

PROPOSING A SOLUTION FOR THE PROBLEM OF CHOOSING ORTHOGONAL CONTRASTS FOR EQUAL SAMPLE SIZES

¹Asst Prof. Mohammed Qadoury Abed
Al-Mansour University College
mohammed.qadaury@muc.edu.iq

²Asst.lecturer.Ahmed Alwan Saleh
Al-Bani University College
ahmedalwan102@gmail.com

³Asst.lecturer.Elaf Bahaa Alwan
part_time.lecturer: Wasit University
leccit2@uowasit.edu.iq

⁴Prof. Abbas Lafta Kneehr
Wasit University- College of Administration & Economics
alafta@uowasit.edu.iq

Abstract

Many scholars as it is stated in the literature faced the problem of finding a suitable solution for finding and selecting orthogonal contrasts with $(k-1)$ number, in terms of having groups (samples, types and treatments) with (k) number, where $(\text{contrasts}=\text{treatments}-1)$. These are based on independence and orthogonalization as two conditions for selecting significant differences between (averages) treatments, and for identifying which one is responsible for these differences. The researchers solve this problem by proposing a method for selecting orthogonal contrasts in terms of equal observations, where independence and orthogonalization exist simply, to let the experimenter to test all averages with no need to test these two conditions if more averages are formulated or added.

Keywords: ANOVA, F-Test, Independence, p value, Stock Exchange, Orthogonal, Contrasts, Treatments, Trading, Rainfall, Annual.

Contribution/Orginality

The proposed method for simply selecting orthogonal contrasts in terms of equal observations where independence and orthogonalization exist is important; especially it is concerned with (k) of averages and treatment. The application of this method is useful in experimental design since it is applied before experimenting, and consequently new results and conclusions may be revealed.

1 - INTRODUCTION

After ANOVA, and if a statistically-significant difference among the squares of treatments, and to identify how these averages are different, two methods are usually followed based on

the time of these orthogonalities, namely before experimenting or after gathering data. One of the tests used before experimenting is orthogonal contrasts.

The literature is rich in studying orthogonal contrasts, especially after 1956. Recently, in 2004, Maria (4) examined these definitions and characteristics of these contrasts. In 2008, Al-Mashhadani & Abdulrazzaq (2) suggested new formulas for counting the values of orthogonal contrasts and the divided – on , and comparing these values with the common ones.

The researchers proposed a new method for solving the problem of selecting orthogonal contrasts with (k) of (averaged) treatments, where independence and orthogonalization as test's two conditions exist. This method works simply and smoothly which in turn gives the experimental the possibility of using it without testing these two conditions.

2 - ORTHOGONAL CONTRASTS⁽¹⁾⁽³⁾⁽⁵⁾⁽⁶⁾

If the decision taken is the selection of averages before experiments, it will be necessary for identifying the comparisons and for not preventing the existence of errors (α) which in turn represents the permitted error in ANOVA. This means the necessity of selecting contrasts carefully; it is also important to state that the number of contrasts is not exceeding the degree of freedom for averaged treatments. This method is called an orthogonal contrast.

Orthogonalization is defined as a linear combination for sums of treatments or averages, where this sum equals zero, and it is symbolized by C. Selecting orthogonal contrasts in terms of equal observations needs the existence of two conditions:

a- Independence, where the sum of parameters of treatments for a single orthogonality equals zero. That is:

$$\sum_{j=0}^k C_{jm} = 0 \dots \dots \dots (1)$$

Where k is the number of treatments (levels of a parameter), m the number of orthogonality (number of orthogonalities), and j=1,2,3,.....,k

b- Orthogonality , where the sum of multiplying the sums of corresponding parameters in each two contrasts equals zero. That is:

$$\sum_{j=1}^k C_{jm} C_{jq} = 0 \dots \dots \dots (2)$$

The sum of each contrast can be counted as follows:

$$C_m = \sum_{j=1}^k C_{jm} T_j \dots \dots \dots (3)$$

Where T_j is the sum of treatment j.

The sum of averages of each contrast can be counted as follows:

$$SSc = \frac{C_m^2}{n \sum_{j=1}^k C_{jm}^2} \dots \dots \dots (4)$$

Where n is the number of observations for each treatment.

3 - THE PROPOSED METHOD

After doing ANOVA for experiments and using F-Test for testing null hypothesis, and if F-Test states the differences or variances among the averaged groups (samples, types or treatments) are statistically significant. But, it does not identify which one of these differences or variances is significant. Therefore, the experimenter will examine which one of the averages is causing these differences or variances. Among these tests are orthogonal contrasts.

Experimenting is a technique used to state these contrasts; this in turn needs an experienced statistician, plus time and efforts. Sometimes results will be misleading or unsuccessful, in both cases leading to wrong orthogonal contrasts or none. Consequently, experimenting before testing will be avoided.

The researchers in this paper propose a solution to this problem for all k of treatments or averages. It is simply, flexibly and smoothly-applicable. It is also possible to have a large number of groups or tables of orthogonal contrasts. It gives the possibility of testing all averages the experimenter wants to examine or use, of course with no prior need to hold independence and orthogonality in forming a new contrast. The procedure of this proposed method is as follows:

- a- having a doubled table, with x -level for sums of treatments, and y -level for sequencing the contrasts, where (number of contrasts= number of treatments-1)
- b- for identifying the parameters of the first contrast, we weigh one of the treatments of the first contrast (one of the cells of the first row) with $(k-1)$, and the other treatments (or the rest of cells of the first row) with (-1)
- c- for the parameters of the second contrast, we use (0) instead of $(k-1)$. We weigh another cell in the second row with $(k-2)$, of course not the one weighed before. Other cells will be weighed with (-1) .
- d- we weigh the treatment with $(k-2)$ by (0) for other contrasts. A new treatment is weighed with $(k-3)$ and put it in a cell in the third row, of course except the weighed ones. Other treatments will be with (0) .
- e- we keep on the same strategy for weighing the other contrasts till the last one where the values of parameters is (1) and (-1) . Those weighed treatments will be (0) . This can be summed up in the following tables:

Table(1): Select m of orthogonal contrasts in case of k treatments

C.I.T.	T_1	T_2	T_3	T_{k-1}	T_k
C_1	$k-1$	-1	-1	-1	-1
C_2	0	$k-2$	-1	-1	-1
C_3 ↓	0 ↓	0 ↓	$k-3$ ↓	-1 ↓	-1 ↓
C_m	0	0	0	$k-m$	-1

Or,

Table(2): Select m of orthogonal contrasts in case of k treatments

C.I.T.	T ₁	T ₂	T ₃	T _{k-1}	T _k
C ₁	-1	-1	k-1	-1	-1
C ₂	-1	-1	0	-1	k-2
C ₃ ↓	k-3 ↓	-1 ↓	0 ↓	-1 ↓	0 ↓
C _m	0	-1	0	k-m	0

By the same procedure, it is possible to suggest many other tables including orthogonal contrasts with independence and orthogonalization for k of treatments or averages.

4 - Applications

The research included two applications. The first application if the number of treatments is small either the second application if the number of treatments is large.

4.1 - First Application

The application was based on real data taken from the annual reports of the Iraq Stock Exchange . The volume representing trading is sectorally divided for the Period (2008 – 2016). The researchers calculated the data (see : table 1 , Appendix) on the bases of annual reports. The volume of trading is the value of shares traded in the stock exchange. It equals the number of shares multiplied by the average price per share during a certain period . It is considered as one of the important indication of stock exchange in order to know of there are

significant differences among the sectors of the volume of trading and which average has contributed to these significant differences . ANOVA table was constituted and then we tested orthogonal contrasts in case if There are significant differences and as follows:

Table (3) : ANOVA table for volume of shares in the Iraqi Stock Exchange.

S.O.V	d.f	SS	MS	F	P
Bet.Treat.	6	8.556e+11	1.426e+11	19.951	0.00
Error	56	4.003e+11	7147535133		
Total	62	1.256e+12			

Comparing calculated F with tabled F, we notice that the value of calculated F is grater in value than tabled F. this means that there are significant differences among the averages (treatments). Later , we have structured a table for orthogonal contrasts and as follows :

Table(4): Selection of orthogonal contrasts of volume of shares by sector

T \ C	T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇
C ₁	6	-1	-1	-1	-1	-1	-1
C ₂	0	5	-1	-1	-1	-1	-1
C ₃	0	0	4	-1	-1	-1	-1
C ₄	0	0	0	3	-1	-1	-1
C ₅	0	0	0	0	2	-1	-1
C ₆	0	0	0	0	0	1	-1

From the table above , we notice that all orthogonal contrasts have the two conditions of independence and orthogonalization. The value of these contrasts was calculated. They were as follows:

Table (5) : The values of the orthogonal contrasts for volume of shares in the Iraqi Stock Exchange.

C ₁ = 17839721	C ₃ = - 782919	C ₅ = 592295
C ₂ = -810349	C ₄ = - 306271	C ₆ = 137185

The following table explains ANOVA table for orthogonal contrasts. The value of calculated F of orthogonal contrasts which were calculated from the following formula:

$$FC_m = \frac{MSC_m}{MSE} \dots \dots (5)$$

with the value of tabled F ($F_{1,56,0.95} = 4$). The asterisk (*) means that the value of calculated F is greater than tabled F and that there is a significant difference among the averages at this contrast and according to its hypothesis.

Table(6) : ANOVA table for orthogonal contrasts for volume of shares in the Iraqi Stock Exchange.

S.O.V	d.f	SS	MS	F
C ₁	1	8.4194615e+11	8.4194615e+11	117.795*
C ₂	1	2432094451	2432094451	0.340
C ₃	1	3405345336	3405345336	0.476
C ₄	1	868536346.6	868536346.6	0.122
C ₅	1	6496543833	6496543833	0.909
C ₆	1	1045540234	1045540234	0.146
Error	56	4.003e+11	7147535133	
Total	62			

We can structure many other groups of orthogonal contrasts which have the two conditions of independence and orthogonalization and as explained in the following tables.

Table(7): Selection of orthogonal contrasts of volume of shares by sector

T \ C	T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇
C ₁	-1	6	-1	-1	-1	-1	-1
C ₂	-1	0	-1	-1	-1	-1	5
C ₃	-1	0	-1	4	-1	-1	0
C ₄	3	0	-1	0	-1	-1	0
C ₅	0	0	-1	0	-1	2	0
C ₆	0	0	1	0	-1	0	0

Or :

Table(8): Selection of orthogonal contrasts of volume of shares by sector

T \ C	T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇
C ₁	-1	-1	-1	-1	2	-1	-1
C ₂	-1	-1	-1	6	0	-1	-1
C ₃	1	-1	-1	0	0	-1	-1
C ₄	0	4	-1	0	0	-1	-1
C ₅	0	0	-1	0	0	5	-1
C ₆	0	0	-1	0	0	0	3

And so on

4.2 - Second Application

We applied real data taken from the statistical annual giving by Iraqi Ministry of Planning – the central statistical body for the year 2016. These data represent the annual quantity of rain (mm) as realized by the governorats from 2008 at 2011 . We would like to know if there are significant differences among the governorats in the annual quantity of rain and also which of these averages have contributed to these significant difference. Later , we have to structure ANOVA table and then to test orthogonal contrasts in case there are significant difference and as follows:

Table (9) : ANOVA table for the total annual rainfall in Iraq.

S.O.V	d.f	SS	MS	F	P
Bet.Treat.	17	1061537.291	62443.370	36.980	0.00
Error	54	91183.615	1688.585		
Total	71	1152720.906			

Comparing the value of calculated F with that of the tabled F, we notice that the value of calculated F is greater than tabled F. this only means that there are significant differences among the averages (treatments). Accordingly, we structure a table for orthogonal contrasts and as follows:

Table(10) : Selection of orthogonal contrasts for the total annual rainfall in Iraq by governorates

	T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇	T ₈	T ₉	T ₁₀	T ₁₁	T ₁₂	T ₁₃	T ₁₄	T ₁₅	T ₁₆	T ₁₇	T ₁₈
C ₁	17	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
C ₂	0	16	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
C ₃	0	0	15	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
C ₄	0	0	0	14	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
C ₅	0	0	0	0	13	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
C ₆	0	0	0	0	0	12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
C ₇	0	0	0	0	0	0	11	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
C ₈	0	0	0	0	0	0	0	10	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
C ₉	0	0	0	0	0	0	0	0	9	-1	-1	-1	-1	-1	-1	-1	-1	-1
C ₁₀	0	0	0	0	0	0	0	0	0	8	-1	-1	-1	-1	-1	-1	-1	-1
C ₁₁	0	0	0	0	0	0	0	0	0	0	7	-1	-1	-1	-1	-1	-1	-1
C ₁₂	0	0	0	0	0	0	0	0	0	0	0	6	-1	-1	-1	-1	-1	-1
C ₁₃	0	0	0	0	0	0	0	0	0	0	0	0	5	-1	-1	-1	-1	-1
C ₁₄	0	0	0	0	0	0	0	0	0	0	0	0	0	4	-1	-1	-1	-1
C ₁₅	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	-1	-1	-1
C ₁₆	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	-1	-1
C ₁₇	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1

Note: we can structure other tables that have the tow conditions of Independence and orthogonalization. From the table above, we notice that all orthogonal contrasts have the two conditions of independence and rthgonalization. the values of these Contrasts are as follows:

Table(11) : The values of the orthogonal contrasts for the total annual rainfall in Iraq.

$C_1 = 17547.2$	$C_7 = -260$	$C_{13} = -170.6$
$C_2 = 8580.5$	$C_8 = -18$	$C_{14} = -357.6$
$C_3 = 21689.9$	$C_9 = -451$	$C_{15} = -375.2$
$C_4 = 6209.9$	$C_{10} = -289$	$C_{16} = -229.4$
$C_5 = 10722.1$	$C_{11} = 1500.6$	$C_{17} = 251.4$
$C_6 = 5063.2$	$C_{12} = 520.6$	

The following table explaining ANOVA table of orthogonal contrasts. Were the value calculated F is compared with that of tabled F.

Table(12) : ANOVA table for orthogonal contrasts for the total annual rainfall in Iraq

S.O.V	d.f	SS	MS	F
C ₁	1	251555.75	251555.75	148.974*
C ₂	1	67670.02	67670.02	40.075*
C ₃	1	490053.92	490053.92	290.216*
C ₄	1	45908.16	45908.16	27.187*
C ₅	1	157916.80	157916.80	93.520*
C ₆	1	41083.32	41083.32	24.330*
C ₇	1	128.03	128.03	0.076
C ₈	1	0.7364	0.7364	0.00044
C ₉	1	565.003	565.003	0.335
C ₁₀	1	290.003	290.003	0.175
C ₁₁	1	10052.68	10052.68	5.953*
C ₁₂	1	1613.24	1613.24	0.955
C ₁₃	1	242.54	242.54	0.144
C ₁₄	1	1598.47	1598.47	0.9466
C ₁₅	1	2932.81	2932.81	1.737
C ₁₆	1	2192.68	2192.68	1.299
C ₁₇	1	7900.245	7900.245	4.679*
Error	54	91183.615	1688.585	
Total	71			

5 - Conclusions

a - The solution (the proposed method) is flexible, easy, simple and rapid . It is considered to be a successful and practical solution for a rather difficult problem that is almost possible for choosing orthogonal contrasts that have the two conditions of dependence and orthogonalization

b - The proposed method could be considered as an important addition and a contribution in the field of statistics and experimental design

c - The simplicity for choosing and testing the orthogonal contrasts when sample Sizes are equal whether the number of treatments is small (first application) Or great (second application).

d - The other researchers can develop this method to cover the case of unequal sample sizes.

e - If we rely or the P value instead of the F value for testing the significant difference , we will find that the P value for the two applications is equal to zero. Because, it is less than the significant level 0.01 we reject the hypothesis of the null significant differences among the averages (treatments) and we accept the alternative hypothesis. Accordingly , the effect of the factor is to considered significant. In order to know which of these averages has caused this significant difference , we conduct the test of orthogonal contrasts.

References

- 1 – Abdi, H. , Edelman, B., valentine, D., & Dowling, w.j.,(2009) Exponential Design and Analysis for Psychology . Oxford: Oxford University Press.
- 2 – Al – Mashhadani, K.A. & Abdulrazzaq, A.S. (2008) " Suggested formulas For calculating the orthogonal contrasts values and the values of the divisor" Journal of Administrative and Economic Sciences , Baghdad, vol.14, no.51, pp(230 – 236).
- 3 - Doncastor, c.p. & Davey,A.J.H.(2007), Analysis of Variance and Covariance How To choose and Construct Models for the Life Science . Combridge Combridge University Press .
- 4 – Maria, C.S.N., (2004) "Orthogonal Contrasts: definitions and concepts. Sei.agrie. (piraccicaba,Braz.), vol.61,No.1.
- 5– Philip.C,Tino.S.,&Mario,S.,(2016), Experimentall Design Research, Siproinger International Publishing , Switzerland.
- 6 – Robert, G.E., (2015) Fundamentals of Statistical Experimental Design And Analysis , John Wiley & Sons , Ltd., New Maxico, U.S.A.

Appendix

Table (1) : A table representing the volume of trading of shares by sectors in the Iraqi Stock Exchange for the period (2008 – 2016).

years	Banks	Insurance	Investment	Services	Industry	Hot. & Tourism	Agriculture
2008	91872	194	384	1902	17379	16035	179
2009	279696	373	1368	14550	61241	8968	613
2010	115988	523	2219	5968	18191	3305	697
2011	407175	203	1285	3360	9774	4949	618
2012	267156	503	6512	2683	16147	13667	540
2013	319235	1025	3461	14543	29934	42635	1093
2014	259411	2293	1141	20606	65530	50217	1158
2015	705455	7830	3149	48946	126222	42560	7033
2016	677332	2031	942	27065	94727	29254	62474

Table (2) : The total annual rainfall (mm) in Iraq by governorates for the period (2008 – 2011)

Governo.	2008	2009	2010	2011	Governo.	2008	2009	2010	2011
Dohouk	370.1	487.2	293.4	434.0	Wasit	76	31.1	84.5	98.2
Ninevah	216.3	223.8	240.6	294.7	Kerbala	87.6	85.3	80.3	120.3
Sulaimaniya	381.2	607.9	385.2	507.4	Salah Al-Deen	158.8	111.7	138.2	104.8
Kirkuk	134.9	225.8	267.2	221.8	Al- Najaf	72.4	64.3	50.3	71.3
Erbil	297	311.6	261.8	301.6	Al- Qadisiya	44.2	46.2	49.1	81.4
Diala	197.9	164.7	206.9	167.2	Al- Muthanna	57.0	54.1	47.0	58.4
Al- anbar	72.9	23.3	109.0	87.9	Thi Qar	65.5	56.9	57.6	85.1
Baghdad	59.1	67.5	92.5	96.0	Maysan	90.6	175.9	128.3	110.7
babylon	51.8	52.4	87.3	80.3	Basrah	67.1	89.8	31.9	65.3