood vessels in the kidneys can cause them to fail. Because tiny blood vessels in the eyes are especially vulnerable to damage, hypertension can lead to vision problems and even blindness.

Many factors can lead to high blood pressure. Clearly, diet plays a role. Too much salt, too little potassium, and too much alcohol have all been found to increase the risk of high blood pressure. Too much stress and too little physical activity both increase the danger of developing high blood pressure, as does being overweight or obese. And as with many chronic illnesses, high blood pressure also tends to run in families, suggesting that genetics plays a role.

This paper is aim at investigating the effect of four different medicines used by patients with high blood pressure (Hypertension) after 7 days. It is really important to know how treatments are reacting on the patients with high blood pressure, so that we can know the appropriate one to use and to have an insight view how strong does the medicines work on the patient of different age group.

MATERIAL AND METHODOLOGY

In this research we have two set of data, one the number of blood pressure both systolic and diastolic of the patients before taking the treatment, while the second set of data is the also the number of systolic and diastolic after taking the treatments for 7 days. Both the data are used for finding the effect of this medicines, Here we have the data of both systolic and diastolic blood pressure before taking the medication and after taking the medication, so to compute the analysis of variance by carefully looking at our goal of testing the effect of four different medicines, we first take the average of systolic and diastolic blood pressure of each treatment to represent the blood pressure in both two tables (i.e. the table of systolic and diastolic before taking the medications as well as after taking the medications). After this, we then proceed to find the difference of the averages of corresponding treatments in both tables.

We finally used our difference table to compute the analysis of variance so as to find the effects this treatments.

COMPUTATION OF THE ANALYSIS OF VARIANCE (ANOVA)

Here we consider the statistical analysis of the completely randomized block design or one way classification and also randomised block designs (two way classifications). In one way ANOVA, we shall supposed that the experiments has available the result of k independent random samples, from k different population (i.e. data concerning k treatments, k groups, k method of production etc.) and he or she is concerned with testing the hypothesis that the mean of these k samples are equal.

In general we denote the jth observation in the ith sampole by y_{ij}, and the scheme for one way ANOVA is classified as follows:

	Observations	**Means**
Sample 1	$y_{11}, y_{12}, \cdots, y_{1j}, \cdots, y_{1n_1}$	$\bar{y}_1$
Sample 2	$y_{21}, y_{22}, \cdots, y_{2j}, \cdots, y_{2n_2}$	$\bar{y}_2$
⋮	⋮⋮⋮	⋮
Sample i	$y_{i1}, y_{i2}, \cdots, y_{ij}, \cdots, y_{in_i}$	$\bar{y}_i$
⋮	⋮⋮⋮	⋮
Sample k	$y_{k1}, y_{k2}, \cdots, y_{kj}, \cdots, y_{kn_k}$	$\bar{y}_k$

To make the calculation simple we use the notation C. for the sum of all the observation and N for the total sample size.

$$C. = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij} \qquad N = \sum_{i=1}^{k} n_i$$

The overall sample mean $\bar{y}$ is

$$\bar{y} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{k} n_i \bar{y}_i}{\sum_{i=1}^{k} n_i} = \frac{T.}{N}$$

Sum square treatment = $SS_{Tr} = \sum_{i=1}^{k} n_i \left(\bar{y}_{i.} - \bar{y}\right)^2$

Degree of freedom for sum square treatment = k-1

Sum square error = $SS_E = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2 - Z$

Degree of freedom for sum square error = nt-k

Sum square total = $SS_T = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}..\right)^2$

Degree of freedom for sum square total = nt-1

ANOVA TABLE

Source of variation	Degree of Freedom (d.f)	Sum squares (SS)	Mean squares (MS)	F-statistics
Treatment	**k-1**	SS_{Tr}	$\frac{SS_{Tr}}{k-1} = MS_{Tr}$	$\frac{MS_{Tr}}{MS_E}$
Error	**nt-k**	SS_E	$\frac{SS_E}{nt-k} = MS_E$	
Total	**nt-1**	SS_T		

While in two way classification, we shall supposed that the experiments has available measurement pertaining to k treatments (medicines) distributed over n blocks (age group). We shall consider the case where there is exactly one observation from each treatment in each block. In this classification, we can see the data are classified into two criteria or factors. There is a little difference in the procedure followed for the analysis of variance in a two classification as compared to one way classification. The following is the layout used for the analysis of two way classification:

Age Group	Blocks	Means
	$A_1, A_2, \cdots, A_j, \cdots, A_{n_1}$	
Sample 1	$y_{11}, y_{12}, \cdots, y_{1j}, \cdots, y_{1n_1}$	$\bar{y}_{1.}$
Sample 2	$y_{21}, y_{22}, \cdots, y_{2j}, \cdots, y_{2n_2}$	$\bar{y}_{2.}$
⋮	⋮⋮⋮	⋮
Sample i	$y_{i1}, y_{i2}, \cdots, y_{ij}, \cdots, y_{in_i}$	$\bar{y}_{i.}$
⋮	⋮⋮⋮	⋮
Sample k	$y_{k1}, y_{k2}, \cdots, y_{kj}, \cdots, y_{kn_k}$	$\bar{y}_{k.}$
Means	$\bar{y}_{.1}, \bar{y}_{.2}, \cdots, \bar{y}_{.j}, \cdots, \bar{y}_{.n_k}$	$\bar{y}_{..}$

We also called this arrangement randomised block design, because our treatments are allocated at random. The sum squares for two way analysis of variance are given below:

Sum of square total (SST) = $SST - SSTr - SSBl$

Sum square treaments (SSTr) = $\dfrac{\sum_{i=1}^{k} X^2_{i.} - Z}{n}$

Sum squeare blocks (SSBl) = $\dfrac{\sum_{j=1}^{n} X^2_{.j} - Z}{k}$

The correction term Z is given by $Z = \dfrac{X_{..}}{nk}$

Sum square error (SSE) = $SST - SSTr - SSBl$

In the above formulas $\mathbf{X_{i.}}$ is the sum of the n observations for the ith treatments, $\mathbf{X_{.j}}$ is the sum of the kth observation in the jth block and $X_{..}$ is the grand total of all observations. The respective mean square error can be obtain by dividing each of the sum squares with their respective degrees of freedom. Now to find F ratio we use the following formulas:

$F_{Tr} = \dfrac{MSTr}{MSE}$, this is F ratio for treatments

$F_{Bl} = \dfrac{MSBl}{MSE}$, this is F ratio for blocks

The result obtained can be summarised in the following analysis of variance table:

ANOVA TABLE

Source of variation	Degree of Freedom (d.f)	Sum squares (SS)	Mean squares (MS)	F-statistics
Treatment	**k-1**	$\mathbf{SS_{Tr}}$	$\dfrac{SS_{Tr}}{k-1} = MS_{Tr}$	$\dfrac{MS_{Tr}}{MS_E}$ $\dfrac{MS_{Bl}}{MS_E}$
Blocks	**n-1**	$\mathbf{SS_{Bl}}$	$\dfrac{SS_{Bl}}{k-1} = MS_{Bl}$	
Error	**(n-1)(k-1)**	$\mathbf{SS_E}$	$\dfrac{SS_E}{nt-k} = MS_E$	
Total	**nk-1**	$\mathbf{SS_T}$		

HYPOTHESIS

Statement of hypothesis for analysis of variance (ANOVA)

H_o: There is no significant difference between the treatments

H_1: At least one treatment differs

Statement of hypothesis for two factor analysis of variance (ANOVA) without replication

H_o: There is no significant difference between the age group

H_1: At least one age group differs

RESULT

One-WayANOVA:A,B,C,D

Method

Null hypothesis	H_0: All means are equal
Alternative hypothesis	H_1: At least one mean is different

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	4	A, B, C, D

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	3	1200.11	400.036	3.39	0.0246
Error	52	6130.61	117.896		
Total	55	7330.71			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
10.8580	16.37%	11.55%	3.01%

Means

Factor	N	Mean	StDev	95% CI
A	14	13.143	10.734	(7.320, 18.966)
B	14	14.107	11.410	(8.284, 19.930)
C	14	2.393	10.472	(-3.430, 8.216)
D	14	8.643	10.794	(2.820, 14.466)

Pooled StDev = 10.8580

Grouping Information Using the Fisher LSD Method and 95% Confidence

Factor	N	Mean	Grouping	
B	14	14.107	A	
A	14	13.143	A	
D	14	8.643	A	B
C	14	2.393		B

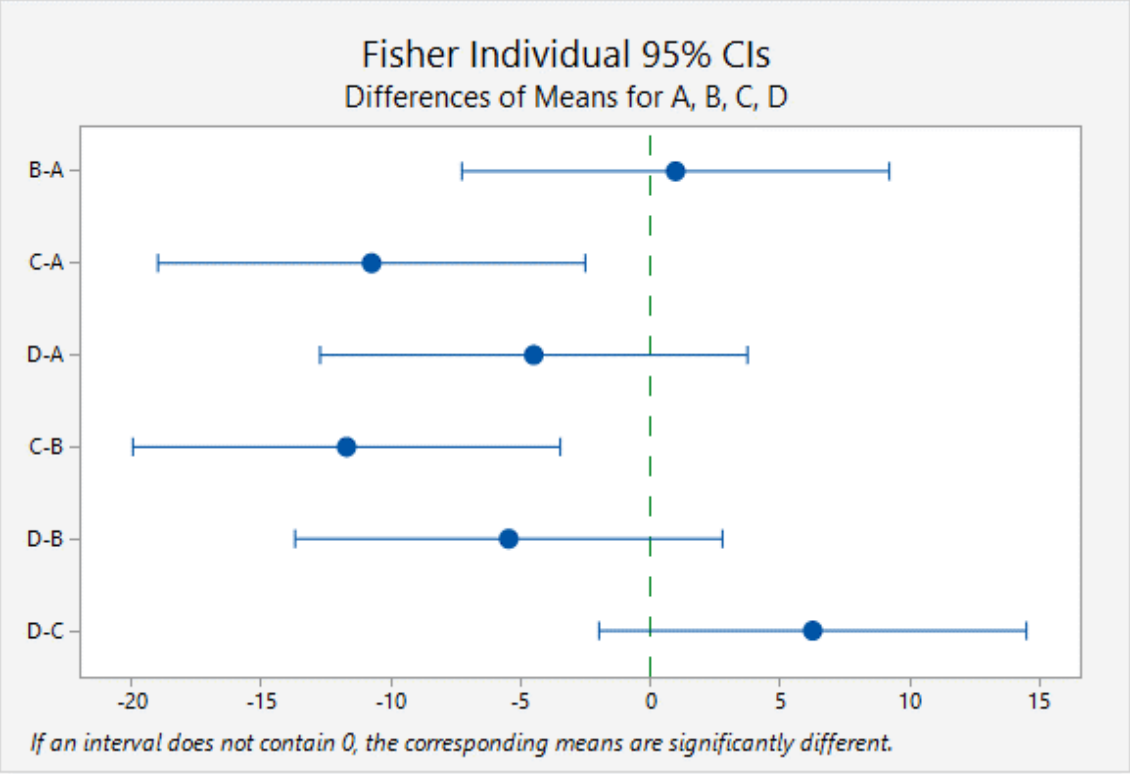
Means that do not share a letter are significantly different.

Fisher Individual Tests for Differences of Means

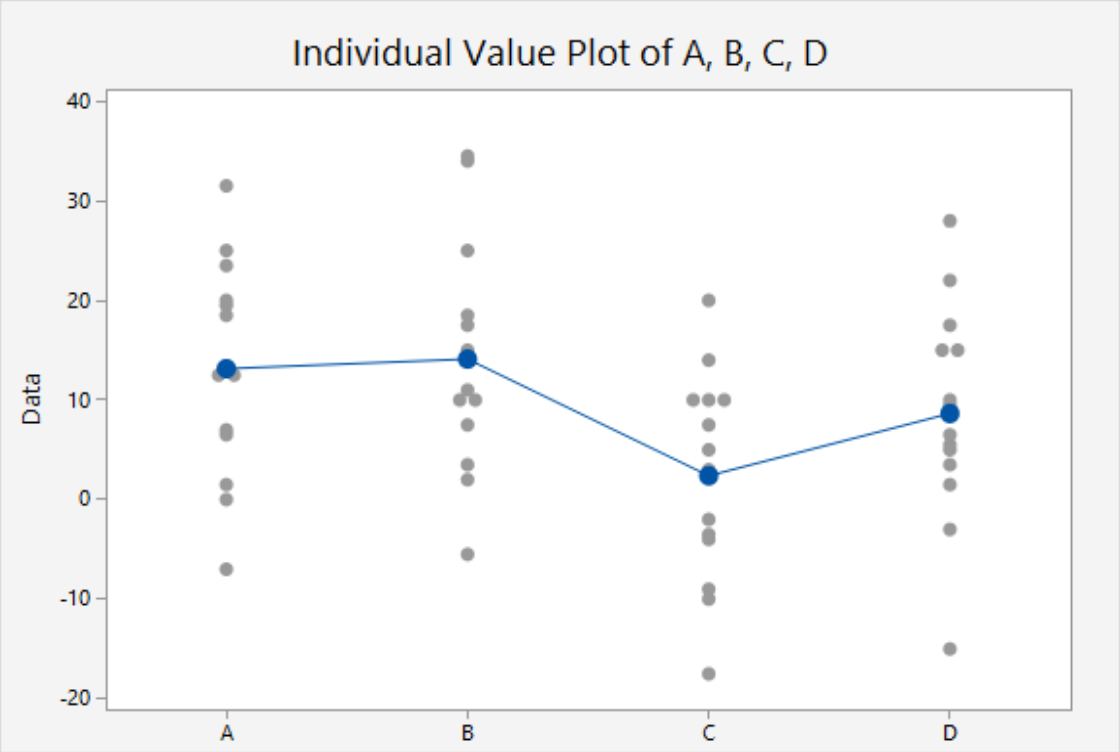
Difference of Levels	Difference of Means	SE of Difference	95% CI	T-Value	Adjusted P-Value
B-A	0.964	4.104	(-7.271, 9.199)	0.23	0.8152

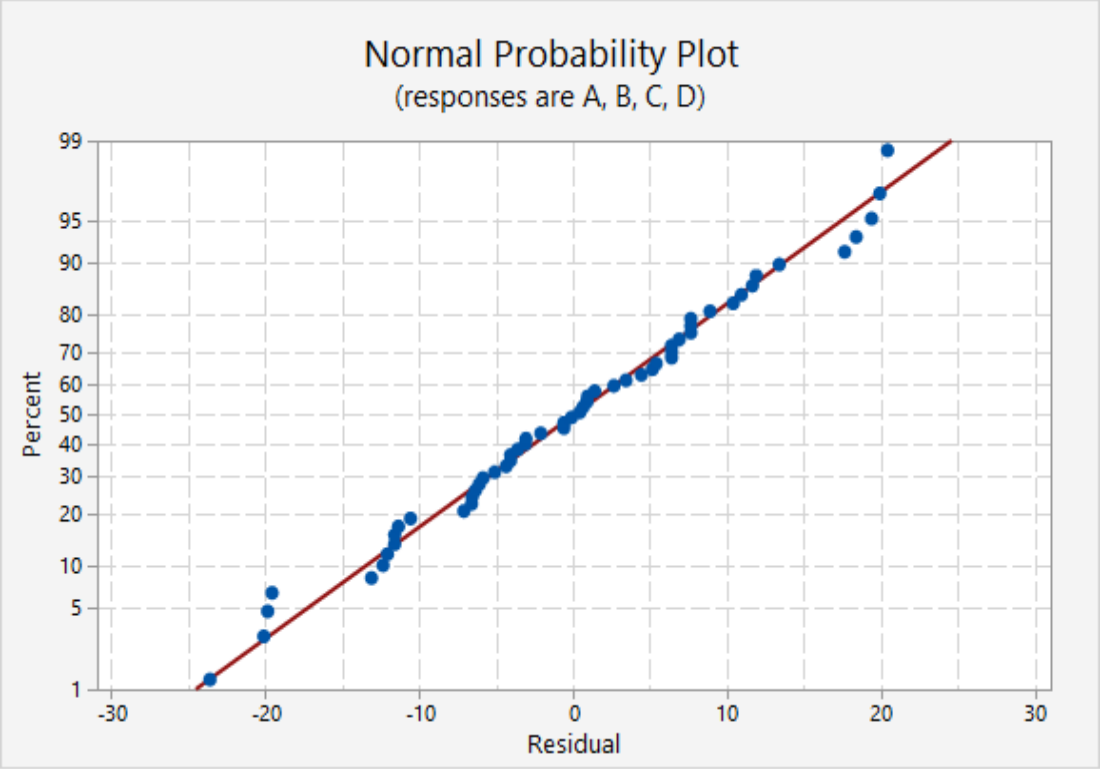
C-A	-10.750	4.104	(-18.985, 2.515)	-2.62	0.0115
D-A	-4.500	4.104	(-12.735, 3.735)	-1.10	0.2779
C-B	-11.714	4.104	(-19.949, 3.479)	-2.85	0.0062
D-B	-5.464	4.104	(-13.699, 2.771)	-1.33	0.1888
D-C	6.250	4.104	(-1.985, 14.485)	1.52	0.1338

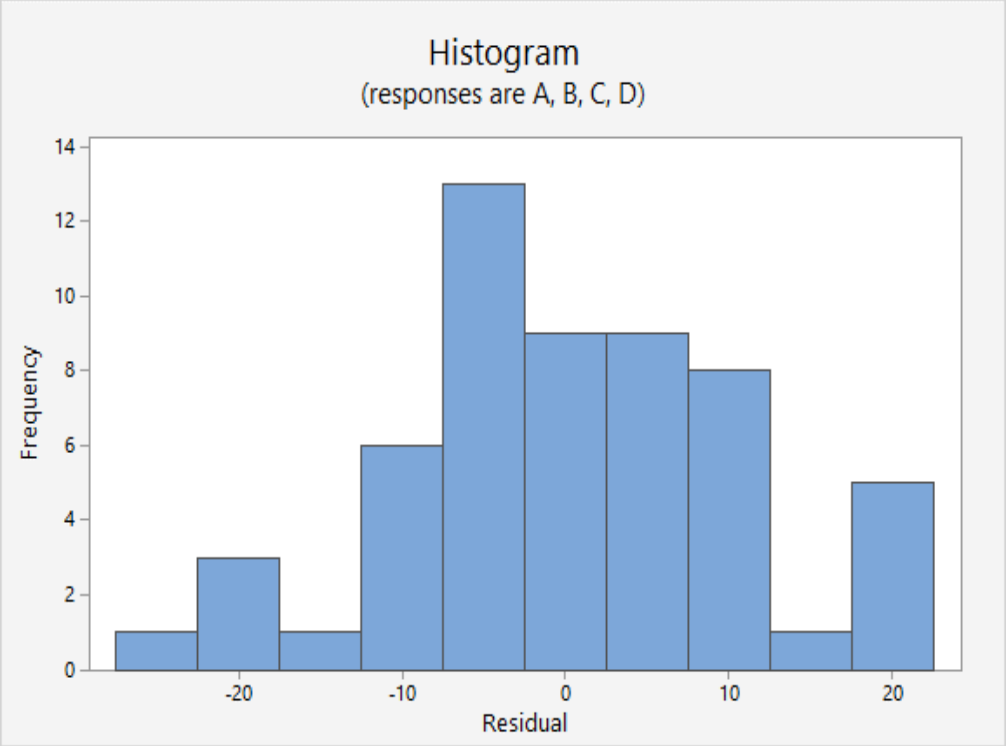
Simultaneous confidence level = 80.12%



ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785







TWO FACTOR ANOVA WITHOUT REPLICATION

Since we have age in the row and treatments in the column, now lets consider two factor ANOVA without replication (i.e. treatments and age) and see if the age is effective on the patients irrespective of their age or not.

Anova: Two-Factor without Replication

SUMMARY	*Count*	*Sum*	*Average*	*Variance*
21-25	4	12.5	3.125	18.72917
26-30	4	27.5	6.875	206.3958
31-35	4	38.5	9.625	52.22917
36-40	4	15	3.75	38.25
41-45	4	70	17.5	367.1667
46-50	4	11	2.75	87.75
51-55	4	41	10.25	68.25
56-60	4	55	13.75	54.08333
61-65	4	75.5	18.875	80.0625
66-70	4	26.5	6.625	96.39583
71-75	4	33.5	8.375	67.72917
76-80	4	58	14.5	17.66667
81-85	4	16.5	4.125	587.2292
86-90	4	55.5	13.875	196.0625
Diastolic A	14	184	13.14286	115.2088
Diastolic B	14	197.5	14.10714	130.1992
Diastolic C	14	33.5	2.392857	109.6607
Diastolic D	14	121	8.642857	116.5165

ANOVA TABLE						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	1516.714	13	116.6703	0.986183	0.481804	1.980528
Columns	1200.107	3	400.0357	3.381395	0.027626	2.845068
Error	4613.893	39	118.3049			
Total	7330.714	55				

DISCUSSION

Since the p-value is less than $\propto$=0.05 (i.e. 0.0246<0.05), we reject the null hypothesis and conclude that at least one means is different. Which means that there is difference between the treatments. The R-square adjusted shows 11.55%. Which means by knowing something about patients taking the treatments, it's explained that 11.55% of the variability in treatments. From Fisher LSD Method we have seen how the treatments differs from each other as we can see in the graph if an interval does not contain , the corresponding means are significantly different between treatment B and A, D and B, D and C all contains zero. So the corresponding means between the treatments are not significantly different but for C and A, C and B this intervals does not contain zero, so the corresponding means are significantly different.

In the way classification of ANOVA (Two-Factor without Replication), P-Value of rows (i.e. age) is showing 0.481804 which is greater than 0.05 that's the value of $\propto$. Hence we accept the null hypothesis and conclude that the treatments are equally effective as according to ages of individuals. So also the P-Value of columns (treatments) is 0.027626 which is also less than $\propto$ value 0.05, therefore we also reject the null hypothesis and conclude that at least one treatment is different as we see in the previous analysis of one way ANOVA by using MATLAB package.

CONCLUSION

In conclusion, we have one way ANOVA and two factor ANOVA without replication. Form the analysis of variance, both one way ANOVA and two factor ANOVA without replication, we have found that the treatments are not equally effective, but with regards to patients age, the treatments are equally effective. We carried out our analysis using Minitab and Excel, this means that there no statistically significant difference between the medicines with regard to age.

We therefore conclude:

1. The three (3) medicines A, B, D are equally effective in managing blood pressure while medicine C was found to be under performing.
2. Age has less influence on the effect of the medicines. That is all medicines A, B, C, D work effectively on all ages.

REFERENCES

[1] R. Johnson and D. Wichern. Applied Multivariate Statistical Analysis. Prentice-Hall, Inc., 5 edition, 2002.
[2] M. Woodward. Epidemiology: Study Design and Data Analysis. Chapman and Hall, 2 edition, 2005.
[3] P. Royston and D. Altman. Stata Technical Bullention, chapter sg26: Using fractional polynomials to model curved regression relationships, pages 11{23. Stata Corporation, September 1994.
[4] S. Weisberg. Applied Linear Regression. John Wiley and Sons Inc., 1 edition, 2005.
[5] Irfan Ali Khan and Atiya Khanum. Biostatistics for pharmacy. Ukaaz publication Hydrabad, second revised edition, 2008.
[6] SHAW AND GINA (2009). "*Prehypertension: Early-Stage High Blood Pressure*" WebMD http://WWW.webmd .com/content/article/73/88927.htm.
[7] SAGNELLA GA AND SWIFT PA (2006). "*The Renal Epithelial Sodium Channel*": Genetic Heterogeneity and implications for the treatment of high blood pressure. Current pharmaceutical design (14): 2221-2234.
[8] MULATERO P, BERTELLO C AND VERHOVEZ A (June 2009) "*Differential Diagnosis of Primary Aldosteronism Subtypes"* Current Hypertension Reports
[9] Mufunda J, Scott LJ, Chifamba J, Matenga J, Sparks B, Cooper R *et al*. Correlates of blood pressure in an urban Zimbabwean population and comparison to other populations of African origins. *J Hum Hypertens* 2000;**14**: 65–73. | Article | PubMed | ISI | ChemPort |
[10] Mufunda J, Chatora R, Ndambakuwa Y, Nyarango P, Kosia A, Chifamba J *et al*. Emerging non

communicable disease epidemic in Africa: preventive measures from the WHO Regional Office for Africa. *Ethn Dis* 2006; **16**: 521–526.

[11] M'buyamba Kabangu JR, Fagard R, Stassen J, Lijnen P, Amery A. Correlates of blood pressure in rural and urban Zaire. *J Hypertens* 1987; **5**: 371–375. | Article | PubMed | ChemPort |

[12] Mufunda J, Fink GD, Sparks HV. Blood pressure responses to dietary salt in rural and urban African men. *Ethn Dis* 1993; **3**: s46–s58. | PubMed