

# Chapter 4

---

## Latin Square Design





## Latin Square Design

Felix Kutsanedzie<sup>1\*</sup>; Sylvester Achio<sup>1</sup>;  
Edmund Ameko<sup>1</sup>; George Kutsanedzie<sup>2</sup>;  
Diaba Kwasi Selassie<sup>3</sup>

<sup>1</sup>Accra Polytechnic, GP 561, Accra, Ghana

<sup>2</sup>Project Office of Electricity Company of Ghana, Accra, Ghana

<sup>3</sup>Anglican University College of Technology, Nkorazan, Sunyani, Ghana

### Abstract

The Latin Square Design (LSD) is one of the known experimental designs used in analysis designed experiments. Though this design exists, most researchers are unfamiliar with its usage. The prime aim of analysis experiment using the various known experimental designs is to reduce error as much as possible, or eliminate them altogether. This paper thus explain the concept of how the LSD is used for analysing experiment as well as how it helps to reduce errors due to the differences that exist in two directions (rows and columns) often referred to as double blocking.

### Keywords

Design, Double Blocking, Experiment, Analysis

## 4.1 Introduction

In an RCBD, blocking is used to place all treatments in groups that are similar or homogenous so as to reduce the error due to variations as much as possible among the various treatments being handled. In RCBD there is single blocking adopted – only treatments are blocked. However with the Latin Square Design (LSD) blocking is done – blocking by row and column. The LSD is set such that each treatment is found in both a row and a column. Thus the design typically assumes a square shape. The number of experimental units is denoted by  $n^2$ , where  $n$  is the number of treatments or blocks or replications. It presupposes that in LSD, the number of blocks or replication used must be equal to the number of treatments being handled in order for each treatment to be found in the each block.

If a researcher needs to analysis an experiment using the Latin Square Design, the treatments must be arranged in the rows and columns in such a way that the major sources come from them.

Below are the types of LSD, and they are limited by the number of treatments.

3 X 3 LSD			
A		C	B
B		A	C
C		B	A

  

4X4 LSD			
A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

5X5 LSD				
A	B	C	D	E
B	A	E	C	D
C	D	A	E	B
D	E	B	A	C
E	C	D	B	A

It can be seen from the designs above that the treatments are all arranged in such a way that each column and row contain all the treatments be experimented on. Thus the number and type of treatments in a row are equal to those in a column.

Assuming a researcher wants to extract milk from five different plants such Tiger nut (T), coconut (Co), soyabean (So), cashew nut (Ca) and shea nut (Sh) to ascertain whether there are significance differences between the volume of milk in cubic centimetres that can be extracted from a kilogramme of each of the fruits; the following 5 x 5 Latin Square Design can be drawn:

5X5 LSD					
Milk Extracted in cm <sup>3</sup> /kg of treatment					
	Column1	Column2	Column3	Column4	Column5
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3
Row2	Co = 5.5	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.9
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2
Row5	Sh = 3.1	Ca = 4.8	So = 4.0	Co = 5.4	T = 5.6

*The first step: Stating of hypothesis*

For Treatments:

H<sub>0</sub>: All the means of the treatments are equal

$H_A$ : At least one of the means of the treatments (t) is different

$$H_0: \mu_{trt_A} = \mu_{trt_B} = \mu_{trt_C} = \mu_{trt_D} = \mu_{trt_E}$$

$$H_1: \mu_{trt_A} \neq \mu_{trt_B} = \mu_{trt_C} = \mu_{trt_D} = \mu_{trt_E}$$

For Columns:

$H_0$ : All the means of the Columns (C) are equal

$H_A$ : At least one of the means of the Columns (C) is different

$$H_0: \mu_{C1} = \mu_{C2} = \mu_{C3} = \mu_{C4} = \mu_{C5}$$

$$H_1: \mu_{C1} \neq \mu_{C2} = \mu_{C3} = \mu_{C4} = \mu_{C5}$$

For Rows:

$H_0$ : All the means of the Rows (R) are equal

$H_A$ : At least one of the means of the Rows (R) is different

$$H_0: \mu_{R1} = \mu_{R2} = \mu_{R3} = \mu_{R4} = \mu_{R5}$$

$$H_1: \mu_{R1} \neq \mu_{R2} = \mu_{R3} = \mu_{R4} = \mu_{R5}$$

*The second step: Calculating the Correction Factor (CF)*

$$GT = \sum_{i=1, j=1}^n (T, Co, Ca, So, Sh) = \text{sum of all observations on experimental units}$$

$$GT = \sum_{i=1, j=1}^{n=5} (5.3 + 5.0 + \dots + 5.6) = 108.6$$

$$CF = \frac{(GT)^2}{N}$$

$$CF = \frac{(108.6)^2}{25} = \frac{11793.96}{25} = 471.76$$

where  $GT$  = grand Total,  $N$  = total number of observation,  $CF$  = Correction factor.

The third step: Calculating the Total Sum of Squares ( $TSS$ )

$$TSS = \sum_{i=1, j=1}^{n=5} [(T_{1,1})^2 + (Co_{1,2})^2 + \dots (T_{5,5})^2] - CF$$

$$TSS = \sum_{i=1, j=1}^{n=5} [(5.3)^2 + (5.0)^2 + \dots (5.6)^2] - 471.96$$

$$TSS = 494.74 - 471.96 = 22.98$$

CT	22	21.3	22.1	20.1	23.1	108.6
RT	21.6	21.6	20.7	21.8	22.9	108.6
$T_T$	5.3	5	4.9	4.5	5.6	25.3
$Co_T$	5.5	5	6	5.4	5.9	27.8
$Ca_T$	3.9	4.8	4.1	3.9	4.2	20.9
$So_T$	4.2	3.6	4	3.9	4.1	19.8
$Sh_T$	3.1	2.9	3.1	2.4	3.3	14.8

The fourth step: Calculating the Treatment Sum of Squares ( $TrtSS$ )

$$Trt SS = \sum \left( \frac{(T_T)^2 + (Co_T)^2 + (Ca_T)^2 + (So_T)^2 + (Sh_T)^2}{t} \right) - CF$$

$$Trt SS = \sum \left( \frac{(25.3)^2 + (27.8)^2 + (20.9)^2 + (19.8)^2 + (14.8)^2}{5} \right) - 471.96$$

$$Trt SS = \sum \left( \frac{640.09 + 772.84 + 436.81 + 392.04 + 219.04}{5} \right) - 471.96$$

$$Trt SS = \sum \left( \frac{2460.82}{5} \right) - 471.96$$

$$Trt SS = 492.16 - 471.96 = 20.41$$

where  $t$ =number of treatments,  $T_T$ =Treatment T total,  $Sh_T$ =Treatment Sh total,  $Co_T$ =Treatment Co total,  $Ca_T$ =Treatment Ca total,  $So_T$ =Treatment So total.

The sixth step: Calculating the Rows Sum of Squares (RSS)

$$RSS = \sum \left( \frac{(R1_T)^2 + (R2_T)^2 + (R3_T)^2 + (R4_T)^2 + (R5_T)^2}{R} \right) - CF$$

$$RSS = \sum \left( \frac{(21.6)^2 + (21.6)^2 + (20.7)^2 + (20.1)^2 + (23.1)^2}{5} \right) - 471.96$$

$$RSS = \sum \left( \frac{466.56 + 466.56 + 428.49 + 475.24 + 524.41}{5} \right) - 471.96$$

$$RSS = \sum \left( \frac{2361.26}{5} \right) - 471.96$$

$$RSS = 472.25 - 471.96$$

$$RSS = 0.49$$

where  $R$ =number of rows,  $R1_T$ =Treatment Row 1 total,  $R2_T$ =Treatment Row 2 total,  $R3_T$ =Treatment Row 3 total,  $R4_T$ =Treatment Row 4 total,  $R5_T$ =Treatment Row 5 total.

The seventh step: Calculating the Column Sum of Squares (CSS)

$$CSS = \sum \left( \frac{(C1_T)^2 + (C2_T)^2 + (C3_T)^2 + (C4_T)^2 + (C5_T)^2}{R} \right) - CF$$

$$CSS = \sum \left( \frac{(22)^2 + (21.3)^2 + (22.1)^2 + (20.1)^2 + (23.1)^2}{R} \right) - 471.96$$

$$CSS = \sum \left( \frac{(484) + (453.69) + (488.41) + (404.01) + (533.61)}{5} \right) - 471.96$$

$$CSS = \sum \left( \frac{2363.72}{5} \right) - 471.96$$

$$CSS = 472.74 - 471.96$$

$$CSS = 0.99$$



where  $C$ =number of rows,  $C1_T$ =Treatment Column 1 total,  $C2_T$ =Treatment Column 2 total,  $C3_T$ =Treatment Column 3 total,  $C4_T$ =Treatment Column 4 total,  $C5_T$ =Treatment Column 5 total.

*The eighth step: Calculating the Error Sum of Squares (ESS)*

$$ESS = TSS - (TrtSS + RSS + CSS)$$

$$ESS = 22.98 - (20.41 + 0.49 + 0.99)$$

$$ESS = (22.98 - 21.89)$$

$$ESS = 1.1$$

where  $ESS$ = Error Sum of Squares or unexplained sum of squares.

*The ninth step: Completing the ANOVA table*

ANOVA TABLE						
Sources of Variation	df	SS	MS	Fcal.	Fcrit. (5%)	
Treatment	$t - 1 = 5 - 1 = 4$	20.41	5.1025	46.39	3.48	
Row	$R - 1 = 5 - 1 = 4$	0.49	0.1225	1.11	3.48	
Column	$C - 1 = 5 - 1 = 4$	0.99	0.2475	2.25	3.48	
Error	$[C \times R - 1] - [(t-1)+(R-1)+(C-1)]$ $=24-16 =10$	1.1	0.11			
Total	$(C \times R) - 1 = 25 - 1 =24$	22.98				
Fcrit (5%) @ df of 4,10	= 3.48					
Fcrit (1%) @ df of 4,10	= 5.99					

*Tenth step: looking up the F-critical table for the critical values*

$$F_{crit}(1\%) \text{ at } df \text{ of } 4,10 \text{ for Treatments} = 5.99$$

$$F_{crit}(1\%) \text{ at } df \text{ of } 4,10 \text{ for Rows} = 5.99$$

$$F_{crit}(1\%) \text{ at } df \text{ of } 4,10 \text{ for Columns} = 5.99$$

$$F_{crit}(5\%) \text{ at } df \text{ of } 4,10 \text{ for Treatments} = 3.48$$

$F_{crit}(5\%)$  at  $df$  of 4,10 for Rows = 3.48

$F_{crit}(5\%)$  at  $df$  of 4,10 for Columns = 3.48

These values for the treatments, rows and columns can be obtained from the F-critical tables below and the values highlighted.

**Critical values of F for the 0.05 significance level:**

	1	2	3	4	5	6
1	161.45	199.50	215.71	224.58	230.16	233.99
2	18.51	19.00	19.16	19.25	19.30	19.33
3	10.13	9.55	9.28	9.12	9.01	8.94
4	7.71	6.94	6.59	6.39	6.26	6.16
5	6.61	5.79	5.41	5.19	5.05	4.95
6	5.99	5.14	4.76	4.53	4.39	4.28
7	5.59	4.74	4.35	4.12	3.97	3.87
8	5.32	4.46	4.07	3.84	3.69	3.58
9	5.12	4.26	3.86	3.63	3.48	3.37
10	4.97	4.10	3.71	3.48	3.33	3.22
11	4.84	3.98	3.59	3.36	3.20	3.10
12	4.75	3.89	3.49	3.26	3.11	3.00
13	4.67	3.81	3.41	3.18	3.03	2.92
14	4.60	3.74	3.34	3.11	2.96	2.85

**Critical values of F for the 0.01 significance level:**

	1	2	3	4	5	6
1	4052.19	4999.52	5403.34	5624.62	5763.65	5858.97
2	98.50	99.00	99.17	99.25	99.30	99.33
3	34.12	30.82	29.46	28.71	28.24	27.91
4	21.20	18.00	16.69	15.98	15.52	15.21
5	16.26	13.27	12.06	11.39	10.97	10.67
6	13.75	10.93	9.78	9.15	8.75	8.47
7	12.25	9.55	8.45	7.85	7.46	7.19
8	11.26	8.65	7.59	7.01	6.63	6.37
9	10.56	8.02	6.99	6.42	6.06	5.80
10	10.04	7.56	6.55	5.99	5.64	5.39
11	9.65	7.21	6.22	5.67	5.32	5.07
12	9.33	6.93	5.95	5.41	5.06	4.82
13	9.07	6.70	5.74	5.21	4.86	4.62
14	8.86	6.52	5.56	5.04	4.70	4.46

*Eleventh Step: Making the decisions and Conclusions.*

Treatments

$$F_{cal}(4,10) = 46.39 > F_{crit}(3.48) \text{ at } 5\%$$

$$F_{cal}(4,10) = 46.39 > F_{crit}(5.99) \text{ at } 1\%$$

Since the F-calculated value (46.39) for the treatments at both 1% and 5% level of significance respectively are greater than the F-critical value (5.99) and

(3.48), the null hypothesis is rejected on the grounds that there is enough evidence to suggest that there is significant difference between the treatments. Hence we fail to reject the alternate hypothesis.

#### Columns

$$F_{cal} (4,10) = 2.25 < F_{crit} (3.48) \text{ at } 5\%$$

$$F_{cal} (4,10) = 2.25 < F_{crit} (5.99) \text{ at } 1\%$$

F calculated value (2.25) for the columns at both 1% and 5% level of significance respectively lesser than the f critical value (3.48) and (5.99), we fail to reject the null hypothesis and say there is no significant difference in the columns, hence no justification for blocking by columns.

#### Rows

$$F_{cal} (4,10) = 1.11 < F_{crit} (3.48) \text{ at } 5\%$$

$$F_{cal} (4,10) = 1.11 < F_{crit} (5.99) \text{ at } 1\%$$

F calculated value (1.11) for the rows at both 1% and 5% level of significance respectively lesser than the F-critical value (3.48) and (5.99), we fail to reject the null hypothesis and say there is no significant difference in the rows, hence no justification for blocking by rows.

Since there is no justification for blocking by rows and columns, it presupposes that the LSD is not the right design for the experiment in question. Since the treatments are significantly different, there is the need to perform the Lsd (Least significance difference) test or the Duncan Multiple Range Test (DMRT).

#### Handling Missing Data under Latin Square Design (LSD).

In every experimental design data can be missing destruction or damage of living things being used as experimental materials, improper allocation of

treatments, loss of materials due to harvesting and processing of materials and illogical data that cannot be considered as reliable results due to how extreme the values are.

For single missing data under the LSD, the formula below can be used:

$$X_o = \left[ \left( \frac{t(R_o + C_o + T_o) - 2G_o}{(t-1)(t-2)} \right) \right]$$

$X_o$ =missing data observed,  $R_o$ =total observed row values with missing data,  $C_o$ =total observed column values with missing data,  $T_o$ =total observed treatment values with missing data,  $G_o$ =grand total of all observed value,  $t$ =number of treatments.

Formula for the Computation of Bias Value

$$\beta_o = \frac{[G_o - R_o - C_o - T_o(t-1)]^2}{[(t-1)(t-2)]^2}$$

$\beta_o$ = bias value

For adjustment of the analysis based on the estimation of the missing data the following is done

1. The value one (1) is subtracted from the total degree of freedom and the unexplained or error degree of freedom.
2. The biased value must be computed and subtracted from the Total Sum of Squares and Treatment Sum of Squares and not that of the column and the row.

Taking the 5 x 5 LSD table below, the missing that can be computed for as follows:

5X5 LSD					
Milk Extracted in cm <sup>3</sup> /kg of treatment					
	Column1	Column2	Column3	Column4	Column5
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3
Row2	Co = Y <sub>21</sub>	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.9
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2
Row5	Sh = 3.1	Ca = 4.8	So = 4.0	Co = 5.4	T = 5.6

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = Y <sub>21</sub>	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	16.1
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.9	20.7
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = 4.8	So = 4.0	Co = 5.4	T = 5.6	22.9
Column Total	16.5	21.3	22.1	20.1	23.1	103.1

$$T_o(Co) = Co_{12} + Co_{35} + Co_{43} + Co_{54}$$

$$T_o(Co) = 5.0 + 6.0 + 5.4 + 5.9 = 22.3$$

$$G_o = \sum_{ij}^{n=5} Y_{ij} = Y_{11} + Y_{12} + \dots + Y_{55} = 5.3 + 5.0 + \dots + 5.6 = 103.1$$

$$R_o = Y_{22} + Y_{23} + \dots + Y_{25} = 5.0 + 3.1 + 3.9 + 4.1 = 16.1$$

$$C_o = Y_{11} + Y_{31} + \dots + Y_{51} = 5.3 + 3.9 + 4.2 + 3.1 = 16.5$$

$$Y_{21} = \left[ \left( \frac{t(R_o + C_o + T_o) - 2G_o}{(t-1)(t-2)} \right) \right]$$

$$Y_{21} = \left[ \left( \frac{5(16.1 + 16.5 + 22.3) - 206.2}{(5-1)(5-2)} \right) \right]$$

$$Y_{21} = \left[ \left( \frac{5(54.4) - 206.2}{12} \right) \right]$$

$$Y_{21} = \left[ \left( \frac{274.5 - 206.2}{12} \right) \right]$$

$$X_o = \left[ \left( \frac{68.3}{12} \right) \right]$$

$$X_o = [5.69]$$

The bias is now computed using the formula:

$$\beta_o = \frac{[G_o - R_o - C_o - T_o(t-1)]^2}{[(t-1)(t-2)]^2}$$

$$\beta_o = \frac{[103.1 - 16.1 - 16.5 - 22.3(5-1)]^2}{[(5-1)(5-2)]^2}$$

$$\beta_o = \frac{[103.1 - 16.1 - 16.5 - 22.3(4)]^2}{[(4)(3)]^2}$$

$$\beta_o = \frac{[103.1 - 16.1 - 16.5 - 89.2]^2}{[12]^2}$$

$$\beta_o = \frac{[-18.7]^2}{[12]^2}$$

$$\beta_o = \frac{[349.69]^2}{[144]^2}$$

$$\beta_o = \frac{[349.69]}{[144]}$$

$$\beta_o = 2.43$$

Now the table can be completed with the estimated missing data in order to do the Analysis of Variance (Anova) table

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.9	20.7
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = 4.8	So = 4.0	Co = 5.4	T = 5.6	22.9
Column Total	22.19	21.3	22.1	20.1	23.1	108.79
<b>CT</b>	<b>22.19</b>	<b>21.3</b>	<b>22.1</b>	<b>20.1</b>	<b>23.1</b>	<b>108.79</b>
RT	21.6	21.79	20.7	21.8	22.9	108.79
T <sub>T</sub>	5.3	5	4.9	4.5	5.6	25.3
Co <sub>T</sub>	5.69	5	6	5.4	5.9	27.99
Ca <sub>T</sub>	3.9	4.8	4.1	3.9	4.2	20.9
So <sub>T</sub>	4.2	3.6	4	3.9	4.1	19.8
Sh <sub>T</sub>	3.1	2.9	3.1	2.4	3.3	14.8

*The first step: Stating of the Hypothesis.*

The hypothesis stated remains the same.

*The second step: Calculating the Correction Factor (CF)*

$GT = \sum_{i=1}^n \sum_{j=1}^n (T, Co, Ca, So, Sh) =$  sum of all observations on experimental units

$$GT = \sum_{i=1, j=1}^{n=5} (5.3 + 5.0 + \dots + 5.6) = 108.79$$

$$CF = \frac{(GT)^2}{N}$$

$$CF = \frac{(108.79)^2}{25} = \frac{11835.26}{25} = 473.41$$

where  $GT$ = Grand Total,  $N$ =total number of observation,  $CF$ =Correction factor.

The third step: Calculating the Total Sum of Squares (TSS)

$$TSS = \sum_{i=1, j=1}^{n=5} [(T_{1,1})^2 + (Co_{1,2})^2 + \dots (T_{5,5})^2] - CF$$

$$TSS = \sum_{i=1, j=1}^{n=5} [(5.3)^2 + (5.0)^2 + \dots (5.6)^2] - 473.41$$

$$TSS = 496.87 - 473.41 = 23.46$$

The fourth step: Calculating the Treatment Sum of Squares (TrtSS)

$$Trt SS = \sum \left( \frac{(T_T)^2 + (Co_T)^2 + (Ca_T)^2 + (So_T)^2 + (Sh_T)^2}{t} \right) - CF$$

$$Trt SS = \sum \left( \frac{(25.3)^2 + (27.99)^2 + (20.9)^2 + (19.8)^2 + (14.8)^2}{5} \right) - 473.41$$

$$Trt SS = \sum \left( \frac{640.09 + 783.44 + 436.81 + 392.04 + 219.04}{5} \right) - 473.41$$

$$Trt SS = \sum \left( \frac{2471.42}{5} \right) - 473.41$$

$$Trt SS = 494.28 - 473.41 = 20.87$$

where  $t$ =number of treatments,  $T_T$ =Treatment T total,  $Sh_T$ =Treatment Sh total,  $Co_T$ =Treatment Co total,  $Ca_T$ =Treatment Ca total,  $So_T$ =Treatment So total.

The sixth step: Calculating the Rows Sum of Squares (RSS)

$$RSS = \sum \left( \frac{(R1_T)^2 + (R2_T)^2 + (R3_T)^2 + (R4_T)^2 + (R5_T)^2}{R} \right) - CF$$

$$RSS = \sum \left( \frac{(21.6)^2 + (21.79)^2 + (20.7)^2 + (21.8)^2 + (22.9)^2}{5} \right) - 473.41$$

$$RSS = \sum \left( \frac{466.56 + 474.80 + 428.49 + 475.24 + 524.41}{5} \right) - 473.41$$



$$\begin{aligned}
 RSS &= \sum \left( \frac{2369.50}{5} \right) - 473.41 \\
 RSS &= 473.90 - 473.41 \\
 RSS &= 0.49
 \end{aligned}$$

where  $R$ =number of rows,  $R1_T$ =Treatment Row 1 total,  $R2_T$ =Treatment Row 2 total,  $R3_T$ =Treatment Row 3 total,  $R4_T$ =Treatment Row 4 total,  $R5_T$ =Treatment Row 5 total.

*The seventh step: Calculating the Column Sum of Squares (CSS)*

$$\begin{aligned}
 CSS &= \sum \left( \frac{(C1_T)^2 + (C2_T)^2 + (C3_T)^2 + (C4_T)^2 + (C5_T)^2}{R} \right) - CF \\
 CSS &= \sum \left( \frac{(22.19)^2 + (21.3)^2 + (22.1)^2 + (20.1)^2 + (23.1)^2}{R} \right) - 473.41 \\
 CSS &= \sum \left( \frac{(492.40) + (453.69) + (488.41) + (404.01) + (533.61)}{5} \right) - 473.41 \\
 CSS &= \sum \left( \frac{2372.12}{5} \right) - 473.41 \\
 CSS &= 474.42 - 473.41 \\
 CSS &= 1.01
 \end{aligned}$$

where  $C$ =number of rows,  $C1_T$ =Treatment Column 1 total,  $C2_T$ =Treatment Column 2 total,  $C3_T$ =Treatment Column 3 total,  $C4_T$ =Treatment Column 4 total,  $C5_T$ =Treatment Column 5 total.

*The eighth step: Calculating the Error Sum of Squares (ESS)*

$$\begin{aligned}
 ESS &= TSS - (TrtSS + RSS + CSS) \\
 ESS &= 23.46 - (20.87 + 0.49 + 1.01) \\
 ESS &= (23.46 - 22.38) \\
 ESS &= 1.1
 \end{aligned}$$

where  $ESS = \text{Error Sum of Squares or unexplained sum of squares}$ .

*The ninth step: Making the necessary Adjustments in the Analysis and Completing the ANOVA table*

*Making the necessary Adjustments in the Analysis*

The following adjustments must be done in completing the Anova table:

1. Deduct the value (1) from the total degree of freedom and the unexplained or error degree of freedom
2. Subtract the computed bias from the Total Sum of Squares (TSS) and Treatment Sum of Squares (TrtSS)

*Adjustments*

$$\text{Total } df = [C \times R] - 1 - 1 = 25 - 1 - 1 = 23$$

$$\begin{aligned} \text{Unexplained or error} &= [C \times R - 1] - [(t - 1) + (R - 1) + (C - 1)] - 1 \\ &= 24 - 16 - 1 = 9 \end{aligned}$$

$$\text{Adjusted TSS} = 23.45 - \beta_0$$

$$\text{where } \beta_0 = 2.43$$

$$\text{Adjusted TSS} = 23.45 - 2.43 = 21.02$$

$$\text{Adjusted TrtSS} = 20.87 - \beta_0$$

$$\text{Adjusted TrtSS} = 20.87 - 2.43 = 18.44$$

ANOVA TABLE					
Sources of Variation	df	SS	MS	Fcal.	Fcrit. (5%)
Treatment	4	18.44	4.61	38.78	3.63
Row	4	0.5	0.13	1.05	3.63
Column	4	1.01	0.25	2.12	3.63
Error	9	1.07	0.12		
Total	23	21.02			
Fcrit (5%) @ df of 4,9	3.63				
Fcrit (1%) @ df of 4,9	6.42				

*Tenth step: looking up the F-critical table for the critical values*

$$F_{crit}(1\%) \text{ at } df \text{ of } 4,9 \text{ for Treatments} = 6.42$$

$$F_{crit}(1\%) \text{ at } df \text{ of } 4,9 \text{ for Rows} = 6.42$$

$$F_{crit}(1\%) \text{ at } df \text{ of } 4,9 \text{ for Columns} = 6.42$$

$$F_{crit}(5\%) \text{ at } df \text{ of } 4,9 \text{ for Treatments} = 3.63$$

$$F_{crit}(5\%) \text{ at } df \text{ of } 4,9 \text{ for Rows} = 3.63$$

$$F_{crit}(5\%) \text{ at } df \text{ of } 4,9 \text{ for Columns} = 3.63$$

*The Eleventh step: Making the Decision and Conclusions*

The values obtained from the F-critical table compared with the F-values calculated for the Treatments, Rows and Columns reveals the decisions and conclusions remain the same as the table without the missing data. This proves the missing data value has been accurately estimated.

## 4.2 Handling More than One Missing Data Under LSD

The procedure is the same as done in the case of RCBD but the formula for estimating the missing data values differ. For LSD, the following formula are used:

$$Y_{ij} = \frac{(T_m + C_m + R_m)}{2}$$

This formula can apply only to LSD when estimating average missing data where  $Y_{ij}$ =estimated average missing data value,  $T_m$ =the mean value of treatment with the missing data,  $C_m$ =the mean value of Column with the missing data,  $R_m$ =the mean value of Row with the missing data.

$$Y_{ij} = \left[ \left( \frac{t(R_o + C_o + T_o) - 2G_o}{(t - 1)(t - 2)} \right) \right]$$

$Y_{ij}$ =missing data observed,  $R_o$ =total observed row values with missing data,  $C_o$ =total observed column values with missing data,  $T_o$ =total observed treatment values with missing data,  $G_o$ =grand total of all observed value,  $t$ =number of treatments.

$$\beta_0 = \frac{[G_o - R_o - C_o - T_o(t - 1)]^2}{[(t - 1)(t - 2)]^2}$$

$\beta_0$ = bias value

For example to compute for the missing data under LSD as seen in the table above, the approach below is used:

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = Y <sub>35</sub>	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = Y <sub>52</sub>	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	17.2	98.09

We first find the average estimated missing data using the formula:

$$Y_{ij} = \frac{(T_m + C_m + R_m)}{3}$$

$$\text{For } Y_{52} = \frac{(T_m + C_m + R_m)}{3}$$

$$C_m = \frac{\sum Y_{11} + Y_{22} + Y_{32} + Y_{42}}{5} = \frac{16.5}{5} = 3.3$$

$$R_m = \frac{\sum Y_{51} + Y_{53} + Y_{54} + Y_{55}}{5} = \frac{18.1}{5} = 3.62$$

$$T_m = \frac{\sum Y_{31} + Y_{13} + Y_{24} + Y_{45}}{5}$$

$$T_m = \frac{\sum(3.9 + 4.1 + 3.9 + 4.2)}{5} = \frac{16.1}{5} = 3.22$$

$$Y_{52} = \frac{(3.22 + 3.3 + 3.62)}{3} = \frac{10.14}{3} = 3.38$$

$$\text{For } Y_{35} = \frac{(T_m + C_m + R_m)}{2}$$

$$C_m = \frac{\sum Y_{15} + Y_{25} + Y_{45} + Y_{55}}{5} = \frac{17.2}{5} = 3.44$$

$$R_m = \frac{\sum Y_{31} + Y_{32} + Y_{33} + Y_{34}}{5} = \frac{14.8}{5} = 2.96$$

$$T_m = \frac{\sum Y_{21} + Y_{12} + Y_{43} + Y_{54}}{5} = \frac{5.69 + 5.0 + 6.0 + 5.4}{5} = 4.42$$

$$Y_{35} = \frac{(4.42 + 3.44 + 2.96)}{2} = \frac{10.82}{2} = 3.61$$

Since the estimated average values of the missing data  $Y_{52}$  and  $Y_{35}$  have been computed, the values of each of this missing data can now be estimated by completing the table with one of the average estimated value where finding the estimated value of the other one.

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 3.61	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = Y <sub>52</sub>	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	20.81	101.7

In order to estimate the value of Y<sub>52</sub>, the average estimated value of Y<sub>35</sub> is substituted in the table as seen.

$$Y_{52} = \left[ \left( \frac{t(R_0 + C_0 + T_0) - 2G_0}{(t-1)(t-2)} \right) \right]$$

$$R_0 = Y_{51} + Y_{53} + \dots + Y_{55}$$

$$R_0 = 3.1 + 4.0 + 5.4 + 5.6 = 18.1$$

$$C_0 = Y_{12} + Y_{22} + \dots + Y_{52}$$

$$C_0 = 5.0 + 5.0 + 3.6 + 2.9 = 16.5$$

$$T_0(Ca) = Y_{51} + Y_{53} + \dots + Y_{55}$$

$$T_0(Ca) = 3.9 + 4.1 + 3.9 + 4.2 = 16.1$$

$$Y_{52} = \left[ \left( \frac{5(18.1 + 16.5 + 16.1) - 2(101.7)}{(5-1)(5-2)} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{5(50.7) - 203.4}{(4)(3)} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{253.5 - 203.4}{12} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{50.1}{12} \right) \right]$$

$$Y_{52} = [4.18]$$

To compute the estimate for  $Y_{35}$ , substitute the value of  $Y_{52} = 4.18$  into the table as shown:

5X5 LSD						
Milk Extracted in $\text{cm}^3/\text{kg}$ of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = $Y_{35}$	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = 4.18	So = 4.0	Co = 5.4	T = 5.6	22.28
Column Total	22.19	20.68	22.1	20.1	17.2	102.27

$$Y_{35} = \left[ \left( \frac{t(R_0 + C_0 + T_0) - 2G_0}{(t-1)(t-2)} \right) \right]$$

$$R_0 = Y_{31} + Y_{32} + \dots + Y_{34}$$

$$R_0 = 3.9 + 3.6 + 4.9 + 2.4 = 14.8$$

$$C_0 = Y_{15} + Y_{25} + \dots + Y_{55}$$

$$C_0 = 3.3 + 4.1 + 4.2 + 5.6 = 17.2$$

$$T_0(Ca) = Y_{21} + Y_{12} + \dots + Y_{54}$$

$$T_0(Ca) = 5.69 + 5.0 + 6.0 + 5.4 = 22.09$$

$$Y_{35} = \left[ \left( \frac{5(14.8 + 17.2 + 22.09) - 2(102.27)}{(5-1)(5-2)} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{5(54.09) - 2(102.27)}{(4)(3)} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{270.45 - 204.54}{12} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{65.91}{12} \right) \right]$$

$$Y_{35} = [5.50]$$

The estimated missing data for  $Y_{35} = 5.50$ . This is replaced in the table and the value of  $Y_{52}$  is estimated again. This step is repeated until a constant value is obtained for all the missing data before the accurate missing data values are obtained.

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.50	20.30
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = $Y_{52}$	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	22.70	103.9

$$Y_{52} = \left[ \left( \frac{t(R_o + C_o + T_o) - 2G_o}{(t-1)(t-2)} \right) \right]$$

$$R_o = 16.5$$

$$C_o = 18.1$$

$$T_o(Ca) = 3.9 + 4.1 + 3.9 + 4.2 = 16.1$$

$$Y_{52} = \left[ \left( \frac{5(16.5 + 18.1 + 16.1) - 2(103.9)}{(5-1)(5-2)} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{5(50.7) - 2(103.9)}{(4)(3)} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{253.5 - 207.8}{12} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{45.70}{12} \right) \right]$$

$$Y_{52} = [3.81]$$



Since the value for the previous estimated of  $Y_{52} = 4.18$  is not the same the 3.81 as obtained now, the process must be repeated. We then substitute the value of  $Y_{52} = 3.81$  into the table to estimate for  $Y_{35}$

$$Y_{35} = \left[ \left( \frac{t(R_o + C_o + T_o) - 2G_o}{(t-1)(t-2)} \right) \right]$$

$$R_o = 14.8$$

$$C_o = 17.2$$

$$T_o(C_o) = 22.09$$

$$Y_{35} = \left[ \left( \frac{5(14.8 + 17.2 + 22.09) - 2(101.27)}{(5-1)(5-2)} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{5(54.09) - 2(101.27)}{(4)(3)} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{270.45 - 207.8}{12} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{67.91}{12} \right) \right]$$

$$Y_{35} = [5.66]$$

Substitute  $Y_{35} = 5.66$  into the table and compute for  $Y_{52}$

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.66	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = $Y_{52}$	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	17.2	103.75

$$Y_{52} = \left[ \left( \frac{t(R_o + C_o + T_o) - 2G_o}{(t-1)(t-2)} \right) \right]$$

$$R_o = 18.1$$

$$C_o = 16.5$$

$$T_o(Ca) = 16.1$$

$$Y_{52} = \left[ \left( \frac{5(18.1 + 16.5 + 16.1) - 2(103.75)}{(5 - 1)(5 - 2)} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{5(50.7) - 2(103.75)}{(4)(3)} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{253.5 - 207.5}{12} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{46}{12} \right) \right]$$

$$Y_{52} = [3.83]$$

Substitute  $Y_{52} = 3.83$  into the table and compute for  $Y_{35}$

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = $Y_{35}$	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = 3.83	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	20.33	22.1	20.1	17.2	103.9

$$Y_{35} = \left[ \left( \frac{t(R_o + C_o + T_o) - 2G_o}{(t - 1)(t - 2)} \right) \right]$$

$$R_o = 14.8$$

$$C_o = 17.2$$

$$T_o(Co) = 22.09$$

$$Y_{35} = \left[ \left( \frac{5(14.8 + 17.2 + 22.09) - 2(101.92)}{(5 - 1)(5 - 2)} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{5(54.09) - 2(101.92)}{(4)(3)} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{270.45 - 203.84}{12} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{66.61}{12} \right) \right]$$

$$Y_{35} = [5.55]$$

Substitute  $Y_{35} = 5.55$  into the table and compute for  $Y_{52}$

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.55	20.35
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = $Y_{52}$	So = 4.0	Co = 5.4	T = 5.6	18.1
Column Total	22.19	16.5	22.1	20.1	22.75	103.64

$$Y_{52} = \left[ \left( \frac{t(R_o + C_o + T_o) - 2G_o}{(t-1)(t-2)} \right) \right]$$

$$R_o = 18.1$$

$$C_o = 16.5$$

$$T_o(C_o) = 16.1$$

$$Y_{52} = \left[ \left( \frac{5(18.1 + 16.5 + 16.1) - 2(103.64)}{(5-1)(5-2)} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{5(50.7) - 2(103.64)}{(4)(3)} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{253.50 - 207.28}{12} \right) \right]$$

$$Y_{52} = \left[ \left( \frac{46.22}{12} \right) \right]$$

$$Y_{52} = [3.85]$$

Substitute  $Y_{52} = 5.55$  into the table and compute for  $Y_{35}$

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = $Y_{52}$	14.8
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = 3.85	So = 4.0	Co = 5.4	T = 5.6	21.95
Column Total	22.19	20.35	22.1	20.1	22.75	101.94

$$Y_{35} = \left[ \left( \frac{t(R_o + C_o + T_o) - 2G_o}{(t-1)(t-2)} \right) \right]$$

$$R_o = 14.8$$

$$C_o = 17.2$$

$$T_o(C_o) = 22.09$$

$$Y_{35} = \left[ \left( \frac{5(14.8 + 17.2 + 22.09) - 2(103.64)}{(5-1)(5-2)} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{5(54.09) - 2(101.94)}{(4)(3)} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{270.45 - 203.88}{12} \right) \right]$$

$$Y_{35} = \left[ \left( \frac{66.57}{12} \right) \right]$$

$$Y_{35} = [5.55]$$

Since the  $Y_{35} = 5.55$  is the same as obtained previously, then it presupposes that the right missing data has been estimated for  $Y_{35}$ . Therefore the estimated missing data values for  $Y_{52}$  and  $Y_{35}$  are 3.85 and 5.55 respectively.

5X5 LSD						
Milk Extracted in cm <sup>3</sup> /kg of treatment						
	Column1	Column2	Column3	Column4	Column5	Row Total
Row1	T = 5.3	Co = 5.0	Ca = 4.1	So = 3.9	Sh = 3.3	21.6
Row2	Co = 5.69	T = 5.0	Sh = 3.1	Ca = 3.9	So = 4.1	21.79
Row3	Ca = 3.9	So = 3.6	T = 4.9	Sh = 2.4	Co = 5.55	20.35
Row4	So = 4.2	Sh = 2.9	Co = 6.0	T = 4.5	Ca = 4.2	21.8
Row5	Sh = 3.1	Ca = 3.85	So = 4.0	Co = 5.4	T = 5.6	21.95
Column Total	22.19	20.35	22.1	20.1	22.75	107.49
<b>CT</b>	<b>22.19</b>	<b>20.35</b>	<b>22.1</b>	<b>20.1</b>	<b>22.75</b>	<b>107.49</b>
RT	21.6	21.79	20.35	21.8	21.95	107.49
T <sub>T</sub>	5.3	5	4.9	4.5	5.6	25.3
Co <sub>T</sub>	5.69	5	6	5.4	5.55	27.64
Ca <sub>T</sub>	3.9	3.85	4.1	3.9	4.2	19.95
So <sub>T</sub>	4.2	3.6	4	3.9	4.1	19.8
Sh <sub>T</sub>	3.1	2.9	3.1	2.4	3.3	14.8

*The first step: Stating of the Hypothesis*

The hypothesis stated remains the same.

*The second step: Calculating the Correction Factor (CF)*

$GT = \sum_{i=1}^n \sum_{j=1}^n (T, Co, Ca, So, Sh) = \text{sum of all observations on experimental units}$

$$GT = \sum_{i=1, j=1}^{n=5} (5.3 + 5.0 + \dots + 5.6) = 107.49$$

$$CF = \frac{(GT)^2}{N}$$

$$CF = \frac{(107.49)^2}{25} = \frac{11554.1}{25} = 462.16$$

where  $GT =$  Grand Total,  $N =$  total number of observation,  $CF =$  Correction factor.

The third step: Calculating the Total Sum of Squares (TSS)

$$TSS = \sum_{i=1, j=1}^{n=5} [(T_{1,1})^2 + (Co_{1,2})^2 + \dots (T_{5,5})^2] - CF$$

$$TSS = \sum_{i=1, j=1}^{n=5} [(5.3)^2 + (5.0)^2 + \dots (5.6)^2] - 462.16$$

$$TSS = 484.64 - 462.16 = 22.48$$

The fourth step: Calculating the Treatment Sum of Squares (TrtSS)

$$Trt SS = \sum \left( \frac{(T_T)^2 + (Co_T)^2 + (Ca_T)^2 + (So_T)^2 + (Sh_T)^2}{t} \right) - CF$$

$$Trt SS = \sum \left( \frac{(25.3)^2 + (27.64)^2 + (19.95)^2 + (19.8)^2 + (14.8)^2}{5} \right) - 462.16$$

$$Trt SS = \sum \left( \frac{640.09 + 763.97 + 398.0 + 392.04 + 219.04}{5} \right) - 462.16$$

$$Trt SS = \sum \left( \frac{2413.14}{5} \right) - 462.16$$

$$Trt SS = 482.63 - 462.16 = 20.47$$

where  $t =$  number of treatments,  $T_T =$  Treatment T total,  $Sh_T =$  Treatment Sh total,  $Co_T =$  Treatment Co total,  $Ca_T =$  Treatment Ca total,  $So_T =$  Treatment So total.

The sixth step: Calculating the Rows Sum of Squares (RSS)

$$RSS = \sum \left( \frac{(R1_T)^2 + (R2_T)^2 + (R3_T)^2 + (R4_T)^2 + (R5_T)^2}{R} \right) - CF$$

$$RSS = \sum \left( \frac{(21.6)^2 + (21.79)^2 + (20.35)^2 + (21.8)^2 + (21.95)^2}{5} \right) - 462.16$$

$$RSS = \sum \left( \frac{466.56 + 474.80 + 414.12 + 475.24 + 481.80}{5} \right) - 462.16$$

$$RSS = \sum \left( \frac{2312.53}{5} \right) - 462.16$$

$$RSS = 462.51 - 462.16$$

$$RSS = 0.35$$

where  $R$ =number of rows,  $R1_T$ =Treatment Row 1 total,  $R2_T$ =Treatment Row 2 total,  $R3_T$ =Treatment Row 3 total,  $R4_T$ =Treatment Row 4 total,  $R5_T$ =Treatment Row 5 total.

The seventh step: Calculating the Column Sum of Squares (CSS)

$$CSS = \sum \left( \frac{(C1_T)^2 + (C2_T)^2 + (C3_T)^2 + (C4_T)^2 + (C5_T)^2}{R} \right) - CF$$

$$CSS = \sum \left( \frac{(22.19)^2 + (20.35)^2 + (22.1)^2 + (20.1)^2 + (22.75)^2}{R} \right) - 462.16$$

$$CSS = \sum \left( \frac{(492.40) + (414.12) + (488.41) + (404.01) + (517.56)}{5} \right) - 462.16$$

$$CSS = \sum \left( \frac{2316.50}{5} \right) - 462.16$$

$$CSS = 463.30 - 462.16$$

$$CSS = 1.14$$

where  $C$ =number of rows,  $C1_T$ =Treatment Column 1 total,  $C2_T$ =Treatment Column 2 total,  $C3_T$ =Treatment Column 3 total,  $C4_T$ =Treatment Column 4 total,  $C5_T$ =Treatment Column 5 total.

The eighth step: Calculating the Error Sum of Squares (ESS)

$$ESS = TSS - (TrtSS + RSS + CSS)$$

$$ESS = 22.48 - (20.47 + 0.35 + 1.14)$$

$$ESS = 22.48 - 21.96)$$

$$ESS = 0.52$$

where  $ESS$ = Error Sum of Squares or unexplained sum of squares.

Now the bias can be computed using the formula:

Calculating bias using Ca as the missing data

$$\beta_o = \frac{[G_o - R_o - C_o - T_o(t - 1)]^2}{[(t - 1)(t - 2)]^2}$$

$\beta_o =$  bias value

$$G_o = 103.64, C_o = 16.5, R_o = 18.1, T_o = 16.1$$

$$\beta_o = \frac{[103.64 - 18.1 - 16.5 - 16.1(5 - 1)]^2}{[(5 - 1)(5 - 2)]^2}$$

$$\beta_o = \frac{[103.64 - 34.6 - 64.4]^2}{[(4)(3)]^2}$$

$$\beta_o = \frac{[103.64 - 99]^2}{[(12)]^2}$$

$$\beta_o = \frac{[4.64]^2}{[12]^2}$$

$$\beta_o = \frac{[21.53]}{[144]} = 0.15$$

Calculating bias using Co as the missing data:

$$\beta_o = \frac{[G_o - R_o - C_o - T_o(t - 1)]^2}{[(t - 1)(t - 2)]^2}$$

$\beta_o =$  bias value

$$G_o = 101.94, C_o = 17.2, R_o = 14.8, T_o = 22.09$$

$$\beta_o = \frac{[101.94 - 14.8 - 17.2 - 22.09(5 - 1)]^2}{[(5 - 1)(5 - 2)]^2}$$

$$\beta_o = \frac{[101.94 - 32 - 88.36]^2}{[(4)(3)]^2}$$

$$\beta_o = \frac{[101.94 - 20.36]^2}{[(12)]^2}$$

$$\beta_o = \frac{[-18.42]^2}{[12]^2}$$



$$\beta_0 = \frac{[339.30]}{[144]} = 2.36$$

*The ninth step: Making the necessary Adjustments in the Analysis and Completing the ANOVA table.*

*Making the necessary Adjustments in the Analysis.*

The following adjustments must be done in completing the Anova table:

1. Deduct the value (1) from the total degree of freedom and the unexplained or error degree of freedom.
2. Subtract the computed bias from the Total Sum of Squares (TSS) and Treatment Sum of Squares (TrtSS).

*Adjustments*

$$\text{Total } df = [C \times R] - 1 - 1 = 25 - 1 - 1 = 23$$

$$\begin{aligned} \text{Unexplained or error} &= [C \times R - 1] - [(t - 1) + (R - 1) + (C - 1)] - 1 \\ &= 24 - 16 - 1 = 9 \end{aligned}$$

*Using Ca as treatment having the missing data*

$$\text{Adjusted TSS} = 22.48 - \beta_0$$

$$\text{where } \beta_0 = 0.15$$

$$\text{Adjusted TSS} = 22.48 - 0.15 = 22.33$$

$$\text{Adjusted TrtSS} = 20.47 - \beta_0$$

$$\text{Adjusted TrtSS} = 20.47 - 0.15 = 20.32$$

*Using Co as treatment having the missing data*

$$\text{Adjusted TSS} = 22.48 - \beta_0$$

$$\text{where } \beta_0 = 2.36$$

$$\text{Adjusted TSS} = 22.48 - 2.36 = 20.12$$

$$\text{Adjusted TrtSS} = 20.47 - \beta_0$$

$$\text{Adjusted TrtSS} = 20.47 - 2.36 = 18.11$$

Anova table when as missing data is from *Ca* treatment:

ANOVA TABLE					
<i>Sources of Variation</i>	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>Fcal.</i>	<i>Fcrit. (5%)</i>
<i>Treatment</i>	4	20.32	5.08	9.77	3.63
<i>Row</i>	4	0.35	0.09	0.17	3.63
<i>Column</i>	4	1.14	0.29	0.55	3.63
<i>Error</i>	9	0.52	0.06		
<i>Total</i>	23	22.33	0.97		
<i>Fcrit (5%) @ df of 4,9</i>	3.63				
<i>Fcrit (1%) @ df of 4,9</i>	6.42				

Anova table when as missing data is from *Co* treatment:

ANOVA TABLE					
<i>Sources of Variation</i>	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>Fcal.</i>	<i>Fcrit. (5%)</i>
<i>Treatment</i>	4	18.11	4.53	8.71	3.63
<i>Row</i>	4	0.35	0.09	0.17	3.63
<i>Column</i>	4	1.14	0.29	0.55	3.63
<i>Error</i>	9	0.52	0.06		
<i>Total</i>	23	20.12	0.87		
<i>Fcrit (5%) @ df of 4,9</i>	3.63				
<i>Fcrit (1%) @ df of 4,9</i>	6.42				

*Tenth step: looking up the F-critical table for the critical values*

$$F_{crit}(1\%) \text{ at } df \text{ of } 4,9 \text{ for Treatments} = 6.42$$

$$F_{crit}(1\%) \text{ at } df \text{ of } 4,9 \text{ for Rows} = 6.42$$

$$F_{crit}(1\%) \text{ at } df \text{ of } 4,9 \text{ for Columns} = 6.42$$

$$F_{crit}(5\%) \text{ at } df \text{ of } 4,9 \text{ for Treatments} = 3.63$$

$F_{crit}(5\%)$  at  $df$  of 4,9 for Rows = 3.63

$F_{crit}(5\%)$  at  $df$  of 4,9 for Columns = 3.63

*The Eleventh step: Making the Decision and Conclusions.*

The values obtained from the F-critical table compared with the F-values calculated for the Treatments, Rows and Columns reveals the decisions and conclusions remain the same as the table without the missing data. This proves the missing data value has been accurately estimated.

## **Bibliography**

- [1] Brownlee, K. A. (1960). *Statistical theory and methodology in science and engineering*. New York: Wiley.
- [2] Campbell, D., Stanley J. (1963). *Experimental and quasi-experimental designs for research and teaching*. In Gage (Ed.), *Handbook on research on teaching*. Chicago: Rand McNally & Co.
- [3] Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.
- [4] Cochran, W. G., Cox, G. M. (1957). *Experimental Designs*. John Willey & Sons, N.Y.
- [5] Le Clerg, E. L., Leonard, W. H., Clark A. G. 1966. *Field Plot Technique*. Burgess Pub. Co., Minn., USA.
- [6] Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.

