

Unit-V	Analysis of Variance (ANOVA)
5.1	Analysis of Variance
5.2	Assumptions
5.3	One way Classification
5.4	Two way Classifications (No Derivations)

## 5.1 Analysis of Variance

The analysis of variance is a powerful statistical tool for tests of significance. The term Analysis of Variance was introduced by Prof. R.A. Fisher to deal with problems in agricultural research. The test of significance based on t-distribution is an adequate procedure only for testing the significance of the difference between two sample means. In a situation where we have three or more samples to consider at a time, an alternative procedure is needed for testing the hypothesis that all the samples are drawn from the same population, i.e., they have the same mean. For example, five fertilizers are applied to four plots each of wheat and yield of wheat on each of the plot is given. We may be interested in finding out whether the effect of these fertilizers on the yields is significantly different or in other words whether the samples have come from the same normal population. The answer to this problem is provided by the technique of analysis of variance. Thus basic purpose of the analysis of variance is to test the homogeneity of several means.

Variation is inherent in nature. The total variation in any set of numerical data is due to a number of causes which may be classified as:

- (i) Assignable causes and
- (ii) Chance causes

The variation due to assignable causes can be detected and measured whereas the variation due to chance causes is beyond the control of human hand and cannot be traced separately.

### Definition:

According to R.A. Fisher, Analysis of Variance (ANOVA) is the “Separation of Variance ascribable to one group of causes from the variance ascribable to other group”. By this technique the total variation in the sample data is expressed as the sum of its nonnegative components where each of these components is a measure of the variation due to some specific independent source or factor or cause.

## 5.2 Assumptions

For the validity of the F-test in ANOVA the following assumptions are made.

- (i) The observations are independent.
- (ii) Parent population from which observations are taken is normal and
- (iii) Various treatment and environmental effects are additive in nature.

### 5.3 One way Classification

Let us suppose that N observations  $x_{ij}$ ,  $i = 1, 2, \dots, k$  ;  $j = 1, 2, \dots, n_i$ ) of a random variable X are grouped on some basis, into k classes of sizes  $n_1, n_2, \dots, n_k$  respectively  

$$N = \sum_{i=1}^k n_i$$
 as exhibited below.

	Mean	Total
$x_{11} \quad x_{12} \quad \dots \quad x_{1n_1}$	$\bar{x}_{1\cdot}$	$T_1$
$x_{21} \quad x_{22} \quad \dots \quad x_{2n_2}$	$x_{2\cdot}$	$T_2$
$\cdot \quad \cdot \quad \dots \quad \cdot$	$\cdot$	$\cdot$
$\cdot \quad \cdot \quad \dots \quad \cdot$	$\cdot$	$\cdot$
$\cdot \quad \cdot \quad \dots \quad \cdot$	$\cdot$	$\cdot$
$\cdot \quad \cdot \quad \dots \quad \cdot$	—	
$x_{i1} \quad x_{i2} \quad \dots \quad x_{in_i}$	$x_{i\cdot}$	$T_i$
$\cdot \quad \cdot \quad \dots \quad \cdot$	$\cdot$	$\cdot$
$\cdot \quad \cdot \quad \dots \quad \cdot$	$\cdot$	$\cdot$
$\cdot \quad \cdot \quad \dots \quad \cdot$	$\cdot$	$\cdot$
$x_{k1} \quad x_{k2} \quad \dots \quad x_{kn_k}$	$\bar{x}_{k\cdot}$	$T_k$
		<b>G</b>

The total variation in the observation  $x_{ij}$  can be split into the following two components :

- (i) The variation between the classes or the variation due to different bases of classification, commonly known as treatments.
- (ii) The variation within the classes i.e., the inherent variation of the random variable within the observations of a class.

The first type of variation is due to assignable causes which can be detected and controlled by human endeavour and the second type of variation due to chance causes which are beyond the control of human hand.

In particular, let us consider the effect of k different rations on the yield in milk of N cows (of the same breed and stock) divided into k classes of sizes  $n_1, n_2, \dots, n_k$  respectively.

$$N = \sum_{i=1}^k n_i$$

Hence the sources of variation are

- (i) Effect of the ratios
- (ii) Error due to chance causes produced by numerous causes that they are not detected and identified.

**Test Procedure:**

The steps involved in carrying out the analysis are:

**1) Null Hypothesis:**

The first step is to set up of a null hypothesis

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

Alternative hypothesis  $H_1$ : all  $\mu_i$  's are not equal ( $i = 1, 2, \dots, k$ )

**2) Level of significance :** Let  $\alpha : 0.05$

**3) Test statistic:**

Various sum of squares are obtained as follows.

- a) Find the sum of values of all the (N) items of the given data. Let this grand total represented by 'G'. Then correction factor (C.F) =  $\frac{G^2}{N}$

- b) Find the sum of squares of all the individual items ( $x_{ij}$ ) and then the Total sum of squares (TSS) is

$$TSS = \sum \sum x_{ij}^2 - C.F$$

- c) Find the sum of squares of all the class totals (or each treatment total)  $T_i$  ( $i:1,2,\dots,k$ ) and then the sum of squares between the classes or between the treatments (SST) is

$$SST = \sum_{i=1}^k \frac{T_i^2}{n_i} - C$$

where  $n_i$  ( $i: 1, 2, \dots, k$ ) is the number of observations in the  $i^{\text{th}}$  class or number of observations received by  $i^{\text{th}}$  treatment

- d) Find the sum of squares within the class or sum of squares due to error (SSE) by subtraction.

$$SSE = TSS - SST$$

**4) Degrees of freedom (d.f):**

The degrees of freedom for total sum of squares (TSS) is (N-1). The degrees of freedom for SST is (k-1) and the degrees of freedom for SSE is (N-k)

5) Mean sum of squares:

The mean sum of squares for treatments is  $\frac{SST}{k-1}$  and mean sum of squares for error is  $\frac{SSE}{N-k}$ .

6) ANOVA Table

The above sum of squares together with their respective degrees of freedom and mean sum of squares will be summarized in the following table.

ANOVA Table for one-way classification

Sources of variation	d.f	S.S	M.S.S	F ratio
Between treatments	K - 1	SST	$\frac{SST}{k-1} = MST$	$\frac{MST}{MSE} = F_T$
Error	N - k	SSE	$\frac{SSE}{N-k} = MSE$	
Total	N - 1			

7) Calculation of variance ratio:

Variance ratio of F is the ratio between greater variance and smaller variance, thus

$$F = \frac{\text{Variance between the treatments}}{\text{Variance within the treatment}} = \frac{MST}{MSE}$$

If variance within the treatment is more than the variance between the treatments, then numerator and denominator should be interchanged and degrees of freedom adjusted accordingly.

8) Critical value of F or Table value of F:

The Critical value of F or table value of F is obtained from F table for (k-1, N-k) d.f at 5% level of significance.

9) Inference:

If calculated F value is less than table value of F, we may accept our null hypothesis  $H_0$  and say that there is no significant difference between treatments.

If calculated F value is greater than table value of F, we reject our  $H_0$  and say that the difference between treatments is significant.

Example 1:

Three processes A, B and C are tested to see whether their outputs are equivalent. The following observations of outputs are made:

A	10	12	13	11	10	14	15	13
B	9	11	10	12	13			
C	11	10	15	14	12	13		

Carry out the analysis of variance and state your conclusion.

Solution:

To carry out the analysis of variance, we form the following tables

									Total	Squares
A	10	12	13	11	10	14	15	13	98	9604
B	9	11	10	12	13				55	3025
C	11	10	15	14	12	13			75	5625
									G = 228	

Squares:

A	100	144	169	121	100	196	225	169		
B	81	121	100	144	169					
C	121	100	225	196	144	169				
									Total = 2794	

**Test Procedure:**

**Null Hypothesis:**  $H_0: \mu_1 = \mu_2 = \mu_3$

i.e., There is no significant difference between the three processes.

**Alternative Hypothesis**  $H_1: \mu_1 \neq \mu_2 \neq \mu_3$

**Level of significance :** Let  $\alpha : 0.05$

Test statistic

$$\begin{aligned}
 \text{Correct factor (C.F)} &= \frac{G^2}{N} \\
 &= \frac{228^2}{19} \\
 &= \frac{51984}{19} \\
 &= 2736
 \end{aligned}$$

$$\begin{aligned} \text{Total sum of squares (TSS)} &= \sum \sum x_{ij}^2 - C.F \\ &= 2794 - 2736 \\ &= 58 \end{aligned}$$

Sum of squares due to processes = (SST)

$$\begin{aligned} &= \sum_{i=1}^3 \frac{T_i^2}{n_i} - C.F \\ &= \frac{9604}{8} + \frac{3025}{5} + \frac{5625}{6} - 2736 \\ &= (1200.5 + 605 + 937.5) - 2736 \\ &= 2743 - 2736 \\ &= 7 \end{aligned}$$

$$\begin{aligned} \text{Sum of squares due to error (SSE)} &= \text{TSS} - \text{SST} \\ &= 58 - 7 = 51 \end{aligned}$$

ANOVA Table

Sources of variation	d.f	S.S	M.S.S	F ratio
Between Processes	3 - 1 = 2	7	$\frac{7}{2} \equiv 3.50$	$\frac{3.5}{3.19} \equiv 1.097$
Error	16	51	$\frac{51}{16} \equiv 3.19$	
Total	19 - 1 = 18			

**Table Value:**

Table value of  $F_e$  for (2,16) d.f at 5% level of significance is 3.63.

*Inference:*

Since calculated  $F_0$  is less than table value of  $F_e$ , we may accept our  $H_0$  and say that there is no significant difference between the three processes.

Example 2:

A test was given to five students taken at random from the fifth class of three schools of a town. The individual scores are

School I	9	7	6	5	8
School II	7	4	5	4	5
School III	6	5	6	7	6

Carry out the analysis of variance.

Solution:

To carry out the analysis of variance, we form the following tables.

						Total	Squares
School I	9	7	6	5	8	35	1225
School II	7	4	5	4	5	25	625
School III	6	5	6	7	6	30	900
					Total	G = 90	2750

**Squares:**

School I	81	49	36	25	64
School II	49	16	25	16	25
School III	36	25	36	49	36
				Total = 568	

Test Procedure:

**Null Hypothesis:**  $H_0: \mu_1 = \mu_2 = \mu_3$  i.e., There is no significant difference between the performance of schools.

**Alternative Hypothesis**  $H_1: \mu_1 \neq \mu_2 \neq \mu_3$

**Level of significance :** Let  $\alpha : 0.05$

Test statistic

$$\begin{aligned} \text{Correct factor (C.F)} &= \frac{G^2}{N} \\ &= \frac{90^2}{15} \\ &= \frac{8100}{15} = 540 \end{aligned}$$

$$\text{Total sum of squares (TSS)} = \sum \sum x_{ij}^2 - C. F$$



$$= 568 - 540 = 28$$

$$\text{Sum of squares between schools} = \frac{\sum T_i^2}{n_i} - \text{C.F}$$

$$= \frac{2750}{5} - 540$$

$$= 550 - 540 = 10$$

Sum of squares due to error (SSE) = TSS – SST

$$= 28-10 = 18$$

ANOVA Table

Sources of variation	d.f	S.S	M.S.S	F ratio
Between Schools	3 – 1 = 2	10	$\frac{10}{2} = 5.0$	$\frac{5}{1.5} = 3.33$
Error	12	18	$\frac{18}{12} = 1.5$	
Total	15 – 1 = 14			

**Table Value:**

Table value of  $F_e$  for (2,12) d.f at 5% level of significance is 3.8853

*Inference:*

Since calculated  $F_0$  is less than table value of  $F_e$ , we may accept our  $H_0$  and say that there is no significant difference between the performance of schools.

### 5.4 Two way Classification

Let us consider the case when there are two factors which may affect the variate values  $x_{ij}$ , e.g the yield of milk may be affected by difference in treatments i.e., rations as well as the difference in variety i.e., breed and stock of the cows. Let us now suppose that the N cows are divided into h different groups or classes according to their breed and stock, each group containing k cows and then let us consider the effect of k treatments (i.e., rations given at random to cows in each group) on the yield of milk.

Let the suffix 'i' refer to the treatments (rations) and j refer to the varieties (breed of the cow), then the yields of milk  $x_{ij}$  (i:1,2, ...,k; j:1,2....h) of  $N = h \times k$  cows furnish the data for the comparison of the treatments (rations). The yields may be expressed as variate values in the following  $k \times h$  two-way table.

		Mean	Total	
$x_{11}$	$x_{12}$	$x_{1j} \dots x_{1h}$	$x_{1\cdot}$	$T_1$
$x_{21}$	$x_{22}$	$x_{2j} \dots x_{2h}$	$x_{2\cdot}$	$T_2$
.	.	.	.	.
.	.	.	.	.
$x_{i1}$	$x_{i2}$	$x_{ij} \dots x_{ih}$	$x_{i\cdot}$	$T_i$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
$x_{k1}$	$x_{k2}$	$x_{kj} \dots x_{kh}$	$\bar{x}_{k\cdot}$	$T_k$
Mean	$\bar{x}_{\cdot 1}$	$\bar{x}_{\cdot 2}$	$\bar{x}_{\cdot j} \dots \bar{x}_{\cdot h}$	$\bar{x}$
Total	$T_{\cdot 1}$	$T_{\cdot 2}$	$\dots T_{\cdot j} \dots T_{\cdot h}$	G

The total variation in the observation  $x_{ij}$  can be split into the following three components:

- (i) The variation between the treatments (rations)
- (ii) The variation between the varieties (breed and stock)
- (iii) The inherent variation within the observations of treatments and within the observations of varieties.

The first two types of variations are due to assignable causes which can be detected and controlled by human endeavour and the third type of variation due to chance causes which are beyond the control of human hand.

**Test procedure for Two - way analysis:**

The steps involved in carrying out the analysis are:

**1. Null hypothesis:**

The first step is to setting up a null hypothesis  $H_0$

$$H_0 : \mu_{1\cdot} = \mu_{2\cdot} = \dots \mu_{k\cdot} = \mu$$

$$H_0 : \mu_{\cdot 1} = \mu_{\cdot 2} = \dots \mu_{\cdot h} = \mu$$

i.e., there is no significant difference between rations (treatments) and there is no significant difference between varieties (breed and stock)

**2. Level of significance:** Let  $\alpha : 0.05$

**3. Test Statistic:**

Various sums of squares are obtained as follows:

- a) Find the sum of values of all the  $N$  ( $k \times h$ ) items of the given data. Let this grand total represented by 'G' Then correction factor (C.F) =  $\frac{G^2}{N}$ .

- b) Find the sum of squares of all the individual items ( $x_{ij}$ ) and then the total sum of squares (TSS)  $\sum_{i=1}^k \sum_{j=1}^k x_{ij}^2 - C.F$

- c) Find the sum of squares of all the treatment (rations) totals, i.e., sum of squares of row totals in the  $h \times k$  two-way table. Then the sum of squares between treatments or sum of squares between rows is

$$SST = SSR = \sum_{i=1}^k \frac{T_i^2}{h} - C.F$$

where  $h$  is the number of observations in each row

- d) Find the sum of squares of all the varieties (breed and stock) totals, in the  $h \times k$  two-way table. Then the sum of squares between varieties or sum of squares between columns is

$$SSV = SSC = \frac{\sum_{j=1}^k T^2_{.j}}{k} - C.F$$

where  $k$  is the number of observations in each column.

- e) Find the sum of squares due to error by subtraction:

$$\text{i.e., } SSE = TSS - SSR - SSC$$

#### 4. Degrees of freedom:

- (i) The degrees of freedom for total sum of squares is  $N - 1 = hk - 1$
- (ii) The degrees of freedom for sum of squares between treatments is  $k - 1$
- (iii) The degree of freedom for sum of squares between varieties is  $h - 1$
- (iv) The degrees of freedom for error sum of squares is  $(k - 1)(h - 1)$

5. Mean sum of squares (MSS)

(i) Mean sum of squares for treatments (MST) is  $\frac{SST}{k-1}$

(i) Mean sum of squares for varieties (MSV) is  $\frac{SSV}{h-1}$

(ii) Mean sum of squares for error (MSE) is  $\frac{SSE}{(h-1)(k-1)}$

**6. ANOVA TABLE**

The above sum of squares together with their respective degrees of freedom and mean sum of squares will be summarized in the following table.

ANOVA Table for Two-way classification

Sources of variation	d.f	SS	MSS	F <sub>0</sub> - ratio
Between Treatments	k - 1	SST	MST	$\frac{MST}{MSE} = F_R$
Between Varieties	h - 1	SSV	MSV	$\frac{MSV}{MSE} = F_c$
Error	(h - 1) (k - 1)	SSE	MSE	
Total	N - 1			

**7. Critical values Fe or Table values of F:**

- (i) The critical value or table value of ‘ F’ for between treatments is obtained from F table for [(k - 1), (k - 1) (h - 1)] d.f at 5% level of significance.
- (ii) The critical value or table value of Fe for between varieties is obtained from F table for [(h - 1), (k - 1) (h - 1)] d.f at 5% level of significance.

**8. Inference:**

- (i) If calculated F<sub>0</sub> value is less than or greater than the table value of F<sub>c</sub> for between treatments (rows) H<sub>0</sub> may be accepted or rejected accordingly.
- (ii) If calculated F<sub>0</sub> value is less than or greater than the table value of F<sub>c</sub> for between varieties (column), H<sub>0</sub> may be accepted or rejected accordingly.

Example 3:

Three varieties of coal were analysed by four chemists and the ash-content in the varieties was found to be as under.

	Chemists			
Varieties	1	2	3	4
A	8	5	5	7

B	7	6	4	4
C	3	6	5	4

Carry out the analysis of variance.

Solution:

To carry out the analysis of variance we form the following tables

Chemists						
Varieties	1	2	3	4	Total	Squares
A	8	5	5	7	25	625
B	7	6	4	4	21	441
C	3	6	5	4	18	324
Total	18	17	14	15	<b>G = 64</b>	1390
Squares	324	289	196	225	1034	

Individual squares

Chemists				
Varieties	1	2	3	4
A	64	25	25	49
B	49	36	16	16
C	9	36	25	16
			Total	= 366

Test Procedure :

**Null hypothesis:**

$$H_0 : \mu_{1.} = \mu_{2.} = \mu_{3.} = \mu$$

$$H_0 : \mu_{.1} = \mu_{.2} = \mu_{.3} = \mu_{.4} = \mu$$

(i) i.e., there is no significant difference between varieties (rows)

(ii) i.e., there is no significant difference between chemists (columns)

Alternative hypothesis  $H_1$ :

(i) not all  $\mu_i$  's equal

(ii) not all  $\mu_j$  's equal

2. Level of significance:

Let  $\alpha : 0.05$

Test statistic:

$$\begin{aligned} \text{Correction factor (C.F)} &= \frac{G^2}{N} = \frac{G^2}{h \times k} \\ &= \frac{(64)^2}{3 \times 4} = \frac{(64)^2}{12} \end{aligned}$$



$$= \frac{4096}{12} = 341.33$$

$$\begin{aligned} \text{Total sum of squares (TSS)} &= \sum_{i=1}^k \sum_{j=1}^k x_{ij}^2 - \text{C.F} \\ &= 366 - 341.33 \\ &= 24.67 \end{aligned}$$

Sum of squares between varieties (Rows)

$$\begin{aligned} &= \frac{\sum T_i^2}{4} - \text{C.F} \\ &= \frac{1390}{4} - 341.33 \\ &= 347.5 - 341.33 \\ &= 6.17 \end{aligned}$$

Sum of squares between chemists (columns)

$$\begin{aligned} &= \frac{\sum T_j^2}{3} - \text{C.F} \\ &= \frac{1034}{3} - 341.33 \\ &= 344.67 - 341.33 \\ &= 3.34 \end{aligned}$$

Sum of square due to error (SSE)

$$\begin{aligned} &= \text{TSS} - \text{SSR} - \text{SSC} \\ &= 24.67 - 6.17 - 3.34 \\ &= 24.67 - 9.51 \\ &= 15.16 \end{aligned}$$

ANOVA TABLE

Sources of variation	d.f	SS	MSS	F - ratio
Between Varieties	$3 - 1 = 2$	6.17	3.085	$\frac{3.085}{2.527} = 1.22$
Between Chemists	$4 - 1 = 3$	3.34	1.113	$\frac{2.527}{1.113} = 2.27$

Error	6	15.16	2.527	
Total	$12 - 1 = 11$			

Table value :

- (i) Table value of  $F_e$  for (2,6) d.f at 5% level of significance is 5.14
- (ii) Table value of  $F_e$  for (6,3) d.f at 5% level of significance is 8.94

Inference:

- (i) Since calculated  $F_0$  is less than table value of  $F_e$ , we may accept our  $H_0$  for between varieties and say that there is no significant difference between varieties.
- (ii) Since calculated  $F_0$  is less than the table value of  $F_e$  for chemists, we may accept our  $H_0$  and say that there is no significant difference between chemists.

## Exercise – 7

I. Choose the best answers:

1. Equality of several normal population means can be tested by
  - (a) Bartlett' s test
  - (b) F - test
  - (c)  $\chi^2$  -test
  - (d) t- test
2. Analysis of variance technique was developed by
  - (a) S. D. Poisson
  - (b) Karl - Pearson
  - (c) R.A. Fisher
  - (d) W. S. Gosset
3. Analysis of variance technique originated in the field of
  - (a) Agriculture
  - (b) Industry
  - (c) Biology
  - (d) Genetics
4. One of the assumption of analysis of variance is that the population from which the samples are drawn is
  - (a) Binomial
  - (b) Poisson
  - (c) Chi-square
  - (d) Normal
5. In the case of one-way classification the total variation can be split into
  - (a) Two components
  - (b) Three components
  - (c) Four components
  - (d) Only one component
6. In the case of one-way classification with N observations and t treatments, the error degrees of freedom is
  - (a) N-1
  - (b) t -1
  - (c) N - t
  - (d) Nt
7. In the case of one-way classification with t treatments, the mean sum of squares for treatment is
  - (a) SST/N-1
  - (b) SST/ t-1
  - (c) SST/N-t
  - (d) SST/t

8. In the case of two-way classification with  $r$  rows and  $c$  columns, the degrees of freedom for error is
- (a)  $(rc) - 1$                       (b)  $(r-1).c$                       (c)  $(r-1)(c-1)$                       (d)  $(c-1).r$
9. In the case of two-way classification, the total variation (TSS) equals.
- (a)  $SSR + SSC + SSE$                       (b)  $SSR - SSC + SSE$   
(c)  $SSR + SSC - SSE$                       (d)  $SSR + SSC$ .
10. With 90, 35, 25 as TSS, SSR and SSC respectively in case of two way classification, SSE is
- (a) 50                      (b) 40                      (c) 30                      (d) 20

**II.** *fill in the blanks*

11. The technique of analysis of variance was developed by\_\_\_\_\_.
12. One of the assumptions of Analysis of variance is: observations are\_\_\_\_\_.
13. Total variation in two – way classification can be split into\_\_\_\_\_components.
14. In the case of one way classification with 30 observations and 5 treatment, the degrees freedom for SSE is\_\_\_\_\_.
15. In the case of two-way classification with 120, 54, 45 respectively as TSS, SSC, SSE, the SSR is\_\_\_\_\_.

**III.** *Answer the following:*

16. What is analysis of variance?
17. Distinguish between t-test for difference between means and ANOVA.
18. State all the assumptions involved in analysis of variance technique.
19. Explain the structure for one-way classification.
20. Write down the ANOVA table for one-way classification.
21. Distinguish between one - way classification and two-way classification.
22. Explain the structure of two-way classification data.
23. Explain the procedure of obtaining various sums of squares in one-way classification.
24. Write down ANOVA table for two-way classification.

25. Explain the procedure of obtaining various sums of squares in two-way classification.
26. A test was given to a number of students taken at random from the eighth class from each of the 5 schools.

The Individual Scores are:

Schools

I	II	III	IV	V
8	9	12	10	12
9	7	14	11	11
10	11	15	9	10
7	12	12	12	9
8	13	11	10	13

Carry out the analysis of variance and give your conclusions.

27. The following figures relate to production in kg of three varieties A, B and C of wheat shown in 12 plots.

A: 20 18 19

B: 17 16 19 18

C: 20 21 20 19 18

Is there any significant difference in the production of the three varieties

28. A special type of fertilizer was used in four agricultural fields A,B,C and D each field was divided into four beds and the fertilizer was applied over them. The respective yields of the beds of four fields are given below. Find whether the difference in mean yields of fields is significant or not?

Plot yield

A	B	C	D
8	9	3	3
12	4	8	7
1	7	2	8
9	1	5	2

29. The following table gives the retail prices of a commodity in (Rs. Per Kg) in some shops selected at random in four cities.

	A	22	24	20	21
CITY	B	20	19	21	22
	C	19	17	21	18
	D	20	22	21	22

Analysis the data to test the significance of the differences between the price of commodity in four cities.

30. For experiments determine the moisture content of sample of a powder, each man taking a sample from each of six consignments Their assessments are:

Consignment

Observer	1	2	3	4	5	6
1	9	10	9	10	11	11
2	12	11	9	11	10	10
3	11	10	10	12	11	10
4	12	13	11	14	12	10

Perform an analysis of variance of these data and discuss if there is any significant difference between consignments or between observers.

31. The following are the defective pieces produced by four operators working in turn, on four different machines:

Operator

Machine	I	II	III	IV
A	3	2	3	2
B	3	2	3	4
C	2	3	4	3
D	3	4	3	2

Perform analysis of variance at 5% level of significance to ascertain whether variability in production is due to variability in operator's performance or variability in machine's performance.

32. Apply the technique of Analysis of variance to the following data relating to yields of 4 varieties of wheat in 3 blocks.

Blocks

Varieties	1	2	3
I	10	9	8
II	7	7	6
III	8	5	4
IV	5	4	4

33. Four Varieties of potato are planted, each on five plots of ground of the same size and type and each variety is treated with five different fertilizers. The yields in tons are as follows.

Fertilizers

Varieties	F1	F2	F3	F4	F5
V1	1.9	2.2	2.6	1.8	2.1
V2	2.5	1.9	2.2	2.6	2.2



V3	1.7	1.9	2.2	2.0	2.1
V4	2.1	1.8	2.5	2.2	2.5

Perform an analysis of variance and test whether there is any significant difference between yields of different varieties and fertilizers.

34. In an experiment on the effects of temperature conditions in human performance, 8 persons were given a test on 4 temperature conditions. The scores in the test are shown in the following table.

Persons

Temperature	1	2	3	4	5	6	7	8
1	70	80	70	90	80	100	90	80
2	70	80	80	90	80	100	90	80
3	75	85	80	95	75	85	95	75
4	65	75	70	85	80	90	80	75

Perform the analysis of variance and state whether there is any significant difference between persons and temperature conditions.

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

Sales Man

Machine	A	B	C	D
May	50	40	48	39
June	46	48	50	45
July	39	44	40	39

Carry out the analysis.

## Answers:

### I.

1. b            2. c            3. a            4. d            5. a  
6. c            7. b            8. c            9. a            10. c

### II.

11. R.A. Fisher      12. Independent      13. Three      14. 25      15. 21

### III.

26. Calculated F = 4.56, Table value of F (4,20) = 2.87

27. Calculated F = 9.11, Table value of F (9,2) = 19.3
28. Calculated F = 1.76, Table value of F (12,3) = 8.74
29. Calculated F = 3.29, Table value of F (3,12) = 3.49
30. Calculated  $F_R$  = 5.03, Table value of F (3,15) = 3.29

- Calculated  $F_C = 2.23$ , Table value of  $F(5,15) = 2.90$
31. Calculated  $F_R = 2.76$ ,  $F_C$  Table value of  $F(9,3) = 8.81$
32. Calculated  $F_R = 18.23$ , Table value of  $F(3,6) = 4.77$
- Calculated  $F_C = 6.4$ , Table value of  $F(2,6) = 5.15$
33. Calculated  $F_R = 1.59$ , Table value of  $F(3,12) = 3.49$
- Calculated  $F_C = 3.53$ , Table value of  $F(4,12) = 3.25$
34. Calculated  $F_R = 3.56$ , Table value of  $F(3,21) = 3.07$
- Calculated  $F_C = 14.79$ , Table value of  $F(7,21) = 2.49$
35. Calculated  $F_R = 3.33$ , Table value of  $F(2,6) = 5.15$
- Calculated  $F_C = 1.02$ , Table value of  $F(3,6) = 4.77$