Module-5: Design of Experiments & ANOVA
Principles of experimentation in design. Analysis of completely randomized design, randomized block design. The ANOVA Technique. Basic Principle of ANOVA, One-way ANOVA, Two-way ANOVA, Latin-square Design, and Analysis of Co-Variance. **(12 Hours)** (RBT Levels: L1, L2 and L3)

5.1 The ANOVA Technique

Introduction:

The analysis of variance (ANOVA) is a statistical technique to test whether the means of three or more populations are equal or not. This technique was developed by R A Fisher. This technique is widely used in Professional Business and Physical Sciences.

In this technique, variance is splitted into two parts:
(i)Variance between samples (Columns) (ii) Variance within samples (Rows)

A table showing the source of variation, the sum of squares, degrees of freedom, mean squares and the formula for the F ratio is called ANOVA table.

If the given data is classified according to one factor, the classification is called one way classification. Then ANOVA table for one-way classification is to be constructed.

If the given data is classified according to two factors, the classification is called two-way classification. Then ANOVA table for two-way classification is to be constructed.

Analysis of variance is based on the following assumptions:

(i) The samples are independently drawn the population.
(ii) Populations from which the sample are selected are normally distributed.
(iii) Each of the population have the same variance.

ANOVA table for one-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between samples	SSC	$c-1$	$MSC = \frac{SSC}{c-1}$	$F = \frac{MSC}{MSE}$
Within samples	SSE	$N-c$	$MSE = \frac{SSE}{N-c}$	
Total	SST	$N-1$	-	-

$$F = \frac{MSC}{MSE}, if\ MSC > MSE, Reciprocate\ otherwise.$$

Expansion of abbreviations:

SSC – Sum of squares between samples (Columns)
SSE – Sum of squares within sample (Rows)
SST – Total sum of squares of variations
MSC – Mean squares of variations between samples (Columns)
MSE - Mean squares of variations within samples (Rows)

Notations:

T – Total sum all the observations
N – Number of observations.
c – Number of columns.

How to find SSC and SSE?

$$SSC = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} + \cdots + \frac{(\Sigma X_k)^2}{n_k} - \frac{T^2}{N}$$

$$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 + \cdots + \Sigma X_k^2 - \frac{T^2}{N}$$

$$SSE = SST - SSC$$

Working rule:

(i) Assume $H_0: \mu_1, \mu_2, \dots, \mu_k$ all are equal.
(ii) Construct ANOVA tale for one-way classification.
(iii) Under H_0, $F = \frac{MSC}{MSE}, if\ MSC > MSE, Reciprocate\ otherwise.$
(iv) If calculated value < tabulated value, accept H_0. Reject otherwise.

1. **Three different machines are used for a production. On the basis of the outputs, test whether the machines are equally effective.**

Output		
Machine 1	Machine 2	Machine 3
10	9	20
5	7	16
11	5	10
10	6	4

Assume $H_0: \mu_1 = \mu_2 = \mu_3$. All the three Machines are equally effective.

To find: SSC and SSE

Output					
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
10	100	9	81	20	400
5	25	7	49	16	256
11	121	5	25	10	100
10	100	6	36	4	16
36	346	27	191	50	772
ΣX_1	ΣX_1^2	ΣX_2	ΣX_2^2	ΣX_3	ΣX_3^2

$$T = \Sigma X_1 + \Sigma X_2 + \Sigma X_3 = 36 + 27 + 50 = 113$$

$$N = 4 + 4 + 4 = 12$$

$$SSC = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} - \frac{T^2}{N} = \frac{36^2}{4} + \frac{27^2}{4} + \frac{50^2}{4} - \frac{113^2}{12} = 67.17$$

$$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 - \frac{T^2}{N} = 346 + 191 + 772 - \frac{113^2}{12} = 244.92$$

$$SSE = SST - SSC = 244.92 - 67.17 = 177.75$$

Construction of ANOVA table for one-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between samples	$SSC = 67.17$	$c - 1 = 3 - 1$ $= 2$	$MSC = \frac{SSC}{c-1}$ $= \frac{67.17}{2} = 33.59$	$F = \frac{MSC}{MSE}$ $= \frac{33.59}{19.75} = 1.7$
Within samples	$SSE = 177.75$	$N - c = 12 - 3$ $= 9$	$MSE = \frac{SSE}{N-c}$ $= \frac{177.75}{9} = 19.75$	

Under H_0, Calculated value of F is 1.7

At $\alpha = 0.05, \nu_1 = 2, \nu_2 = 9$, Tabulated value of F is 4.26

Calculated value $<$ Tabulated value

Accept H_0. All the three Machines are equally effective.

2. **Three samples each of size 5 were drawn from three uncorrelated normal populations with equal variances. Test the hypothesis that the population means are equal at 5% level.**

Sample 1	10	12	9	16	13
Sample 2	9	7	12	11	11
Sample 3	14	11	15	14	16

Assume $H_0: \mu_1 = \mu_2 = \mu_3$. All the three samples have equal population means.

To find: SSC and SSE

Output					
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
10	100	9	81	14	196
12	144	7	49	11	121
9	81	12	144	15	225
16	256	11	121	14	196
13	169	11	121	16	256
60	750	50	516	70	994
ΣX_1	ΣX_1^2	ΣX_2	ΣX_2^2	ΣX_3	ΣX_3^2

$T = \Sigma X_1 + \Sigma X_2 + \Sigma X_3 = 60 + 50 + 70 = 180$

$N = 5 + 5 + 5 = 15$

$SSC = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} - \frac{T^2}{N} = \frac{60^2}{5} + \frac{50^2}{5} + \frac{70^2}{5} - \frac{180^2}{15} = 40$

$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 - \frac{T^2}{N} = 750 + 516 + 994 - \frac{180^2}{15} = 100$

$SSE = SST - SSC = 100 - 40 = 60$

Construction of ANOVA table for one-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between samples	$SSC = 40$	$c - 1 = 3 - 1 = 2$	$MSC = \frac{SSC}{c-1} = \frac{40}{2} = 20$	$F = \frac{MSC}{MSE} = \frac{20}{5} = 4$
Within samples	$SSE = 60$	$N - c = 15 - 3 = 12$	$MSE = \frac{SSE}{N-c} = \frac{60}{12} = 5$	

Under H_0, Calculated value of F is 4

At $\alpha = 0.05, \nu_1 = 2, \nu_2 = 12$, Tabulated value of F is 3.89

Calculated value > Tabulated value

Reject H_0. All the three samples have **not** equal population means.

3. A Manager of a merchandizing firm wishes to test whether its three salesmen A, B, C tend to make sales of the same size or whether they differ in their selling abilities. During a week there have been 14 sales calls, A made 5 calls, B made 4 calls and C made 5 calls. Following are the weekly sales record (in rupees) of the three salesmen:

A	500	400	700	300	600
B	300	700	400	600	–
C	500	300	500	400	300

Perform the analysis of variance and draw your conclusions.

The sales data have a common factor 100. Divide all the above values by 100.

X_1	5	4	7	3	6
X_2	3	7	4	6	–
X_3	5	3	5	4	3

Assume H_0: $\mu_1 = \mu_2 = \mu_3$. All the three salesmen tend to make sales of the same size.

To find: SSC and SSE

Output					
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
5	25	3	9	5	25
4	16	7	49	3	9
7	49	4	16	5	25
8	64	6	36	4	16
6	36	-	-	3	9
30	190	20	110	20	84
ΣX_1	ΣX_1^2	ΣX_2	ΣX_2^2	ΣX_3	ΣX_3^2

$$T = \Sigma X_1 + \Sigma X_2 + \Sigma X_3 = 30 + 20 + 20 = 70$$

$$N = 5 + 4 + 5 = 14$$

$$SSC = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} - \frac{T^2}{N} = \frac{30^2}{5} + \frac{20^2}{4} + \frac{20^2}{5} - \frac{70^2}{14} = 10$$

$$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 - \frac{T^2}{N} = 190 + 110 + 84 - \frac{70^2}{14} = 34$$

$$SSE = SST - SSC = 34 - 10 = 24$$

Construction of ANOVA table for one-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F −Ratio
Between samples	$SSC = 10$	$c - 1 = 3 - 1 = 2$	$MSC = \frac{SSC}{c-1} = \frac{10}{2} = 5$	$F = \frac{MSC}{MSE} = \frac{5}{2.18} = 2.29$
Within samples	$SSE = 24$	$N - c = 14 - 3 = 11$	$MSE = \frac{SSE}{N-c} = \frac{24}{11} = 2.18$	

Under H_0, Calculated value of F is 2.29

At $\alpha = 0.05, \nu_1 = c - 1 = 2, \nu_2 = N - c = 11$, Tabulated value of F is 3.98

Calculated value < Tabulated value. Accept H_0.

All the three salesmen tend to make sales of the same size.

4. Three samples of five, five and four car tyres are drawn respectively from three brands A, B, and C manufactured by three machines. The lifetime of these tyres (per 1000 miles) is given below. Test whether the average life time of the three brands of tyres are equal or not.

A	35	40	33	36	31
B	30	25	34	28	33
C	28	24	30	26	-

Assume $H_0: \mu_1 = \mu_2 = \mu_3$. The average lifetime of three brands of tyres are equal.

Subtract 30 from each of the given values.

X_1	5	10	3	6	1
X_2	0	-5	4	-2	3
X_3	-2	-6	0	-4	$-$

To find: SSC and SSE

Output					
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
5	25	0	0	-2	4
10	100	-5	25	-6	36
3	9	4	16	0	0
6	36	-2	4	-4	16
1	1	3	9	$-$	$-$
25	171	0	54	-12	56
ΣX_1	ΣX_1^2	ΣX_2	ΣX_2^2	ΣX_3	ΣX_3^2

$$T = \Sigma X_1 + \Sigma X_2 + \Sigma X_3 = 25 + 0 - 12 = 13$$

$$N = 5 = 5 + 4 = 14$$

$$SSC = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} - \frac{T^2}{N} = \frac{25^2}{5} + \frac{0^2}{5} + \frac{(-12)^2}{4} - \frac{13^2}{14} = 148.93$$

$$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 - \frac{T^2}{N} = 171 + 54 + 56 - \frac{13^2}{14} = 268.93$$

$$SSE = SST - SSC = 268.93 - 148.93 = 120$$

Construction of ANOVA table for one-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between samples	$SSC = 148.93$	$c-1=2$	$MSC = \frac{SSC}{c-1} = 74.465$	$F = \frac{MSC}{MSE} = 6.83$
Within samples	$SSE = 120$	$N-c=11$	$MSE = \frac{SSE}{N-c} = 10.9$	

Under H_0, Calculated value of F is 6.83

At $\alpha = 0.05, \nu_1 = c-1 = 2, \nu_2 = N-c = 11$, Tabulated value of F is 3.98

Calculated value > Tabulated value. Reject H_0.

The average lifetime of three brands of tyres are not equal.

5. To assess the significance of possible variation in performance in a certain test between the grammar school of a city, a common test was given to a number of students taken at random from the senior fifth class of each of the four schools concerned. The results are given below. Make an analysis of variance data.

Schools			
A	*B*	*C*	*D*
8	12	18	13
10	11	12	9
12	9	16	12
8	14	6	16
7	4	8	15

Assume $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$. Samples have come from the same universe.

Subtract 10 from each of the given values.

X_1	−2	0	2	−2	−3
X_2	2	1	−1	4	−6
X_3	8	2	6	−4	−2
X_4	3	−1	2	6	5

To find: SSC and SSE

Output							
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2	X_4	X_4^2
−2	4	2	4	8	64	3	9
0	0	1	1	2	4	−1	1
2	4	−1	1	6	36	2	4
−2	4	4	16	−4	16	6	36
−3	9	−6	36	−2	4	5	25
−5	21	0	58	10	124	15	75
ΣX_1	ΣX_1^2	ΣX_2	ΣX_2^2	ΣX_3	ΣX_3^2	ΣX_4	ΣX_4^2

$$T = \Sigma X_1 + \Sigma X_2 + \Sigma X_3 = -5 + 0 + 10 + 15 = 20$$

$$N = 5 + 5 + 5 + 5 = 20$$

$$SSC = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} - \frac{T^2}{N}$$

$$= \frac{(-5)^2}{5} + \frac{(0)^2}{5} + \frac{(10)^2}{5} + \frac{(15)^2}{5} - \frac{20^2}{20}$$

$$= 5 + 0 + 20 + 45 - 20 = 50$$

$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 + \Sigma X_4^2 - \frac{T^2}{N} = 21 + 58 + 124 + 75 - \frac{20^2}{20} = 258$

$SSE = SST - SSC = 258 - 50 = 208$

Construction of ANOVA table for one-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between samples	$SSC = 50$	$c - 1 = 4 - 1 = 3$	$MSC = \frac{SSC}{c-1} = \frac{50}{3} = 16.7$	$F = \frac{MSC}{MSE} = 1.285$
Within samples	$SSE = 208$	$N - c = 20 - 4 = 16$	$MSE = \frac{SSE}{N-c} = \frac{208}{16} = 13$	

Under H_0, Calculated value of F is 1.285

At $\alpha = 0.05, \nu_1 = c - 1 = 3, \nu_2 = N - c = 16$, Tabulated value of F is 3.24

Calculated value < Tabulated value. Accept H_0.

Samples have come from the same universe.

6. **The three samples below have been obtained from normal populations with equal variances. Test the hypothesis that the sample means are equal.**

8	**7**	**12**
10	**5**	**9**
7	**10**	**13**
14	**9**	**12**
11	**9**	**14**

Assume $H_0: \mu_1 = \mu_2 = \mu_3$. All the three samples have taken from the same population.

Subtract 10 from each of the given values.

X_1	-2	0	-3	4	1
X_2	-3	-5	0	-1	-1
X_3	2	-1	3	2	4

To find: SSC and SSE

X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
-2	4	-3	9	2	4
0	0	-5	25	-1	1
-3	9	0	0	3	9
4	16	-1	1	2	4
1	1	-1	1	4	16
0	30	-10	36	10	34
ΣX_1	ΣX_1^2	ΣX_2	ΣX_2^2	ΣX_3	ΣX_3^2

$$T = \Sigma X_1 + \Sigma X_2 + \Sigma X_3 = 0 \pm 10 + 10 = 0$$

$$N = 5 + 5 + 5 = 15$$

$$SSC = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} - \frac{T^2}{N}$$

$$= \frac{(0)^2}{5} + \frac{(-10)^2}{5} + \frac{(10)^2}{5} - 0$$

$$= 0 + 20 + 20 = 40$$

$$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 - \frac{T^2}{N} = 30 + 36 + 34 - 0 = 100$$

$$SSE = SST - SSC = 100 - 40 = 60$$

Construction of ANOVA table for one-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between samples	$SSC = 40$	$c - 1 = 3 - 1 = 2$	$MSC = \frac{SSC}{c-1} = \frac{40}{2} = 20$	$F = \frac{MSC}{MSE} = 4$
Within samples	$SSE = 60$	$N - c = 15 - 3 = 12$	$MSE = \frac{SSE}{N-c} = \frac{60}{12} = 5$	

Under H_0, Calculated value of F is 4.

At $\alpha = 0.05, \nu_1 = c - 1 = 2, \nu_2 = N - c = 12$, Tabulated value of F is 3.88

Calculated value > Tabulated value. Reject H_0.

All the three samples have **not** taken from the same population.

7. The following table gives the yields on 15 sample plots under three varieties of seeds:

A	B	C
20	18	25
21	20	28
23	17	22
16	15	28
20	25	32

Find out if the average yields of land under different varieties of seeds show significant differences.

Assume $H_0: \mu_1 = \mu_2 = \mu_3$. The average yields of land under different varieties of seeds do not show significant differences.

Subtract 20 from each of the given values.

X_1	0	1	3	-4	0
X_2	-2	0	-3	-5	5
X_3	5	8	2	8	12

To find: SSC and SSE

X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
0	0	-2	4	5	25
1	1	0	0	8	64
3	9	-3	9	2	4
-4	16	-5	25	8	64
0	0	5	25	12	144
0	26	-5	63	35	301
ΣX_1	ΣX_1^2	ΣX_2	ΣX_2^2	ΣX_3	ΣX_3^2

$$T = \Sigma X_1 + \Sigma X_2 + \Sigma X_3 = 0 - 5 + 35 = 30$$

$$N = 5 + 5 + 5 = 15$$

$$SSC = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} - \frac{T^2}{N}$$

$$= \frac{(0)^2}{5} + \frac{(-5)^2}{5} + \frac{(35)^2}{5} - \frac{30^2}{15}$$

$$= 0 + 5 + 245 - 60 = 190$$

$$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 - \frac{T^2}{N} = 26 + 63 + 301 - \frac{30^2}{15} = 390 - 60 = 330$$

$$SSE = SST - SSC = 330 - 190 = 140$$

Construction of ANOVA table for one-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F −Ratio
Between samples	$SSC = 190$	$c - 1 = 3 - 1 = 2$	$MSC = \frac{SSC}{c-1} = \frac{190}{2} = 95$	$F = \frac{MSC}{MSE} = \frac{95}{11.67}$ $= 8.14$
Within samples	$SSE = 140$	$N - c = 15 - 3 = 12$	$MSE = \frac{SSE}{N-c} = \frac{140}{12} = 11.6$	

Under H_0, Calculated value of F is 8.14

At $\alpha = 0.05, \nu_1 = c - 1 = 2, \nu_2 = N - c = 12$, Tabulated value of F is 3.88

Calculated value > Tabulated value. Reject H_0.

The average yields of land under different varieties of seeds show significant differences.

8. **Test the significance of the variation of the retail prices of a commodity in three cities Mumbai, Chennai and Bengaluru. Four shops were chosen at random in each city and prices observed in rupees were as follows:**

Bengaluru	*Chennai*	*Mumbai*
16	14	4
8	10	10
12	10	8
14	6	8

Do the data indicate that the prices in the three cities are significantly different?

Assume $H_0: \mu_1 = \mu_2 = \mu_3$.

There is no significant difference in the prices in the three cities.

Subtract 10 from each of the given values.

X_1	6	-2	2	4
X_2	4	0	0	-4
X_3	-6	0	-2	-2

To find: SSC and SSE

X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
6	36	4	16	-6	36
-2	4	0	0	0	0
2	4	0	0	-2	4
4	16	-4	16	-2	4
10	60	0	32	-10	44
ΣX_1	ΣX_1^2	ΣX_2	ΣX_2^2	ΣX_3	ΣX_3^2

$$T = \Sigma X_1 + \Sigma X_2 + \Sigma X_3 = 10 + 0 - 10 = 0$$

$$N = 4 + 4 + 4 = 12$$

$$SSC = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \frac{(\Sigma X_3)^2}{n_3} - \frac{T^2}{N}$$

$$= \frac{(10)^2}{4} + \frac{(0)^2}{4} + \frac{(-10)^2}{4} - 0$$

$$= 25 + 0 + 25 - 0 = 50$$

$$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 - \frac{T^2}{N} = 60 + 32 + 44 - 0 = 136$$

$$SSE = SST - SSC = 136 - 50 = 86$$

Construction of ANOVA table for one-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between samples	$SSC = 50$	$c - 1 = 3 - 1 = 2$	$MSC = \frac{SSC}{c-1} = \frac{50}{2} = 25$	$F = \frac{MSC}{MSE} = \frac{25}{9.56} = 2.62$
Within samples	$SSE = 86$	$N - c = 12 - 3 = 9$	$MSE = \frac{SSE}{N-c} = \frac{86}{9} = 9.56$	

Under H_0, Calculated value of F is 2.62

At $\alpha = 0.05, \nu_1 = c - 1 = 2, \nu_2 = N - c = 9$, Tabulated value of F is 4.26

Calculated value < Tabulated value. Accept H_0.

The prices in the three cities are not significantly different.

ANOVA for two-way classification

In a two-way classification, the data are classified according to two different criteria or factors.

Expansion of abbreviations:

SSC – Sum of squares between columns	CF – Correction Factor
SSR – Sum of squares between rows	MSC – Mean squares of variations between columns
SST – Total sum of squares of variations	MSR – Mean squares of variations between rows
SSE – Sum of squares due to errors	MSE - Mean squares of variations between rows

Notation:

T_1, T_2, T_3, T_4 – Row totals	T – Grand total
T_5, T_6, T_7 – Column Totals	N – Total number of elements

ANOVA table for two-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between columns	SSC	$c-1$	$MSC = \frac{SSC}{c-1}$	$F_C = \frac{MSC}{MSE}$
Between rows	SSR	$r-1$	$MSR = \frac{SSR}{r-1}$	$F_R = \frac{MSR}{MSE}$
Error	SSE	$(c-1)(r-1)$	$MSE = \frac{SSE}{(c-1)(r-1)}$	

$$F_C = \frac{MSC}{MSE}, if\ MSC > MSE. Reciprocate\ otherwise.$$

$$F_R = \frac{MSR}{MSE}, if\ MSR > MSE. Reciprocate\ otherwise.$$

How to find SSC, SSE and SST from the following table?

	R_1	R_2	R_3	R_4	Total
C_1	a_1	b_1	c_1	d_1	T_5
C_2	a_2	b_2	c_2	d_2	T_6
C_3	a_3	b_3	c_3	d_3	T_7
Total	T_1	T_2	T_3	T_4	T

$$CF = \frac{T^2}{N}$$

$$SSC = \frac{T_1^2}{3} + \frac{T_2^2}{3} + \frac{T_3^2}{3} + \frac{T_4^2}{3} - CF$$

$$SSR = \frac{T_5^2}{4} + \frac{T_6^2}{4} + \frac{T_7^2}{4} - CF$$

$$SST = \Sigma a_i^2 + \Sigma b_i^2 + \Sigma c_i^2 + \Sigma d_i^2 - CF$$

$$SSE = SST - SSC$$

Working rule:

(v) Assume H_0:There is no significant difference between rows and between columns.

(vi) Construct ANOVA table for two-way classification.

Under H_0,

$$F_C = \frac{MSC}{MSE}, if\ MSC > MSE. Reciprocate\ otherwise.$$

$$F_R = \frac{MSR}{MSE}, if\ MSR > MSE. Reciprocate\ otherwise.$$

(vii) Find tabulated value at α level at ν_1, ν_2 degrees of freedom.

Where ν_1 - Degrees of freedom of the numerator

ν_2 - Degrees of freedom of the denominator

(viii) If calculated value < tabulated value, accept H_0. Reject otherwise.

1. **In a certain factory, production can be accomplished by four different workers on five different machines. A simple study, in context of a two way design without repeated values, is being made with two fold objectives of examining whether the four workers differ with respect to mean productivity and whether the mean productivity is the same for the five different machines. The researcher involved in this study reports while analyzing the gathered data as under:**

(i) Sum of squares of variance between machines is 35.2

(ii) Sum of squares of variance between workmen is 53.8

(iii) Sum of squares for total variance is 174.2

Setup ANOVA table for the given information and draw the inference about variance at 5% level of significance.

Assume H_0: There is no significant difference between the four workers with respect to mean productivity and mean productivity is the same for five different machines.

By data, $SSC = 35.2$, $SSR = 53.8$, $SST = 174.2$

Therefore, $SSE = SST - (SSC + SSR) = 174.2 - (35.2 + 53.8) = 85.2$

ANOVA table for two-way classification:

	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between columns	SSC	$c-1$	$MSC = \frac{SSC}{c-1}$	$F_C = \frac{MSC}{MSE}$
Between rows	SSR	$r-1$	$MSR = \frac{SSR}{r-1}$	$F_R = \frac{MSR}{MSE}$
Error	SSE	$(c-1)(r-1)$	$MSE = \frac{SSE}{(c-1)(r-1)}$	----

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between workers	35.2	4	$\frac{35.2}{4} = 8.8$	$F_C = \frac{8.8}{7.1} = 1.24$
Between machines	53.8	3	$\frac{53.8}{3} = 17.93$	$F_R = \frac{17.93}{7.1} = 2.53$
Error	85.2	12	$\frac{85.2}{12} = 7.1$	----

Conclusion:

(i) From the above table, calculated value of $F_C = 1.24$

The tabulated value of F_C for (4, 12) at 5% level of significance is 3.259

Calculated value < Tabulated value. Accept H_0.

Therefore, there is no significant difference between the four workers with respect to mean productivity.

(ii) From the above table, calculated value of $F_R = 2.53$

The tabulated value of F_R for (3, 12) at 5% level of significance is 3.49

Calculated value < Tabulated value. Accept H_0.

Therefore, mean productivity is the same for five different machines.

2. **The following table gives the number of refrigerators sold by 4 salesmen in three months May, June and July:**

Month	**Salesmen**			
	A	**B**	**C**	**D**
May	**50**	**40**	**48**	**39**
June	**46**	**48**	**50**	**45**
July	**39**	**44**	**40**	**39**

(i) Is there any significant difference in the sales made by the four salesmen?
(ii) Is there any significant difference in the sales made during different months?

Assume H_0: There is no significant difference in the sales made by the four salesmen and during different months.
The given data are coded by subtracting 40 from each observation.

	A	B	C	D	Total
Summer	10	0	8	−1	17
Winter	6	8	10	5	29
Monsoon	−1	4	0	−1	2
Total	15	12	18	3	48

To find: SSC, SSR, SSE

$$CF = \frac{T^2}{N} = \frac{48^2}{12} = 192$$

$$SSC = \frac{T_1^2}{3} + \frac{T_2^2}{3} + \frac{T_3^2}{3} - CF = \frac{15^2}{3} + \frac{12^2}{3} + \frac{18^2}{3} + \frac{3^2}{3} - 192 = 42$$

$$SSR = \frac{T_4^2}{4} + \frac{T_5^2}{4} + \frac{T_6^2}{4} - CF = \frac{17^2}{4} + \frac{29^2}{4} + \frac{2^2}{4} - 192 = 91.5$$

$$SST = \Sigma a_i^2 + \Sigma b_i^2 + \Sigma c_i^2 + \Sigma d_i^2 - CF$$

$$= (100 + 36 + 1) + (0 + 64 + 16) + (64 + 100 + 0) + (1 + 25 + 1) - 192$$

$$= 137 + 80 + 164 + 27 - 192 = 216$$

$$SSE = SST - (SSC + SSR) = 216 - (42 + 91.5) = 82.5$$

ANOVA table for two-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between columns	SSC	$c-1$	$MSC = \frac{SSC}{c-1}$	$F_C = \frac{MSC}{MSE}$
Between rows	SSR	$r-1$	$MSR = \frac{SSR}{r-1}$	$F_R = \frac{MSR}{MSE}$
Error	SSE	$(c-1)(r-1)$	$MSE = \frac{SSE}{(c-1)(r-1)}$	

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between Salesmen	42	3	$\frac{42}{3} = 14$	$F_C = \frac{14}{13.75} = 1.018$
Between months	91.5	2	$\frac{91.5}{2} = 45.75$	$F_R = \frac{45.75}{13.75} = 3.33$
Error	82.5	6	$\frac{82.5}{6} = 13.75$	

Conclusion:

(iii) From the above table, calculated value of $F_C = 1.018$
The tabulated value of F_C for (3, 6) at 5% level of significance is 4.76
Calculated value < Tabulated value. Accept H_0.
Therefore, there is no significant difference in the sales made by the four salesmen.

(iv) From the above table, calculated value of $F_R = 3.33$
The tabulated value of F_R for (2, 6) at 5% level of significance is 5.14
Calculated value < Tabulated value. Accept H_0.
Therefore, there is no significant difference in the sales made during different months.

3. **Perform a two-way ANOVA on the data given below:**

Plot of land	Treatment			
	A	B	C	D
I	38	40	41	39
II	45	42	49	36
III	40	38	42	42

(i) **Is there any significant difference between the treatments?**

(ii) **Is there any significant difference between the Plots?**

Assume H_0: There is no significant difference between treatments and plots.
The given data are coded by subtracting 40 from each observation.

	A	B	C	D	Total
Summer	−2	0	1	−1	−2
Winter	5	2	9	−4	12
Monsoon	0	−2	2	2	2
Total	3	0	12	−3	12

To find: SSC, SSR, SSE

$$CF = \frac{T^2}{N} = \frac{12^2}{12} = 12$$

$$SSC = \frac{T_1^2}{3} + \frac{T_2^2}{3} + \frac{T_3^2}{3} - CF = \frac{3^2}{3} + \frac{0^2}{3} + \frac{12^2}{3} + \frac{(-3)^2}{3} - 12 = 42$$

$$SSR = \frac{T_4^2}{4} + \frac{T_5^2}{4} + \frac{T_6^2}{4} - CF = \frac{(-2)^2}{4} + \frac{12^2}{4} + \frac{2^2}{4} - 12 = 26$$

$$SST = \Sigma a_i^2 + \Sigma b_i^2 + \Sigma c_i^2 + \Sigma d_i^2 - CF$$

$$= (4 + 25 + 0) + (0 + 4 + 4) + (1 + 81 + 4) + (1 + 16 + 4) - 12$$

$$= 29 + 8 + 86 + 21 - 12 = 132$$

$$SSE = SST - (SSC + SSR) = 132 - (42 + 26) = 64$$

ANOVA table for two-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between columns	SSC	$c-1$	$MSC = \frac{SSC}{c-1}$	$F_C = \frac{MSC}{MSE}$
Between rows	SSR	$r-1$	$MSR = \frac{SSR}{r-1}$	$F_R = \frac{MSR}{MSE}$
Error	SSE	$(c-1)(r-1)$	$MSE = \frac{SSE}{(c-1)(r-1)}$	

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between treatments	42	3	$\frac{42}{3} = 14$	$F_C = \frac{14}{10.67} = 1.312$
Between plots	26	2	$\frac{26}{2} = 13$	$F_R = \frac{13}{10.67} = 1.218$
Error	64	6	$\frac{64}{6} = 10.67$	

Conclusion:

(v) From the above table, calculated value of $F_C = 1.312$
The tabulated value of F_C for (3, 6) at 5% level of significance is 4.76
Calculated value < Tabulated value. Accept H_0.
Therefore, there is no significant difference between treatments.

(vi) From the above table, calculated value of $F_R = 1.218$
The tabulated value of F_R for (2, 6) at 5% level of significance is 5.14
Calculated value < Tabulated value. Accept H_0.
Therefore, there is no significant difference between plots.

4. **A tea company appoints 4 salesmen A, B, C and D and observes their sales in 3 seasons- Summer, Winter and Monsoon. The figures (in lakhs) are given in the following table:**

Seasons	Salesman A	Salesman B	Salesman C	Salesman D
Summer	36	36	21	35
Winter	28	29	31	32
Monsoon	26	28	29	29

(i) **Do the salesmen significantly differ in performance?**

(ii) **Is there significant difference between the seasons?**

Assume H_0: The salesmen and seasons are alike so far as the sales are concerned.

The given data are coded by subtracting 30 from each observation.

Seasons	Salesman A	Salesman B	Salesman C	Salesman D	Total
Summer	6	6	−9	5	8
Winter	−2	−1	1	2	0
Monsoon	−4	−2	−1	−1	−8
Total	0	3	−9	6	0

To find: SSC, SSR, SSE

$$CF = \frac{T^2}{N} = \frac{(0)^2}{9} = 0$$

$$SSC = \frac{T_1^2}{3} + \frac{T_2^2}{3} + \frac{T_3^2}{3} - CF$$

$$= \frac{0^2}{3} + \frac{3^2}{3} + \frac{(-9)^2}{3} + \frac{6^2}{3} - 0 = 42$$

$$SSR = \frac{T_4^2}{4} + \frac{T_5^2}{4} + \frac{T_6^2}{4} - CF$$

$$= \frac{8^2}{4} + \frac{0^2}{4} + \frac{(-8)^2}{4} - 0 = 32$$

$$SST = \Sigma a_i^2 + \Sigma b_i^2 + \Sigma c_i^2 + \Sigma d_i^2 - CF$$

$$= (6^2 + (-2)^2 + (-4)^2) + (6^2 + (-1)^2 + (-2)^2)$$

$$+((-9)^2 + 1^2 + (-1)^2 + (5^2 + 2^2 + (-1)^2)) - 0$$

$$= 56 + 41 + 83 + 30 = 210$$

$$SSE = SST - (SSC + SSR) = 210 - (42 + 32) = 136$$

ANOVA table for two-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between columns	SSC	$c-1$	$MSC=\dfrac{SSC}{c-1}$	$F_C=\dfrac{MSC}{MSE}$
Between rows	SSR	$r-1$	$MSR=\dfrac{SSR}{r-1}$	$F_R=\dfrac{MSR}{MSE}$
Error	SSE	$(c-1)(r-1)$	$MSE=\dfrac{SSE}{(c-1)(r-1)}$	

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between salesmen	42	3	$\frac{42}{3}=14$	$F_C=\dfrac{22.67}{14}=1.619$
Between seasons	32	2	$\frac{32}{2}=16$	$F_R=\dfrac{22.67}{16}=1.417$
Error	136	6	$\frac{136}{6}=22.67$	

Conclusion:

(i) From the above table, calculated value of $F_C=1.619$
The tabulated value of F_C for $(6,3)$ at 5% level of significance is 8.94
Calculated value < Tabulated value. Accept H_0.
Therefore, the salesmen are alike so far as the sales are concerned.

(ii) From the above table, calculated value of $F_R=2.38$
The tabulated value of F_R for $(6,2)$ at 5% level of significance is 19.33
Calculated value < Tabulated value. Accept H_0.
Therefore, the seasons are alike so far as the sales are concerned.

5. **To study the performance of three detergents and three different water temperatures the following whiteness readings were obtained with specially designed equipment.**

Water temperature	Detergent A	Detergent B	Detergent C
Cold water	57	55	67
Warm water	49	52	68
Hot water	54	46	58

Perform a two-way analysis using 5% level of significance. (Given $F = 6.94$)

Assume H_0: The performance of three detergents are equal and the performance of three different temperature of waters are equal.

The given data are coded by subtracting 50 from each observation.

Water temperature	Detergent A	Detergent B	Detergent C	Total
Cold water	7	5	17	29
Warm water	−1	2	18	19
Hot water	4	−4	8	8
Total	10	3	43	56

To find: SSC, SSR, SSE

$$CF = \frac{T^2}{N} = \frac{56^2}{9} = 348.44$$

$$SSC = \frac{T_1^2}{3} + \frac{T_2^2}{3} + \frac{T_3^2}{3} - CF$$

$$= \frac{10^2}{3} + \frac{3^2}{3} + \frac{43^2}{3} - 348.44 = 304.22$$

$$SSR = \frac{T_4^2}{4} + \frac{T_5^2}{4} + \frac{T_6^2}{4} - CF$$

$$= \frac{29^2}{4} + \frac{19^2}{4} + \frac{8^2}{4} - 348.44 = 73.55$$

$$SST = \Sigma a_i^2 + \Sigma b_i^2 + \Sigma c_i^2 + \Sigma d_i^2 - CF$$

$$= (7^2 + (-1)^2 + 4^2) + (5^2 + 2^2 + (-4)^2) + (17^2 + 18^2 + 8^2) - 348.44$$

$$= 439.56$$

$$SSE = SST - (SSC + SSR) = 439.56 - (304.22 + 73.55) = 61.79$$

ANOVA table for two-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between columns	SSC	$c-1$	$MSC = \frac{SSC}{c-1}$	$F_C = \frac{MSC}{MSE}$
Between rows	SSR	$r-1$	$MSR = \frac{SSR}{r-1}$	$F_R = \frac{MSR}{MSE}$
Total	SSE	$(c-1)(r-1)$	$MSE = \frac{SSE}{(c-1)(r-1)}$	

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between detergents	304.22	2	$\frac{304.22}{2} = 152.11$	$F_C = \frac{152.11}{15.45} = 9.85$
Between types of waters	73.55	2	$\frac{73.55}{2} = 36.78$	$F_R = \frac{36.78}{15.45} = 2.38$
Error	61.79	4	$\frac{61.79}{4} = 15.45$	

Conclusion:

(i) From the above table, calculated value of $F_C = 9.85$

The tabulated value of F_C for (2, 4) at 5% level of significance is 6.94

Calculated value > Tabulated value. Reject H_0.

Therefore, performance of three detergents are not equal.

(ii) From the above table, calculated value of $F_R = 2.38$

The tabulated value of F_R for (2, 4) at 5% level of significance is 6.94

Calculated value < Tabulated value. Accept H_0.

Therefore, performance of three different temperature of waters are equal.

6. A Farmer applies three types of fertilizers on 4 separate plots. The figure on yield per square acre are tabulated below:

Plots	Yield				Total
Fertilizers	A	B	C	D	
Nitrogen	6	4	8	6	24
Potash	7	6	6	9	28
Phosphates	8	5	10	9	32
Total	21	15	24	24	84

Find out if the plots are materially different in fertility as also, if three fertilizers make any material difference in yields.

Assume H_0: Plots are equally fertile; fertilizers are equally effective.

To find: SSC, SSR, SSE

$$CF = \frac{T^2}{N} = \frac{84^2}{12} = 588$$

$$SSC = \frac{T_1^2}{3} + \frac{T_2^2}{3} + \frac{T_3^2}{3} + \frac{T_4^2}{3} - CF$$

$$= \frac{21^2}{3} + \frac{15^2}{3} + \frac{24^2}{3} + \frac{24^2}{3} - 588 = 18$$

$$SSR = \frac{T_5^2}{4} + \frac{T_6^2}{4} + \frac{T_7^2}{4} - CF$$

$$= \frac{24^2}{4} + \frac{28^2}{4} + \frac{32^2}{4} - 588 = 8$$

$$SST = \Sigma a_i^2 + \Sigma b_i^2 + \Sigma c_i^2 + \Sigma d_i^2 - CF$$

$$= (6^2 + 7^2 + 8^2) + (4^2 + 6^2 + 5^2) + (8^2 + 6^2 + 10^2) + (6^2 + 9^2 + 9^2) - 588$$

$$= 36$$

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

$$= 36 - (18 + 8) = 10$$

ANOVA table for two-way classification:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between columns	SSC	$c-1$	$MSC = \frac{SSC}{c-1}$	$F_C = \frac{MSC}{MSE}$
Between rows	SSR	$r-1$	$MSR = \frac{SSR}{r-1}$	
Error	SSE	$(c-1)(r-1)$	$MSE = \frac{SSE}{(c-1)(r-1)}$	$F_R = \frac{MSR}{MSE}$

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F –Ratio
Between plots	18	3	$\frac{18}{3} = 6$	
Between fertilizers	8	2	$\frac{8}{2} = 4$	$F_C = \frac{6}{1.67} = 3.6$
Error	10	6	$\frac{10}{6} = 1.67$	$F_R = \frac{4}{1.67} = 2.4$

Conclusion:

(iii) From the above table, calculated value of $F_C = 3.6$
The tabulated value of F_C for (3, 6) at 5% level of significance is 7.76
Calculated value < Tabulated value. Accept H_0.
Therefore, plots are equally fertile.

(iv) From the above table, calculated value of $F_R = 2.4$
The tabulated value of F_R for (2, 4) at 5% level of significance is 5.14
Calculated value < Tabulated value. Accept H_0.
Therefore, fertilizers are equally effective.

Critical Values of the F-Distribution: α = 0.05

Denom. d.f.	Numerator Degrees of Freedom 1	2	3	4	5	6	7	8	9	10
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883	240.543	241.882
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165
31	4.160	3.305	2.911	2.679	2.523	2.409	2.323	2.255	2.199	2.153
32	4.149	3.295	2.901	2.668	2.512	2.399	2.313	2.244	2.189	2.142
33	4.139	3.285	2.892	2.659	2.503	2.389	2.303	2.235	2.179	2.133
34	4.130	3.276	2.883	2.650	2.494	2.380	2.294	2.225	2.170	2.123
35	4.121	3.267	2.874	2.641	2.485	2.372	2.285	2.217	2.161	2.114