ANOVA – Analysis of Variance

Working Rule (One – Way Classification)

Set the null hypothesis H_0: There is no significance difference between the treatments.

Set the alternative hypothesis H_1: There is a significance difference between the treatments.

Step: 1 Find N = number of observations

Step: 2 Find T = The total value of observations

Step: 3 Find the Correction Factor C . F $= \frac{T^2}{N}$

Step: 4 Calculate the total sum of squares and find the total sum of squares

$$\text{TSS} = \left(\sum {X_1}^2 + \sum {X_2}^2 + \sum {X_3}^2 + \ . \ . \ .\right) - C.F$$

Step: 5 Column sum of squares SSC $\left(\frac{(\sum X_1)^2}{N_1} + \frac{(\sum X_2)^2}{N_2} + \frac{(\sum X_3)^2}{N_3} + \ldots\right) - C.F$

Where N_i= Total number of observation in each column $(i = 1, 2, 3, \ldots)$

Step: 6 Prepare the ANOVA to calculate F – ratio

Source of variation	Sum of Degrees	Degrees of Freedom	Mean Square	F - Ratio
Between Samples	SSC	K - 1	MSC = $\frac{SSC}{K-1}$	$F_c = \frac{MSC}{MSE}$ if MSC$>$ MSE
Within Samples	SSE	N - K	MSE = $\frac{SSE}{N-K}$	$F_c = \frac{MSE}{MSC}$ if MSE$>$ MSC

Step: 7 Find the table value (use chi square table)

Step: 8 Conclusion:

Calculated value $<$ Table value, then we accept null hypothesis.

Calculated value $>$Table value, then we reject null hypothesis.

PROBLEMS ON ONE WAY ANOVA

1.A completely randomised design experiment with 10 plots and 3 treatments gave the following results.

Plot No	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Treatment	**A**	**B**	**C**	**A**	**C**	**C**	**A**	**B**	**A**	**B**

Yield	**5**	**4**	**3**	**7**	**5**	**1**	**3**	**4**	**1**	**7**

Analyse the result for treatment effects.

Solution:

Set the null hypothesis H_0: There is no significance difference between the treatments.

Set the alternative hypothesis H_1: There is a significance difference between the treatments.

Treatments	Yields from plots			
A	5	7	3	1
B	4	4	7	-
C	3	5	1	-

TABLE:

Treatment A		Treatment B		Treatment C	
X_1	${X_1}^2$	X_2	${X_2}^2$	X_3	${X_3}^2$
5	25	4	16	3	9
7	49	4	16	5	25
3	9	7	49	7	7
1	1	-	-	-	-
$\sum X_1$ =16	$\sum {X_1}^2$ =84	$\sum X_2$ =5	$\sum {X_2}^2$ =81	$\sum X_3$ =9	$\sum {X_3}^2$ =35

Step: 1 N =10

Step: 2 Sum of all the items (T) $= \sum X_1 + \sum X_2 + \sum X_3 = 16 + 15 + 9 = 40$

Step: 3 Find the Correction Factor C . F $= \frac{T^2}{N} = \frac{(40)^2}{10} = 160$

Step: 4 TSS = Total sum of squares

= sum of squares of all the items – C. F

TSS =$\left(\sum {X_1}^2 + \sum {X_2}^2 + \sum {X_3}^2 + \ . \ . \ .\right) - C.F$

=$(84 + 81 + 35) - 160 = 40$

Step: 5 SSC = Sum of squares between samples

SSC = $\left(\frac{(\sum X_1)^2}{N_1} + \frac{(\sum X_2)^2}{N_2} + \frac{(\sum X_3)^2}{N_3} + \dots\right) - C.F$

SSC = $\left(\frac{(16)^2}{4} + \frac{(15)^2}{3} + \frac{(9)^2}{3} + \dots\right) - 160$

$$= 64 + 75 + 27 - 160 = 6$$

Step: 6 MSC = Mean squares between samples

$$= \frac{Sum\ of\ squares\ between\ samples}{d.f}$$

$$= \frac{6}{2} = 3$$

SSE = Sum of squares within samples

= Total sum of squares – Sum of squares between samples

$$= 40 - 6 = 34$$

Step:7 MSE = Mean squares within samples

$$= \frac{Sum\ of\ squares\ within\ samples}{d.f}$$

$$= \frac{34}{7} = 4.86$$

ANOVA TABLE

Source of variation	Sum of Degrees	Degrees of Freedom	Mean Square	F - Ratio
Between Samples	SSC = 6	K – 1 = 3 – 1 =2	MSC = $\frac{SSC}{K-1} = 3$	
Within Samples	SSE = 34	N – K =10 – 3 = 7	MSE = $\frac{SSE}{N-K} =$ 4.86	$F_c = \frac{MSE}{MSC}$= 1.62

d.f for (7, 2) at 5% level of significance is 19.35

Step: 8 Conclusion:

Calculated value $<$ Table value, then we accept null hypothesis.

2. Three different machines are used for a production. On the basis of the outputs, set up one – way ANOVA table and test whether the machines are equally effective.

Outputs		
Machine I	**Machine II**	**Machine III**
10	**9**	**20**
15	**7**	**16**
11	**5**	**10**

10	**6**	**14**

Given that the value of F at 5% level of significance for (2, 9) d. f is 4.26

Solution:

Set the null hypothesis H_0: The machines are equally effective.

TABLE:

Treatment A		Treatment B		Treatment C	
X_1	${X_1}^2$	X_2	${X_2}^2$	X_3	${X_3}^2$
10	100	9	81	20	400
15	225	7	49	16	256
11	121	5	25	10	100
20	400	6	36	14	196
$\sum X_1$ =56	$\sum {X_1}^2$ =846	$\sum X_2$ =27	$\sum {X_2}^2$ =191	$\sum X_3$ =60	$\sum {X_3}^2$ =952

Step: 1 N =12

Step: 2 Sum of all the items (T) $= \sum X_1 + \sum X_2 + \sum X_3 = 56 + 27 + 60 = 143$

Step: 3 Find the Correction Factor C . F $= \frac{T^2}{N} = \frac{(143)^2}{12} = 1704.08$

Step: 4 TSS = Total sum of squares

= sum of squares of all the items – C. F

$$\text{TSS} = \left(\sum {X_1}^2 + \sum {X_2}^2 + \sum {X_3}^2 + \ .\ .\ .\right) - C.F$$

$$=(846 + 191 + 952) - 1704.08 = 284.92$$

Step: 5 SSC = Sum of squares between samples

$$\text{SSC} = \left(\frac{(\sum X_1)^2}{N_1} + \frac{(\sum X_2)^2}{N_2} + \frac{(\sum X_3)^2}{N_3} + \dots\right) - C.F$$

$$\text{SSC} = \left(\frac{(56)^2}{4} + \frac{(27)^2}{4} + \frac{(60)^2}{4} + \dots\right) - 1704.08$$

$$= 784 + 182.25 + 900 - 1704.08 = 162.17$$

Step: 6 MSC = Mean squares between samples

$$= \frac{Sum\ of\ squares\ between\ samples}{d.f}$$

$$= \frac{162.17}{2} = 81.085$$

SSE = Sum of squares within samples

= Total sum of squares – Sum of squares between samples

$= 284.92 - 162.17 = 122.75$

Step: 7 MSE = Mean squares within samples

$= \frac{Sum\ of\ squares\ within\ samples}{d.f}$

$= \frac{122.75}{9} = 13.63$

ANOVA TABLE

Source of variation	Sum of Degrees	Degrees of Freedom	Mean Square	F - Ratio
Between Samples	SSC = 162.17	K – 1 = 3 – 1 =2	MSC = $\frac{SSC}{K-1} = 81.085$	
Within Samples	SSE = 122.75	N – K =12 – 3 = 9	MSE = $\frac{SSE}{N-K} = 13.63$	$F_c = \frac{MSE}{MSC}$= 5.95

d.f for (2, 9) at 5% level of significance is 4.26.

Step: 8 Conclusion:

Calculated value $>$ Table value, then we reject the null hypothesis.

i.e., the three machines are not equally effective.

Randomised Block Design (RBD)

Working Rule:

Set the null hypothesis H_0: There is no significance difference between the treatments.

Step: 1 Find T = The total value of observations

Step: 2 Find the Correction Factor C . F $= \frac{T^2}{N}$

Step: 3 Calculate the total sum of squares and find the total sum of squares

$$\text{TSS} = \left(\sum {X_1}^2 + \sum {X_2}^2 + \sum {X_3}^2 + \ ...\right) - C.F$$

Step: 4 Find column sum of squares SSC= $\left(\frac{(\sum X_1)^2}{N_1} + \frac{(\sum X_2)^2}{N_2} + \frac{(\sum X_3)^2}{N_3} + \dots\right) - C.F$

Where N_i= Total number of observation in each column ($i = 1, 2, 3, \dots$)

Step: 5 Find Column sum of squares SSR = $\left(\frac{(\sum Y_1)^2}{N_1} + \frac{(\sum Y)^2}{N_2} + \frac{(\sum Y_3)^2}{N_3} + \dots\right) - C.F$

Where N_J= Total number of observation in each ROW ($j = 1, 2, 3, \dots$)

Step: 6 SSE = TSS – (SSC + SSR)

Step: 7 Prepare the ANOVA to calculate F – ratio

Source of variation	Sum of Degrees	Degrees of Freedom	Mean Square	F - Ratio
Between Columns	SSC	c - 1	MSC = $\frac{SSC}{c-1}$	$F_c = \frac{MSC}{MSE}$ if MSC> MSE $F_c = \frac{MSE}{MSC}$ if MSE> MSC
Between Rows	SSR	r – 1	MSR = $\frac{SSE}{r-1}$	$F_c = \frac{MSR}{MSE}$ if MSR> MSE $F_c = \frac{MSE}{MSR}$ if MSE> MSR
Error	SSE	(r – 1)(c – 1)	MSE = $\frac{SSE}{(r-1)(c-1)}$	

Step: 8 Find the table value (use chi square table)

Step: 9 Conclusion:

Calculated value < Table value, then we accept null hypothesis.

Calculated value >Table value, then we reject null hypothesis.

PROBLEMS ON TWO WAY ANOVA TABLE

1. Three varieties A, B, C of a crop are tested in a randomized block design with four replication. The plot yields in pounds as follows.

A6	**C5**	**A8**	**B9**
C8	**A4**	**B6**	**C9**
B7	**B6**	**C10**	**A6**

Analysis the experiment yield and state your conclusion.

Solution:

Set the null hypothesis H_0: There is no significance difference between the rows and columns.

Varieties	Yields				
	1	2	3	4	Total
A	6	4	8	6	24
B	7	6	6	9	28
C	8	5	10	9	32
Total	21	15	24	24	84

TEST STATISTIC:

Varieties		1	2	3	4	Total	${X_1}^2$	${X_2}^2$	${X_3}^2$	${X_4}^2$
		X_1	X_2	X_3	X_4					
Y_1	A	6	4	8	6	24	36	16	64	36
Y_2	B	7	6	6	9	28	49	36	36	81
Y_3	C	8	5	10	9	32	64	25	100	81
Total		21	15	24	24	84	149	77	200	198

Step:1 Grand Total T = 84

Step: 2 Correction Factor C . F $= \frac{T^2}{N} = \frac{(84)^2}{12} = 588$

Step: 3 Calculate the total sum of squares and find the total sum of squares

$$\text{TSS} = \left(\sum {X_1}^2 + \sum {X_2}^2 + \sum {X_3}^2 + \ ...\right) - C.F$$

$$= (149 + 77 + 200 + 198) - 588$$

$$= 624 - 588 = 36$$

Step: 4 Find column sum of squares SSC= $\left(\frac{(\sum X_1)^2}{N_1} + \frac{(\sum X_2)^2}{N_2} + \frac{(\sum X_3)^2}{N_3} + \dots\right) - C.F$

$$\text{SSC} = \left(\frac{(21)^2}{3} + \frac{(15)^2}{3} + \frac{(24)^2}{3} + \frac{(24)^2}{3}\right) - 588 = 18$$

Step: 5 Find Row sum of squares SSR = $\left(\frac{(\sum Y_1)^2}{N_1} + \frac{(\sum Y)^2}{N_2} + \frac{(\sum Y_3)^2}{N_3} + \dots\right) - C.F$

$$\text{SSR} = \left(\frac{(24)^2}{4} + \frac{(28)^2}{4} + \frac{(32)^2}{4} + \dots\right) - 588 = 8$$

Step: 6 SSE = Residual sum of squares

= TSS – (SSC + SSR)

= 36 – (18 + 8) 10

Step: 7 Prepare the ANOVA to calculate F – ratio

Source of variation	Sum of Degrees	Degrees of Freedom	Mean Square	F - Ratio
Between Columns	SSC=18	c – 1 = 4 – 1 =3	MSC = $\frac{SSC}{c-1}$ = 6	$F_c = \frac{MSC}{MSE}$ =3.6
Between Rows	SSR=8	r – 1 =3 – 1 = 2	MSR = $\frac{SSR}{r-1}$ = 4	$F_R = \frac{MSR}{MSE}$ = 2.4
Error	SSE = 10	(r – 1)(c – 1) 2 x 3 = 6	MSE = $\frac{SSE}{(r-1)(c-1)}$ = 1.667	

Step: 8 d.f for (3, 6) at 5% level of significance is 4.76

d.f for (2, 6) at 5% level of significance is 5.14

Step: 9 Conclusion:

Calculated value $F_c <$ Table value, then we accept null hypothesis.

There is no significance difference between the columns.

Calculated value $F_R <$ Table value, then we accept null hypothesis.

There is no significance difference between the rows.

2. **Four varieties A, B, C, D of a fertilizer are tested in a randomized block design with four replication. The plot yields in pounds as follows.**

A 12	**D 20**	**C 16**	**B 10**
D 18	**A 14**	**B 11**	**C 14**
B 12	**C 15**	**D 19**	**A 13**
C 16	**B 11**	**A 15**	**D 20**

Analysis the experimental yield.

Solution:

Set the null hypothesis H_0: There is no significance difference between the rows and columns.

Varieties	Yields				
	1	2	3	4	Total
A	12	14	15	13	54
B	12	11	11	10	44
C	16	15	16	14	61

D	18	20	19	20	77
Total	58	60	61	57	236 (T)

TEST STATISTIC:

Varieties		1	2	3	4	Total	${X_1}^2$	${X_2}^2$	${X_3}^2$	${X_4}^2$
		X_1	X_2	X_3	X_4					
Y_1	A	12	14	15	13	54	144	196	225	169
Y_2	B	12	11	11	10	44	144	121	121	100
Y_3	C	16	15	16	14	61	256	225	256	196
Y_4	D	18	20	19	20	77	324	400	361	400
Total		58	60	61	57	236	868	942	963	865

Step:1 Grand Total T = 236

Step: 2 Correction Factor C . F $= \frac{T^2}{N} = \frac{(236)^2}{16} = 3481$

Step: 3 Calculate the total sum of squares and find the total sum of squares

$$\text{TSS} = \left(\sum {X_1}^2 + \sum {X_2}^2 + \sum {X_3}^2 + \ ...\right) - C.F$$

$$= (868 + 942 + 963 + 865) - 3481$$

$$= 3638 - 3481 = 157$$

Step: 4 Find column sum of squares SSC= $\left(\frac{(\sum X_1)^2}{N_1} + \frac{(\sum X_2)^2}{N_2} + \frac{(\sum X_3)^2}{N_3} + \ldots\right) - C.F$

$$\text{SSC} = \left(\frac{(58)^2}{4} + \frac{(60)^2}{4} + \frac{(61)^2}{4} + \frac{(57)^2}{4}\right) - 3481$$

$$= 841 + 900 + 930 + 812 - 3481 = 2$$

Step: 5 Find Row sum of squares SSR = $\left(\frac{(\sum Y_1)^2}{N_1} + \frac{(\sum Y)^2}{N_2} + \frac{(\sum Y_3)^2}{N_3} + \ldots\right) - C.F$

$$\text{SSR} = \left(\frac{(54)^2}{4} + \frac{(44)^2}{4} + \frac{(61)^2}{4} + \frac{(77)^2}{4}\right) - 3481$$

$$= 729 + 484 + 930.25 + 1482.25 - 3481 = 144.5$$

Step: 6 SSE = Residual sum of squares

= TSS – (SSC + SSR)

= 157 – (2 +144.5)= 10.5

Step: 7 Prepare the ANOVA to calculate F – ratio

Source of variation	Sum of Degrees	Degrees of Freedom	Mean Square	F - Ratio
Between Columns	SSC=2	c – 1 = 4 – 1 =3	MSC = $\frac{SSC}{c-1}$ =0.666	$F_c = \frac{MSE}{MSC}$ =1.74
Between Rows	SSR=144.5	r – 1 =4 – 1 = 3	MSR = $\frac{SSR}{r-1}$ = 48.16	$F_R = \frac{MSR}{MSE}$ = 41.51
Error	SSE = 10.5	(r – 1)(c – 1) =3 x 3 = 9	MSE = $\frac{SSE}{(r-1)(c-1)}$ = 1.6	

Step: 8 d.f for (9, 3) at 5% level of significance is 8.82

d.f for (3, 9) at 5% level of significance is 3.86

Step: 9 Conclusion:

Calculated value $F_c <$ Table value, then we accept null hypothesis.

There is no significance difference between the columns.

Calculated value $F_R >$Table value, then we reject null hypothesis.

There is a significance difference between the rows.

LATIN SQUARE:

Steps in constructing Latin Square

Step:1

Square the Grand total (T) and divide it by the number of observations (N).

i.e), Find $\frac{T^2}{N}$ which is called the correction factor (C.F)

Step:2

Add the squares of the individual observations ($X_i's$) and substract the C.F from it to get the total sum of squares. i.e)., Find Total sum of squares TSS

i.e)., TSS $=\sum_i (X_i)^2 - \frac{T^2}{N}$

Step:3

Add the squares of the row sums (R_i) divide it by the number of items in a row and substract the C.F from the result to get the row sum of squares.

Row sum of squares $SSR = \frac{(\sum R_i)^2}{n_1} - C.F$

Where n_1 is the number of items in a row.

Step:4

Add the squares of the columns sums (C_i) divide it by the number of items and substract the C.F from the result to get the column sum of squares.

Column sum of squares $SSC = \frac{(\sum C_j)^2}{n_2} - C.F$

Where n_2 is the number of items in a column.

Step:5

Sum of the squares of the treatment sums (T_i) divide it by the number of treatments and substract the C.F from it to get the treatment sum of squares, i.e., Treatment sum of squares.

$$SST = \frac{(\sum T_i)^2}{n_i} - C.F$$

Where n_i is the number of treatments.

Step:6

Substract the sum obtained in steps 3, 4, and 5 from 2 we get residual.

i.e)., Residual $SSE = TSS - (SSR + SSC + SST)$

Step:7

Prepare the ANOVA table using all these and calculate the various mean squares as follows.

Source of variation	Sum of Degrees	Degrees of Freedom	Mean Square	F - Ratio
Between Rows	SSR	n – 1	MSR = $\frac{SSR}{n-1}$	$F_R = \frac{MSR}{MSE}$ if MSR> MSE $F_R = \frac{MSE}{MSR}$ if MSE> MSR
Between Columns	SSC	n - 1	MSC = $\frac{SSC}{n-1}$	$F_c = \frac{MSC}{MSE}$ if MSC> MSE $F_c = \frac{MSE}{MSC}$ if MSE> MSC

Treatments	SST	n - 1	MST = $\frac{SST}{n-1}$	$F_T = \frac{MST}{MSE}$ if MST> MSE $F_T = \frac{MSE}{MST}$ if MSE> MST
Residual or Error	SSE	(n – 1)(n – 2)	MSE = $\frac{SSE}{(n-1)(n-2)}$	

Step:8

Compute the F-ratio and find out whether the differences are significant or not according to the given level of significance.

1. Set up the analysis of variance for the following results of a Latin square design.

A 12	**C 19**	**B 10**	**D 8**
C 18	**B 12**	**D 6**	**A 7**
B 22	**D 10**	**A 5**	**C 21**
C 12	**A 7**	**C 27**	**B 17**

Solution:

Set the null hypothesis H_0: There is no significance difference between the rows, columns and treatments.

Table I (To find TSS, SSR and SSC)

	C_1	C_2	C_3	C_4	Row Total R_i	${R_i}^2/4$
R_1	12	19	10	8	49	600.25
R_2	18	12	6	7	43	462.25
R_3	22	10	5	21	58	841
R_4	12	7	27	17	63	992.25
Column Total	64	48	48	53	213 (T)	2895.75

C_j						$\sum R_i^2/4$
$C_j^2/4$	1024	576	576	702.25	2895.75 $\sum C_j^2/4$	

Table II (To find SST)

	1	2	3	4	Row Total T_i	$T_i^2/4$
A	12	7	5	7	31	240.25
B	10	12	22	17	61	930.25
C	19	18	21	27	85	1806.25
D	8	6	10	12	36	324
						3300.75= $\sum T_i^2/4$

Step:1

Grand total (T) =213

Step:2

Correction factor (C.F)= $\frac{T^2}{N} = \frac{(213)^2}{16} = 2835.56$

Step:3

Sum of squares of individual observations

$= (\mathbf{12})^2 + (\mathbf{7})^2 + (\mathbf{5})^2 + (\mathbf{7})^2 + (\mathbf{10})^2 + (\mathbf{12})^2 + (\mathbf{22})^2 + (\mathbf{17})^2 + (\mathbf{19})^2 + (\mathbf{18})^2 + (\mathbf{21})^2 + (\mathbf{27})^2 + (\mathbf{8})^2 + (\mathbf{6})^2 + (\mathbf{10})^2 + (\mathbf{12})^2$

$= \mathbf{3483}$

Step:4

TSS $= sum\ of\ squares\ of\ individual\ observations - C.F$

$$= \sum_i (X_i)^2 - \frac{T^2}{N} = 3486 - 2835.56 = 647.44$$

Step:5

Row sum of squares $SSR = \frac{(\sum R_i)^2}{4} - C.F = 2895.75 - 2835.56 = 60.19$

Step:6

Column sum of squares $SSC = \frac{(\sum C_j)^2}{4} - C.F = 2878.25 - 2835.56$

$$= 42.69$$

Step:7

Sum of squares of Treatment

$$SST = \frac{(\sum T_i)^2}{n_i} - C.F = 3300.75 - 2835.56 = 465.19$$

Step:8

Residual $SSE = TSS - (SSR + SSC + SST)$

$$= 647.44 - (60.19 + 42.69 + 465.19) = 79.37$$

Step:9

Prepare the ANOVA table using all these and calculate the various mean squares as follows.

Source of variation	Sum of Degrees	Degrees of Freedom	Mean Square	F - Ratio
Between Rows	SSR=60.19	4 – 1 =3	MSR = $\frac{SSR}{n-1}$ =20.06	$F_R = \frac{MSR}{MSE}$=1.52
Between Columns	SSC=42.69	4 – 1 =3	MSC = $\frac{SSC}{n-1}$ =14.23	$F_c = \frac{MSC}{MSE}$ =1.08
Treatments	SST=465.19	4 – 1 =3	MST = $\frac{SST}{n-1}$ =155.06	$F_T = \frac{MST}{MSE}$ =11.73
Residual or Error	SSE=79.37	(4 – 1)(4 – 2) =6	MSE = $\frac{SSE}{(n-1)(n-2)}$ =13.22	

Step: 10 d.f for (3, 6) at 5% level of significance is 4.76

Step: 9 Conclusion:

Calculated value $F_c <$ Table value, then we accept null hypothesis.

There is no significance difference between the columns.

Calculated value $F_R <$Table value, then we accept null hypothesis.

There is no significance difference between the rows.

Calculated value $F_T >$Table value, then we reject null hypothesis.

There is a significance difference between the rows.