

# DESIGN OF EXPERIMENTS

M.Sc., Statistics, First year, Paper- IV

## Lesson Writer

**Prof. K.L.A.P. Sarma M.Sc.,Ph.D.**

Department of Statistics,  
Sri Krishnadevaraya University,  
Anantapuram – 515 003

## *Editor*

**Prof. R.R.L. Kantam, M.Sc.,M.Phil.,Ph.D.**

Department of Statistics,  
Acharya Nagarjuna University,  
Nagarjuna Nagar- 522 510



## Director

**Prof. D. Krishna Rao, M.Sc., Ph.D.**

**Centre for Distance Education  
Acharya Nagarjuna University  
Nagarjuna Nagar - 522 510**

Ph: 0863-2293299, 2293356, 08645-211023, Cell:98482 85518

0863 - 2346259 ( Study Material )

Website : [www.anucde.ac.in](http://www.anucde.ac.in) or [www.anucde.info](http://www.anucde.info)

e-mail : [info@anucde.ac.in](mailto:info@anucde.ac.in)

**M.Sc. (STATISTICS), FIRST YEAR**  
**Paper – 4: DESIGN OF EXPERIMENTS**  
**SYLLABUS**

**UNIT – I**

Matrix Algebra : Fundamental definitions, determinants, rank of a matrix, inverse of a matrix, orthogonal matrix, idempotent matrix, characteristic roots and vectors of a matrix, Cauley – Hamilton theorem, trace of a matrix, differentiation using matrices, statement of Coebran's theorem for quadratic forms.

**UNIT – II**

Theory of Linear estimation, linear models, estimability of linear parametric function, best linear unbiased estimates, Gauss- Markov set-up, Gauss-Markov theorem, Generalised linear models, generalized Gauss-Markov theorem (Atken's theorem)

**UNIT – III**

Analysis of Variance of one way, two way and higher way classification with equal and unequal number of observations, fixed, random and mixed effect models. Analysis of Co-variance of one way and two way classification.

**UNIT –IV**

Need for design of experiments, principles of experimental designs, completely randomized design, (CRD). Randomised block design (RBD), Latin square design (LSD), Gracco Latin square design (GLSD), mutually orthogonal Latin squares, missing plot technique general theory and applications.

**UNIT – V**

Factorial experiments –  $2^2$ ,  $2^3$ ,  $3^2$ ,  $3^3$ , Balanced incomplete block designs and their analysis.

Note: Two Questions are to be set from each unit

**Books for Study :**

1. Theory of Linear Estimation : D.D – Joshi : Woley Easten Ltd, New Delhi.
2. Design and Analysis of Experiments, M.N. Das & N.C. Giri : Woley Easten Ltd, New Delhi.
3. Experimental Designs, Federer
4. Fundamentals of Applied Statistics, Gupta & Kapoor : Sultan Chand & Sons New Delhi.

(DMSTT 04)

M.Sc. First Year, Statistics, DEGREE EXAMINATION, MAY 2011.

Paper IV — DESIGN OF EXPERIMENTS

Time : Three hours

Maximum : 100 marks

Answer any FIVE questions.

All questions carry equal marks.

1. (a) Define the terms Rank, inverse, determinant of a matrix.  
(b) Explain the concepts of orthogonal matrix and characteristics roots and vectors of a matrix with suitable examples.
2. (a) State and Prove the Cayley-Hamilton Theorem.  
(b) State and Prove coeheran's theorem for quadratic form.
3. (a) Explain general linear model and give its assumptions.  
(b) State and Prove Gauss – Markov Theorem.
4. (a) Explain the concepts of
  - (i) Estimability of linear parametric functions
  - (ii) BLUE.  
(b) State and Prove Aitken's Theorem.
5. (a) Describe the analysis of variance for a two-way classified data with multiple but equal number of observations per cell.  
(b) Explain and random effect model.
6. (a) Explain the necessity of Covariance technique work out Analysis of covariance one way classification.  
(b) Explain the analysis of covariance with two-way classification.
7. (a) Explain the basic principles of Design of Experiments.  
(b) Explain the analysis of RBD with one missing plot Technique.
8. (a) Explain the analysis of LSD.  
(b) What are Mutually orthogonal Latin squares?
9. (a) Describe the Intra block analysis of BIBD.  
(b) Explain the analysis of  $2^3$  factorial Experiments.
10. Write short notes on any TWO of the following :
  - (a) Transpose of a matrix
  - (b) Gauss-Markov setup
  - (c) Multiple comparisons
  - (d) CRD
  - (e)  $3^2$  Factorial Experiments.

(DMSTT 04)

M.Sc. DEGREE EXAMINATION, DECEMBER 2011.

First Year

Statistics

Paper IV — DESIGN OF EXPERIMENTS

Time : Three hours

Maximum : 100 marks

Answer any FIVE questions.

All questions carry equal marks.

1. (a) Explain the role of Quadratic forms in design of experiments.  
(b) State and prove Cochran's theorem.
2. (a) Explain the methodology of deriving Eigen values and Eigen vector associated with a matrix. What is the importance of it?  
(b) Define trace of a matrix. Explain its relationship in finding the derivative of a matrix with respect to its elements.
3. (a) Explain the limitations of ordinary least square estimators in linear models.  
(b) Stating the assumptions clearly and bring out Gauss–Markov set up of a linear model.
4. (a) State and prove Aitken's theorem.  
(b) Explain the importance of  $G$ -inverse in estimating the parameters of a linear model.
5. (a) Explain the procedure of analysing a two way classified data.  
(b) Discuss the uses of analysis of variance with two way classification.
6. (a) Derive the  $F$ -statistics for testing the equality of  $K$  means in one way classified data.  
(b) With the usual notation show that the degrees of freedom associated error in two-way classified ANOVA Test is  $(n - 1)(K - 1)$ .
7. (a) With the usual notations derive the least square estimates and expectations of sum of squares of a Randomised Block design.  
(b) Describe the uses of R.B.D.
8. (a) Explain the layout of a L.S.D. with an example.  
(b) Compare and contrast the Latin Square design with R.B.D.
9. (a) What is a treatment contrast? When are two such contrasts orthogonal? Show that in  $2^3$  experiment main effects and interaction effects are orthogonal.  
(b) Explain the philosophy of Factorial experiments.
10. (a) State and prove Fisher's inequality in  $2^k$  factorial design.  
(b) When is  $SSB/BD$  called symmetric?

**Lesson – 1**

# Fundamentals of Matrix Algebra

**1.0 Objectives :**

After going through this lesson, you should be able to:

- Understand the need, concepts of Matrix Algebra, definitions and applications.
- The concepts of determinant, rank and inverse of matrix.
- The applications and examples of matrices.
- Calculation of Inverses of matrices and their properties.
- Applications of matrix algebra.

**Structure**

- 1.1 Introduction to Matrix Algebra.**
- 1.2 Fundamental concepts and definitions of Matrix Algebra.**
- 1.3 Determinants, Inverse and Rank of a matrix and their properties.**
- 1.4 Applications and Worked out examples.**
- 1.5 Summary**
- 1.6 Self Assessment Questions**
- 1.7 Further Readings.**

**1.1. Introduction to Matrix Algebra**

In real life there are good number of situations, where we have to consider many things, to take decision on a single concept. For example, health of a person is depending on many quantities like body temperature, B.P.,(systolic and diastolic), blood sugar, urine sugar, pulse rate, R.B.C, W.B.C, cholesterol, triglycerides, proteins, L.D.L, V.L.D.L and so on. All these factors determine the health condition of the person. Similarly, in agriculture the yield of a crop is mainly depends on the rainfall, soil temperature, seed quality, type of soil, type of water, type of cultivation, climatic conditions and so on. Thus there is a need to deal many quantities simultaneously at a time and to analyze these factors. A mathematical tool that helps to analyze such quantities is known as “**Matrix Algebra**” or “**Matrix Analysis**”. Matrices are useful in almost every branch of Science, Engineering, Medicine and Design of Experiments.

Thus there is a need to study some definitions, fundamental concepts, theorems, properties and applications of matrix algebra in the present unit, so that these results can be

applied in forth coming units. In this lesson, we proceed to learn some definitions, fundamental concepts of matrix algebra.

## 1.2 Fundamental concepts and definitions of Matrix Algebra

A matrix is defined as an arrangement of  $m \cdot n$  numbers (real or complex) arranged in the rectangular form in **m-rows** and **n-columns** enclosed by pair of brackets is known as a “**Matrix**” and is denoted by capital letters A,B,C, M or X. These numbers are called elements of the matrix and are denoted by lower case letters like a,b,c,m or x. For example, a matrix A with m-rows and n-columns is written as:

$$A_{mn} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & a_{i3} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}$$

In the above matrix,  $a_{ij}$  where,  $i$  ranges from  $1,2,\dots,m$  and  $j$  ranges from  $1,2,\dots,n$  represent  $m \cdot n$  elements of the matrix .

A compact form of the above matrix is represented as follows:

$$A_{mn} = \|\|a_{ij}\| \text{ or } [a_{ij}] \text{ or } (a_{ij}) \text{ or } \llbracket a_{ij} \rrbracket \quad i = 1,2,3,\dots,m \text{ and } j = 1,2,3,\dots,n.$$

In the above notation, if  $m = 1$ , then the matrix is called “**Row Matrix**” and if  $n = 1$ , then the matrix is called “**Column Matrix**”. If  $m=n$ , then the matrix is called “**Square Matrix**” and if  $m \neq n$ , then the matrix is called “**Rectangular Matrix**”. If all the elements are zeros then the matrix is called “**Null Matrix**” or “**Zero Matrix**”.

All the elements of the matrix where  $i = j$  are called “**Diagonal**” or “**Principal Diagonal**” or “**Main Diagonal**” elements. All the elements where  $i \neq j$  are called “**Non-diagonal**” elements. If the matrix is a square matrix, with one as diagonal elements and zeros as non-diagonal elements, then the matrix is known as “**Unit Matrix**” or “**Identity Matrix**”. A square matrix in which all non-diagonal elements are zeros is known as “**Diagonal Matrix**”. In a diagonal matrix if all the diagonal elements are equal is called “**Scalar Matrix**”. In a square matrix, if all the elements below the principal diagonal are zero is known as “**Upper Triangular**” matrix. Similarly, in a square matrix, if all the elements above the principal diagonal are zeros then the matrix is known as “**Lower Triangular**” matrix. In a given matrix, if one or more rows or columns or both are deleted, then the resultant matrix is called as “**Sub-matrix**”.

### 1.2.1. Operations on Matrices

Following are the basic operations of matrix algebra.

**1.Equality of matrices:** Two matrices are said to be equal if they are of the same order and the corresponding elements must be equal. That is if  $A = B$  and  $A = \begin{bmatrix} a & b \\ 4 & c \end{bmatrix}$  and  $B = \begin{bmatrix} 6 & 3 \\ d & -2 \end{bmatrix}$ , then  $a = 6$ ,  $b = 3$ ,  $c = -2$  and  $d = 4$ .

**2. Addition or subtraction of matrices:** The sum or difference of two matrices is defined only for the matrices of the same order. To add two matrices, the corresponding elements are to be added. Similarly, to subtract two matrices, the corresponding elements are to be subtracted. For example, to add two matrices A and B; Where  $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$  and  $B = \begin{bmatrix} 6 & 3 \\ 5 & -2 \end{bmatrix}$ ,

$$A + B = \begin{bmatrix} 1+6 & 2+3 \\ 4+5 & 3+(-2) \end{bmatrix} = \begin{bmatrix} 7 & 5 \\ 9 & 1 \end{bmatrix}. \text{ Similarly, } A - B = \begin{bmatrix} 1-6 & 2-3 \\ 4-5 & 3-(-2) \end{bmatrix} = \begin{bmatrix} -5 & -1 \\ -1 & 5 \end{bmatrix}.$$

**Properties of Matrix Addition:** If A, B, C and O are matrices of the same order, then:

- $A + B = B + A$  that is Matrices of same order are commutative.
- $A + (B + C) = (A + B) + C$  that is matrices of same order are associative.
- $K(A + B) = KA + KB$ , where K is any constant.
- If O is the null matrix then  $O + A = A + O = A$  hence O is called additive identity.
- If  $-A$  is the negative of the matrix A then  $(-A) + A = A + (-A) = O$ . Hence  $-A$  is called additive inverse of A.

**3. Multiplication or Product of two matrices:** The product of two matrices A and B is defined only if the columns of matrix A must be equal to the rows of matrix B. That is if the matrix is of order  $m \times n$ , then the matrix B must be of order  $n \times k$ . After multiplication the product matrix A.B will be of order  $m \times k$ . To get  $(i, j)^{\text{th}}$  element in the product matrix A.B, elements of  $i^{\text{th}}$  row of matrix A are to be multiplied with the elements of  $j^{\text{th}}$  column of matrix B and

their product sum is taken. That is if  $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 3 & 4 \end{bmatrix}$  and  $B = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 1 & 1 \end{bmatrix}$ , then the product A.B

is of order  $2 \times 2$  and is to be calculated as follows:

$$A.B = \begin{bmatrix} 1.2 + 0.3 + 2.1 & 1.1 + 0.0 + 2.1 \\ 2.2 + 3.3 + 4.1 & 2.1 + 3.0 + 4.1 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 17 & 6 \end{bmatrix}.$$

Mathematically, matrix multiplication is defined as follows:

**Definition:** If  $A = [a_{ik}]$  where  $i = 1, 2, \dots, m$ ,  $k = 1, 2, \dots, n$  and  $B = [b_{kj}]$  where  $k = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, p$  then the product matrix A.B is of order  $m \times p$  and is defined as:

$$A.B = [a_{ij}] = \sum_{k=1}^n (\sum a_{jk} b_{kj}) \text{ where } i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, p \quad (1.2.1)$$

**Properties of Matrix Multiplication:**

- Matrix multiplication is not commutative in general, that is  $A.B \neq B.A$ .
- If the matrix multiplication is defined between A, B and C then matrix multiplication is associative. That is  $(AB)C = A(BC)$ .
- Multiplication is distributive over addition in matrices. That is (i)  $A(B+C) = AB + AC$

- (ii)  $(A + B)C = AC + BC$ , where all the above products are defined.
- (d) If  $A, B, C$  are three matrices such that  $AB = AC$  then in general  $B \neq C$ .
- (e) If  $A$  is an  $m \times n$  matrix and  $O$  is  $n \times p$  null matrix then  $A.O = O.A = O$ .
- (f) If  $A$  is a square matrix and  $I$  is a unit matrix of the same order, then  $AI = IA = A$ .
- (g) Product of two non-zero matrices may be a zero matrix.

### 1.2.2. Types of Matrices

**(1) Transpose of a Matrix:** Let  $A$  is a given matrix of size  $m \times n$ . If we inter change rows and columns of the given matrix, we obtain an  $n \times m$  matrix and the resultant matrix is known as transpose of the given matrix  $A$  and is denoted by  $A^T$  or  $A^l$ . Symbolically, if  $A = [a_{ij}]$

$$\text{and } A^l = [b_{ij}] \text{ then } a_{ij} = b_{ji}. \text{ For Example, } A = \begin{bmatrix} 2 & 4 & 9 \\ -5 & 7 & -2 \end{bmatrix} \text{ then } A^l = \begin{bmatrix} 2 & -5 \\ 4 & 7 \\ 9 & -2 \end{bmatrix}$$

**Properties of Transpose of a matrix:**

- (a)** A matrix is transpose of its transpose. That is  $A = (A^l)^l$ .
- (b)** Transpose of sum of two matrices is the sum of their transpose. That is  $(A+B)^l = A^l + B^l$ .
- (c)** Transpose of multiplication of a matrix with a constant  $K$  is equal to the multiplication of the constant number by the transpose of that matrix. That is  $(KA)^l = K(A^l)$ .
- (d)** The transpose of the product of two matrices is the product of their transposes in reverse order. That is  $(AB)^l = B^l . A^l$ .

**(2) Symmetric Matrix:** Any matrix  $A$  is symmetric if  $A = A^l$ . This implies that a symmetric matrix is always a square matrix. For example  $A = \begin{bmatrix} 1 & 3 & 7 \\ 3 & 4 & -2 \\ 7 & -2 & 6 \end{bmatrix} = A^l$ . Hence  $A$  is called symmetric.

**(3) Skew-symmetric Matrix:** Any matrix  $A$  is called skew-symmetric if  $A^l = -A$ . That is if  $A = [a_{ij}]$  then  $a_{ij} = -a_{ji}$ . Further it implies that in a skew-symmetric matrix all the diagonal elements must be equal to zero and must be a square matrix.

**(4) Orthogonal Matrix:** Any matrix  $A$  is said to be orthogonal matrix if  $A . A^l = I$ . For example, if  $A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$  then  $A . A^l = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$ . Hence  $A$  is an orthogonal matrix.

Similarly, is  $A = \frac{1}{3} \begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}$  Then  $A . A^l = I$ . Hence  $A$  is orthogonal matrix.

**Remark:**

Orthogonal matrices have special applications in “**Design of Experiments**”, will be discussed at appropriate place in the last Unit while discussing properties of Balanced Incomplete Block Designs (BIBD).

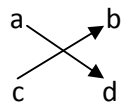


### 1.3 Determinants, Inverse and Rank of a matrix and their properties.

Now we proceed to explain other important concepts of matrix analysis.

**1.3.1 Determinants:** Determinant of a matrix is made by the same array of numbers of which the matrix is made. However, unlike a matrix, the determinant is not just arrangement of numbers but also a numerical value calculated in a particular procedure. Further, it is important to note that only square matrices have determinant value. The determinants of non-square matrices are not defined.

A determinant having one row and one column is called a determinant of order 1. A determinant having two rows and two columns is called determinant or order 2 and so on. In general, a determinant having  $n$  rows and  $n$  columns is called a determinant of order  $n$ . The determinant of a matrix  $A$  is denoted by  $|A|$  or  $\det. A$  or  $\Delta$ . If  $A$  has only one row and one column, that is if  $A = [K]$  then  $|A| = K$  or  $\det. A = K$ . For a  $2 \times 2$  matrix  $A$ , where  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ ,  $|A|$  or  $\det. A$  is defined as  $a d - c b$ . The multiplication rule is illustrated by arrows as follows:



The determinant of order 3 is calculated as follows: Let  $A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$  then,  $|A|$

is defined as:  $a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$  that is given as follows:

$$a_1 (b_2 c_3 - c_2 b_3) - a_2 (b_1 c_3 - c_1 b_3) + a_3 (b_1 c_2 - c_1 b_2). \quad (1.3.1)$$

For example,  $|A|$  where  $A = \begin{bmatrix} 6 & 5 \\ 4 & 8 \end{bmatrix} = 6 \cdot 8 - 4 \cdot 5 = 48 - 20 = 28$ .

Similarly is  $A = \begin{bmatrix} 1 & 3 & 7 \\ 3 & 4 & -2 \\ 7 & -2 & 6 \end{bmatrix}$  Then  $|A| = 1[24 - (-4)] - 3[18 - (-14)] + 7[(-6) - 28]$   
 $= 28 - 96 - 238 = -306$ .

**Definition of Singular and Non-singular Matrices:** Any square matrix is known as “Singular Matrix” if the  $|A| = 0$  and is known as “Non-singular Matrix” if the  $|A| \neq 0$ .

**Properties of Determinants:**

- The value of the determinant remains un-changed if its rows and columns are interchanged. That is  $|A| = |A^t|$ . For example,  $|A^t| = \begin{vmatrix} 6 & 4 \\ 5 & 8 \end{vmatrix} = 48 - 20 = 28$ .
- The value of the determinant changes its sign if we interchange any two rows (or any two columns). That is  $\begin{vmatrix} 4 & 8 \\ 6 & 5 \end{vmatrix} = 20 - 48 = -28$ . Similarly  $\begin{vmatrix} 4 & 6 \\ 8 & 5 \end{vmatrix} = 20 - 48 = -28$ .
- The value of the determinant is zero, if any two rows (or any two columns) are identical.

- (d) The value of the determinant becomes C times (where C is a constant) if any one row (or column) is multiplied by the constant C.
- (e) The value of the determinant is zero if any two of its rows (or columns) are proportional.
- (f) The value of the determinant remains un-changed if any to row (or column) a constant multiple of row (column) is added or subtracted.
- (g) If each element of any row (or column) of a determinant is sum of two numbers, the determinant can be expressed as the sum of the two determinants.

### 1.3.2 Minors and Co-factors of the elements of a determinant:

The “**Minor**” of the  $(i,j)^{\text{th}}$  element of a determinant is defined the determinant  $M_{ij}$  obtained by removing  $i^{\text{th}}$  row and  $j^{\text{th}}$  column in which the element is existing. Similarly, the “**Co-factor**” of the  $(i,j)^{\text{th}}$  element of the determinant is defined as  $(-1)^{i+j}M_{ij}$  and is denoted by  $C_{ij}$ .

For example consider  $|A| = \begin{vmatrix} 1 & 3 & 7 \\ 3 & 4 & -2 \\ 7 & -2 & 6 \end{vmatrix}$  then  $M_{11} = \begin{vmatrix} 4 & -2 \\ -2 & 6 \end{vmatrix}$  and  $C_{11} = (-1)^2 M_{11}$ .

Similarly,  $M_{23} = \begin{vmatrix} 1 & 3 \\ 7 & -2 \end{vmatrix}$  and  $C_{23} = (-1)^5 M_{23}$  and so on.

The value of the determinant  $|A|$ , given in the equation (1.3.1) can be expressed in terms of co-factors as:

$$|A| = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \quad (1.3.2)$$

**1.3.3 Ad-joint of a Matrix:** The ad-joint of the matrix A is defined as the transpose of the matrix formed by co-factors of the corresponding elements of the determinant A. Ad-joint of a matrix A is denoted by Adj. A.

**1.3.4 Inverse of a square matrix:** If two matrices A and B are such that  $AB = BA = I$ , where I is a “Unit Matrix” or “Identity Matrix” then the matrix B is defined as the inverse of the matrix A and is denoted by  $A^{-1}$ . That is  $B = A^{-1}$ . The “Inverse” of a square matrix is also defined as  $A^{-1} = \frac{1}{|A|} (\text{Adj. A})$  or  $(\text{Adj. A}) / |A|$ . (1.3.3)

#### Properties:

- (a) If A is an invertible matrix, then A has unique inverse.
- (b) If A and B are invertible square matrices of the same order, then  $(AB)^{-1} = B^{-1}A^{-1}$ .
- (c) If A is an invertible matrix, then  $(A^T)^{-1} = (A^{-1})^T$ .
- (d) If A is a non-singular matrix, then A is invertible matrix. That is if  $|A| \neq 0$  then  $A^{-1}$  exists.

**1.3.5 Rank of a Matrix:** A number ‘r’ is said to be the rank of a matrix A if it posses the following two properties:

- (i) There is at least at least one square sub-matrix of A of order ‘r’ whose determinant is not equal to zero.
- (ii) If the matrix contains any square matrix or order ‘(r+1)’ then the determinant of every such square matrix of A of order (r+1) should be equal to zero.

**Definition of rank of a matrix:** The rank of a matrix is the order of any highest order non-vanishing minor of the matrix. We shall denote the rank of the matrix  $A$  by the symbol  $\rho(A)$ .

Properties of Rank of a matrix:

- Since the rank of every non zero matrix  $r \geq 1$ , we assign the rank of zero or null matrix as zero. That is the rank of null matrix = 0.
- The rank of a matrix is  $\leq r$ , if all  $(r+1)$ -rowed minors of the matrix vanish.
- The rank of a matrix is  $\geq r$ , if there is at least one  $r$ -rowed minor of the matrix is not equal to zero.
- The rank of the transpose of a matrix is the same as that of the original matrix. That is  $\rho(A) = \rho(A')$ .
- An  $n$ -rowed square matrix of rank  $(n-1)$ , then  $\text{Adj. } A \neq 0$ .
- The rank of a matrix does not alter on affixing any number of rows or columns of zeros.

**1.3.6. Elementary Operations or Transformations of a matrix:** An “Elementary transformation or (E-transformation)” is an operation of any one of the following types:

(1) The interchange of any two rows (or Columns). Symbolically this transformation is denoted by  $R_i \leftrightarrow R_j$  (interchange if  $i^{\text{th}}$  row and  $j^{\text{th}}$  row. Similarly, inter change of  $i^{\text{th}}$  column and  $j^{\text{th}}$  column is denoted by  $C_i \leftrightarrow C_j$ .

(2) The multiplication of the elements of any row (or column) by any non-zero number. Symbolically this transformation is denoted by  $R_i \rightarrow KR_i$  (or  $C_i \rightarrow KC_i$ ), that is the multiplication of the elements of  $i^{\text{th}}$  row (or column) by non-zero number  $K$ .

(3) The addition to the elements of any other row (or column) the corresponding elements of any other row (or column) multiplied by any number. Symbolically,  $R_i \rightarrow R_i + KR_j$  ( or  $C_i \rightarrow C_i + KC_j$ ) implies that addition of  $K$  times  $j^{\text{th}}$  row (or column) to the  $i^{\text{th}}$  row (or column).

**Definition:** A matrix obtained from a unit matrix by a single elementary transformation is called “Elementary matrix” or “E-matrix”

**Result (1):** Every elementary row (or column) transformation of a matrix can be obtained by pre-multiplication (or post-multiplication) with corresponding elementary matrix.

**Result (2):** Elementary transformations do not alter the rank of the matrix.

**Result (3):** The rank of the matrix remains un-altered by a finite chain of elementary operations.

Basically there are two forms namely (1) Echelon form and (2) Normal form or Canonical form of a matrix. They are defined as follows:

Definition (1): A matrix  $A$  is said to be in Echelon form if:

- Every row of  $A$  which has all its entries 0 occurs below every row which has a non zero entry.
- First non-zero entry in each non-zero row is equal to 1.
- The number of zeros before the first non-zero element in a row is less than the number of such zeros in the next row.

For example the matrix  $A = \begin{bmatrix} 0 & 1 & 4 & 5 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$  is in Echelon form.

**Definition(2):** Every  $m \times n$  matrix with rank  $r$  can be reduced to the form  $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ , which is known as “Normal or Canonical” form by a finite chain of E-operations, where  $I_r$  is the  $r$ -rowed unit matrix.

**Result (1):** The rank of  $m \times n$  matrix  $A$  is  $r$  if and only if (iff) it can be reduced to the normal form  $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ , by finite chain of E-operations.

**Result (2):** If  $A$  be an  $m \times n$  matrix with rank  $r$ , then there exist non-singular matrices  $P$  and  $Q$  such that  $PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ .

**Definition: Equivalence of Matrices:** If  $B$  be an  $m \times n$  matrix obtained from an  $m \times n$  matrix  $A$  by finite number of elementary transformations of  $A$ , then  $A$  is called equivalent to  $B$  and is denoted by  $A \sim B$ .

**Properties:**

(a) **Reflexivity:** Equivalence property is reflexive. That is  $A \sim A$ .

(b) **Symmetry:** Equivalence is symmetric. That is if  $A \sim B$ , then it implies that  $B \sim A$ .

(c) **Transitivity:** Equivalence is transitive. That is if  $A \sim B, B \sim C$ , then,  $A \sim C$ .

#### 1.4. Applications and Worked out examples.

Matrix theory has lot of applications in many fields where there is a need to analyze many variables at a time. For example in solving system of linear equations, Linear Programming Problems, Transportation Problems, Game theory, Design of experiments, Multivariate analysis, Discriminate Analysis, Factor Analysis, Business Forecasting, Marketing, Medicine, meteorology, weather forecasting and so on. To demonstrate, first we discuss matrix Analysis in solving system of linear equations in the following section.

##### 1.4.1 Solving a System of Linear Equations.

A system of equations in two variables (or unknowns) is of the type:

$$\begin{aligned} a_1 x + b_1 y &= c_1 \\ a_2 x + b_2 y &= c_2 \end{aligned} \quad (1.4.1)$$

Similarly, a system of three unknowns (or variables) is of the form:

$$\begin{aligned} a_1 x + b_1 y + c_1 z &= d_1 \\ a_2 x + b_2 y + c_2 z &= d_2 \\ a_3 x + b_3 y + c_3 z &= d_3 \end{aligned} \quad (1.4.2)$$

**Homogeneous System:** A system of linear equations is said to be homogeneous if the constant term in each equation is equal to zero. That is  $c_1 = c_2 = 0$  in equation (1.4.1) and  $d_1 = d_2 = d_3 = 0$  in equation (1.4.2).

**Consistent and In-consistent System of equations:** A system of linear equations is said to be “consistent” if it poses at least one solution. Otherwise, the system is known as “In-consistent” system.

A consistent system of linear equations can be solved by using matrix theory by the following two methods. Namely:

- (1) The method Matrix Inversion and
- (2) The Method of Reduction using elementary row transformations.

These methods with illustrated examples are explained in the following section.

**(1) The method Matrix Inversion:** Let us consider a system of 2-linear equations in 2-variables (unknowns) as follows:

$$\begin{aligned} a_1 x + b_1 y &= c_1 \\ a_2 x + b_2 y &= c_2 \end{aligned} \quad (1.4.3)$$

In the matrix notation these equations can be written as:

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad \text{or} \quad \mathbf{AX} = \mathbf{B} \quad (1.4.4)$$

Where  $\mathbf{A} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$ ,  $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}$  and  $\mathbf{B} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ .

Multiply both sides of equation (1.4.4) by  $\mathbf{A}^{-1}$ , then we have:

$$\mathbf{A}^{-1}(\mathbf{A})\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}. \quad (1.4.5)$$

By the property of  $\mathbf{A}^{-1}$  we know that  $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ . Hence we have,

$$\mathbf{IX} = \mathbf{A}^{-1}\mathbf{B}. \quad (1.4.6)$$

Since,  $\mathbf{IX} = \mathbf{X}$  itself, we have:

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}. \quad (1.4.7)$$

Thus by finding  $\mathbf{A}^{-1}$  and multiplying this inverse with  $\mathbf{B}$ , the values of  $\mathbf{X}$  can be obtained, which is solution to the given system of linear equations. This method of analysis can be extended to K- number of linear equations on similar lines.

**Example (1.4.1):** Solve the system of linear equations using matrix inversion method.

$$\begin{aligned} 2x - y + 2z &= 6 \\ x - 2y + 3z &= 6 \\ 3x - 3y - z &= -6. \end{aligned}$$

**Solution:** Given system of linear equations can be expressed in matrix notation as follows:

$$\begin{bmatrix} 2 & -1 & 2 \\ 1 & -2 & 3 \\ 3 & -3 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ -6 \end{bmatrix} \quad \text{or} \quad \mathbf{AX} = \mathbf{B}. \quad (1.4.8)$$

Where,  $\mathbf{A} = \begin{bmatrix} 2 & -1 & 2 \\ 1 & -2 & 3 \\ 3 & -3 & -1 \end{bmatrix}$ ,  $\mathbf{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$  and  $\mathbf{B} = \begin{bmatrix} 6 \\ 6 \\ -6 \end{bmatrix}$ . The solution of the given system of linear equations is:

$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$  or  $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 & -1 & 2 \\ 1 & -2 & 3 \\ 3 & -3 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 6 \\ -6 \end{bmatrix}$ . Hence we have to find  $\mathbf{A}^{-1}$  first to get the

solution of the equations. To find the inverse, we have to find  $|\mathbf{A}|$  and  $\mathbf{Adj.}(\mathbf{A})$  as follows:

$$|A| = \begin{vmatrix} 2 & -1 & 2 \\ 1 & -2 & 3 \\ 3 & -3 & -1 \end{vmatrix} = 2(2+9) + 1(-1-9) + 2(-3+6) = 22-10+6 = 18.$$

$$\text{Similarly, Adj. (A)} = \begin{bmatrix} 11 & -7 & 1 \\ 10 & -8 & -4 \\ 3 & 3 & -3 \end{bmatrix}. \text{ Hence, } A^{-1} = \text{Adj. (A)} \setminus |A| = (1/18) \begin{bmatrix} 11 & -7 & 1 \\ 10 & -8 & -4 \\ 3 & 3 & -3 \end{bmatrix}.$$

$$\text{Therefore, } \begin{bmatrix} x \\ y \\ z \end{bmatrix} = (1/18) \begin{bmatrix} 11 & -7 & 1 \\ 10 & -8 & -4 \\ 3 & 3 & -3 \end{bmatrix} \begin{bmatrix} 6 \\ 6 \\ -6 \end{bmatrix} = (1/18) \begin{bmatrix} 18 \\ 36 \\ 54 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}. \text{ Hence, the solution to the given system of equations is: } x = 1, y = 2 \text{ and } z = 3.$$

**Important Remark:** This method of solution can be applied only when the coefficient matrix in (1.4.4) is a non-singular matrix. That is when  $|A|$  is not equal to zero.

## (2) The Method of Reduction using elementary row transformations.

In order to find the solution of a given system of linear equations, using elementary row transformations, we use any one of the following elementary row transformations. Namely:

- (1) interchange of two rows.
- (2) multiplication of a row by a non-zero number and
- (3) addition of a product of the elements of any row by any constant to the corresponding elements of the other row.

In this method, we use another concept known as “Augmented Matrix” which is defined as follows:

**Definition:** a matrix is said to be augmented matrix if it is enlarged by introducing one or more new columns.

Let us consider the following system of Linear equations.

$$\begin{aligned} a_1 x + b_1 y + c_1 z &= d_1 \\ a_2 x + b_2 y + c_2 z &= d_2 \\ a_3 x + b_3 y + c_3 z &= d_3 \end{aligned} \quad (1.4.9)$$

First of all we express (1.4.9) as the augmented matrix **B** as follows:

$$B = \left[ \begin{array}{ccc|c} a1 & b1 & c1 & d1 \\ a2 & b2 & c2 & d2 \\ a3 & b3 & c3 & d3 \end{array} \right] \quad (1.4.10)$$

By elementary row transformations, this augmented matrix B will be reduced to the following form:

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & \alpha \\ 0 & 1 & 0 & \beta \\ 0 & 0 & 1 & \gamma \end{array} \right]. \text{ Then, the solution to the given system of linear equations is:}$$

$x = \alpha, y = \beta$  and  $z = \gamma$ . Now we proceed to explain this method of reduction using elementary row transformations, with an example.

**Example (1.4.2):** Solve the following system of linear equations:

$$4x + 2y = 3$$

$$3x - 4y = 5$$

Solution: First we have to form the augmented matrix as follows:

$\left[ \begin{array}{cc|c} 4 & 2 & 3 \\ 3 & -4 & 5 \end{array} \right]$ . To make the first element in the first row 4 as 1 we subtract elements of 2<sup>nd</sup> row from the corresponding elements of 1<sup>st</sup> row. This operation is denoted by  $R_1 - R_2$ . Then we have :

$\left[ \begin{array}{cc|c} 1 & 6 & -2 \\ 3 & -4 & 5 \end{array} \right]$ . To make first element in the 2<sup>nd</sup> row as zero, we multiply the elements of 1<sup>st</sup> row by 3 and subtract the corresponding elements from the elements of 2<sup>nd</sup> row. This operation is denoted by  $R_2 - 3R_1$ . Thus we have:

$\left[ \begin{array}{cc|c} 1 & 6 & -2 \\ 0 & -22 & 11 \end{array} \right]$ . To make second element in 2<sup>nd</sup> row, that is -22 as 1, we multiply the entire elements of the row by  $-(1/22)$ . This operation is denoted by  $-(1/22)R_2$ . Then we have:

$\left[ \begin{array}{cc|c} 1 & 6 & -2 \\ 0 & 1 & -1/2 \end{array} \right]$ . Finally, we have to make second element in the 1<sup>st</sup> row, that is 6 as zero. For this, we multiply the elements of 2<sup>nd</sup> row by -6 and add these elements to the corresponding elements of 1<sup>st</sup> row. This operation is denoted by  $R_1 + (-6)R_2$  or  $R_1 - 6R_2$ .

Then we have:  $\left[ \begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -1/2 \end{array} \right]$ .

Hence the solution for the given system of linear equations is  $\mathbf{x} = \mathbf{1}$  and  $\mathbf{y} = -\mathbf{1/2}$ .

On similar lines, matrix analysis has lot of other applications in Operations Research, Linear Programming problems, Transportation Problems, Game theory, Design of experiments and so on.

## 1.5. Summary

In this lesson the concept of matrix is introduced along with different type of matrices and their properties. Various operations of matrices, namely addition, subtraction and multiplication of matrices along with their properties are discussed. Determinants, Rank and inverse of the matrix are introduced and also discussed various properties and their applications. Matrix analysis has lot of applications in various fields like solving system of equations, Linear Programming Problems, Transportation Problems, Game theory and so on. Various Elementary operations applicable on matrix analysis along with various forms of matrices and reduction of "Normal form" or "Canonical form" are discussed along with their properties and applications. Finally, "Equality of matrices" and application of matrix analysis in solving system of linear equations are discussed. Application of matrix analysis in design of experiments will be discussed in Unit – V.

### 1.6. Self assessment questions

1. Define a matrix and discuss various types of matrices along with their properties.
2. Explain various operations of matrix analysis along with their properties.
3. Find the determinants of the following matrices:

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 4 & -1 & 2 \\ 3 & 5 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix} \quad C = \begin{bmatrix} 4 & -6 & 1 \\ 1 & 3 & 6 \\ 5 & 7 & 9 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 2 & 7 \\ 2 & 0 & 2 \\ 3 & -4 & 6 \end{bmatrix}$$

{Answers:  $|A| = 28$ ;  $|B| = -1$   $|C| = -194$   $|D| = -180$  }.

4. Find Ad-joints of the following matrices:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 0 \\ 2 & 4 & 3 \end{bmatrix}; B = \begin{bmatrix} 1 & 2 \\ 3 & -5 \end{bmatrix}; C = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}; D = \begin{bmatrix} 1 & -3 & 2 \\ 4 & 0 & 6 \\ 5 & 7 & 3 \end{bmatrix}$$

$$\{\text{Answers: Adj. (A)} = \begin{bmatrix} 15 & 6 & -15 \\ 0 & -3 & 0 \\ -10 & 0 & 5 \end{bmatrix}; \text{Adj. (B)} = \begin{bmatrix} -5 & -2 \\ -3 & 1 \end{bmatrix};$$

$$\text{Adj. (C)} = \begin{bmatrix} -1 & 1 & -1 \\ 8 & -6 & 2 \\ 5 & 3 & -1 \end{bmatrix}; \text{Adj. (D)} = \begin{bmatrix} -42 & 23 & -18 \\ 18 & -7 & 2 \\ 28 & -22 & 12 \end{bmatrix}\}$$

5. Find the rank of the following matrices:

$$A = \begin{bmatrix} 2 & 4 & 3 & 2 \\ 3 & 5 & 1 & 4 \end{bmatrix}, B = \begin{bmatrix} 3 & 1 & 2 \\ 6 & 2 & 4 \\ 3 & 1 & 2 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 0 & 2 & 2 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

{Solution:  $\rho(A) = 2$ ;  $\rho(B) = 1$ ;  $\rho(C) = 3$ ;  $\rho(D) = 0$ ;  $\rho(I) = 3$  }.

6. Find the inverse of the following matrices:

$$A = \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}; B = \begin{bmatrix} 4 & -6 & 1 \\ 1 & 3 & 6 \\ 5 & 7 & 9 \end{bmatrix}; C = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}; D = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$\{\text{Answers: } A^{-1} = (-1/8) \begin{bmatrix} 7 & -3 \\ -5 & 1 \end{bmatrix} \quad B^{-1} = (-1/194) \begin{bmatrix} -15 & 61 & -39 \\ 21 & 31 & -23 \\ -8 & -58 & 18 \end{bmatrix};$$

$$C^{-1} = (-1/2) \begin{bmatrix} -1 & 1 & -1 \\ 8 & -6 & 2 \\ -5 & 3 & -1 \end{bmatrix}; \text{Since, } |D| = \cos^2 \alpha + \sin^2 \alpha = 1, D^{-1} \text{ exists; where}$$

$$D^{-1} = \begin{bmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

7. Show that  $AA^{-1} = A^{-1}A = I$  by considering all the matrices given in the above question 5.
8. Solve the following system of linear equations by using (i) matrix inversion method and (ii) : Reduction method using elementary row transformations.



$$\begin{array}{lll}
 \text{(a) } 2x - y + 3z = 9 & \text{(b) } 2x - 3y + 3z = 1 & \text{(c) } 2x - 3y + z = 7 \\
 x + y + z = 6 & 2x + 2y + 3z = 2 & 2x + y - z = 1 \\
 x - y + z = 2 & 3x - 2y + 2z = 3 & 4y + 3z = -11
 \end{array}$$

{ Solutions: (a)  $x = 1$ ;  $y = 2$  and  $z = 3$ ; (b)  $x = 7/5$ ,  $y = 1/5$ ,  $z = -2/5$ ; (c)  $x = 1$ ,  $y = -2$ ,  $z = -1$ . }

9. A company manufactures two types of lap-tops which are assembled and finished in two workshops  $W_1$  and  $W_2$ . Each type of lap-top takes 20 hours and 10 hours for assembly and 5 hours and 3 hours for finishing in the respective workshops. If total number of hours available are 450 and 230 in workshops  $W_1$  and  $W_2$  respectively, calculate the number of lap-tops of each type, in both workshops using (i) matrix inversion method and (ii) Reduction method using elementary row transformations.

{Solution: Let  $x$  represents number of lap-tops of the first type and  $y$  represent number of lap-tops of the second type. Then we have  $\begin{bmatrix} 20 & 5 \\ 10 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 450 \\ 230 \end{bmatrix}$ .  $x = 20$ ,  $y = 10$ . }

10. Let  $A = \begin{bmatrix} 2 & -1 & 4 \\ 4 & 0 & 2 \\ 3 & -2 & 7 \end{bmatrix}$ , find  $A^{-1}$  using elementary row operations.

{Solution:  $A^{-1} = \begin{bmatrix} -2 & 1/2 & 1 \\ 11 & -1 & -6 \\ 4 & -1/2 & -2 \end{bmatrix}$ }

### 1.7. Further readings

1. A.R. Vasishtha "Matrices" Krishna Prakashan Media (p) Ltd.; Thirty fourth Edn., Meerut – 250 001; 2003.
2. A. RamachandraRao and P. BhimaShankaram "Linear Algebra" Tata McGraw-Hill Publishing Company Limited; 1992.
3. D.D. Joshi "Theory of Linear Estimation" Woley Eastern Ltd. New Delhi. 2005.
4. Gupta and Kapoor "fundamentals of Applied Statistics" Sultan Chand & sons, New Delhi. 2008.

**Lesson – 2****Vectors of Matrix Algebra****2.0 Objectives :**

After going through this lesson, you should be able to:

- Understand the need, concepts of Vector analysis of Matrix Algebra, definitions and applications.
- The concepts of orthogonal matrix and Idempotent Matrix.
- The characteristic equations and characteristics roots.
- Calculation of Eigen Values and Eigen Vectors.
- Cayley-Hamilton theorem and its Applications.

**Structure**

- 2.1 Introduction to Vectors and vector Spaces**
- 2.2 Fundamental concepts and definitions of Matrix Polynomials**
- 2.3 Characteristic roots and characteristic Vectors of a matrix**
- 2.4 Cayley-Hamilton theorem and its applications**
- 2.5 Summary**
- 2.6 Self Assessment Questions**
- 2.7 Further Readings**

**2.1. Introduction to Vectors and vector Spaces**

In the previous lesson we have discussed some fundamental concepts of matrix analysis, where the elements of the matrix are real numbers. We can extend this analysis in many directions. For example, one can extend this to **complex numbers** by considering the elements of a matrix as complex numbers. Similarly one can view the elements in a row or column as a row or column **Vector**. Those concepts relating to these extensions are discussed in this lecture. First we discuss some fundamental concepts of Vectors and vector space of the given matrix. A vector is defined as follows:

**Definition:** Any ordered n-tuple of numbers is called an **n-vector**. By an ordered n-tuple we mean a set consisting 'n' numbers, in which the place of a number is fixed. For example elements of matrix in a row or in a column, where 1<sup>st</sup>, 2<sup>nd</sup> and n<sup>th</sup> element's places are fixed and are denoted by  $x_1, x_2, \dots, x_n$ , where, these numbers are any real or complex numbers. If **A** is a matrix of size  $m \times n$ , then each row of **A** contains 'n' elements and hence is an n-vector and is known as "**Row Vector**" of **A**. Similarly, each column of the matrix **A** consists of 'm' elements hence, is a m-vector and is known as a "**Column Vector**" of **A**. This ordered n-tuple is called an **n-vector** and is denoted by  $\mathbf{X} = (x_1, x_2, \dots, x_n)$ . For example, (1,0,-1,3,5,6) is a 6-vector. Similarly,  $(x_1, x_2, x_3)$  is a 3-vector and (2, -8, 11, 3) is 4-vector and so on.

**2.1.1. Algebra of Vectors:** Since, an n-vector is nothing but a row matrix or a column matrix, one can develop an algebra of vectors in the same manner as the algebra of matrices or operations of matrices discussed in the previous lecture.

**(a) Equality of two Vectors:** Two n-vectors  $\mathbf{X}$  and  $\mathbf{Y}$ , where  $\mathbf{X} = (x_1, x_2, \dots, x_n)$  and  $\mathbf{Y} = (y_1, y_2, \dots, y_n)$  are said to be **equal** if and only if their corresponding components are equal. For example if  $\mathbf{X} = (4, -10, 12)$ ,  $\mathbf{Y} = (4, -10, 12)$  and  $\mathbf{Z} = (12, -10, 4)$ , then  $\mathbf{X} = \mathbf{Y}$  but  $\mathbf{X} \neq \mathbf{Z}$ .

**(b) Addition of two Vectors:** If  $\mathbf{X} = (x_1, x_2, \dots, x_n)$  and  $\mathbf{Y} = (y_1, y_2, \dots, y_n)$  are two n-vectors, then by definition  $\mathbf{X} + \mathbf{Y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$ , which is a n-vector whose elements are the sum of the corresponding elements  $\mathbf{X}$  and  $\mathbf{Y}$ .

**(c) Multiplication of a Vector by a Scalar:** Let  $C$  is any number and  $\mathbf{X} = (x_1, x_2, \dots, x_n)$  a n-vector, then by definition the product  $C\mathbf{X} = (Cx_1, Cx_2, \dots, Cx_n)$ .

**Properties of addition and scalar multiplication of Vectors:** If  $\mathbf{X}$ ,  $\mathbf{Y}$ ,  $\mathbf{Z}$  are any three n-vectors and  $a, b$  are any two numbers, then the following properties are true for addition and scalar multiplication of vectors.

- (1)  $\mathbf{X} + \mathbf{Y} = \mathbf{Y} + \mathbf{X}$ .
- (2)  $(\mathbf{X} + \mathbf{Y}) + \mathbf{Z} = \mathbf{X} + (\mathbf{Y} + \mathbf{Z})$ .
- (3)  $a(\mathbf{X} + \mathbf{Y}) = a\mathbf{X} + a\mathbf{Y}$ .
- (4)  $(a + b)\mathbf{X} = a\mathbf{X} + b\mathbf{X}$ .
- (5)  $a(b\mathbf{X}) = (ab)\mathbf{X} = b(a\mathbf{X})$

**(a) Linearly dependent and linearly independent set of Vectors:** Let us consider a set 'r' n-vectors  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_r$ . These set of vectors is said to be "**Linearly Dependent Vectors**" if there exists 'r' scalars  $C_1, C_2, \dots, C_r$  not all zeros, such that:

$$\mathbf{X} = C_1\mathbf{X}_1 + C_2\mathbf{X}_2 + \dots + C_r\mathbf{X}_r = \mathbf{0}, \quad (2.1.1)$$

where  $\mathbf{0}$  denote a null vector or n-vector whose elements are all zeros.

In particular equation (2.1.1) holds good when  $C_1 = C_2 = \dots = C_r = 0$ , then the set of vectors is said to be "**Linearly Independent Vectors**". The vector  $\mathbf{X}$  expressed in the form  $\mathbf{X} = C_1\mathbf{X}_1 + C_2\mathbf{X}_2 + \dots + C_r\mathbf{X}_r$  is said to be a "**Linear Combination**" of the Set of Vectors  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_r$ . Here  $C_1, C_2, \dots, C_r$  are any numbers. The following two results of linear combinations are obvious.

**Result 1:** If a set of vectors is linearly dependent, then at least one member of the set can be expressed as a linear combination of the remaining members.

**Result 2:** If a set of vectors is linearly independent, then no member of the set can be expressed as a linear combination of the remaining members.

**(b) The n-vector Space and sub-space:** The set of all n-vectors of a field  $F$  is called the n vector space over  $F$ . It is usually denoted by  $V_n(F)$  or simply by  $V_n$ . Similarly, a non-empty set,  $S$  of vectors of  $V_n$  is called a "**vector sub-space**" of  $V_n$  if  $\mathbf{a} + \mathbf{b}$  belongs to  $S$  whenever  $\mathbf{a}, \mathbf{b}$

belong to  $\mathbf{S}$  and  $C\mathbf{a}$  belongs to  $\mathbf{S}$  whenever  $C$  belongs to  $\mathbf{S}$ , where  $C$  is any scalar. It is important to note that every sub-space of  $\mathbf{V}_n$  contains the zero vector, which is being the scalar product of any vector with the scalar zero.

## 2.2. Fundamental concepts and definitions of Matrix Polynomials

Now we proceed to explain some fundamental concepts of “Eigen values” and “Eigen Vectors” along with polynomials of matrices known as “Matrix Polynomials”. First we define the concept of “Matrix Polynomial” as follows:

**Definition:** An expression of the form:

$$\mathbf{F}(\lambda) = \mathbf{A}_0 + \mathbf{A}_1\lambda + \mathbf{A}_2\lambda^2 + \dots + \mathbf{A}_{m-1}\lambda^{m-1} + \mathbf{A}_m\lambda^m \quad (2.2.1)$$

Where  $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$ , are all square matrices of the same order is called as “Matrix Polynomial” of degree ‘m’ provided  $\mathbf{A}_m$  is not a null matrix. The symbol  $\lambda$  is called “Indeterminate”. If the order of each of the matrix coefficients  $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$  is ‘n’, then we say that the matrix polynomial is n-rowed. According to this definition of matrix polynomial, every square matrix can be expressed as a matrix polynomial with zero degree. For example, let  $\mathbf{A}$  is any square matrix then we can write  $\mathbf{A} = \lambda^0 \mathbf{A}$ .

**2.2.1. Equality of polynomials:** Two matrix polynomials are equal if and only if the coefficients of the like powers of  $\lambda$  are the same.

**Theorem(2.2.1):** Every square matrix, whose elements are ordinary polynomials in  $\lambda$ , can easily be expressed as a matrix polynomial in  $\lambda$  of degree ‘m’, where m is the highest power of  $\lambda$  occurring in any element of the matrix.

## 2.3. Characteristic roots and characteristic Vectors of a matrix.

Let  $\mathbf{A} = [a_{ij}]_{n \times n}$  be an n-rowed square matrix. Let  $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$  be a column vector.

Consider the vector equation:  $\mathbf{AX} = \lambda\mathbf{X}$ , (2.3.1)

where  $\lambda$  is a scalar, that is any number.

It is observed that the vector equation (2.3.1), has an obvious solution  $\mathbf{X} = \mathbf{0}$ , that is a zero vector is an obvious solution for any value of  $\lambda$ . But an interesting question is that: whether there exists a scalar  $\lambda$  and a non-zero vector  $\mathbf{X}$  satisfying vector equation (2.3.1)? To answer this question we proceed as follows:

Consider a unit matrix  $\mathbf{I}$  of order n, and then the vector equation (2.3.1) can be written as:  $\mathbf{AX} = \lambda\mathbf{I}\mathbf{X}$  or  $(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0}$  (2.3.2)

The above matrix equation represents the following system of ‘n’ homogeneous linear equation in ‘n’ unknowns.

$$(a_{11} - \lambda) x_1 + a_{12} x_2 + \dots + a_{1n} x_n = 0$$

$$a_{21} x_1 + (a_{22} - \lambda) x_2 + \dots + a_{2n} x_n = 0$$

.....  
 .....

$$a_{n1} x_1 + a_{n2} x_2 + \dots + (a_{nn} - \lambda) x_n = 0. \quad (2.3.3)$$

The matrix  $(\mathbf{A} - \lambda \mathbf{I})$  in the vector equation (2.3.2) is called “Coefficient Matrix” and is given as follows:

$$\begin{bmatrix} (a_{11} - \lambda) & a_{12}, \dots, & a_{1n} \\ a_{21} & (a_{22} - \lambda), \dots, & a_{2n} \\ a_{n1} & a_{n2} & (a_{nn} - \lambda) \end{bmatrix} \quad (2.3.4)$$

**Result:** The necessary and sufficient condition for the system of equations  $(\mathbf{A} - \lambda \mathbf{I}) \mathbf{X} = \mathbf{0}$  to possess a non-zero solution ( $\mathbf{X} \neq \mathbf{0}$ ) is that the coefficient matrix  $(\mathbf{A} - \lambda \mathbf{I})$  should be of the rank less than the number of unknowns ‘n’.

Above result will be true if and only if the coefficient matrix  $(\mathbf{A} - \lambda \mathbf{I})$  is singular. That is if and only if the determinant of the coefficient matrix must be equal to zero. That is

$$|(\mathbf{A} - \lambda \mathbf{I})| = 0 \quad (2.3.5)$$

Thus the scalars  $\lambda$  for which  $|(\mathbf{A} - \lambda \mathbf{I})| = 0$ , the roots of the equation are of special importance in matrix algebra and are known as “Eigen Values” or “Characteristic Roots” or “Characteristic Values” or “latent roots” or “proper values” of the matrix  $\mathbf{A}$ . Based on the above discussion, now we proceed to give definitions of Eigen values and Eigen vectors as follows:

**Definition (1):** Let  $A = [a_{ij}]_{n \times n}$  be any n rowed square matrix and  $\lambda$  is an indeterminate. The matrix  $(A - \lambda I)$  is called the “Characteristic Matrix” of  $A$ , where  $I$  is the unit matrix of order  $n$ .

**Definition (2):** The determinant  $|A - \lambda I| = \begin{vmatrix} (a_{11} - \lambda) & a_{12}, \dots, & a_{1n} \\ a_{21} & (a_{22} - \lambda), \dots, & a_{2n} \\ a_{n1} & a_{n2} & (a_{nn} - \lambda) \end{vmatrix}$

is an ordinary polynomial in  $\lambda$  of degree  $n$ , is called the “Characteristic Polynomial” of  $A$ .

**Definition (3):** The equation  $|A - \lambda I| = 0$  is called “Characteristic Equation” of  $A$  and the roots of this equation are called “Characteristic roots” or “Characteristic values” or “Eigen values” or “Latent roots” or “proper values” of the matrix  $A$ .

**Definition (4):** The set of all the Eigen values or characteristic roots of the matrix  $A$  is called the “Spectrum” of  $A$ .

**Definition (5):** If  $\lambda$  is the characteristic root of an  $n \times n$  matrix  $A$ , then a non-zero vector  $X$  such that  $AX = \lambda X$  is called the “Characteristic vector” or “Eigen vector” or “Latent vector” of the matrix  $A$  corresponding to the characteristic root  $\lambda$ .

Now we proceed to discuss some results relating to characteristic roots and characteristic vectors of matrices.

**Result (1):** The number  $\lambda$  is the characteristic root of a matrix  $\mathbf{A}$ , if and only if there exists a non-zero vector  $\mathbf{X}$  such that  $\mathbf{AX} = \lambda\mathbf{X}$ .

**Result (2):** If  $\mathbf{X}$  is a characteristic vector of a matrix  $\mathbf{A}$ , corresponding to the characteristic value  $\lambda$ , then  $C\mathbf{X}$  is also a characteristic vector of  $\mathbf{A}$  corresponding to the same characteristic root  $\lambda$ , where  $C$  is any non-zero scalar.

**Result (3):** If  $\mathbf{X}$  is a characteristic vector of a matrix  $\mathbf{A}$ , then  $\mathbf{X}$  cannot correspond to more than one characteristic value of  $\mathbf{A}$ .

**Result (4):** The characteristic vectors corresponding to distinct characteristic roots of a matrix are linearly independent.

Now we proceed to explain the process or procedure of finding Characteristic roots and characteristic vectors of a given matrix  $\mathbf{A}$ .

### 2.3.1. The procedure of finding Eigen values and Eigen vectors:

Let  $A = [a_{ij}]_{n \times n}$  be the given square matrix of order  $n$ . The procedure of calculating characteristic roots and characteristic vectors is explained as follows:

Step – 1: Write the characteristic equation of the given matrix. That is  $|\mathbf{A} - \lambda\mathbf{I}| = 0$  is to be formed.

Step – 2: Form the polynomial of degree  $n$  in  $\lambda$ .

Step – 3: The equation formed in step 2 will have  $n$  roots. These  $n$  roots denoted by  $\lambda_1, \lambda_2, \dots, \lambda_n$  are the characteristic roots or Eigen values of the given matrix  $\mathbf{A}$ .

Step – 4: If  $\lambda_1$  is the Eigen value of the matrix  $\mathbf{A}$ , then the corresponding Eigen vector of the matrix  $\mathbf{A}$  will be the non zero vector  $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ , satisfying the equation  $\mathbf{AX} = \lambda_1\mathbf{X}$  or  $(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{X} = \mathbf{0}$ .

Using the above procedure, now we proceed to work out some examples of finding characteristic roots and characteristic vectors of a given matrix.

**Example (2.3.1):** Find the characteristic roots of the matrix  $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$ .

**Solution:** The characteristic equation of the given matrix is:  $|\mathbf{A} - \lambda\mathbf{I}| = 0$ . Thus we have:

$$\begin{vmatrix} 1 - \lambda & 4 \\ 2 & 3 - \lambda \end{vmatrix} = 0 \text{ or } (1 - \lambda)(3 - \lambda) - 8 = 0 \text{ or } \lambda^2 - 4\lambda - 5 = 0 \text{ or } (\lambda - 5)(\lambda + 1) = 0.$$

Hence,  $\lambda = 5$  and  $\lambda = -1$  are the roots of the equation and hence are called characteristic roots or Eigen values of the given matrix  $\mathbf{A}$ .

**Example (2.3.2):** Find the characteristic roots of the matrix  $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & 2 \\ 0 & 0 & 7 \end{bmatrix}$ .

**Solution:** The characteristic equation of the given matrix is:  $|\mathbf{A} - \lambda\mathbf{I}| = 0$ . Thus we have:

$$\begin{vmatrix} 1-\lambda & 2 & 3 \\ 0 & -4-\lambda & 2 \\ 0 & 0 & 7-\lambda \end{vmatrix} = 0 \text{ or } (1-\lambda)(-4-\lambda)(7-\lambda) = 0. \text{ Hence } \lambda = 1, -4 \text{ and } 7 \text{ are the}$$
characteristic roots of the matrix  $\mathbf{A}$ .

**Example (2.3.3):** Find the characteristic roots of the matrix  $\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -1 \\ 2 & -1 & 0 \end{bmatrix}$ .

**Solution:** The characteristic equation of the given matrix is:  $|\mathbf{A} - \lambda\mathbf{I}| = 0$ . Thus we have:

$$\begin{vmatrix} 0-\lambda & 1 & 2 \\ 1 & 0-\lambda & -1 \\ 2 & -1 & 0-\lambda \end{vmatrix} = \begin{vmatrix} -\lambda & 1 & 2 \\ 1 & -\lambda & -1 \\ 2 & -1 & -\lambda \end{vmatrix} = -\lambda(\lambda^2 - 1) - 1(-\lambda + 2) + 2(-1 + 2\lambda) = 0.$$

$= \lambda^3 - 6\lambda + 4 = 0$  or  $(\lambda - 2)(\lambda^2 + 2\lambda - 2) = 0$ . The roots of the equation are 2,  $-1 \pm \sqrt{3}$ . Hence, the characteristic roots of the matrix are 2,  $(-1 + \sqrt{3})$  and  $(-1 - \sqrt{3})$ .

**Example (2.3.4):** Determine the Eigen values and Eigen vectors of the matrix  $\mathbf{A} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$ .

**Solution:** The characteristic equation of the given matrix is:  $|\mathbf{A} - \lambda\mathbf{I}| = 0$ . Thus we have:

$$\begin{vmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{vmatrix} = 0 \text{ or } (5-\lambda)(2-\lambda) - 4 = 0 \text{ or } \lambda^2 - 7\lambda + 6 = 0 \text{ or } (\lambda - 6)(\lambda - 1) = 0.$$

Hence,  $\lambda = 6$  and  $\lambda = 1$  are the roots of the equation and hence are called characteristic roots or Eigen values of the matrix  $\mathbf{A}$ . Now we proceed to calculate Eigen vectors  $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$  of the matrix  $\mathbf{A}$  corresponding Eigen values 6 and 1. The Eigen vector for the Eigen value 6 is given by:

$(\mathbf{A} - 6\mathbf{I})\mathbf{X} = \mathbf{0}$ . That is  $\begin{bmatrix} 5-6 & 4 \\ 1 & 2-6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  or  $\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ . Applying elementary row operation  $R_2 = R_2 + R_1$ , that is adding 2<sup>nd</sup> row and 1<sup>st</sup> row to get second row, we have:  $\begin{bmatrix} -1 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ . Since the rank of the coefficient matrix is 1. Therefore, these equations have 2 -1 = 1 linearly independent solution. These equations reduce to a single equation namely,  $-x_1 + 4x_2 = 0$ . Thus we have  $x_1 = 4$ , and  $x_2 = 1$ . Thus the Eigen vector of the matrix  $\mathbf{A}$ , corresponding to  $\lambda = 6$  is given by  $\mathbf{X}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ . Now we proceed to find the Eigen vector for  $\lambda = 1$  is given by:  $(\mathbf{A} - 1\mathbf{I})\mathbf{X} = \mathbf{0}$ .

$$\text{That is } \begin{bmatrix} 5-1 & 4 \\ 1 & 2-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ or } \begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus we have:  $4x_1 + 4x_2 = 0$  and  $x_1 + x_2 = 0$ .

Solving these equations we have  $x_1 = 1$  and  $x_2 = -1$ . Hence the second Eigen vector for  $\lambda = 1$

$$\mathbf{X}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

## 2.4. Cayley-Hamilton theorem and its applications.

Now we proceed to prove an important theorem known as “Cayley-Hamilton Theorem” which has lot of applications in matrix analysis particularly in obtaining the inverse a square matrix.

### 2.4.1. Cayley-Hamilton theorem:

**Statement:** Every Square matrix satisfies its characteristic equation.

**Explanation:** If  $A$  is a square matrix of order  $n$ , then the characteristic equation  $|\mathbf{A} - \lambda \mathbf{I}|$  can be expressed as  $(-1)^n [\lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + \dots + a_n \lambda^0]$ . Then the matrix equation:

$\mathbf{X}^n + a_1 \mathbf{X}^{n-1} + a_2 \mathbf{X}^{n-2} + \dots + a_n \mathbf{I} = \mathbf{0}$  is satisfied by  $\mathbf{X} = \mathbf{A}$ . That is  $\mathbf{A}^n + a_1 \mathbf{A}^{n-1} + a_2 \mathbf{A}^{n-2} + \dots + a_n \mathbf{I} = \mathbf{0}$ .

**Proof:** We know that the elements of the matrix  $(\mathbf{A} - \lambda \mathbf{I})$  are at most the first degree in  $\lambda$ , the elements of  $\text{Adj.}(\mathbf{A} - \lambda \mathbf{I})$  are ordinary polynomials in  $\lambda$  of degree  $(n-1)$  or less. Therefore,

$\text{Adj.}(\mathbf{A} - \lambda \mathbf{I})$  can be written as a matrix polynomial in  $\lambda$  and is given as follows:

$$\text{Adj.}(\mathbf{A} - \lambda \mathbf{I}) = \mathbf{B}_0 \lambda^{n-1} + \mathbf{B}_1 \lambda^{n-2} + \dots + \mathbf{B}_{n-2} \lambda + \mathbf{B}_{n-1}. \quad (2.4.1)$$

Where,  $\mathbf{B}_0, \mathbf{B}_1, \dots, \mathbf{B}_{n-1}$  are matrices of order  $n \times n$ , whose elements are functions of  $a_{ij}$ 's.

We know that  $\mathbf{A} \text{Adj.} \mathbf{A} = |\mathbf{A}| \mathbf{I}_n$ . Hence using the equation (2.4.1) we have:

$$(\mathbf{A} - \lambda \mathbf{I}) \text{Adj.}(\mathbf{A} - \lambda \mathbf{I}) = |\mathbf{A} - \lambda \mathbf{I}| \mathbf{I}_n.$$

Therefore we have:  $(\mathbf{A} - \lambda \mathbf{I}) (\mathbf{B}_0 \lambda^{n-1} + \mathbf{B}_1 \lambda^{n-2} + \dots + \mathbf{B}_{n-2} \lambda + \mathbf{B}_{n-1})$

$$= (-1)^n [\lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + \dots + a_n] \mathbf{I}. \quad (2.4.2)$$

Now, comparing coefficients of like powers of  $\lambda$  on both sides of equation (2.4.2) we have:

$$\begin{aligned} -\mathbf{I} \mathbf{B}_0 &= (-1)^n \mathbf{I}, \\ \mathbf{A} \mathbf{B}_0 - \mathbf{I} \mathbf{B}_1 &= (-1)^n a_1 \mathbf{I} \\ \mathbf{A} \mathbf{B}_1 - \mathbf{I} \mathbf{B}_2 &= (-1)^n a_2 \mathbf{I} \\ &\dots \\ &\dots \\ \mathbf{A} \mathbf{B}_{n-1} &= (-1)^n a_n \mathbf{I}. \end{aligned} \quad (2.4.3)$$

Pre-multiplying above set of equations given in (2.4.3) successively by  $\mathbf{A}^n, \mathbf{A}^{n-1}, \dots, \mathbf{I}$  and adding, we obtain:  $\mathbf{0} = (-1)^n [\mathbf{A}^n + a_1 \mathbf{A}^{n-1} + a_2 \mathbf{A}^{n-2} + \dots + a_n \mathbf{I}]$ .

Thus we have:  $\mathbf{A}^n + a_1 \mathbf{A}^{n-1} + a_2 \mathbf{A}^{n-2} + \dots + a_n \mathbf{I} = \mathbf{0}$ . Hence the theorem.

### 2.4.2. Applications of Cayley-Hamilton Theorem:

The Cayley-Hamilton theorem can be applied to find the determinant of a given square matrix and also its inverse. These applications are discussed as the following two results.

**Result (1):** If  $\mathbf{A}$  be a non-singular matrix,  $|\mathbf{A}| \neq 0$ , then  $|\mathbf{A}| = (-1)^n a_n$  and hence  $a_n \neq 0$ .



**Result (2):** Pre-multiplying  $[A^n + a_1A^{n-1} + a_2A^{n-2} + \dots + a_nI = 0]$  by  $A^{-1}$  we have:

$$A^{n-1} + a_1A^{n-2} + a_2A^{n-3} + \dots + a_{n-1}I + a_nA^{-1} = 0. \quad (2.4.4)$$

$$\text{Hence we have } A^{-1} = -(1/a_n) [A^{n-1} + a_1A^{n-2} + a_2A^{n-3} + \dots + a_{n-1}I]. \quad (2.4.5)$$

**Example (2.4.1):** Find the characteristic equation of the matrix  $A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$ .

Show that  $A$  satisfies the Cayley-Hamilton theorem and hence find  $A^{-1}$ .

**Solution:** : The characteristic equation of the given matrix is:  $|A - \lambda I| = 0$ . Thus we have:

$$\begin{vmatrix} 2-\lambda & -1 & 1 \\ -1 & 2-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{vmatrix} = (2-\lambda)\{(2-\lambda)^2 - 1\} + 1\{-1(2-\lambda) + 1\} + 1\{1 - (2-\lambda)\} \\ = (2-\lambda)(3-4\lambda+\lambda^2) + (\lambda-1) + (\lambda-1) = -\lambda^3 + 6\lambda^2 - 9\lambda + 4.$$

Thus the characteristic equation of the matrix  $A$  is given by  $\lambda^3 - 6\lambda^2 + 9\lambda - 4I = 0$ .

$$\text{Now we have } I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}, A^2 = AxA = \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix}$$

$$A^3 = A^2xA = \begin{bmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{bmatrix}.$$

Now we can verify that  $A^3 - 6A^2 + 9A - 4I = 0$ . That is

$$\begin{bmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{bmatrix} - 6 \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix} + 9 \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} - 4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus Cayley-Hamilton theorem is satisfied. Now we proceed to calculate  $A^{-1}$ , where:

$A^{-1} = (1/4) [A^2 - 6A + 9I]$ . Now we have:

$$A^2 - 6A + 9I = \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix} - 6 \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} + 9 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}$$

Therefore  $A^{-1} = (1/4) \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}$  or  $\begin{bmatrix} 3/4 & 1/4 & -1/4 \\ 1/4 & 3/4 & 1/4 \\ -1/4 & 1/4 & 3/4 \end{bmatrix}$ . Further one can verify the result

that:  $AxA^{-1} = I$ .

## 2.5 Summary

In this lesson first we have understood the concepts of “**Row vector**”, “**Column vector**” and the “**matrix polynomial**”. Then, we have discussed the properties of polynomials like equality of polynomials, addition and scalar multiplication of polynomials and their properties. Then we have discussed the concepts like, “**Characteristic Matrix**”, “**Characteristic Equation**” “**Eigen values**” or “**Latent roots**”, or “**Characteristic roots**” and their vectors and discussed the procedure of obtaining them along with some worked out examples. Finally, an important

theorem, due to “Cayley and Hamilton” ,known as “Cayley-Hamilton Theorem” was proved and discussed their special applications in finding out the determinant and the inverse of a given square matrix.

## 2.6 Self Assessment Questions

1. Define row vector, column vector and equality of two vectors, with suitable examples.
2. Explain various operations of vectors and discuss their properties. Also explain linearly dependent and independent vectors.
3. Define n-vector space and its sub-space and matrix polynomials with suitable examples.
4. Define Eigen values and Eigen vectors and calculate the same for the matrix A, where

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}.$$

5. Find the characteristic equation of the matrix  $A = \begin{bmatrix} 1 & 3 & 7 \\ 4 & 2 & 3 \\ 0 & 2 & 1 \end{bmatrix}$  and show that it is satisfied by A. Hence, obtain  $A^{-1}$ .

6. State and prove “Cayley-Hamilton” theorem along with their applications.

7. Find Characteristic roots and characteristic vectors of the matrix  $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}$  and show that A satisfies the Cayley-Hamilton theorem. Also find  $A^{-1}$ .

8. Define the characteristic equation of a square matrix and show that the matrix A

$$= \begin{bmatrix} 1 & 2 & 3 \\ 3 & -2 & 1 \\ 4 & 2 & 1 \end{bmatrix} \text{ satisfies the characteristic equation.}$$

9. Verify whether the Cayley-Hamilton theorem is true for  $A = \begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix}$ . If it is true,

Find  $A^{-1}$ . Also verify whether  $AxA^{-1} = I$  or not.

10. Show that the matrix  $A = \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix}$  satisfies Cayley-Hamilton theorem.

## 2.7 Further Readings

1. A.R. Vasishtha "Matrices" Krishna Prakashan Media (p) Ltd.; Thirty fourth Edn., Meerut – 250 001; 2003.
2. A. RamachandraRao and P. BhimaShankaram "Linear Algebra" Tata McGraw-Hill Publishing Company Limited; 1992.
3. D.D. Joshi "Theory of Linear Estimation" Woley Eastern Ltd. New Delhi. 2005.
4. Gupta and Kapoor "fundamentals of Applied Statistics" Sultan Chand & sons, New Delhi. 2008.

**Lesson – 3****Orthogonal Vectors and Matrices****1.0 Objectives :**

After going through this lesson, you should be able to:

- Understand the concepts of inner product, orthogonality, definitions and their applications.
- The concepts of orthogonal vectors, matrices, their properties and applications.
- The definition and application of trace of matrices.
- Differentiation using matrices and their properties.
- Applications of above concepts in matrix algebra.

**Structure**

- 3.1. Introduction to inner product and its properties
- 3.2. Fundamental concepts and definitions of orthogonal vectors
- 3.3. Fundamental concepts and definitions of Unitary and Orthogonal matrices
- 3.4. Definition and properties of trace of a matrix and their properties
- 3.5. Summary
- 3.6. Self Assessment Questions
- 3.7. Further Readings

**3.1. Introduction to inner product and its properties.**

In the earlier lessons, we have discussed some fundamental definition of matrix algebra and introduced vectors, vector spaces, sub-spaces and polynomials. We have also introduced Eigen values and Eigen vectors, their properties and applications. Now we proceed to discuss some advanced topics like inner product and orthogonal vectors and matrices, which have further applications in matrix analysis, where the elements of the matrix can be complex numbers. Now we proceed to introduce the inner product of two vectors, in the following section.

**3.1.1. Inner Product of two vectors:**

**Definition (1):** Let  $\mathbf{X}$  and  $\mathbf{Y}$  are two complex column vectors of order  $n \times 1$ , where  $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$  and

$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ . The “inner product” of the vectors  $\mathbf{X}$  and  $\mathbf{Y}$ , denoted by  $(\mathbf{X}, \mathbf{Y})$  and is defined as:

$$(\mathbf{X}, \mathbf{Y}) = \bar{x}_1 y_1 + \bar{x}_2 y_2 + \bar{x}_3 y_3 + \dots + \bar{x}_n y_n. \quad (3.1.1)$$

Here  $\bar{x}_i$ 's denote the conjugate complex of the complex number  $x_i$ .

For all practical purposes, generally we define a  $1 \times 1$  matrix with its single element. That is if  $[a]_{1 \times 1}$  is a  $1 \times 1$  matrix, then we shall simply regard it as a scalar  $a$ . With this definition, the inner product of the vectors  $\mathbf{X}$  and  $\mathbf{Y}$  may be conveniently defined as:  $(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^o \mathbf{Y}$ . Here  $\mathbf{X}^o$  is a  $1 \times n$  matrix and  $\mathbf{Y}$  is a  $n \times 1$  matrix. Hence,  $\mathbf{X}^o \mathbf{Y}$  is a  $1 \times 1$  matrix and it has been taken equal to its element. Similarly, if  $\mathbf{X}$  and  $\mathbf{Y}$  are real  $n$ -vector written as column vectors, then their inner product will be written as:

$$(\mathbf{X}, \mathbf{Y}) = \mathbf{X}^T \mathbf{Y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n. \quad (3.1.2)$$

If  $\mathbf{X}$  and  $\mathbf{Y}$  are complex  $n$ -vectors, written as row vectors, then their inner product is defined as:  $(\mathbf{X}, \mathbf{Y}) = \mathbf{X} \mathbf{Y}^o = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n$ . (3.1.3)

**Remark:** Usually in all the "inner products" discussions,  $\mathbf{X}$  and  $\mathbf{Y}$  are assumed as column vectors of size  $1 \times n$  unless otherwise stated.

**3.1.1. Properties of inner product:** Let  $\mathbf{X}, \mathbf{Y}$  and  $\mathbf{Z}$  are any three complex  $n$ -vectors and  $C$  is any complex number, then the following results are true.

(1)  $(\mathbf{X}, \mathbf{X}) \geq 0$ , and  $(\mathbf{X}, \mathbf{X}) = 0$  if and only if  $\mathbf{X} = \mathbf{0}$ .

(2)  $(\mathbf{X}, \mathbf{Y}) = \overline{(\mathbf{Y}, \mathbf{X})}$ .

(3)  $(\mathbf{X}, \mathbf{Y} + \mathbf{Z}) = (\mathbf{X}, \mathbf{Y}) + (\mathbf{X}, \mathbf{Z})$ .

(4)  $(\mathbf{X} + \mathbf{Y}, \mathbf{Z}) = (\mathbf{X}, \mathbf{Z}) + (\mathbf{Y}, \mathbf{Z})$ .

(5)  $(\mathbf{X}, C\mathbf{Y}) = C(\mathbf{X}, \mathbf{Y})$  and (6)  $(C\mathbf{X}, \mathbf{Y}) = \bar{C}(\mathbf{X}, \mathbf{Y})$ .

### 3.2. Fundamental concepts and definitions of Orthogonal Vectors

Now we proceed to define the orthogonal vectors and matrices along with their properties, related results and applications in this section.

**Definition 1: Orthogonal Vectors:** Let  $\mathbf{X}$  and  $\mathbf{Y}$  be two complex  $n$ -vectors, then  $\mathbf{X}$  is said to be orthogonal to  $\mathbf{Y}$  if:

$$(\mathbf{X}, \mathbf{Y}) = 0. \quad \text{That is if } \mathbf{X}^o \mathbf{Y} = 0. \quad (3.2.1)$$

Where  $\mathbf{X}^o$  is the transpose conjugate of  $\mathbf{X}$ .

**Result 1:** The relation of orthogonality is symmetric, that is if  $\mathbf{X}$  is orthogonal to  $\mathbf{Y}$  then  $\mathbf{Y}$  is orthogonal to  $\mathbf{X}$ . This is because of the fact that: if  $\mathbf{X}$  is orthogonal to  $\mathbf{Y} \Rightarrow (\mathbf{X}, \mathbf{Y}) = 0$ .

This implies that  $\overline{(\mathbf{X}, \mathbf{Y})} = \overline{0} \Rightarrow (\mathbf{Y}, \mathbf{X}) = 0 \Rightarrow \mathbf{Y}$  is orthogonal to  $\mathbf{X}$ .

**Result 2:** If  $\mathbf{X}$  is orthogonal to  $\mathbf{Y}$ , then every scalar multiple of  $\mathbf{X}$  is orthogonal to every scalar multiple of  $\mathbf{Y}$ . That is if  $a, b$  are any two scalars, then  $(a\mathbf{X}, b\mathbf{Y}) = \bar{a} b (\mathbf{X}, \mathbf{Y}) = 0$ . This is because of the fact that  $(\mathbf{X}, \mathbf{Y}) = 0$ . Thus  $a\mathbf{X}$  and  $b\mathbf{Y}$  are orthogonal vectors.

**Result 3:** The zero vector is the only vector which is orthogonal to itself. We know that  $\mathbf{X}$  is orthogonal to  $\mathbf{X}$  itself. That is  $(\mathbf{X}, \mathbf{X}) = 0 \Rightarrow \mathbf{X} = \mathbf{0}$ .

**Result 4:** Two real  $n$ -vectors  $\mathbf{X} = [x_1, x_2, \dots, x_n]^T$  and  $\mathbf{Y} = [y_1, y_2, \dots, y_n]^T$  are orthogonal if and only if  $(\mathbf{X}, \mathbf{Y}) = 0$ . That is true if and only if  $\mathbf{X}^T \mathbf{Y} = 0$ . That is true if and only if  $x_1 y_1 + x_2 y_2 + \dots + x_n y_n = 0$ .

**Definition 2: Orthogonal Set:** A set  $\mathbf{S}$  of complex  $n$ -vectors  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k$  is said to be orthogonal Set if any two distinct vectors in  $\mathbf{S}$  are orthogonal.

**Definition 3: Orthonormal Set:** A set  $\mathbf{S}$  of complex  $n$ -vectors  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k$  is said to be orthonormal Set if:

(i) Each vector in  $\mathbf{S}$  is a unit vector and

(ii) any two distinct vectors in  $\mathbf{S}$  are orthogonal.

**Definition 4: Kronecker Delta:** The symbol  $\delta_{ij}$  is said to be Kronecker delta if

$$\delta_{ij} = 0 \text{ when } i \neq j \text{ and}$$

$$\delta_{ij} = 1 \text{ when } i = j.$$

In terms of Kronecker delta, the unit matrix  $\mathbf{I}_n$  can be written as:

$$\mathbf{I}_n = [\delta_{ij}]_{n \times n}.$$

In terms of Kronecker delta an orthogonal set may be defined as follows:

**Definition 5:** A set  $\mathbf{S}$  of complex  $n$ -vectors  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k$  is said to be orthonormal Set if:

$$(\mathbf{X}_i, \mathbf{X}_j) = \delta_{ij}, \quad i = 1, 2, 3, \dots, k \text{ and } j = 1, 2, 3, \dots, k.$$

**Definition 6: Orthogonal basis:** If an orthogonal set  $\mathbf{S}$  is a basis of  $V_n$ , then it is called an orthogonal basis of  $V_n$ .

**Definition 7: Orthonormal basis:** If an orthonormal set  $\mathbf{S}$  is a basis of  $V_n$ , then it is called an orthonormal basis of  $V_n$ .

### 3.2.1. Properties of Orthogonal sets:

(1) Every orthogonal set of non-zero vectors is linearly independent.

(2) Every orthonormal set of vectors is linearly independent.

(3) If  $\mathbf{S} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k\}$  is an orthogonal set of non-zero complex  $n$ -vectors and  $\mathbf{Y}$  is any complex  $n$ -vector, then:

$$\mathbf{Z} = \mathbf{Y} - \left\{ \frac{(\mathbf{X}_1, \mathbf{Y})}{(\mathbf{X}_1, \mathbf{X}_1)} \mathbf{X}_1 + \frac{(\mathbf{X}_2, \mathbf{Y})}{(\mathbf{X}_2, \mathbf{X}_2)} \mathbf{X}_2 + \frac{(\mathbf{X}_3, \mathbf{Y})}{(\mathbf{X}_3, \mathbf{X}_3)} \mathbf{X}_3 + \dots + \frac{(\mathbf{X}_k, \mathbf{Y})}{(\mathbf{X}_k, \mathbf{X}_k)} \mathbf{X}_k \right\} \text{ is orthogonal to each of the vectors } \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k.$$

Further it is important to note that we can always construct an orthogonal basis of the vector space  $V_n$  from a given basis. Orthogonal vectors have lot of applications in design of experiments particularly in the analysis of Balanced incomplete Block designs.

### 3.3. Fundamental concepts and definitions of Unitary and Orthogonal matrices

Now we proceed to explain the concepts of unitary matrices and orthogonal matrices which are frequently used in design of experiments.

**Definition 1: Unitary matrix:** A square matrix  $\mathbf{P}$  with complex elements is said to be unitary matrix if  $\mathbf{P}^{\circ}\mathbf{P} = \mathbf{I}$ . Here  $\mathbf{P}^{\circ}$  is the transpose conjugate of  $\mathbf{P}$ .

**Result:** If  $\mathbf{P}$  is a unitary matrix, then by definition  $\mathbf{P}^{\circ}\mathbf{P} = \mathbf{I} \Rightarrow |\mathbf{P}^{\circ}\mathbf{P}| = |\mathbf{I}| \Rightarrow |\mathbf{P}^{\circ}| |\mathbf{P}| = 1$ .

$\Rightarrow |\mathbf{P}| \neq 0$ . Hence,  $\mathbf{P}$  is invertible. Further,  $\mathbf{P}^{\circ}\mathbf{P} = \mathbf{I} \Rightarrow \mathbf{P}^{\circ} = \mathbf{P}^{-1}$  which in turn implies

that  $\mathbf{P} \mathbf{P}^{\circ} = \mathbf{I}$ . Finally we conclude that if  $\mathbf{P}$  is a unit matrix if and only if,  $\mathbf{P}^{\circ}\mathbf{P} = \mathbf{I} = \mathbf{P} \mathbf{P}^{\circ}$ . That is if  $\mathbf{P}^{\circ} = \mathbf{P}^{-1}$ .

**Definition 2: Orthogonal Matrix:** A square matrix  $\mathbf{P}$  with real elements is said to be orthogonal if and only if  $\mathbf{P}^T \mathbf{P} = \mathbf{I}$ .

**Result:** As in the case of unitary matrix, it can be easily seen that a real matrix  $\mathbf{P}$  is orthogonal if and only if  $\mathbf{P}^T \mathbf{P} = \mathbf{I} = \mathbf{P} \mathbf{P}^T$ . That is if and only if  $\mathbf{P}^T = \mathbf{P}^{-1}$ .

#### 3.3.1: Properties of Orthogonal and Unitary matrices:

- (1) A real matrix is unitary if and only if it is orthogonal.
- (2) If  $\mathbf{P}$  is an Unitary matrix, then  $\mathbf{P}^T, \mathbf{P}^{\circ}, \bar{\mathbf{P}}$  and  $\mathbf{P}^{-1}$  are all unitary matrices.
- (3) If  $\mathbf{P}$  and  $\mathbf{Q}$  are unitary matrices then  $\mathbf{PQ}$  is also unitary matrix.
- (4) If  $\mathbf{P}$  is unitary, then  $|\mathbf{P}|$  is of unit modulus.
- (5) Any two Eigen vectors corresponding to the distinct Eigen values of a unitary matrix are orthogonal.

Above properties holds good for orthogonal matrices also. That is:

- (1) If  $\mathbf{P}$  is orthogonal,  $\mathbf{P}^T$  and  $\mathbf{P}^{-1}$  are also orthogonal.
- (2) If  $\mathbf{P}$  and  $\mathbf{Q}$  are orthogonal matrices, then  $\mathbf{PQ}$  is also orthogonal matrix.
- (3) If  $\mathbf{P}$  is orthogonal, then  $|\mathbf{P}| = \pm 1$ .

### 3.4. Definition and properties of trace of a matrix and their properties

Now we proceed to discuss another important of matrix algebra, namely trace of a matrix, which is defined as follows:

**Definition 1: Trace of a matrix:** Let  $\mathbf{A}$  be a square matrix of order  $n$ . The sum of the elements of  $\mathbf{A}$  lying along the principal diagonal is called the trace of  $\mathbf{A}$  and is denoted by  $\text{tr}(\mathbf{A})$ .

Thus if  $\mathbf{A} = [a_{ij}]_{n \times n}$  then  $\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \dots + a_{nn}$ . (3.4.1)

**3.4.1. Properties of trace of a matrix:** Let  $\mathbf{A}$  and  $\mathbf{B}$  be two square matrices of order  $n$  and  $K$  be the scalar. Then the following properties are true:

- (1)  $\text{tr}(\mathbf{KA}) = K \text{tr}(\mathbf{A})$ .

$$(2) \operatorname{tr}(\mathbf{A}+\mathbf{B}) = \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B}).$$

$$(3) \operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA}).$$

$$(4) \operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{A}^T)$$

### 3.5. Summary

In this lesson some advanced topics of matrix when the elements of matrix are complex numbers are discussed. The concept of inner product is discussed first along with its properties. Then the concept of orthogonal vectors is discussed along with their properties. Then the concepts of unitary and orthogonal matrices were discussed some relating results and properties are discussed. Finally trace of a matrix and its properties were discussed. The concepts of orthogonal set, orthonormal set, orthogonal basis and orthonormal basis are also defined. Application of these concepts in design of experiments is discussed.

### 3.6. Self Assessment Questions

1. Define inner product of two vectors and discuss its properties.
2. Define orthogonal vectors and discuss their properties.
3. Define (i) orthogonal set, (ii) orthonormal set and (iii) Kronecker delta.
4. Define unitary matrix and discuss its properties.
5. Define Orthogonal matrix and discuss its properties.
6. If  $P$  is an unitary matrix, show that  $P^0 = P^{-1}$ .
7. If  $P$  is an orthogonal matrix show that  $P^T = P^{-1}$ .
8. Define trace of a matrix  $A$  and discuss its properties.
9. Stating the conditions, show that  $\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{A}^T)$
10. Define orthogonal basis and orthonormal basis of vectors.

### 3.7. Further Readings.

1. A.R. Vasishtha "Matrices" Krishna Prakashan Media (p) Ltd.; Thirty fourth Edn., Meerut – 250 001; 2003.
2. A. RamachandraRao and P. BhimaShankaram "Linear Algebra" Tata McGraw-Hill Publishing Company Limited; 1992.
3. D.D. Joshi "Theory of Linear Estimation" Woley Eastern Ltd. New Delhi. 2005.
4. Gupta and Kapoor "fundamentals of Applied Statistics" Sultan Chand & sons, New Delhi. 2008.



**Lesson – 4****Differentiation and Quadratic forms of Matrices****4.0 Objectives :**

After going through this lesson, you should be able to:

- Understand the concepts of differentiation using matrices and their applications.
- The concepts of quadratic forms, their properties and applications.
- The concepts of linear transformations and congruence of matrices.
- Orthogonal reduction of a real quadratic form and its properties.
- Statement of Cochran's theorem for quadratic forms and its Applications.

**Structure**

- 4.1. Introduction to differentiation using matrices and Jacobian of transformation
- 4.2. Fundamental concepts of quadratic forms and their properties
- 4.3. Classification of quadratic forms
- 4.4. Statement of Cochran's theorem and its applications
- 4.5. Summary
- 4.6. Self Assessment Questions
- 4.7. Further Readings

**4.1. Introduction to differentiation using matrices and Jacobian of transformation.**

There are numerous applications of matrix algebra in many fields particularly in Statistics where we deal many variables at a time. In such situations, to determine the joint distribution of these variables, we need to differentiate functions involving many variables. Hence, there is a need to study the concept of differentiation using matrices. In order to find the distribution of general functions of two random variables, we use the method of transformation of variables which is popularly known as "Jacobian transformation of variables", which is explained in the following section.

**4.1.1. Jacobian of Transformation of variables:**

Let us consider functions of two variables  $(X_1, X_2)$ , with the joint density function  $f_{x_1, x_2}(x_1, x_2)$ . Let  $Y_1 = g_1(X_1, X_2)$  and  $Y_2 = g_2(X_1, X_2)$  two functions of  $(X_1, X_2)$ . Now we want to find  $F_{Y_1, Y_2}(y_1, y_2)$ ,

Which is the joint distribution of  $(Y_1, Y_2)$ , which is given as follows:

$$F_{Y_1, Y_2}(y_1, y_2) = \Pr(Y_1 \leq y_1, Y_2 \leq y_2) = \iint_{R_{Y_1, Y_2}} f_{x_1, x_2}(x_1, x_2) dx_1 dx_2, \quad (4.1.1)$$

where,  $R_{Y_1, Y_2} = \{(X_1, X_2) : Y_1 \leq y_1, Y_2 \leq y_2\}$ .

Now, define:  $|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} \\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$  as the Jacobian of transformation.

To find the joint probability density function of  $(Y_1, Y_2)$ , solve the following simultaneous equations:  $g_1(x_1, x_2) = y_1$ ;  $g_2(x_1, x_2) = y_2$  (4.1.2)

Which gives  $(x_1, x_2)$  in terms of  $(y_1, y_2)$ . Let  $x_{1i} = g^{-1}_{1i}(y_1, y_2)$ ,  $x_{2i} = g^{-1}_{2i}(y_1, y_2)$ ,  $i=1, 2, \dots, m$

be  $m$  set of solutions which are all real quantities. Now calculate Jacobins as follows:

$$J_i = \begin{vmatrix} \frac{\partial g_{-11i}}{\partial y_1} & \frac{\partial g_{-12i}}{\partial y_1} \\ \frac{\partial g_{-11i}}{\partial y_2} & \frac{\partial g_{-12i}}{\partial y_2} \end{vmatrix}. \text{ Then the joint probability density function of } (Y_1, Y_2) \text{ is given by:}$$

$$f_{Y_1, Y_2}(y_1, y_2) = \sum_{i=1}^m f_{x_1, x_2}(x_{1i}, x_{2i}) |J_i|. \quad (4.1.3)$$

Here, we assume that the first partial derivatives  $g^{-1}_{1i}$ ,  $g^{-1}_{2i}$  are continuous functions on the domain of  $(Y_1, Y_2)$  and  $|J_i| \neq 0$  on the domain of  $(Y_1, Y_2)$   $i=1, 2, \dots, m$ . In any case, if for certain  $(y_1, y_2)$ , the equations (4.1.2) have no real solution, then  $f(y_1, y_2)$  is considered as zero. Similarly, if  $(x_{1i}, x_{2i}) = (x_1, x_2)$  for all  $i=1, 2, \dots, m$ , then  $f_{Y_1, Y_2}(y_1, y_2) = f(x_1, x_2) |J|$ . (4.1.4)

In particular, if we require only the distribution of  $Y_1$  is required,  $Y_2 = g_2(X_1, X_2)$  is suitably chosen, and the joint distribution of  $(Y_1, Y_2)$  is obtained and distribution of  $Y_1$  is found out by integrating over  $Y_2$ .

Above explained transformation of variables using Jacobian transformation is explained with the following example.

**Example (4.1.1):** If  $X_1$  and  $X_2$  are independent standard normal variables, then, show that

$(X_1^2 + X_2^2)$  is a chi-square variable, with 2 degree of freedom.

Solution: Consider the differential joint probability distribution function of  $X_1$  and  $X_2$  as follows:

$$dP(x_1, x_2) = f(x_1, x_2) dx_1 dx_2 = f_1(x_1) f_2(x_2) dx_1 dx_2. \quad (4.1.5)$$

$$= \frac{1}{2\pi} \exp \left\{ -\frac{(x_1^2 + x_2^2)}{2} \right\} dx_1 dx_2. \quad -\infty < (x_1, x_2) \leq \infty. \quad (4.1.6)$$

Let us now transform to polar co-ordinates by transforming the two variables by considering:

$x_1 = r \cos \theta$  and  $x_2 = r \sin \theta$ . Then the Jacobian of transformation  $|J|$  is given by:

$$|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} \\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{vmatrix} = r. \quad (4.1.7)$$

Further, we know that  $r^2 = x_1^2 + x_2^2$  and  $\tan \theta = x_2 / x_1$ . Further, as  $x_1$  and  $x_2$  varies from  $-\infty$  to  $\infty$ ,  $\theta$  varies from 0 to  $2\pi$ . Now the joint probability distribution of  $r$  and  $\theta$  will become:

$$dG(r, \theta) = \frac{1}{2\pi} \exp \left\{ -\frac{r^2}{2} \right\} r dr d\theta. \quad 0 \leq r \leq \infty \text{ and } 0 \leq \theta \leq 2\pi. \quad (4.1.8)$$

$$dG_1(r) = \int_0^{2\pi} dG(r, \theta) = r \exp(-r^2/2) dr \left[ \frac{\theta}{2\pi} \right]_0^{2\pi} = \exp(-r^2/2) r dr. \quad (4.1.9)$$

$$\text{This implies that: } dG_1(r^2) = \frac{1}{2} \exp(-r^2/2) dr^2. \quad (4.1.10)$$

$$= \frac{1}{\Gamma(1)} \exp(-r^2/2)^{1-1} d(r^2/2) \quad (4.1.11)$$

Thus,  $(r^2/2) = (X_1^2 + X_2^2)/2$  is  $\gamma(1)$ , that is Gamma variate with 1 degree of freedom. Further,  $r^2 = X_1^2 + X_2^2$  is a chi-square variate with 2 degrees of freedom.

Now we proceed to explain some fundamental concepts of quadratic forms of matrix algebra in the following section.

## 4.2. Fundamental concepts of quadratic forms and their properties.

**Definition(1) : Quadratic forms:** An expression of the form

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad (4.2.1)$$

where,  $a_{ij}$ 's are elements of a field F, is called a **quadratic form** in the n variables,  $x_1, x_2, \dots, x_n$  over a field F.

**Definition (2) :Real Quadratic forms :** An expression of the form

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad (4.2.2)$$

where,  $a_{ij}$ 's are all real numbers, is called a **real quadratic form** in the n variables,  $x_1, x_2, \dots, x_n$ . For example,  $4x^2 + 8xy - 6y^2$  is a real quadratic form in two variables x and y.

Similarly,  $2x^2 - 4y^2 - 6z^2 + 2yz - 4zx + 8xy$  is another real quadratic form in 3-variables. On similar lines, we can give examples of 4-variable or more than 4-variable quadratic forms.

**Definition (3) : Matrix of a quadratic form:** If  $\Phi = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$  is a quadratic form in n-variables,  $x_1, x_2, \dots, x_n$ , then there exists a unique symmetric matrix **B** of order n such that  $\Phi = \mathbf{X}^T \mathbf{B} \mathbf{X}$  where  $\mathbf{X} = [x_1, x_2, \dots, x_n]^T$ . The symmetric matrix **B** is known as "the matrix of the quadratic form"  $\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$ .

**Definition (4): Rank of a quadratic form:** Let  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  be a quadratic form over a field F. The rank of the matrix **A** is called the rank of the quadratic form  $\mathbf{X}^T \mathbf{A} \mathbf{X}$ .

Since every quadratic form can always be so written that matrix of its coefficients is a symmetric matrix. Therefore, we shall be considering quadratic forms, which are so adjusted that the coefficient matrix is symmetric.

**Example (4.2.1):** Write down the matrix of the following quadratic form in 3-variables:

$$x_1^2 + 2x_2^2 - 5x_3^2 - x_1x_2 + 5x_2x_3 - 3x_1x_3. \quad (4.2.3)$$

Solution: Let B be the matrix of the given quadratic form. Then, we have:

$$\mathbf{B} = \begin{bmatrix} 1 & -1/2 & -3/2 \\ -1/2 & 2 & 5/2 \\ -3/2 & 5/2 & -5 \end{bmatrix} \text{ and } \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \text{ Then } \mathbf{X}^T = [x_1 \quad x_2 \quad x_3].$$

Now we can observe that  $\mathbf{X}^T \mathbf{B} \mathbf{X}$  will give the quadratic form given in equation (4.2.3).

**Example (4.2.2):** Write down the quadratic form for the matrix  $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 5 \\ 3 & 5 & -5 \end{bmatrix}$ .

**Solution:** Let  $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ . Then  $\mathbf{X}^T = [x_1 \quad x_2 \quad x_3]$ .

$$\text{Then, } \mathbf{X}^T \mathbf{A} \mathbf{X} = [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 5 \\ 3 & 5 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1^2 - 5x_3^2 + 4x_1x_2 + 6x_1x_3 + 10x_2x_3.$$

#### 4.2.1. Properties of quadratic forms:

1. Every quadratic form over a field  $F$  in  $n$ -variables  $x_1, x_2, \dots, x_n$  can be expressed in the form  $\mathbf{X}^T \mathbf{B} \mathbf{X}$

Where,  $\mathbf{X} = [x_1, x_2, \dots, x_n]^T$  is a column vector and  $\mathbf{B}$  is a symmetric matrix of order  $n$  over the field  $F$ .

2. Let  $\mathbf{A} = [a_{ij}]_{n \times n}$  be a symmetric matrix over the field  $F$ . and let  $\mathbf{X} = [x_1, x_2, \dots, x_n]^T$  be a column vector. Then,  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  determines a unique quadratic form  $\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$  in  $n$ -variables  $x_1, x_2, \dots, x_n$  over a field  $F$ .

3. If  $\mathbf{A}$  be any  $n$ -rowed non-zero symmetric matrix of rank ' $r$ ', over field  $F$ , then, there exists an  $n$ -rowed non-singular matrix  $\mathbf{P}$  over a field  $F$  such that:

$$\mathbf{P}^T \mathbf{A} \mathbf{P} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$
 where,  $\mathbf{A}_1$  is a non-singular diagonal matrix of order ' $r$ ', over  $F$  and each  $\mathbf{0}$  is a null matrix of suitable order.

4. Every symmetric matrix of rank ' $r$ ' equivalent to a diagonal matrix, ' $r$ ', of whose diagonal elements only are non-zero.

### 4.3. Classification of quadratic forms.

In the last section we have seen that there exists one-to-one correspondence between the set of all quadratic forms in  $n$ -variables over a field  $F$  and the set of all  $n$ -rowed symmetric matrix. Before we proceed to discuss the classification of quadratic forms, now we define canonical form or normal form, Signature and Index of a real quadratic form as follows:

**Definition (1): Canonical form or Normal form of a real quadratic form:** If  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  is a real quadratic form in  $n$ -variables. Then there exists a real non singular linear transformation  $\mathbf{X} = \mathbf{P} \mathbf{Y}$  which transforms  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  to the form:

$$\mathbf{Y}^T \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{Y} = y_1^2 + y_2^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2 \quad (4.3.1)$$

**Definition (2): Signature and Index of a real quadratic form:** Let,  $y_1^2 + y_2^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2$  be the normal form of the real quadratic form  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  of rank ' $r$ '. The number ' $p$ ' of the positive terms in a normal form of  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  is called '**Index**' of the quadratic form. The excess of the number of positive terms in a normal form of a quadratic form  $\mathbf{X}^T \mathbf{A} \mathbf{X}$ , that is  $p - (r - p) = 2p - r$  is called the '**Signature**' of the quadratic form and is usually denoted by the letter ' $s$ '.

**Results:**

1. Two real quadratic forms in  $n$ -variables are real equivalent if and only if they have the same **rank** and **index** (or **signature**).

2. **Reduction of a real quadratic form in the complex field:** If  $\mathbf{A}$  be any  $n$ -rowed real symmetric matrix of rank  $r$ , then, there exists a non-singular matrix  $\mathbf{P}$  whose elements may be any complex numbers such that  $\mathbf{P}^T \mathbf{A} \mathbf{P} = \text{diag}[1, 1, \dots, 1, 0, \dots, 0]$ , where 1 appears  $r$  times.

3. **Orthogonal reduction of a real quadratic form:** if  $\Phi = \mathbf{X}^T \mathbf{A} \mathbf{X}$  be a real quadratic form of rank  $r$ , in  $n$ -variables, then there exists a real orthogonal transformation  $\mathbf{X} = \mathbf{P} \mathbf{Y}$  which transforms  $\Phi$  to the form:  $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_r y_r^2$ , where,  $\lambda_1, \lambda_2, \dots, \lambda_r$  are the  $r$  non-zero Eigen values of  $\mathbf{A}$ , and  $n - r$  Eigen values of  $\mathbf{A}$ , being equal to zero.

3. Every real quadratic form  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  in  $n$ -variables is real equivalent to the form:

$$y_1^2 + y_2^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2.$$

Let  $\Phi = \mathbf{X}^T \mathbf{A} \mathbf{X}$  is a real quadratic form in  $n$ -variables  $x_1, x_2, \dots, x_n$ . There are basically three types of forms of these real quadratic forms. Namely: (1) Definite, (2) Semi-definite and (3) Indefinite forms. Definite forms are further classified in to two categories, namely (i) Positive definite forms and (ii) Negative definite forms. Similarly, Semi-definite forms are also two types, namely, (i) Positive semi-definite form and (ii) Negative semi-definite form. These forms are defined as follows:

**Definition (3): Positive Definite form:** If  $\Phi \geq 0$  for all real real values of the variables,  $x_1, x_2, \dots, x_n$  and  $\Phi = 0$  only if  $\mathbf{X} = \mathbf{0}$ . That is  $x_1 = x_2 = \dots = x_n = 0$ .

For example,  $x_1^2 + x_2^2 + x_3^2$  in three variables is a positive definite form.

**Definition (4) : Negative definite form :** If  $\Phi \leq 0$  for all real real values of the variables,  $x_1, x_2, \dots, x_n$  and  $\Phi = 0$  only if  $\mathbf{X} = \mathbf{0}$ . That is  $x_1 = x_2 = \dots = x_n = 0$ .

For example,  $-x_1^2 - x_2^2 - x_3^2$  in three variables is a negative definite form.

**Definition (5) : Positive semi-definite form:** If  $\Phi \geq 0$  for all real real values of the variables,  $x_1, x_2, \dots, x_n$  and  $\Phi = 0$  for some non-zero real vector  $\mathbf{X}$ , that is  $\Phi = 0$  for some real values of the variables  $x_1, x_2, \dots, x_n$  not all zero.

For example,  $x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_3 - 2x_2x_3$ . Is positive semi definite form because it can be written in the form  $(x_1 - x_3)^2 + (x_2 - x_3)^2$  which is  $\geq 0$  for all real values of  $x_1, x_2, x_3$  but is zero for non zero values also. For example,  $x_1 = x_2 = x_3 = 1$ .

**Definition (6) : Negative semi-definite form:** If  $\Phi \leq 0$  for all real values of the variables,  $x_1, x_2, \dots, x_n$  and  $\Phi = 0$  for some values of the variables  $x_1, x_2, \dots, x_n$  not all zero.

For example, the quadratic form  $-x_1^2 - x_2^2 - 0x_3^2$  in three variables  $x_1, x_2, x_3$  is negative semi-definite form.

**Definition (7) : Indefinite form:** If  $\Phi$  takes positive as well as negative values for real values of the variables  $x_1, x_2, \dots, x_n$  then, we say that the form of  $\Phi$  is indefinite form.

For example,  $x_1^2 - x_2^2 + x_3^2$  in three variables  $x_1, x_2, x_3$  is indefinite form because it takes  $+1$  if

$x_1 = 1, x_2 = 1$  and  $x_3 = 1$  and it takes -1 if  $x_1 = 0, x_2 = 1$  and  $x_3 = 0$ .

It is important to note that the above five forms of quadratic forms are mutually exclusive and are called **as value classes of real quadratic forms**. Every real quadratic form must belong to one and only one value class.

**Result:** A real quadratic form  $\Phi = \mathbf{X}^T \mathbf{A} \mathbf{X}$  in  $n$ -variables is:

- (1) Positive definite if and only if all the Eigen values of  $\mathbf{A}$  are positive.
- (2) Negative-definite if and only if all the Eigen values of  $\mathbf{A}$  are negative.
- (3) Positive semi-definite if and only if all the Eigen values of  $\mathbf{A}$  are  $\geq 0$  and at least one Eigen value of  $\mathbf{A}$  is 0.
- (4) Negative semi-definite if and only if all the Eigen values of  $\mathbf{A}$  are  $\leq 0$  and at least one Eigen value of  $\mathbf{A}$  is 0.
- (5) Indefinite if and only if  $\mathbf{A}$  has positive as well as negative Eigen values.

**Definition (8):** A real symmetric matrix  $\mathbf{A}$  is said to be definite, semi-definite or indefinite if the corresponding quadratic form  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  is definite, semi-definite or indefinite respectively.

On similar lines, we conclude that the symmetric matrix  $\mathbf{A}$ , is positive definite if the corresponding form  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  is positive definite and is negative definite if  $\mathbf{X}^T \mathbf{A} \mathbf{X}$  is negative definite.

#### 4.4. Statement of Cochran's theorem and its applications

Now we proceed to discuss an important and most widely used theorem in design of experiments due to Cochran (1934) along with its applications.

**Statement of Cochran's Theorem:** If  $\mathbf{q}_i = \sum_{\alpha=1}^N (a_i) \alpha \beta \quad y \alpha y \beta^i = 1, 2, \dots, m$ , and is of rank ' $r_i$ ' and  $\sum_{i=1}^m q_i = \sum_{\alpha=1}^N (y)_{\alpha}^2$  then, a necessary and sufficient condition that there exists an orthogonal transformation of  $\{y_{\alpha}\}$  to  $\{z_{\alpha}\}$  such that:

$$q_i = \sum_{\alpha=r_1+r_2+\dots+r_{i-1}+1}^{r_1+r_2+\dots+r_i} z_{\alpha}^2, \quad \text{is that } r_1+r_2+\dots+r_m = N.$$

**Applications of Cochran's theorem:** Cochran's theorem is useful in proving that certain "Vector quadratic forms" are distributed as sums of "Vector squares". This type of result has very important application in the "Analysis of Variance" (ANOVA) which is basically used in the analysis of design of experiments, and is discussed in the forthcoming units. It is important to note that Cochran's theorem ensures that the transformation is orthogonal that is,

$\mathbf{I} = \mathbf{B}^T \mathbf{I} \mathbf{B} = \mathbf{B}^T \mathbf{B}$ . An alternative statement of Cochran's theorem is that:

Let the rank of the  $N$ -order square matrix  $\mathbf{A}_i$  be  $r_i$  ( $i=1, 2, \dots, q$ ) and suppose,  $\sum_{i=1}^q \mathbf{A}_i = \mathbf{I}$ .

A necessary and sufficient condition that there exists an orthogonal matrix  $\mathbf{B} = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix}$  such that

$A_i = B_i^T B_i$  is that  $\sum_{i=1}^m r_i = N$ . More applications are discussed in detail in the forthcoming units, while discussing ANOVA and Design of experiments.

#### 4.5. Summary

In this lesson first we have discussed the differentiation using matrices particularly Jacobian of transformations in obtaining the joint distribution of function of two variables. Then, we have introduced the concept of quadratic forms their types along with properties of quadratic forms are discussed. Classification of real quadratic forms are discussed the relation between quadratic forms and ranks, are discussed. Finally, Statement of Cochran's theorem and its applications are discussed.

#### 4.6. Self Assessment Questions

1. Explain the need of differentiation using matrices and also define the Jacobian of transformation.
2. Define quadratic forms and discuss their classification.
3. What is a real quadratic form? Also define the matrix of a quadratic form.
4. Obtain the matrices corresponding to the following quadratic forms:

(i)  $x_1^2 - 2x_2^2 + 4x_3^2 - 4x_1x_2 + 6x_1x_3 + 4x_2x_3$ .

(ii)  $2x_1^2 + 3x_2^2 + 4x_3^2 + 2x_1x_2 + 10x_1x_3 - 4x_2x_3$ .

[ answers: (i)  $\begin{bmatrix} 1 & -2 & 3 \\ -2 & -2 & 2 \\ 3 & 2 & 4 \end{bmatrix}$ . (ii)  $\begin{bmatrix} 2 & 2 & 5 \\ 2 & 3 & -2 \\ 5 & -2 & 4 \end{bmatrix}$  ]

4. Write down the quadratic forms of the following matrices:

$A = \begin{bmatrix} 2 & 1 & 5 \\ 1 & 3 & -2 \\ 5 & -2 & 4 \end{bmatrix}$   $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -7 \end{bmatrix}$

{[Answers:  $A = 2x_1^2 + 3x_2^2 + 4x_3^2 + 2x_1x_2 + 10x_1x_3 - 4x_2x_3$ .  $B = (i) x_1^2 + 5x_2^2 - 7x_3^2$ ]

5. Define various forms of real quadratic forms.
6. Discuss various properties of real quadratic forms.
7. Define canonical form of quadratic forms. Also define index and signature of quadratic forms.
8. Define orthogonal reduction of a real quadratic form and discuss its properties.
9. State Cochran's theorem and discuss its applications.
10. Define (i) Canonical form of a Real quadratic form  
(ii) Signature and Index of a xd quadratic form  
(iii) Positive definite & positive be semi – definite form  
(iv) Negative definite & Negative semi definite form

**4.7. Further Readings.**

1. A.R. Vasishtha "Matrices" Krishna Prakashan Media (p) Ltd.; Thirty fourth Edn.,Meerut – 250 001; 2003.
2. A. RamachandraRao and P. BhimaShankaram "Linear Algebra" Tata McGraw-Hill Publishing Company Limited; 1992.
3. D.D. Joshi "Theory of Linear Estimation" Woley Eastern Ltd. New Delhi. 2005.
4. Gupta and Kapoor "fundamentals of Applied Statistics" Sultan Chand & sons, New Delhi. 2008.



**Lesson – 5.****Fundamental Concepts of Linear Models****5.0 Objectives :**

After going through this lesson, you should be able to:

- Understand the need, concepts of linear models, definitions and applications.
- The concepts of linear estimation and properties.
- The concepts of best linear estimation.
- Calculation of linear parametric function and its properties.
- Applications of linear unbiased estimates.

**Structure**

- 5.1 Introduction to Linear Models**
- 5.2 Fundamental concepts and definitions of Linear Models**
- 5.3. Estimation of a linear parametric function and its properties**
- 5.4. Best Linear Un-biased estimates and properties**
- 5.5 Summary**
- 5.6. Self Assessment Questions**
- 5.7. Further Readings**

**5.1. Introduction to Linear Models**

Linear models have lot of applications in many fields of real life. For example, Linear Programming Problems (LPP) in Operations research, Analysis of Variance (ANOVA) in Agricultural Research, Linear regression Models in Regression Analysis or Time series and so on. First of all, a “**Model**” is a replica of some real life situation, which reacts like a real one but not real. For example, to learn piloting, it is not possible to try directly in Airplane, but to be practiced on a similar system like an airplane and not a real one. Such a system which is like a real airplane but not real, and should serve the purpose of practice the piloting by the beginner. Such a system which is used by the beginner is called a “**Model**”. In particular, this is an “**Iconic Model**”. Similarly, if we represent the real system using mathematical equation by using variables, then it is called the “**Mathematical Model**”. For example regression models representing relation between rainfall and yield of a crop or relation between intake of food and calories of energy gained, number of hours of study by a student and the marks obtained by him in an examination, and so on. There are many types of models and basic purpose of any model is the amicability of the model for further analysis. It is very difficult to tackle the real life situation. First we convert the situation to a model and for the model, one can get solutions. The obtained solution of the model is to be applied to the real situation, so that the problem of the real situation can be solved. Thus to obtain the solution to the problem of a real situation, we use the suitable model and obtain the solution to the model, which in turn can be applied in real situation so that the existing problem can be solved. This is the

procedure to be adopted in many research problems to solve many practical problems, in Agriculture, Medicine, Marketing, meteorology, predictions of the future behavior of a random variable.

On similar lines, to obtain the solutions for the problems relating to agriculture, we have to frame the suitable model, after forming the model, we can discuss about the method of obtaining solution to the model. Thus first, we are discussing the linear model along with its fundamental concepts in the following section.

**5.2. Fundamental concepts and definitions of Linear Models**

First we define the “Linear Model”, its parameters, “Linear Parametric Function” and the method of estimation of parameters of the linear model. We also discuss the properties of these estimates and related results in this section.

**Definition (1): Linear Model or Gauss-Markov Setup:** Let  $Y_1, Y_2, \dots, Y_n$  be n-random variables or observations collected from a real situation. Let  $\beta_1, \beta_2, \dots, \beta_m$  be ‘m’ unknown parameters.

Let  $X_{ij}, i=1,2,\dots,n; j=1,2,\dots,m$  be mn known variables. Let:

$$\begin{aligned}
 E(y_1) &= X_{11} \beta_1 + X_{12} \beta_2 \dots, X_{1m} \beta_m. \\
 E(y_2) &= X_{21} \beta_1 + X_{22} \beta_2 \dots, X_{2m} \beta_m. \\
 &\dots\dots\dots \\
 &\dots\dots\dots \\
 &\dots\dots\dots \\
 E(y_n) &= X_{n1} \beta_1 + X_{n2} \beta_2 \dots, X_{nm} \beta_m. \tag{5.2.1}
 \end{aligned}$$

Expectations of the random variables  $Y_1, Y_2, \dots, Y_n$  given in equations (5.2.1) are linear functions of the un-known parameters  $\beta_1, \beta_2, \dots, \beta_m$ . The set of equations given in (5.2.1) is called “Gauss-Markov setup” or “Linear Model” and the assumptions of linear model as follows:

**Assumptions of Gauss-Markov Setup:**

- (1)  $V(Y_i) = \sigma^2$ , for  $i=1,2,\dots,n$ . Here we assume that the variance  $\sigma^2$  is an un-known parameter.
- (2)  $Cov(Y_i Y_j) = 0, \forall i \neq j$ . That is any pair of variables  $Y_i, Y_j \forall i \neq j$  are un-correlated or independent.
- (3) Random variables  $Y_i$  are Normally distributed with mean 0 and variance  $\sigma^2$  for all values of  $i = 1,2,\dots,n$ . That is  $Y_i \sim N(0, \sigma^2)$ .

Alternatively, a linear model can also be written as:

$$\begin{aligned}
 y_1 &= X_{11} \beta_1 + X_{12} \beta_2 \dots, X_{1m} \beta_m + \epsilon_1. \\
 y_2 &= X_{21} \beta_1 + X_{22} \beta_2 \dots, X_{2m} \beta_m + \epsilon_2. \\
 &\dots\dots\dots \\
 &\dots\dots\dots \\
 &\dots\dots\dots \\
 y_n &= X_{n1} \beta_1 + X_{n2} \beta_2 \dots, X_{nm} \beta_m + \epsilon_n. \tag{5.2.2}
 \end{aligned}$$

with the assumptions that:

- (1)  $E(\epsilon_i) = 0, \forall i$  and  $\epsilon_i$ 's  $\sim N(0, \sigma^2)$ .
- (2)  $V(Y_i) = \sigma^2, \text{ for } i=1,2,\dots,n$ . Or  $V(\epsilon_i) = \sigma^2, \text{ for } i=1,2,\dots,n$ .
- (3)  $\text{Cov}(Y_i Y_j) = 0, \forall i \neq j$ . Or  $\text{Cov}(\epsilon_i \epsilon_j) = 0, \forall i \neq j$ .

Any set of linear equations with the above assumptions is known as “**Gauss Markov Set-up**”. Above defined linear model in equations (5.2.2) can be expressed in matrix notation as follows:

**Definition(2): Linear Model or Gauss Markov model in Matrix notations:**

Let  $\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$  be a column vector of n observations,  $\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}$  be another column vector

of m un-known parameters and  $\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$  be another column vector of n errors.

Let  $\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} = [x_{ij}]_{n \times m}$   $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, m$ , is known as

“**Coefficient Matrix**”. Then the set of linear equations given in (5.2.2) can be written in the

matrix notation as  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ .  $E(\boldsymbol{\epsilon}) = \mathbf{0}$  (zero vector).  $D(\boldsymbol{\epsilon}) = D(\mathbf{Y}) = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix} = \sigma^2$

$\mathbf{I}_{n \times n} \sigma^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \dots & \dots & 1 \end{bmatrix} \sigma^2$  is an nxn matrix known as “**Dispersion Matrix**”.

Let  $\sigma^2 > 0$  is known as Variance of  $Y_i$  or  $\epsilon_i$ .

Above linear model or gauss Markov model or Gauss Markov Set-up is denoted by the notation  $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$ .

In general if  $\mathbf{I} = \boldsymbol{\Sigma}$  is any nxn is a positive-definite matrix then the model is known as “**General Linear Model**”. The general linear model is denoted by  $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{\Sigma})$ . Thus Gauss Markov model is a particular case of General Linear model when  $\boldsymbol{\Sigma} = \mathbf{I}$ .

**Definition(3): Linear Parametric Function:**

Consider a Gauss Markov Set-up  $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$ . That is  $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$  with  $D(\mathbf{Y}) = \mathbf{I}\sigma^2$  with the

parameters  $\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}$ . Any linear function of the parameters  $\mathbf{P}^T \boldsymbol{\beta} = p_1 \beta_1 + p_2 \beta_2 + \dots + p_m \beta_m$  is

called a “**Linear Parametric Function**” (l.p.f). Here,  $\mathbf{P} = [p_1, p_2, \dots, p_m]$  is a row vector of real numbers.

Linear functions are two types namely: (1) Estimable linear parametric functions and (2) Non-estimable linear parametric functions. Now we proceed to discuss definitions, properties of these functions in the following section.

### 5.3. Estimation of a linear parametric function and its properties

Now, we proceed to define the “**Linear Unbiased Estimate**” as follows:

#### Definition (1): Linear Unbiased Estimate:

A linear function of observations  $y_1, y_2, \dots, y_n$  say  $\sum_{i=1}^n b_i y_i$  is said to be linear unbiased estimate of  $P^T \beta$  if  $E[\sum_{i=1}^n b_i y_i] = \sum_{i=1}^n b_i \beta_i = P^T \beta$ .

In other words, if  $B = [b_1, b_2, \dots, b_n]^T$  and  $\sum_{i=1}^n b_i y_i = B^T Y$ ,  $B^T Y$  is said to be linear unbiased estimate of  $P^T \beta$  if  $E(B^T Y) = P^T \beta$ . This means that a linear function of observations  $B^T Y$  is linear unbiased estimate of the linear parametric function  $P^T \beta$  if  $E(B^T Y) = P^T \beta$ .

A linear parametric function is said to be **estimable**, in the linear model  $(Y, X\beta, I\sigma^2)$  if there exists at least one linear unbiased estimate of  $P^T \beta$ .

**Theorem (5.3.1):** A necessary and sufficient condition for a linear parametric function to be estimable in a linear model  $(Y, X\beta, I\sigma^2)$  is that  $\text{rank} \left[ \begin{matrix} X \\ P \end{matrix} \right] = \text{Rank}(X)$ .

Proof: Consider a linear model  $E(Y) = X\beta$  with  $D(Y) = I\sigma^2$ .

By definition of estimability of a l.p.f.,  $P^T \beta$  is estimable, if there exists a  $B$  such that  $E(B^T Y) = P^T \beta$ . This implies that  $B^T E(Y) = P^T \beta$ . That is  $B^T X \beta = P^T \beta$ .

That is  $B^T X = P^T$  or  $X^T B = P$ .

That is  $P^T \beta$  is estimable, if and only if (iff) there exists a  $B$  such that  $X^T B = P$ .

Therefore,  $P^T \beta$  is estimable, implies that the simultaneous equations  $X^T B = P$  is solvable in  $B$ .

That is  $P^T \beta$  is estimable, iff  $\text{rank}[X:P] = \text{rank}[X]$  or  $\text{rank} \left[ \begin{matrix} X \\ P \end{matrix} \right] = \text{Rank}(X)$ .

Hence the proof.

**Result(1):** Let  $E(Y) = X\beta$  and  $D(Y) = I\sigma^2$ . The l.p.f.  $P^T \beta$  is estimable, iff  $P^T$  is dependent on row vectors of  $X$ .

**Remark:** It is important to note that if  $P$  is such that  $P^T$  is not linearly dependent on row vectors of  $X$  than such  $P^T \beta$  cannot be estimable and hence is known as “**Non-estimable Linear Parametric Function**”.

**Result (2):** In any Gauss Markov set sup  $E(Y) = X\beta$  with  $D(Y) = I\sigma^2$  and if rank of the coefficient matrix  $X$  is full that is  $\text{rank}(X) = m$ ,  $m \leq n$ , the number of observations, then, every linear parametric function (l.p.f.)  $P^T \beta$ , for any  $P^T$  is estimable.

From the above result we conclude that in particular, each one of the parameter in a Gauss Markov set sup  $E(Y) = X\beta$  with  $D(Y) = I\sigma^2$  are estimable, that is  $\beta_1$  ( $1\beta_1 + 0\beta_2 + \dots + 0\beta_m$ ),  $\beta_2$  ( $0\beta_1 + 1\beta_2 + \dots + 0\beta_m$ ),  $\dots$ ,  $\beta_m$  ( $0\beta_1 + 0\beta_2 + \dots + 1\beta_m$ ) are all individually estimable, if and only if (iff) the rank of  $X$  is **full**. If the rank of  $X$  is not full, then every parameter cannot be estimable individually but, the sum of them may be estimable. This is explained with the following examples.

**Example (5.3.1):** Consider a two parametric model with parameters  $\omega_1$  and  $\omega_2$  such that:

$$Y_1 = \omega_1 + \varepsilon_1$$

$$Y_2 = \omega_2 + \varepsilon_2$$

$$Y_3 = \omega_1 + \omega_2 + \varepsilon_3$$

$$Y_4 = \omega_1 - \omega_2 + \varepsilon_4.$$

Represent the above model in the Gauss Markov Set up- and comment on the estimability of the linear parametric function  $\rho \omega_1 + \rho \omega_2$ .

Solution: Let  $\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$ ,  $\mathbf{X} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}$  and  $\boldsymbol{\beta} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$  and  $\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$  thus we have  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ .

Since the rank of the coefficient matrix is 2, that is the rank is full, every linear parametric function  $\rho \omega_1 + \rho \omega_2$  is always individually estimable.

**Example (5.3.2):** Consider a three parametric model with parameters  $\omega_1, \omega_2$  and  $\omega_3$  such that:

$$Y_1 = \omega_1 + \omega_2 + \omega_3 + \varepsilon_1$$

$$Y_2 = \omega_1 - \omega_2 + \omega_3 + \varepsilon_2$$

$$Y_3 = \omega_1 + \varepsilon_3$$

Represent the above model in the Gauss Markov Set up- and comment on the estimability of the linear parametric function  $\rho \omega_1 + \rho \omega_2$ .

Solution: Let  $\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ ,  $\mathbf{X} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$  and  $\boldsymbol{\beta} = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}$  and  $\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$ .

Thus we have  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ .

Since the rank of the coefficient matrix is 2, that is the rank is not full. Therefore, the

Parameters  $\omega_1, \omega_2$  and  $\omega_3$  are not individually estimable. Now we proceed to discuss another important result which states that there will be only one or infinitely many linear unbiased estimates.

**Result (3):** The number of linear unbiased estimates (l.u.e.s) of a linear parametric function (l.p.f) is either only one or infinitely many.

Proof: Let there are two l.u.e.s, say  $\mathbf{B}^T \mathbf{Y}$  and  $\mathbf{C}^T \mathbf{Y}$  for an l.p.f.  $\mathbf{P}^T \boldsymbol{\beta}$ . That is  $\mathbf{B}^T \mathbf{Y} = \sum_{i=1}^n b_i y_i$ ,  $\mathbf{C}^T \mathbf{Y} = \sum_{i=1}^n c_i y_i$  and  $\mathbf{P}^T \boldsymbol{\beta} = \sum_{i=1}^n p_i \beta_i$ . Further, by the definition of l.u.e.s we have:

$$E[\sum_{i=1}^n b_i y_i] = \sum_{i=1}^n p_i \beta_i \text{ and } E[\sum_{i=1}^n c_i y_i] = \sum_{i=1}^n p_i \beta_i.$$

Now consider a linear combination  $g\mathbf{B}^T \mathbf{Y} + (1-g)\mathbf{C}^T \mathbf{Y}$  for  $0 < g < 1$ . Then we have:

$$E[g\mathbf{B}^T \mathbf{Y} + (1-g)\mathbf{C}^T \mathbf{Y}] = E[g\sum_{i=1}^n b_i y_i + (1-g)\sum_{i=1}^n c_i y_i] = gE[\sum_{i=1}^n b_i y_i] + (1-g)E[\sum_{i=1}^n c_i y_i]$$

$= gP^T\beta + (1-g)P^T\beta = P^T\beta = \sum_{i=1}^n p_i\beta_i$ . Therefore  $E[gB^TY + (1-g)C^TY] = P^T\beta$ .

Hence,  $gB^TY + (1-g)C^TY$  is also another linearly unbiased estimate of  $P^T\beta$ . Since,  $g$  can take infinitely many values in the interval  $(0,1)$ ,  $P^T\beta$  there exists infinite number of l.u.e.s for  $P^T\beta$ .

Therefore, there will be only one or infinitely many l.u.e.s for a l.p.f.  $P^T\beta$ .

Hence the proof.

#### 5.4. Best Linear Un-biased Estimates and properties

From the result (3) discussed in section 5.3, there will be only one or infinitely many l.u.e.s for an l.p.f.  $P^T\beta$ . If there is only one l.u.e. then, that itself is the “**best linear un-biased estimate**”(BLUE) for that l.p.f.  $P^T\beta$ . If there are many, among the class of all linear un-biased estimators, of an estimable linear parametric function, that linear unbiased estimator which has minimum variance is known as “**Best Linear Un-biased Estimates**” (BLUE). Thus BLUEs are unbiased estimates as well as the variance of the estimates are smallest when compared to other estimates. Thus BLUEs are most popularly used in linear estimation of a linear parametric function. But, finding BLUEs form a large group of estimates can be done easily through a short-cut method discussed in a theorem known as “Gauss Markov’s Theorem” which is proved in the next lesson.

#### 5.5 Summary.

In this lesson, first a linear model is defined and discussed its assumptions and parameters. Then representation of a linear model in matrix notation is discussed along with Gauss Markov set up or Gauss Markov Model. Then the concept of linear parametric function and its estimation is discussed and proved the necessary and sufficient condition for estimability of a linear parametric function. The concept of linear unbiased estimates (l.u.e.s) their properties and results relating to l.u.e. are discussed. Finally, the concept of Best Linear Unbiased Estimates (BLUEs) their properties are discussed.

#### 5.6. Self Assessment Questions

1. Define a linear model and discuss the assumptions and parameters of a linear model.
2. What is meant by a Gauss Markov Set up? Discuss its assumptions.
3. Explain the Gauss Markov model with an example.
4. Define the linear parametric function and Dispersion matrix.
5. Explain the method of representing the linear model in matrix notation.
6. Define the concept of estimability of a linear parametric function and state the conditions under which the parameters of a linear model are estimated individually .
7. State and prove the necessary and sufficient conditions for estimability of a linear parametric function.

8. Show that the number of l.u.e.'s of a estimable l.p.f. is either one or infinitely many.
9. Define BLUEs and discuss their properties.
10. Explain the problems in determining BLUEs for a given linear model.

### **5.7. Further Readings**

1. P. RamachandraRao and P. Bhimasankaram (1996) "Linear Algebra" Tata McGraw-Hill Publishing Company Limited, New Delhi.
2. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons., New Delhi.
3. D.D. Joshi "Theory of Linear Estimation" Wiley Eastern Ltd. New Delhi. 2005.
4. C. R. Rao (1965) "Linear Statistical inference and its applications" John Wiley, New Delhi.

## Lesson – 6

# Gauss Markov set up and Gauss Markov Theorem

### 6.0 Objectives :

After going through this lesson, you should be able to:

- Understand the problems and method of obtaining BLUEs.
- The concept and importance of Gauss Markov theorem.
- The Statement and proof of Gauss Markov theorem.
- Calculation of BLUEs and some examples of obtaining BLUEs.
- Applications of Gauss Markov theorem.

### Structure

- 6.1. Importance of Gauss Markov set up
- 6.2 Gauss Markov Theorem and its special cases
- 6.3 Estimation of Variance of BLUEs
- 6.4. Estimation of  $\sigma^2$  in Gauss Markov set up
- 6.5 Summary
- 6.6. Self Assessment Questions
- 6.7. Further Readings

### 6.1. Importance of Gauss Markov Set up

**Gauss Markov Setup** or **Gauss Markov Model** explained in the last lesson play predominant role in the Analysis Of Variance (ANOVA) and Design of experiments particularly in estimating various effects of different factors in many real life problems. In all these problems, we usually want to study the effects of different factors on the main variable. For example, in agricultural problems, yield of the crop is considered as the main variable and rainfall, manure, type of soil, seed quality, type of water, fertilizers used; are different factors influencing the main variable. Similarly, in industry, there are various factors influencing the quality of the product which is the main variable. The raw material used, design, production process, worker's skills producing the product are the various factors influencing the main variable. In these problems, we want to study significant effects of various factors on the main variable. In all these problems, we assume that the Gauss Markov model is the best suitable model.

Further, we want to estimate the Linear Parametric function, which is the function of various effects which have linear effects on the main variable 'Y'. These effects are to be estimated and these estimates must have basic properties like Un-biasedness, efficient, sufficient and minimum variance. Such estimates are called '**Best**' estimates. In the last lesson, we have introduced such estimates, known as "**Best Linear Unbiased Estimates**" (BLUEs). In this lesson, we discuss the method of obtaining BLUEs for a Gauss Markov Model or Gauss Markov Setup or Standard Gauss Markov's setup.



## 6.2. Gauss Markov Theorem and its special cases

We know that in a Gauss Markov setup, linear unbiased estimates of linear parametric function (l.p.f)  $P^T\beta$  is either one or infinitely many in number. If there is one, we consider that itself as **Best Linear Unbiased estimate** (BLUE) for the l.p.f.  $P^T\beta$ . If, there are two, then we have proved in the lesson 5 that there will be infinitely many in number, and selecting the best one having minimum variance will become big problem, because, there are infinitely many in number. Thus searching the Best Linear Unbiased from all the available ones will become practically impossible. At this juncture, **Gauss Markov theorem** will help us to identify the BLUE if there are many. This theorem provides a systematic procedure to obtain the BLUE for the given l.p.f.  $P^T\beta$ . Now we proceed to state and prove the Gauss Markov theorem as follows:

**Statement:** In a Gauss Markov's set up  $(Y, X\beta, I\sigma^2)$ , the Best Linear Unbiased Estimate (BLUE) of an estimable linear parametric function (l.p.f)  $P^T\beta$  is given by  $P^T\hat{\beta}$  where  $\hat{\beta}$  is a solution of the "Normal Equations"  $X^T X\beta = X^T Y$ .

**Note:** The solution of normal equations  $X^T X\beta = X^T Y$ , namely  $\hat{\beta} = [X^T X]^{-1} X^T Y$  is any least square estimate of  $\beta$ . Hence,  $\hat{\beta}$  is the BLUE of  $\beta$ .

**Proof:** Consider the Gauss Markov set up  $(Y, X\beta, I\sigma^2)$ , that is  $E(Y) = X\beta$  with  $D(Y) = I\sigma^2$  and  $P^T\beta$  is the estimable linear parametric function to be estimated using  $n$ - observations  $Y = [y_1, y_2, \dots, y_n]$ .

$$\text{Consider } B = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{bmatrix} \text{ then } B^T B = \sum_{i=1}^n (b_i)^2 \text{ and } \frac{\partial}{\partial B} (B^T B) = 2B. \quad (6.2.1)$$

$$\text{Let } A = \begin{bmatrix} a_1 \\ a_1 \\ \cdot \\ \cdot \\ a_1 \end{bmatrix}, \text{ then } A^T B = \sum_{i=1}^n (a_i b_i) \text{ and } \frac{\partial}{\partial B} (A^T B) = A. \quad (6.2.2)$$

Let  $P^T\beta$  is the estimable l.p.f. Let  $B^T Y$  be any l.u.e. of  $P^T\beta$ . We have to prove that BLUE of  $P^T\beta$  is  $P^T\hat{\beta}$ . That is we have to find a  $B^T Y$  such that  $B^T Y$  is the BLUE of  $P^T\beta$ .

Using the definition of l.u.e. we have:

$E[B^T Y] = P^T\beta$ . That is  $B^T E[Y] = P^T\beta$ . Since  $E(Y) = X\beta$ , we have:

$$B^T X\beta = P^T\beta \Rightarrow B^T X = P^T \text{ or } P = X^T B. \quad (6.2.3)$$

The set of equations (6.2.3) will have one or infinite number of solutions.

If it has **only one** solution, that is  $B = B^T Y$ , which is the required BLUE for the l.p.f.  $P^T\beta$ .

If it has **infinite number** of solutions, we have to identify the l.p.f. with minimum variance.

$$\text{Thus consider } V(B^T Y) = V[\sum_{i=1}^n (b_i y_i)] = \sum_{i=1}^n (b_i)^2 V(y_i) = \sum_{i=1}^n (b_i)^2 \sigma^2 = B^T B \sigma^2. \quad (6.2.4)$$

Therefore, the required  $\mathbf{B}$  for the BLUE of  $\mathbf{P}^T\boldsymbol{\beta}$  can be obtained by minimizing  $\mathbf{B}^T\mathbf{B}$  with respect to  $\mathbf{B}$  subjected to the condition that  $\mathbf{X}^T\mathbf{B} = \mathbf{P}$ .

Using **Lagrangian** method of multipliers, for the expression given in equation (6.2.3) we have:

$$\mathbf{L} = \mathbf{B}^T\mathbf{B} - 2\boldsymbol{\Lambda}^T(\mathbf{X}^T\mathbf{B} - \mathbf{P}) \quad (6.2.4)$$

Where  $\boldsymbol{\Lambda}$  is the Lagrangian multiplier.

$$\frac{\partial \mathbf{L}}{\partial \mathbf{B}} = 2\mathbf{B} - 2\mathbf{X}\boldsymbol{\Lambda} = \mathbf{0}. \Rightarrow \mathbf{B} = \mathbf{X}\boldsymbol{\Lambda} \quad (6.2.5)$$

Substituting (6.2.5) in equation (6.2.3) we have:

$$\mathbf{X}^T\mathbf{X}\boldsymbol{\Lambda} = \mathbf{P}. \quad (6.2.6)$$

Solve the set of equations in (6.2.6) for  $\boldsymbol{\Lambda}$  and substitute the same in equation (6.2.5), we get the required  $\mathbf{B}$  to find the BLUE  $\mathbf{B}^T\mathbf{Y}$ .

$$\text{Therefore, the required BLUE of } \mathbf{P}^T\boldsymbol{\beta} = \mathbf{B}^T\mathbf{Y} = \boldsymbol{\Lambda}^T\mathbf{X}^T\mathbf{Y}. \quad (6.2.7)$$

In equation (6.2.7) put  $\mathbf{X}^T\mathbf{Y} = \mathbf{Q}$ . Then we have:  $\boldsymbol{\Lambda}^T\mathbf{Q}$  is the required BLUE, where  $\boldsymbol{\Lambda}^T$  came from equation (6.2.6).

Consider the “Normal Equations”  $\mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^T\mathbf{Y}$ .

$$\text{BLUE of } \mathbf{P}^T\boldsymbol{\beta} = \boldsymbol{\Lambda}^T\mathbf{Q} = \boldsymbol{\Lambda}^T\mathbf{X}^T\mathbf{Y} = \boldsymbol{\Lambda}^T\mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} = [\mathbf{X}^T\mathbf{X}\boldsymbol{\Lambda}]^T\hat{\boldsymbol{\beta}} = \mathbf{P}^T\hat{\boldsymbol{\beta}}. \quad (6.2.8)$$

Hence the theorem.

### 6.2.1. Application of Gauss Markov's Theorem in practice.

Now we proceed to explain the method of finding BLUE using Gauss Markov Theorem in practice:

Let  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$  is the given linear model such that  $E(\mathbf{Y}) = \mathbf{0}$  and  $D(\boldsymbol{\epsilon}) = \mathbf{I}\sigma^2$ .

**Main problem here is to find the BLUE for the l.p.f.  $\mathbf{P}^T\boldsymbol{\beta}$ .** To do this, first form the normal equations in matrix notations as follows:

$$\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} = \mathbf{X}^T\mathbf{Y} \quad (6.2.9)$$

Using the normal equations given in (6.2.9), solve for  $\boldsymbol{\beta}$ . This solution obtained the required BLUE of  $\boldsymbol{\beta}$  denoted by  $\hat{\boldsymbol{\beta}}$ . **Thus BLUE of  $\mathbf{P}^T\boldsymbol{\beta} = \mathbf{P}^T\hat{\boldsymbol{\beta}}$ . It is important to note that this BLUE is nothing but the Least Square solution  $\tilde{\boldsymbol{\beta}}$  obtained from the method of least squares.**

**That is  $\mathbf{P}^T\hat{\boldsymbol{\beta}} = \mathbf{P}^T\tilde{\boldsymbol{\beta}}$ .**

### 6.2.2. Some important Special cases of Gauss Markov's Theorem.

From the discussions in Section (6.2.1), we observe that the method of obtaining BLUE for an estimable l.p.f.,  $\mathbf{P}^T\boldsymbol{\beta}$ , there is need to solve the set of normal equations  $\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} = \mathbf{X}^T\mathbf{Y}$ . In this method, the coefficient matrix  $\mathbf{X}$  play a vital role. Thus the method of finding BLUE differs depending on the nature of the coefficient matrix  $\mathbf{X}$  particularly, based on the Rank ( $\mathbf{X}$ ). We know that the rank of any given matrix may be (i) full or (ii) less or not full. That is the Rank( $\mathbf{X}$ )

is equal to the number of parameters, then we say that the matrix has full rank. Otherwise, the matrix is known as less rank matrix. Based on the nature of the coefficient matrix  $X$ , the methods of finding BLUE are of two types. Namely (i) Method of finding BLUE in full rank case and (ii) Method of finding BLUE when the rank of  $X$  is not full. These methods are explained as follows:

(i) **Method of finding BLUE when the rank of the coefficient matrix  $X$  is full:**

Consider the linear model  $Y = X\beta + \epsilon$ . In this model number of observations are ' $n$ ' and number of parameters of the model are ' $m$ ' and usually  $n \geq m$  and hence, the size of the coefficient matrix  $X$  is of size  $n \times m$ . Suppose the rank of the coefficient matrix  $X$  denoted by  $\rho(X) = m$  (number of parameters), then we say that the coefficient matrix  $X$  has **full rank**. That is in full rank case, the rank of the coefficient matrix  $X$  is exactly equal to the number of parameters of the model namely ' $m$ '. In this case we have one important result which is stated as follows:

**Result (1): If the rank of the coefficient matrix  $X$  is full, every linear parametric function  $P^T\beta$  is estimable.**

**Proof:** Let the rank of  $X = \rho(X) = m$  (number of parameters). Thus  $X$  has full rank. Using the result (2) in lesson 5.3, we know that given any l.p.f.  $P^T\beta$  and  $P^T\beta$  is estimable if and only if (iff)

$\rho(X) = \rho\left(\frac{X}{P^T}\right)$ . This is always true because the rank of the matrix  $X$ ,  $\rho(X) = m$  (full).

Thus in full rank case, we know that  $\rho(X) = \rho(X^T X)$ . Thus  $X^T X$  is also full rank matrix. Further,  $X^T X$  is a square matrix with full rank. Hence  $X^T X$  has unique inverse. Now consider the normal equations  $X^T X\beta = X^T Y \Rightarrow \hat{\beta} = [X^T X]^{-1} X^T Y$ .

(6.2.10)

Put  $X^T X = S_{m \times m}$  and  $X^T Y = Q$  then we have  $\hat{\beta} = S^{-1}Q = CQ$ , where  $C = S^{-1}$ . (6.2.11)

Here the  $C$  is called the **C-matrix** of the Linear Model and the solution  $\hat{\beta}$  in this case is **Unique**.

(ii) **Method of determining BLUE when the rank of the matrix  $X$  is not full.**

When the rank of the coefficient matrix  $X$ , that is  $\rho(X) = r < m$ . To obtain the BLUE of an estimable l.p.f.,  $P^T\beta$ , solve the normal equations:  $X^T X\beta = X^T Y$ . Let  $\hat{\beta}$  be any one solution. Then,  $P^T\hat{\beta}$  is the required BLUE of the l.p.f.  $P^T\beta$ .

### 6.3. Estimation of Variance of BLUEs

Variance of the estimates is the most important measure, which explains the efficiency of the estimate. Best estimate means that estimate which has minimum variance and should be un-biased. Any l.p.f. is estimable, means it is a linear unbiased estimate but there may be many l.u.e.s to select one among those group of l.u.e.s, one must concentrate on its variance and select that estimate which has minimum variance as BLUE. Hence, there is a need to study and measure the variance of the BLUE. To do this, we proceed as follows:

Consider a Gauss Markov's set up  $(Y, X\beta, I\sigma^2)$ . To obtain the BLUE of an l.p.f.  $P^T\beta$ , through the Gauss Markov Theorem (GMT), after solving normal equations, we have  $P^T\hat{\beta}$  as

the BLUE of  $P^T\beta$ . That is  $P^T\hat{\beta} = \Lambda^T Q = \Lambda^T X^T Y = (\Lambda X)^T Y$  (6.3.1)

$$\begin{aligned} \text{Thus } V(P^T\hat{\beta}) &= V(\Lambda^T Q) = V(\Lambda^T X^T Y) = V((\Lambda X)^T Y) = (\Lambda X)^T (\Lambda X) V(Y) \\ &= (\Lambda X)^T (\Lambda X) \sigma^2 = \Lambda^T X^T X \Lambda \sigma^2. \end{aligned} \quad (6.3.2)$$

We know that  $X^T X \Lambda = P$ . Hence we have:  $V(P^T\hat{\beta}) = \Lambda^T P \sigma^2$  (6.3.3)

Here,  $\sigma^2$  is the  $V(\epsilon_i)$  or  $V(Y_i)$  as discussed in lesson 5.

### 6.3.1. Estimation of the Variance of BLUES in full rank case.

Consider a Gauss Markov's set up  $(Y, X\beta, I\sigma^2)$  and the rank of the coefficient matrix  $X$  is full. Hence, every  $\beta_i$ ,  $i = 1, 2, \dots, m$  are individually estimable. Let the BLUE of  $\beta_i$  is denoted by  $\hat{\beta}_i$ , where  $\hat{\beta}_i$  is obtained from normal equations  $X^T X \beta = X^T Y$ . That is  $\hat{\beta} = (X^T X)^{-1} X^T Y = CQ$ .

Let  $D$  is the Dispersion or Variance co-variance matrix of  $\hat{\beta}$ .

$$\text{That is } D(\hat{\beta}) = D(CQ) = D(CX^T Y) = D[(XC^T)^T Y]. \quad (6.3.4)$$

We know that the matrix  $X^T X$  is always symmetric and hence,  $C$  is also symmetric.

$$\begin{aligned} \text{Therefore } D(\hat{\beta}) &= D[(CX)^T Y] = (XC)^T XC D(Y) = (XC)^T XC I\sigma^2 \quad (\text{because } D(Y) = I\sigma^2) \\ &= C^T X^T X C I\sigma^2 = IC I\sigma^2 = C\sigma^2 \quad (\text{because } C = (X^T X)^{-1}). \end{aligned} \quad (6.3.5)$$

### 6.4. Estimation of $\sigma^2$ in Gauss Markov set up

We know from (6.3.5) that  $D(\hat{\beta}) = C\sigma^2$ . Here,  $C$  is known and  $\sigma^2$  is un-known. Hence, there is a need to estimate  $\sigma^2$  to determine  $D(\hat{\beta})$  completely. This is to be estimated from the sample data collected. Let the coefficient matrix  $X$  of the linear model formed can be written

$$\text{as: } X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1m} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{nm} \end{bmatrix}. \quad \text{The } i^{\text{th}} \text{ column vector of } X \text{ is represented by } \alpha_i, i=1, 2, \dots, m.$$

Let the rank of  $X = \rho(X) = r$ . This implies that only  $r$  columns out of  $m$  columns are

independent. Let  $D = [\beta_1, \beta_2, \dots, \beta_{n-r}]$  be orthogonal vectors and orthogonal vector space, orthogonal to  $X$ . This implies that:  $\alpha_i \beta_j = 0$  for all values of  $i, j$  or  $\alpha_i^T \beta_j = 0$  for all  $i, j$  and  $\alpha_i \beta_i = 1$  for all  $i$ . The matrix  $D$  is called the "Deficiency Matrix" of the coefficient matrix  $X$ . Now we proceed to discuss some important properties of this deficiency matrix  $D$  in the following section.

#### 6.4.1. Properties of Deficiency or $D$ matrix.

1. The product of  $X^T D = [0]$  or  $D^T X = [0]$ , where  $[0]$  is known as "Zero Matrix".
2. The product  $D^T D = I$ , where  $I$  is a "Unit Matrix".

3. Let  $\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$  is the vector of n-observations. Then the n-component vector  $\mathbf{Y}$  can be expressed in terms of  $\alpha_1, \alpha_2, \dots, \alpha_m$  and  $\beta_1, \beta_2, \dots, \beta_{n-r}$  as follows:

$$\mathbf{Y} = \mathbf{XC} + \mathbf{DL}, \text{ where } \mathbf{C} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \text{ and } \mathbf{L} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix} \quad (6.4.1)$$

$$\text{That is } \mathbf{Y} = \sum_{i=1}^n c_i \alpha_i + \sum_{i=1}^n l_i \beta_i \quad (6.4.2)$$

Pre-multiplying the equation (6.4.1), by  $\mathbf{D}^T$ , we have:

$$\mathbf{D}^T \mathbf{Y} = \mathbf{D}^T \mathbf{XC} + \mathbf{D}^T \mathbf{DL} = [\mathbf{0}] + \mathbf{IL} = \mathbf{L}. \quad (\text{using properties 1 and 2}) \quad (6.4.3)$$

Thus  $\mathbf{L}$  is a random vector with the following properties:

$$E(\mathbf{L}) = E(\mathbf{D}^T \mathbf{Y}) = \mathbf{D}^T E(\mathbf{Y}) = \mathbf{D}^T \mathbf{X} \boldsymbol{\beta} = \mathbf{0}. \quad (6.4.4)$$

This is because  $\mathbf{X}^T \mathbf{D} = [\mathbf{0}]$  or  $\mathbf{D}^T \mathbf{X} = [\mathbf{0}]$ .

Now we proceed to obtain the formula for estimating  $\sigma^2$ , in the following section.

#### 6.4.2. Estimation of $\sigma^2$ .

Consider the dispersion matrix  $\mathbf{D}$  and we know that  $\mathbf{D}^T \mathbf{Y} = \mathbf{L}$ , where,  $\mathbf{L}$  is a random vector such that  $E(l_i) = 0$  for all values of  $i$  and  $V(l_i) = \sigma^2$  and  $\text{Cov}(l_i, l_j) = 0$  for all values of  $i, j$ .

$$\text{Therefore, } E(l_i^2) = V(l_i) + [E(l_i)]^2 = \sigma^2 + 0 = \sigma^2. \quad (6.4.5)$$

$$\text{Thus we have: } E\{\sum_{i=1}^{n-r} (l_i^2)\} = (n-r) \sigma^2 \Rightarrow \sigma^2 = E\{\sum_{i=1}^{n-r} (l_i^2) / (n-r)\}. \quad (6.4.6)$$

**Therefore,  $\sum_{i=1}^{n-r} (l_i^2) / (n-r)$  is an un-biased estimate of  $\sigma^2$ .**

#### 6.4.3. Estimation of $\sigma^2$ in practice:

Consider the Residual Sum of Squares (RSS) or Error Sum of squares (ESS) of the linear model, which is given as follows:

$$\text{RSS} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = (\mathbf{XC} + \mathbf{DL} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{XC} + \mathbf{DL} - \mathbf{X}\boldsymbol{\beta}) \quad [\text{using (6.4.1)}] \quad (6.4.7)$$

$$= [\mathbf{X}(\mathbf{C} - \boldsymbol{\beta}) + \mathbf{DL}]^T [\mathbf{X}(\mathbf{C} - \boldsymbol{\beta}) + \mathbf{DL}]$$

$$= (\mathbf{C} - \boldsymbol{\beta})^T \mathbf{X}^T \mathbf{X} (\mathbf{C} - \boldsymbol{\beta}) + \mathbf{L}^T \mathbf{D}^T \mathbf{DL} + \mathbf{0}. \quad [\text{because } \mathbf{X}^T \mathbf{D} = \mathbf{0}] \quad (6.4.8)$$

$$= (\mathbf{C} - \boldsymbol{\beta})^T \mathbf{X}^T \mathbf{X} (\mathbf{C} - \boldsymbol{\beta}) + \sum_{i=1}^{n-r} (l_i^2).$$

$$\text{Therefore } \text{RSS} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = \sum_{i=1}^{n-r} (l_i^2) + (\mathbf{C} - \boldsymbol{\beta})^T \mathbf{X}^T \mathbf{X} (\mathbf{C} - \boldsymbol{\beta}) \quad (6.4.9)$$

Expression given in (6.4.9) is minimum when  $\mathbf{C} = \boldsymbol{\beta}$ . Thus substituting  $\mathbf{C} = \boldsymbol{\beta}$  in (6.4.9)

$$\text{We have: } \text{Min RSS} = \sum_{i=1}^{n-r} (l_i^2) = \mathbf{R}_0^2.$$

$$\text{But, we know that } \mathbf{R}_0^2 = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = \sum_{i=1}^{n-r} (l_i^2) \quad (6.4.10)$$

Let  $\widehat{\sigma}^2$  is the unbiased estimator of  $\sigma^2$  and is given by:

$$\widehat{\sigma}^2 = [(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}})^T (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}})] / (n-r) = \mathbf{R}_0^2 / (n-r). \quad (6.4.11)$$

Remark: (1) In general, we cannot say  $\widehat{\sigma^2}$  is the Minimum variance Unbiased (MVUB) of  $\sigma^2$ .

$$(2) \widehat{\sigma^2} = \mathbf{R}_0^2 / (\mathbf{n}-r) = [(\mathbf{Y} - \mathbf{X} \widehat{\boldsymbol{\beta}})^T (\mathbf{Y} - \mathbf{X} \widehat{\boldsymbol{\beta}})] / (\mathbf{n}-r)$$

$$= [\mathbf{Y}^T \mathbf{Y} - 2\mathbf{Y}^T \mathbf{X} \widehat{\boldsymbol{\beta}} + \widehat{\boldsymbol{\beta}}^T \mathbf{X}^T \mathbf{X} \widehat{\boldsymbol{\beta}}] / (\mathbf{n}-r) = [\mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{X} \widehat{\boldsymbol{\beta}}] / (\mathbf{n}-r)$$

This is because  $\mathbf{X}^T \mathbf{X} \widehat{\boldsymbol{\beta}} = \mathbf{X}^T \mathbf{Y}$ .

$$\text{Thus } \widehat{\sigma^2} = (\mathbf{Y}^T \mathbf{Y} - \mathbf{Q}^T \widehat{\boldsymbol{\beta}}) / (\mathbf{n}-r) = [\sum_{i=1}^n (y_i^2) - \sum_{i=1}^n \widehat{\beta}_i q_i] / (\mathbf{n}-r) \quad (6.4.12)$$

$$= [\text{Sum of squares of } y_i - \sum_{i=1}^n \widehat{\beta}_i (\text{RHS of the } i\text{th Normal equation})] / (\mathbf{n}-r).$$

Here  $n$  = number of observations and  $r$  = rank of the coefficient matrix.

**Example (6.4.1):** Consider a simple linear regression model  $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i, i=1,2,\dots,n$  with the assumptions  $E(\varepsilon_i) = 0$ ,  $V(\varepsilon_i) = \sigma^2$  and  $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$ . Determine BLUES for  $\beta_1$  and  $\beta_2$  by using Gauss Markov theorem.

Solution: In order to represent the given model in Gauss Markov's setup, consider

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \text{ and } \mathbf{X}^T \mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}. \text{ Similarly,}$$

$$\mathbf{X}^T \mathbf{Y} = \mathbf{Q} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix}. \text{ Let } \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \text{ and the "Normal Equations" for the G.M. model is given by:}$$

$$\mathbf{X}^T \mathbf{X} \boldsymbol{\beta} = \mathbf{X}^T \mathbf{Y}. \text{ That is } \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix}. \text{ Thus we have:}$$

$$\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix}$$

$$= \{1/n(\sum_{i=1}^n x_i^2 - n \bar{x}^2)\} \begin{bmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix}.$$

$$\text{Thus we have } \widehat{\beta}_1 = [\sum_{i=1}^n x_i \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i] / n(\sum_{i=1}^n x_i^2 - n \bar{x}^2) \quad (6.4.12)$$

$$\text{And } \widehat{\beta}_2 = -[\sum_{i=1}^n x_i \sum_{i=1}^n y_i + n \sum_{i=1}^n x_i y_i] / n(\sum_{i=1}^n x_i^2 - n \bar{x}^2). \quad (6.4.13)$$

$$\text{Further, } V(\widehat{\boldsymbol{\beta}}) = (\mathbf{X}^T \mathbf{X})^{-1} \sigma^2 = [1/n S_{xx}] \begin{bmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{bmatrix} \sigma^2. \quad (6.4.14)$$

Where,  $S_{xx} = (\sum_{i=1}^n x_i^2 - n \bar{x}^2)$ .

This implies that :  $V(\widehat{\beta}_1) = [\sum_{i=1}^n x_i^2 / n S_{xx}] \sigma^2$  and  $V(\widehat{\beta}_2) = \sigma^2 / S_{xx}$  and

Cov  $(\hat{\beta}_1 \hat{\beta}_2) = - [\sum_{i=1}^n xi / n S_{xx}] \sigma^2$ . The unbiased estimate of  $\sigma^2$  is denoted by  $\hat{\sigma}^2$  which is given by:  $[S_{yy} - \hat{\beta}_2 S_{xy}] / (n-r)$ , where  $S_{yy} = \sum_{i=1}^n yi^2 - n \bar{y}^2$ ,  $S_{xy} = \sum_{i=1}^n xi yi - n \bar{x} \bar{y}$ ,  $n$  = number of observations and  $r$  = rank of the coefficient matrix  $\mathbf{X}$ .

**Example (6.4.2):** Let the linear model be:

$$y_1 = \beta_1 + \beta_2 + \varepsilon_1$$

$$y_2 = \beta_1 - \beta_2 + \varepsilon_2$$

$y_3 = \beta_1 + \varepsilon_3$ . Examine whether  $\beta_1$  and  $\beta_2$  are individually estimable? If so obtain their BLUE's along with their standard errors. Also obtain the BLUE of  $(\beta_1 + \beta_2)$  if possible.

Solution: For the given problem  $Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ ,  $X = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix}$  and  $X^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}$ .

Since the rank of the coefficient matrix  $X = 2$ , it is a full rank case and hence the parameters are individually estimable. The given model is a non-singular Gauss Markov setup.

Determine  $S = X^T X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$ . Let  $C = S^{-1} = (X^T X)^{-1} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/2 \end{bmatrix}$ .

Further,  $X^T Y = \begin{bmatrix} y_1 + y_2 + y_3 \\ y_1 - y_2 \end{bmatrix}$  and  $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$ .

For the above G.M. model, normal equations are:  $X^T X \beta = X^T Y$ . Thus we have:

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = (X^T X)^{-1} X^T Y = C Q = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} y_1 + y_2 + y_3 \\ y_1 - y_2 \end{bmatrix}. \text{ Thus we have:}$$

$$\hat{\beta}_1 = (y_1 + y_2 + y_3) / 3 \text{ and } \hat{\beta}_2 = (y_1 - y_2) / 2.$$

Therefore, BLUE of  $\beta_1$  is  $\hat{\beta}_1 = (y_1 + y_2 + y_3) / 3$  and BLUE of  $\beta_2$  is  $\hat{\beta}_2 = (y_1 - y_2) / 2$ .

Further, the dispersion matrix of  $\hat{\beta} = C \sigma^2 = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/2 \end{bmatrix} \sigma^2$ .

Therefore,  $V(\hat{\beta}_1) = \sigma^2 / 3$ ,  $V(\hat{\beta}_2) = \sigma^2 / 2$  and  $\text{Cov}(\hat{\beta}_1 \hat{\beta}_2) = 0$ . The un-biased estimate of  $\sigma^2$  is given by  $\hat{\sigma}^2 = [\sum_{i=1}^n (y_i^2) - \sum_{i=1}^n \hat{\beta}_i q_i] / (n - r)$ . Here  $n = 3$  and  $r = 2$ .

$$= [y_1^2 + y_2^2 + y_3^2 - ((y_1 + y_2 + y_3) / 3) (y_1 + y_2 + y_3) - ((y_1 - y_2) / 2) (y_1 - y_2)] / 3 - 2.$$

Unbiased estimate of  $V(\hat{\beta}_1) = \hat{\sigma}^2 / 3$  and Unbiased estimate of  $V(\hat{\beta}_2) = \hat{\sigma}^2 / 2$ .

BLUE of  $(\beta_1 + \beta_2) = \hat{\beta}_1 + \hat{\beta}_2 = [(y_1 + y_2 + y_3) / 3] + [(y_1 - y_2) / 2] = (5y_1 + 2y_2 - y_3) / 6$ .

And  $V(\hat{\beta}_1 + \hat{\beta}_2) = V(\hat{\beta}_1) + V(\hat{\beta}_2) + \text{Cov}(\hat{\beta}_1 \hat{\beta}_2) = \hat{\sigma}^2 / 3 + \hat{\sigma}^2 / 2 = 5 \hat{\sigma}^2 / 6$ .

Result (1): The normal equations are always consistent.

Proof: Consider the normal equations  $X^T X \beta = X^T Y$ .

Normal equations will have at least one solution. For this the necessary and sufficient condition is that  $X^T Y$  must be dependent on the column vectors of  $X^T X$  matrix.

That is  $X^T Y$  is dependent on the columnvectors  $X^T$ .

This means that  $X^T Y$  is dependent on the columnvectors  $X$ .

Therefore,  $X^T Y$  is dependent on the columnvectors  $X^T X$ .

Hence the proof.

#### 6.4.4. Alternative procedure to for obtaining BLUEs:

Now we proceed to explain an alternative procedure to determine BLUEs for the parameters of the given linear model. For this first we define the concept of “Zero Function” as follows:

##### Definition: Zero Function or Linear functions with zero expectations.

Let  $Y = X\beta + \varepsilon$  be the linear model. Let  $C^T Y$  be a linear function of observations  $y_1, y_2, \dots, y_n$  such that  $E(C^T Y) = 0$ , then  $C^T Y$  is called “Zero Function” or Linear functions with zero expectations.

**Remark:** There will be  $n - r$  independent zero functions in a given linear model, where  $n$ = number of observations and  $r$ =rank of the coefficient matrix of the model.

##### An alternative method for obtaining BLUE:

Let  $B^T Y$  be the un-known BLUE of  $P^T \beta$  in a given linear model  $Y = X\beta + \varepsilon$ . Let  $C^T Y$  be any zero function of the linear model. Then the  $Cov(C^T Y, B^T Y) = C^T B \sigma^2 = B^T C \sigma^2$ .

But  $B$  is the BLUE which satisfies that  $B = X\lambda$ .

Therefore,  $Cov(C^T Y, B^T Y) = \lambda^T X^T C \sigma^2 = 0$ . This is because of the fact that  $X^T C = 0$ .

Hence, BLUE should be un-correlated with every zero function of the linear model.

## 6.5. Summary

In this lesson first Gauss Markov’s set up is explained and an important theorem known as Gauss Markov’s theorem is proved, which is useful to obtain Best Linear Un-biased estimates (BLUE) of a linear parametric function of the Gauss Markov’s model. Further we have discussed various special cases of G.M. theorem. We have discussed the method of obtaining BLUE in full rank case and not full rank case. Further we have discussed the method of obtaining variance of the BLUEs in full rank non full rank cases. Then we have defined the Deficiency Matrix  $D$  and discussed the properties of  $D$  matrix. . Latter, we have discussed the method of estimation of  $\sigma^2$  from the sample data. After explaining the method of obtaining BLUEs with two examples, “Zero Function” or Linear functions with zero expectations of a linear model is defined. Finally an alternative method of obtaining BLUEs using zero functions is discussed.



### 6.6. Self Assessment Questions

1. Explain the Gauss Markov set up along with its assumptions.
2. Define BLUE and discuss its properties.
3. State and prove the Gauss Markov theorem.
4. Explain the use of Gauss Markov theorem with an example.
5. What are Normal Equations? Explain their uses.
6. Explain the method of obtaining BLUE in full rank case.
7. Derive the expression for the variance of BLUE in full rank case.
8. Explain the method of estimating  $\sigma^2$  in practice from the given data.
9. Given the G.M. model

$$Y_1 = \beta_1 + \beta_2 + \beta_3 + \varepsilon_1.$$

$$Y_2 = \beta_1 + \beta_2 - \beta_3 + \varepsilon_2.$$

$$Y_3 = \beta_1 - \beta_2 + \beta_3 + \varepsilon_3.$$

$$Y_4 = -\beta_1 + \beta_2 + \beta_3 + \varepsilon_4.$$

10. Define zero function and explain the alternative procedure of obtaining BLUE.

### 6.7. Further Readings.

1. P. RamachandraRao and P. Bhimasankaram (1996) "Linear Algebra" Tata McGraw-Hill Publishing Company Limited, New Delhi.
2. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons., New Delhi.
3. D.D. Joshi "Theory of Linear Estimation" Wiley Eastern Ltd. New Delhi. 2005.
4. C. R. Rao (1965) "Linear Statistical inference and its applications" John Wiley, New Delhi.

**Lesson – 7.****Generalization of G. M. set up and G.M. Theorem****7.0 Objectives :**

After going through this lesson, you should be able to:

- Understand the need for the generalization of G.M. Model.
- The concept and generalization of Gauss Markov theorem.
- The Statement and proof of Aitkin's theorem.
- Method of calculation of BLUEs in Generalized Linear Models.
- Standardization of Gauss Markov set up and its applications.

**Structure**

- 7.1. **Need for the Generalization of Gauss Markov set up**
- 7.2. **Generalization of Gauss Markov set up and Aitkin's Theorem**
- 7.3. **Estimation of  $\sigma^2$  in General Gauss Markov set up**
- 7.4. **Summary**
- 7.5. **Self Assessment Questions**
- 7.6. **Further Readings**

**7.1. Need for the Generalization of Gauss Markov set up**

In the previous lesson Linear model or Gauss Markov's Model (G.M.M) is introduced which explains the relation between the main variable, depending on various factors. These factors are independently acting on the main variable and hence, in the assumptions, we assume the dispersion matrix  $\mathbf{D}(\mathbf{Y})$  or  $\mathbf{D}(\boldsymbol{\epsilon})$  as  $\mathbf{I}\sigma^2$ , where  $\mathbf{I}$  is the Unit matrix.

This implies that  $V(y_i) = \sigma^2$  covariance between any two pairs  $y_i$  and  $y_j$  is zero. That is  $\text{Cov}(y_i, y_j) = 0$ . That  $y_i$  and  $y_j$  are independent for all values of  $i \neq j = 1, 2, \dots, n$ . This implies that we are assuming that various factors are independently acting on the main variable and have no "**Interaction Effects**". Interaction Effects means, combined effects. That is the effects one factor on the main variable in the presence of another factor. Such combined effects are known as "**Interaction effects**" and in many real life problems; these **Interaction effects** play a **vital role**.

For example, medicine dosage and age are having interaction effects. Doctor cannot prescribe same medicine for all patients, ignoring the age of the patient. He has to prescribe one dosage for children, one dosage for adults and one dosage for old people. Similarly, in industry, experiences of the employee and the efficiency or quality of the work are depending on each other, in the production process. Thus, there are many real life situations where, such

interaction effects are to be considered and propose models suitable to such situations. Ignoring such effects lead to increase the error in our estimates or predictions. Hence there is a necessity to modify the model suitable to situations where interaction effects play a predominant role. That is  $\text{Cov}(y_i, y_j) > 0$ , which in turn means that the factors are not independent but are dependent on each other and acts collectively on the main variable. Thus there is a need to modify or generalize the proposed linear model or G.M. Model. Thus in this lesson we discuss the “**Generalized Linear Model**”(GLM) or “**Generalized Gauss Markov’s Model**”(GGMM).

## 7.2. Generalization of Gauss Markov set up and Aitkin’s Theorem

In the last lesson, we have defined a linear model as  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$  with the assumptions that  $E(\boldsymbol{\varepsilon}) = \mathbf{0}$  and the dispersion or variance co-variance matrix  $\mathbf{D}(\boldsymbol{\varepsilon}) = \mathbf{I}\sigma^2$  where  $\mathbf{I}$  is the Unit matrix. In order to generalize the above model, we assume a general scalar matrix  $\boldsymbol{\Sigma}$ , in the place of  $\mathbf{I}$ . That is  $\text{Cov}(y_i, y_j)$  takes some scalar quantity which is greater than 0. This implies that  $y_i, y_j$  are not independent but they are dependent for all values of  $i \neq j = 1, 2, \dots, m$ . In particular if  $\text{Cov}(y_i, y_j) = 0$ , the general linear model reduces to simple linear model, discussed earlier. In other words, if we assume that interaction effects are absent, GGMM reduces to GMM. Thus “**General Linear Model**” or “**Aitkin’s Model**” is defined as follows:

**Definition (1): A General linear model or Aitkin’s model.**

A General Gauss Markov’s Model or Aitkin’s Model is defined as  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$  with the assumptions that  $E(\boldsymbol{\varepsilon}) = \mathbf{0}$  and the dispersion or variance co-variance matrix  $\mathbf{D}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}\sigma^2$  where  $\boldsymbol{\Sigma}$  is the scalar matrix of size  $n \times n$  and  $\sigma^2$  is un-known.

Further, we know that Gauss Markov set up is denoted as  $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{I}\sigma^2)$ , on similar lines, a General Gauss Markov set up is denoted as  $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}\sigma^2)$ , where  $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix}$ .

**Remark:** In particular if  $\boldsymbol{\Sigma} = \mathbf{I}$  the General Gauss Markov set up ( Model) or Aitkin’s set up (Model) will reduce to the Standard G.M. set up (Model).

Now we proceed to state and prove Aitkin’s Theorem which is the Generalization of Gauss Markov’s Theorem proved in the last lesson.

**Statement:** Consider the General Gauss Markov’s set up  $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}\sigma^2)$ . That is  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$  with

Here  $\boldsymbol{\Sigma}$  is assumed as known and non-singular matrix. The BLUE of any l.p.f.  $\mathbf{P}^T\boldsymbol{\beta}$  is given by  $\mathbf{P}^T\hat{\boldsymbol{\beta}}$ , where  $\hat{\boldsymbol{\beta}}$  is a solution of the “**Modified Normal Equations**”

$$(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})\boldsymbol{\beta} = (\mathbf{X}^T\boldsymbol{\Sigma}^{-1})\mathbf{Y} \quad (7.2.1)$$

**Proof:** It is given in the General Gauss Markov set up that:  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$  is the general linear model, with the assumptions that  $E(\boldsymbol{\varepsilon}) = \mathbf{0}$  and the dispersion or variance co-variance matrix  $\mathbf{D}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}\sigma^2$  where  $\boldsymbol{\Sigma}$  is the scalar matrix of size  $n \times n$  and  $\sigma^2$  is un-known. Since  $\boldsymbol{\Sigma}$  is non-singular matrix, there exists a triangle matrix  $\mathbf{T}$  such that  $\boldsymbol{\Sigma} = \mathbf{T}^T\mathbf{T}$ . This implies that:

$$(\Sigma)^{-1} = T^{-1} (T^{-1})^T. \quad (7.2.2)$$

Consider the transformation  $Y$  to  $Z$  by transforming  $Z = T^{-1}Y$ .

With this transformation, we show that Generalized G.M. set up will become the Standard G.M. set up (discussed earlier in lesson 6). That is, we show that  $Z$  has the standard G.M. setup. This is done as follows:

$$\text{Consider } Y = X\beta + \epsilon \text{ such that } E(\epsilon) = 0 \text{ and } D(\epsilon) = \Sigma\sigma^2. \quad (7.2.3)$$

Pre-multiplying both sides of the model given in (7.2.3) by  $T^{-1}$  we have:

$$T^{-1}Y = T^{-1}X\beta + T^{-1}\epsilon \quad (7.2.4)$$

Put  $T^{-1}Y = Z$ ,  $T^{-1}X = B$  and  $T^{-1}\epsilon = e$  in equation (7.2.4). Then we have:

$$Z = B\beta + e \quad (7.2.5)$$

Now we shall show that equation (7.2.5) is a Standard G.M. set up as follows:

Consider  $E(e) = E(T^{-1}\epsilon) = T^{-1}E(\epsilon) = 0$ . Similarly consider:

$$\begin{aligned} D(e) &= D(T^{-1}\epsilon) = (T^{-1})^T T^{-1} D(\epsilon) = (T^{-1})^T \Sigma T^{-1} \sigma^2 = (T^{-1})^T (T^T T) T^{-1} \sigma^2. \text{ Because } \Sigma = (T^T T) \\ &= \Sigma^{-1} \Sigma \sigma^2 = I \sigma^2. \end{aligned}$$

Therefore  $Z = B\beta + e$  is a standard G.M. Set up. Thus applying G.M. theorem on  $Z$  we obtain the BLUE for the l.p.f.  $P^T\beta$  in routine manner.

The BLUE of  $P^T\beta = P^T\hat{\beta}$  where,  $\hat{\beta}$  is a solution of the normal equations:

$$B^T B \hat{\beta} = B^T Z \quad (7.2.6)$$

Now transforming back from  $Z$  to  $Y$  equation (7.2.6) can be written as:

$$(T^{-1}X)^T (T^{-1}X) \hat{\beta} = (T^{-1}X)^T T^{-1} Y. \quad (7.2.7)$$

$$X^T (T^{-1})^T T^{-1} X \hat{\beta} = X^T (T^{-1})^T T^{-1} Y \text{ or } (X^T \Sigma^{-1} X) \hat{\beta} = X^T \Sigma^{-1} Y. \quad (7.2.8)$$

$$\text{Therefore, } \hat{\beta} = (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1} Y. \quad (7.2.9)$$

Thus the required BLUE of  $P^T\beta$  is  $P^T\hat{\beta}$ , where  $\hat{\beta}$  is obtained from the equation (7.2.9).

Hence the proof.

### 7.3. Estimation of $\sigma^2$ in General Gauss Markov set up

Earlier we have estimated  $\sigma^2$  in the standard G.M. set up, because the value of  $\sigma^2$  is the un-known quantity. This basically, represents the population variance. In some cases the value of  $\sigma^2$  will be known from the past experience. This will happen when we deal known or standard populations. Majority times, we deal populations where the variance is not known and hence, is to be estimated from the sample data collected. In the last lesson we have derived the formula and explained the method of estimating  $\sigma^2$ . Further this estimate an unbiased estimate. On similar lines, now we proceed to derive the formula for  $\sigma^2$  in the general G.M. set up.

To obtain an Un-biased estimate of  $\sigma^2$  is  $\hat{\sigma}^2$  derived in the equation (6.4.12) for the variable Z in G.G.M. set up we have:

$$\hat{\sigma}^2 = [\mathbf{Z}^T \mathbf{Z} - \hat{\boldsymbol{\beta}}^T (\mathbf{X}^T \boldsymbol{\Sigma}^{-1}) \mathbf{Y}] / \mathbf{n} - \mathbf{r}. \quad (7.3.1)$$

Converting back from  $\mathbf{Z}$  to  $\mathbf{Y}$  and using the results of Standard G.M. set up we have:

$$\hat{\sigma}^2 = [(\mathbf{T}^{-1} \mathbf{Y})^T \mathbf{T}^{-1} \mathbf{Y} - \hat{\boldsymbol{\beta}}^T (\mathbf{X}^T \boldsymbol{\Sigma}^{-1}) \mathbf{Y}] / \mathbf{n} - \mathbf{r}. \quad (7.3.2)$$

After substituting the corresponding terms in the inverse transformation for  $\boldsymbol{\Sigma}^{-1} = \mathbf{T}^{-1} (\mathbf{T}^{-1})^T$  finally, we have the following expression for  $\hat{\sigma}^2$  for the General Gauss Markov's set up:

$$\hat{\sigma}^2 = [\mathbf{Y}^T \boldsymbol{\Sigma}^{-1} \mathbf{Y} - \hat{\boldsymbol{\beta}}^T (\mathbf{X}^T \boldsymbol{\Sigma}^{-1}) \mathbf{Y}] / \mathbf{n} - \mathbf{r}. \quad (7.3.3)$$

The results obtained in the last lesson and the present one have lots of applications particularly in the Analysis Of Variance (ANOVA) and Design Of Experiments (DOE), the concepts which we are going to discuss in the forthcoming units. Gauss Markov's theorem is the base for both the topics, which will be re-collected at appropriate places in those units.

#### 7.4. Summary

In this lesson, we have primarily discussed the need for generalizing the Gauss Markov's set up and its applications. The Standard G.M. set up introduced in the previous lesson is applicable only when the factors influencing the main variable are independent. But this assumption may go wrong in many practical problems where, factors have interaction effects. Interaction effects play a significant role on the main variable this type of studies are concentrated in the last unit. For these discussions, we need to generalize the standard G.M. set up introduced in lesson 6. After giving the introduction to generalization of G.M. set up, an important theorem namely Aitkin's theorem or Generalization of Gauss Markov's theorem is stated and proved. A general method of obtaining BLUE for the l.p.f.  $\mathbf{P}^T \boldsymbol{\beta}$  is discussed after converting the General Gauss Markov's set up into Standard Gauss Markov's set up. Finally, method of obtaining un-biased estimate for the population variance  $\sigma^2$  is discussed and obtained the formula for the un-biased estimate in the General Gauss Markov's set up. Applications of the results obtained in this lesson are discussed at the end.

#### 7.5. Self Assessment Questions.

1. Explain the need for generalization of Linear models.
2. Distinguish between Standard G.M. Model and Aitkin's Model.
3. State and prove Aitkin's theorem.
4. Explain how Aitkin's model is a Generalization to Standard Gauss Markov Model.
5. Explain when a generalized G.M. set up will become a particular case.
6. Explain the need for estimating the population variance  $\sigma^2$ .
7. Derive the expression for the un-biased estimate of the population variance  $\sigma^2$ .

8. Explain the importance of BLUEs in Statistical Inference.
9. Define Generalization of G.M. model along with its assumptions.
10. Explain various applications of General G.M. model with suitable examples.

### **7.6. Further Readings.**

1. P. RamachandraRao and P. Bhimasankaram (1996) "Linear Algebra" Tata McGraw-Hill Publishing Company Limited, New Delhi.
2. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons., New Delhi.
3. D.D. Joshi "Theory of Linear Estimation" Wiley Eastern Ltd. New Delhi. 2005.
4. C. R. Rao (1965) "Linear Statistical inference and its applications" John Wiley, New Delhi.

**Lesson – 8.**

## **Analysis of Variance (ANOVA)-One Way Classification**

**8.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the need, concepts and definition of Analysis of variance.
- The concepts of ANOVA – one way classification.
- Mathematical model of ANOVA – one way classification.
- Assumption of ANOVA – one way classification.
- Various other models of ANOVA.
- Relation between student's t-test and ANOVA .

**Structure**

- 8.1 Introduction to ANOVA**
- 8.2 Fundamental concepts of ANOVA**
- 8.3. ANOVA – one way classification**
- 8.4. Mathematical model of ANOVA – one way classification**
- 8.5. Other model of ANOVA**
- 8.6. Summary**
- 8.7. Self Assessment Questions**
- 8.8. Further Readings**

**8.1. Introduction to ANOVA.**

**Analysis of Variance (ANOVA)** is the most powerful and widely used as a statistical tool for “**Tests of Significance**”. ANOVA is considered as the generalization of student's t-test. Student's t- test is mainly used to test the significant difference between two sample means. There are many real life situations, where we have to test more than two means simultaneously at a time. For example, we have to test the average effects of five fertilizers applied to four plots each of the plot has the same crop say paddy or wheat or Red Gram or cotton. Thus we have to study the average effects of these five types fertilizers on the yield of a crop. Hence, there is a need to compare five average effects these fertilizers simultaneously at a time. If we have only two means, we can apply student's t-test, but if we have three or four or more means, we need a statistical procedure, to be developed to test all the means simultaneously at a time. The statistical tool used to deal such situation is known as “**Analysis**”

of Variance" (ANOVA). Thus the basic purpose of Analysis of Variance is to test the homogeneity of **several means** (more than two means) at a time.

## 8.2. Fundamental concepts of ANOVA.

The term "**Analysis of Variance**" (ANOVA) was introduced by Prof. R.A. Fisher in 1920's to deal problems relating to agriculture and to analyze the agronomical data. It is important to note that there will be "**Natural Variation**" in any data. For example, if we sign 10 times, there will be difference between signatures. This is un-avoidable. Such natural variations are called "**Ignorable Variations**" or "**Random Variations**" or "**Chance Variations**". However careful we are, it is impossible to remove such variations. Such random variations are small in size and causes for the variation are many in number. On the other hand there will be some variations which are large in size and cannot be ignored. For example if someone does your Signature, the difference will be large and called as "**Forgery**" which is to be controlled. Such considerable or **significant variations** are called "**Assignable causes of Variation**". Such assignable causes of variations are to be controlled can be removed from the data. Assignable causes of variation if significant and causes for the variation are very few in number. Thus the variation due to **assignable causes** can be detected and measured whereas, variation due to **chance causes** beyond the control of human hand and cannot be detected separately. According to Prof. R. A. Fisher, Analysis of variance (ANOVA) is defined as follows:

**Definition:** Analysis of Variance (ANOVA) is the technique of "**Separation of variance ascribable to one group of causes from the variance ascribable to another group**".

By using the technique of ANOVA, total variation in sample data collected is expressed as the sum of its non-negative components, where, each of these components is a measure of the variation due to some specified independent factor or treatment or cause or some source of variation. The ANOVA technique consists of estimating the amount of variation due to each of these independent factors or the treatments or causes separately and comparing these estimates due to assignable causes or factors with the estimates due to chance factors or causes. Chance causes or factors effect is also known as "**Experimental Error**" or "**Random Errors**" or simply "**Error**". Basically ANOVA is used to for testing the homogeneity of several means. But now a days, ANOVA is frequently used for testing the linearity of the fitted regression line or to test the significance of observed correlation ratio '**p**'. Now we proceed to discuss various assumptions underline the ANOVA technique.

### 8.2.1. Assumptions and Classification of ANOVA.

In ANOVA we have to use F-test. To do this, we require some assumptions which are explained as follows:

- (1) The observations are independent.
- (2) Parent population from which observations are taken is normal and
- (3) Various treatment effects and environmental effects are additive in nature.

Under the above three assumptions, we can use F-test because of one important theorem known as "**Cochran's theorem**" which is stated as follows:



**Statement of Cochran’s Theorem:** Let  $X_1, X_2, \dots, X_n$  denote a random sample from normal population  $N(0, \sigma^2)$ . Let the sum of squares of these values can be written in the form:

$\sum_{i=1}^n X_i^2 = Q_1 + Q_2 + \dots + Q_k$ , where  $Q_i$  is a quadratic form in  $X_1, X_2, \dots, X_n$  with rank (degrees of freedom)  $r_i, j = 1, 2, \dots, k$ . The random variables  $Q_1, Q_2, \dots, Q_k$  are mutually independent and  $Q_j / \sigma^2$  is a chi-square variate with  $r_j$  degrees of freedom if and only if  $\sum_{j=1}^k r_j = n$ .

Above stated Cochran’s theorem play a vital role in the Analysis of Variance (ANOVA) and is considered as the most fundamental theorem of ANOVA. Using the assumptions and the Cochran’s theorem, we can use F-test in ANOVA. ANOVA is of three types. Namely:

- (1) ANOVA – one-way classification.
- (2) ANOVA – two-way classification.
- (3) ANOVA – three or multi- way classification.

Now we proceed to discuss in detail about the ANOVA one-way classification in the following sections.

### 8.3. ANOVA – one way classification

When all the experimental units are homogeneous, and k-different treatments are applied on these homogeneous units, we can expect the variation is due to difference between the treatments and the units within each treatment are homogeneous and hence, the variation within the treatment is insignificant and ignorable. That is the variation within the treatment units is due to random causes of variation and variation between units under different treatments is considered as assignable causes of variation, which is assigned due to the difference among the treatments. This type of splitting the total variation in the collected data due to “**random causes of variation**” and “**assignable causes of variation**” (due to treatments) is called as ANOVA – one way classification. In other words, in ANOVA-one way classification, experimental units are homogeneous and any significant difference among the treatment means is due to the difference among the treatments. Hence this is called as one-way classification. The data under the ANOVA one way classification is given as follows:

**Structure of the Data under ANOVA-one way classification:**

Let suppose that N observations  $x_{ij}$  ( $i = 1, 2, 3, \dots, k; j = 1, 2, \dots, n_i$ ) of random variable X grouped on some basis (usually difference among treatments) into k-classes of sizes  $n_1, n_2, \dots, n_k$  respectively such that  $\sum_{i=1}^k n_i = n$  as explained in the following table:

Treatments/rows	Data				Means	Total
Treatment – 1: $t_1$	$X_{11}$	$X_{12}$	$X_{13}$ . . . . .	$X_{1n_1}$	$\bar{x}_1$	<b>T<sub>1</sub></b>
Treatment – 2: $t_2$	$X_{21}$	$X_{22}$	$X_{23}$ . . . . .	$X_{2n_2}$	$\bar{x}_2$	<b>T<sub>2</sub></b>
Treatment – 3: $t_3$	$X_{31}$	$X_{32}$	$X_{33}$ . . . . .	$X_{3n_3}$	$\bar{x}_3$	<b>T<sub>3</sub></b>
.	.....				.	.
.	.....				.	.

Treatment – k: $t_k$	..... $x_{k1}$ $x_{k2}$ $x_{k3}$ . . . . . $x_{knk}$	$\bar{x}_k$	$T_k$
			<b>G</b> = Grand total

The total variation in the observations  $x_{ij}$ 's can be split into the following two components, namely:

- (1) The variation '**between the classes**' or variation due to different treatments or variation due to different rows. We expect Assignable causes of variation here, that is **due to treatments**. Here we measure the average effects of various treatments under consideration. This variation is known as "**Assignable Causes of Variation**".
- (2) The variation within the classes or treatments that is the inherent variation of the random variable within the observations of a class or treatment. This variation is usually known as "**Random Causes of Variation**".

The first type of variation, namely, **assignable causes of variation** can be detected and can be controlled, by removing causes causing the variation and hence controlled by human efforts. Whereas, the second type of variation, namely random causes of variation is beyond the control of human hand. This is an uncontrollable variation cannot be controlled by any human effort. This is the natural variation present in any data. For example, if we put our signature 10 or 20 times slight variation will automatically present in between signatures. However carefully we sign, naturally variation will be there and it is impossible to do two identical signatures. In repetitive processes, such variations are inevitable and hence such random or natural variations are ignorable. Thus, we conclude that these signatures, consists of in-significant variation or ignorable variation and hence will be considered as same. Now we proceed to explain the mathematical model of ANOVA one-way classification along with its assumptions in the following section.

#### 8.4. Mathematical model and assumptions of ANOVA – one way classification

The mathematical model used under ANOVA – one way classification is as follows:

$$X_{ij} = \mu_i + \varepsilon_{ij} = \mu + (\mu_i - \mu) + \varepsilon_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad (8.4.1)$$

$$i = 1, 2, \dots, K \text{ and } j = 1, 2, \dots, n_i$$

Where:  $x_{ij}$  represents the observation of  $j^{\text{th}}$  experimental unit receiving  $i^{\text{th}}$  treatment.

$$\mu \text{ is the general mean effect given by } \mu = \sum_{i=1}^k n_i \mu_i / N. \quad (8.4.2)$$

$$\alpha_i \text{ is the effect if } i^{\text{th}} \text{ treatment given by } (\mu_i - \mu), \text{ for } i=1, 2, \dots, K. \quad (8.4.3)$$

$\varepsilon_{ij}$  is the error effect due to chance or random causes.

Here,  $\mu$  and  $\alpha_i$   $i = 1, 2, \dots, k$  are called the parameters of the model. Thus ANOVA one way classification model has  $(K+1)$  un-known parameters, which are to be estimated from the Sample data collected.

$$\begin{aligned} \text{Now consider } \sum_{i=1}^k n_i \alpha_i &= \sum_{i=1}^k n_i (\mu_i - \mu) = \sum_{i=1}^k n_i \mu_i - \mu \sum_{i=1}^k n_i . \\ &= N \mu - \mu N = 0. \end{aligned} \quad (8.4.4)$$

#### 8.4.1. Assumptions of ANOVA one way classification:

ANOVA one way classification model discussed in 8.4. has the following assumptions:

- (1) All the observations  $x_{ij}$  are independent.
- (2) Different treatment effects are additive in nature and
- (3) Errors  $\varepsilon_{ij}$ 's are independently and identically distributed normally with mean 0 and variance  $\sigma^2$ . That is  $\varepsilon_{ij} \sim N(0, \sigma^2)$ .

Under the above assumption (3), the model given in equation (8.4.1) will become:

$$E(x_{ij}) = \mu_i = \mu + \alpha_i \quad i=1,2,\dots,K \text{ and } j=1,2,\dots,n_i. \quad (8.4.5)$$

Using the above model along with the above assumptions, we want to test the equality of the population means. That is the homogeneity of different treatment mean effects. Hence, we have the following Null Hypothesis  $H_0$ :

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k = \mu \text{ or } V(\mu_i) = 0. \quad (8.4.6)$$

$$\text{Thus (8.4.6) implies that } H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_k = 0. \quad (8.4.7)$$

and the following Alternative Hypothesis  $H_1$

$$H_1 : \mu_1 \neq \mu_2 \neq \dots \neq \mu_k . = V(\mu_i) > 0. \quad (8.4.8)$$

To test above hypothesis, we use ANOVA one way classification. It is important to note that if  $k = 2$ , above explained problem reduces the **student's t-test** problem. Thus ANOVA is considered as a generalization to student's t-test in the sense that Student's test is used to test the **significance difference** between two sample means where as ANOVA is used to test the difference among several means (variance among the sample means). It is important to note that when we have only two comparable quantities, we use the term "**Difference**". For example, testing the **difference** between A and B or testing the **difference** between Treatment -1 and Treatment -2 or testing the significant difference between first sample mean and the second sample mean and so on. If we have more than two, the term **difference** is not an appropriate term and in fact we have to use the term "**Variance**" or "**Variation**". For example, **Variation** among the class of students or **Variation** among different employees, or **Variation** among different sectors of people or **Variation** among several sample means and so on. Thus in ANOVA we consider the Variation among several means and tested for its significance. Hence ANOVA is also known as "**Variance Ratio Test**". This is because of the fact known that here we are testing  $H_0: V(\mu_i) = 0$  against the alternative hypothesis  $H_1: V(\mu_i) > 0$ . In student's t-test we consider the hypothesis as:  $H_0$  : There is no significant difference between two sample means. or  $H_0: \mu_1 = \mu_2$  and the alternative hypothesis  $H_1: \mu_1 \neq \mu_2$ . This implies that Student's t- test is a "**Two tailed test**" procedure. Since the variance cannot take

negative values, that is  $V(\mu_i) < 0$  is not possible, the test procedure is a “**One tail test**” procedure. In particular, this is a **Right tail** test procedure. Thus we conclude that ANOVA is a Right tail test procedure. Now we proceed to explain Various other models which can be considered in ANOVA in the following section.

### 8.5. Other model of ANOVA

We usually find it useful to describe the observations  $x_{ij}$  from an experiment with a model. One method of writing this model is

$$x_{ij} = \mu_i + \varepsilon_{ij} ; i=1,2,\dots,K \text{ and } j = 1,2,\dots,n \quad (8.5.1)$$

where,  $x_{ij}$  is the  $j$ th observation receiving  $i^{\text{th}}$  treatment ;  $\mu_i$  is the mean of the  $i^{\text{th}}$  treatment and  $\varepsilon_{ij}$  is the random error component with the following assumptions.

$$(1) E(\varepsilon_{ij}) = 0 \quad (2) V(\varepsilon_{ij}) = \sigma^2 \text{ and } (3) \text{Cov}(\varepsilon_i \varepsilon_j) = 0. \quad (8.5.2)$$

These errors are usually known as experimental errors which include measurement errors, variability arising from uncontrollable factors, differences between experimental units, (such as test material or body conditions or resistance power and so on) to which the treatments are applied. This error is also known as white noise in the process, such as variability over time, effect of environmental variables, climatic conditions and so on. With the assumptions given in (8.5.2) we have  $E(x_{ij}) = \mu_i$ . the model given in (8.5.1) is known as “**Means Model**”. This model can also be written alternatively as follows:

$$\mu_i = \mu + \tau_i , i=1,2,\dots,K \quad (8.5.3)$$

Substituting (8.5.3) in (8.5.1) we have:

$$x_{ij} = \mu + \tau_i + \varepsilon_{ij} ; i=1,2,\dots,K \text{ and } j = 1,2,\dots,n. \quad (8.5.4)$$

In the above model,  $\mu$  is a parameter common to all treatments called “Overall Mean” and  $\tau_i$  is a parameter unique to the  $i^{\text{th}}$  “**treatment effect**”. Model given in equation (8.5.4) is usually known as “**Effects Model**”. These models are called one way or single factor ANOVA models because, only one factor (treatments) effect is investigated in these models. These models are to be applied on uniform experimental units and treatments are to be applied randomly on these uniform units. Such type of models are ANOVA one factor (one way classification) models.

The mathematical model given in equation (8.5.4) explains two different situations with respect to the treatment effects, namely,(1) **Fixed effect models** and (2)**Random effect models**.

If the  $K$  treatments are specifically chosen by the experimenter to test the hypothesis about the treatment means, and the conclusions will be applied to the factor levels considered in the analysis and cannot be extended to similar treatments that were not explicitly considered. We can also estimate the model parameters like  $\mu$ ,  $\tau_i$  and  $\sigma^2$ . Such models are called “**Fixed effects models**”. On the other hand, if these  $K$  treatments are selected randomly, from a larger population treatments. In such situations, we will be in a position to extend conclusions to all treatments in that population of treatments. This is because of the fact the selected treatments are random sample from the population treatments. In such experiments, the treatments  $\tau_i$ 's are random variables and knowledge about a particular treatment

investigated is relatively useless. Such models are known as “**Random effects models**” or “**components of variance models**”. In these situations, the experimenter is frequently interested in a factor that has large number of possible levels. If the experimenter randomly selects ‘a’ of these levels, from the population factor levels, we call such experiments as random effects model. Here we assume that the levels of the factor actually used in the experiment were chosen randomly and inferences are made about the entire population of factor levels. We further assume that these factor levels are either of finite size or large enough to be considered as infinite. Usually populations of factor levels of smaller sizes are not considered here.

One can extend above models to two or three factors models. In two factors A, B models, we can consider Factor A as fixed effects and factor B as random effects. Such models are called “**Mixed Effects Models**”. Discussion and analysis of such models are out of scope of the syllabus of this book. Now we proceed to discuss ANOVA one way classification model in the following lesson.

### **8.6. Summary.**

In this lesson, we have introduced the concept of Analysis of variance (ANOVA), its assumptions, and some fundamental concepts of ANOVA. The classification of ANOVA was discussed and stated an important theorem applicable in the analysis of ANOVA, namely Cochran’s theorem. The structure of data under ANOVA one way classification the corresponding model involved in ANOVA one way classification its assumptions the corresponding hypotheses to be tested are discussed. Further, how ANOVA is considered as generalization to student’s t-test is also discussed. It is important to note that the ANOVA procedure is a one tail test procedure. In particular, ANOVA is a Right tail test procedure. These points are discussed with examples at the end of the lesson.

### **8.7. Self Assessment Questions**

1. Define ANOVA along with its assumptions.
2. Explain different types of variation present in the data with suitable examples.
3. Distinguish between Assignable causes of variation and random causes of variation with suitable examples.
4. Explain the structure of the data under ANOVA one way classification.
5. Explain the mathematical model of ANOVA one way classification.
6. Explain various parameters of the model of ANOVA one way classification.
7. Explain various hypotheses considered under ANOVA one way classification.
8. Explain various assumptions of ANOVA one way classification model.
9. Explain how ANOVA is a generalization to Student’s t-test.
10. Show that ANOVA is a Right tail test procedure.

**8.8. Further Readings.**

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons., New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 9.**

## Statistical Analysis of ANOVA-One Way Classification

**9.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the statistical analysis of Analysis of variance – one way classification.
- Method of derivation and expectations of various sums of squares.
- ANOVA table of ANOVA – one way classification.
- Short cut procedure for the analysis of ANOVA – one way classification.
- Further Statistical Analysis – CD test.

**Structure**

- 9.1 Statistical Analysis of ANOVA one way classification
- 9.2 Fundamental concepts of ANOVA one way classification
- 9.3. ANOVA – Table of one way classification
- 9.4. Further Statistical Analysis of ANOVA – one way classification
- 9.5. Summary
- 9.6. Self Assessment Questions
- 9.7. Further Readings

**9.1. Statistical Analysis of ANOVA – one way classification**

Now we proceed to explain the Statistical analysis of ANOVA model in traduced in Section 8.4 of the previous lesson. In order to test the hypotheses  $H_0: \mathbf{V}(\mu_i) = \mathbf{0}$  against the alternative hypothesis  $H_1: \mathbf{V}(\mu_i) > \mathbf{0}$ , for the model considered in (8.4.1) and to estimate the parameters of the model, we consider the following Statistical analysis:

$$\text{Let the mean of } i^{\text{th}} \text{ class} = \bar{x}_i = \sum_{j=1}^{n_i} x_{ij} / n_i. \quad i=1,2,\dots,k. \quad (9.1.1)$$

$$\text{and the overall mean } \bar{x}_{..} = [ \sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij} ] / N = \sum_{i=1}^k n_i \bar{x}_i / k. \quad (9.1.2)$$

The  $(k + 1)$  parameters of the given in (8.4.1), namely  $\mu, \alpha_1, \alpha_2, \dots, \alpha_k$  are to be estimated by using the principle of least squares, which minimizes the residual or error sum of squares as follows:

$$E = [ \sum_{i=1}^k \sum_{j=1}^{n_i} \epsilon_{ij}^2 ] = \{ \sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \mu - \alpha_i]^2 \} \quad (9.1.3)$$

$$\partial E / \partial \mu = -2 \sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \mu - \alpha_i] = 0. \quad (9.1.4)$$

$$\partial E / \partial \mu_i = -2 \sum_{j=1}^{n_i} [x_{ij} - \mu - \alpha_i] = 0. \quad i=1,2,\dots,k. \quad (9.1.5)$$

Simplifying (9.1.4), we have:  $\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij} - n\mu - \sum_{i=1}^k n_i \alpha_i = 0$ . (9.1.6)

Thus we have:  $\hat{\mu} = [\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}] / N = \bar{x}_{..}$ . (9.1.7)

This is because  $\sum_{i=1}^k n_i \alpha_i = 0$  using (9.1.4).

Similarly, simplifying (9.1.5), we have:  $\sum_{j=1}^{n_i} x_{ij} - n_i \hat{\mu} - n_i \alpha_i = 0, i=1,2,\dots,k$ . (9.1.8)

Solving the above set of equations we have:

$$\hat{\alpha}_i = (1/n_i) \sum_{j=1}^{n_i} x_{ij} - \hat{\mu} = \bar{x}_{i.} - \hat{\mu} = \bar{x}_{i.} - \bar{x}_{..} \quad i=1,2,\dots,k. \quad (9.1.9)$$

The set of equations given in (9.1.6) and (9.1.8) are called “Normal Equations”.

Now, substituting (9.1.7) and (9.1.9) in the model given in (8.4.1), we have:

$$x_{ij} = \bar{x}_{..} + (\bar{x}_{i.} - \bar{x}_{..}) + (x_{ij} - \bar{x}_{i.}). \quad (9.1.10)$$

We have to introduce the error component  $\epsilon_{ij}$ , on RHS of (9.1.10) so that LHS = RHS.

The error term is the deviation within the class which is due to randomization. Transposing  $\bar{x}_{..}$  to the LHS of (8.5.10) and squaring both sides and summing over i and j, we have:

$$\begin{aligned} \sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \bar{x}_{..}]^2 &= \sum_{i=1}^k \sum_{j=1}^{n_i} [(x_{ij} - \bar{x}_{i.}) + (\bar{x}_{i.} - \bar{x}_{..})]^2 \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} [(x_{ij} - \bar{x}_{i.})]^2 + \sum_{i=1}^k n_i (\bar{x}_{i.} - \bar{x}_{..})^2 + 2[\sum_{i=1}^k \{(\bar{x}_{i.} - \bar{x}_{..}) \sum_{j=1}^{n_i} [(x_{ij} - \bar{x}_{i.})]\}]. \end{aligned} \quad (9.1.11)$$

We know that the  $\sum_{j=1}^{n_i} [(x_{ij} - \bar{x}_{i.})] = 0$ , because, the sum of the deviations of the observations from their mean is always zero. Therefore the third term on the RHS of (9.1.11) will become zero and hence (9.1.11) will become:

$$\sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \bar{x}_{..}]^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} [(x_{ij} - \bar{x}_{i.})]^2 + \sum_{i=1}^k n_i (\bar{x}_{i.} - \bar{x}_{..})^2. \quad (9.1.12)$$

Now, let the terms in (9.1.12) will be denoted as follows:

$$S_T^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \bar{x}_{..}]^2 \text{ is known as Total Sum of Squares (TSS)}$$

$$S_E^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} [(x_{ij} - \bar{x}_{i.})]^2 \text{ is known as Error Sum of Squares (ESS) or Residual Sum of Squares (RSS) which represents the natural variation and}$$

$$S_t^2 = \sum_{i=1}^k n_i (\bar{x}_{i.} - \bar{x}_{..})^2 \text{ is known as Sum of squares due to treatments (SST). Thus (9.1.12) can be written as: } \quad \text{Total SS} = \text{ESS} + \text{SST} \text{ or } S_T^2 = S_E^2 + S_t^2. \quad (9.1.13)$$

Now we proceed to explain some important concepts and finding of expectations of various sums of squares from the sample data collected in the following section.

## 9.2 Fundamental concepts of ANOVA one way classification

In ANOVA we use two important concepts, namely (1) Degrees of freedom and (2) Mean Sum of squares. Before explaining the method of finding expectations of various sum of squares, we explain these concepts first.



### 9.2.1. Degrees freedom of Various Sums of Squares:

The total Sum of Squares (TSS) is computed as  $S_T^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \bar{x}_{..}]^2$ , which contains N quantities. Hence, has (N-1) degrees of freedom (dof). This one degree of freedom is lost because of the linear constraint, namely  $\sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \bar{x}_{..}]^2 = 0$ .

Similarly, the degrees of freedom (dof) of the treatment Sum of Squares which is given by  $S_T^2 = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x}_{..})^2$  will have k terms and has (k-1) degrees of freedom (dof), because,  $\sum_{i=1}^k n_i (\bar{x}_i - \bar{x}_{..}) = 0$ .

Further, the error Sum of Square (ESS) is given by:

$S_E^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \bar{x}_i]^2$  will have (N-k) degrees of freedom (dof). This is because ESS is based on N quantities which are subjected to k linear constraints  $\sum_{j=1}^{n_i} [x_{ij} - \bar{x}_i] = 0$  for  $i=1,2,\dots,k$ . Further, we see that the degrees freedoms for various SS are additive because:

$$N - 1 = (N - k) + (K - 1).$$

### 9.2.2. Mean Sum of Squares (MSS):

**Definition:** The sum of Squares divided by its degrees of freedom gives the corresponding variance or the Mean Sum of Squares (MSS).

Thus we have:  $S_t / (k - 1) = SST/(k-1) = [S_t^2]$  is the MSS due to treatments.

Similarly,  $S_E^2 / (N-k) = ESS/(N - k) = [S_E^2]$  is the MSS due to Error or Residues.

**9.2.3. Expectations of Various SS:** In order to develop the test procedure to test the Null Hypothesis  $H_0: \mu_1 = \mu_2 = \dots, \mu_k = \mu$  or  $V(\mu_i) = 0$ . This is equivalent as  $H_0: \alpha_1 = \alpha_2 = \dots, \alpha_k = 0$ , we require expectations of various SS due to each of the independent factors. Now we proceed to calculate expectations of various SS as follows.

Consider the linear model of ANOVA explained in Section 8.4.

$$X_{ij} = \mu_i + \epsilon_{ij} = \mu + (\mu_i - \mu) + \epsilon_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad (9.2.1)$$

$i = 1, 2, \dots, K$  and  $j = 1, 2, \dots, n_i$

Summing (9.2.1) over j and dividing by  $n_i$  we have:

$$[\sum_{j=1}^{n_i} X_{ij}] / n_i = (1/n_i) [n_i \mu + n_i \alpha_i + \sum_{j=1}^{n_i} \epsilon_{ij}]. \quad \text{This implies that: } \bar{x}_i = \mu + \alpha_i + \bar{\epsilon}_i. \quad (9.2.2)$$

Now summing (9.2.1) over i and j and dividing by  $N = \sum_{i=1}^k n_i$ , we have:

$$(1/N) \sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij} = (1/N) [N \mu + \sum_{i=1}^k n_i \alpha_i + \sum_{i=1}^k \sum_{j=1}^{n_i} \epsilon_{ij}].$$

$$\text{That is } \bar{x}_{..} = \mu + \bar{\epsilon}_{..} \quad (9.2.3)$$

### Expectation of Treatment Sum of Square (SST):

$$E(SST) = E(S_T^2) = E [\sum_{i=1}^k n_i (\bar{x}_i - \bar{x}_{..})^2]. \quad (9.2.4)$$

Substituting (9.2.2) and (9.2.3) in (9.2.4) we have:

$$E(SST) = E \sum_{i=1}^k [n_i (\mu + \alpha_i + \bar{\epsilon}_i) - (\mu + \bar{\epsilon}_{..})]^2 = E \sum_{i=1}^k [n_i (\alpha_i + \bar{\epsilon}_i - \bar{\epsilon}_{..})^2].$$

$$\begin{aligned}
&= E \sum_{i=1}^k n_i [(\alpha_i)^2 + (\bar{\epsilon}_i - \bar{\epsilon}_{..})^2 + 2 \alpha_i (\bar{\epsilon}_i - \bar{\epsilon}_{..})] \\
&= \sum_{i=1}^k E [n_i (\alpha_i)^2 + n_i (\bar{\epsilon}_i - \bar{\epsilon}_{..})^2] + 2 \sum_{i=1}^k [n_i (\alpha_i) E(\bar{\epsilon}_i - \bar{\epsilon}_{..})] \quad (9.2.5)
\end{aligned}$$

$$= \sum_{i=1}^k n_i (\alpha_i)^2 + E \sum_{i=1}^k n_i (\bar{\epsilon}_i - \bar{\epsilon}_{..})^2 \quad (9.2.6)$$

(The second term on the rhs of (9.2.5) will become zero because  $E(\bar{\epsilon}_i - \bar{\epsilon}_{..}) = 0$ , as  $\epsilon_{ij}$ 's are independently and identically distributed Normal variables with mean 0 variance  $\sigma_e^2$ .)

Now equation (9.2.6) can be written as:

$$\begin{aligned}
&\sum_{i=1}^k n_i (\alpha_i)^2 + E \left[ \sum_{i=1}^k n_i (\bar{\epsilon}_i)^2 - N (\bar{\epsilon}_{..})^2 \right] \\
&= \sum_{i=1}^k n_i (\alpha_i)^2 + \sum_{i=1}^k n_i E(\bar{\epsilon}_i^2) - N E(\bar{\epsilon}_{..}^2) \quad (9.2.7)
\end{aligned}$$

Since,  $\epsilon_{ij}$ 's are independently and identically distributed Normal variables with mean 0 variance  $\sigma_e^2$ . Further we know that:  $E(\bar{\epsilon}_j) = 0$  and  $E(\bar{\epsilon}_{..}) = 0$ , we have:

$$\text{Var}(\bar{\epsilon}_j) = E(\bar{\epsilon}_j^2) - E(\bar{\epsilon}_j)^2 = E(\bar{\epsilon}_j^2) \Rightarrow E(\bar{\epsilon}_j^2) = \text{Var}(\bar{\epsilon}_j) = \sigma_e^2 / n_j. \quad (9.2.8)$$

This is because of the fact that the variance of the mean of a random sample of size  $n$  from a population with variance  $\sigma^2$  is  $\sigma^2 / n$ .

$$\text{Similarly, we have } E(\bar{\epsilon}_{..}^2) = \text{Var}(\bar{\epsilon}_{..}) = \sigma_e^2 / N. \quad (9.2.9)$$

Substituting (9.2.8) and (9.2.9) in (9.2.7) we have:

$$\begin{aligned}
E(\text{SST}) &= \sum_{i=1}^k n_i (\alpha_i)^2 + \sum_{i=1}^k n_i (\sigma_e^2 / n_i) - N \sigma_e^2 / N = \sum_{i=1}^k n_i (\alpha_i)^2 + \sum_{i=1}^k \sigma_e^2 - \sigma_e^2 \\
&= \sum_{i=1}^k n_i (\alpha_i)^2 + K \sigma_e^2 - \sigma_e^2 = \sum_{i=1}^k n_i (\alpha_i)^2 + (K-1) \sigma_e^2.
\end{aligned}$$

$$\text{Thus we have: } E[\text{SST}/(K-1)] = \sigma_e^2 + (1/(K-1)) \sum_{i=1}^k n_i (\alpha_i)^2 \quad (9.2.10)$$

Hence, Under the  $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$  Mean Sum of Squares (MSS) due to treatments provides an unbiased estimate of  $\sigma_e^2$ .

Now we proceed to derive the Expectation of error sum of squares (ESS) as follows:

$$\begin{aligned}
\text{Consider } E(\text{ESS}) &= E(S_E^2) = \sum_{i=1}^k \sum_{j=1}^{n_i} [x_{ij} - \bar{x}_i.]^2 \\
&= E \left[ \sum_{i=1}^k \sum_{j=1}^{n_i} \{ \mu + \alpha_i + \epsilon_{ij} - (\mu + \alpha_i + \bar{\epsilon}_i.) \}^2 \right]. \quad [\text{Using (9.2.1) and (9.2.2)}] \\
&= E \sum_{i=1}^k \sum_{j=1}^{n_i} (\epsilon_{ij} - \bar{\epsilon}_i.)^2 = E \left[ \sum_{i=1}^k \left( \sum_{j=1}^{n_i} (\epsilon_{ij} - \bar{\epsilon}_i.)^2 \right) \right]. \\
&= E \left[ \sum_{i=1}^k \left( \sum_{j=1}^{n_i} \epsilon_{ij}^2 - n_i \bar{\epsilon}_i.^2 \right) \right] = E \left[ \sum_{i=1}^k \left( \sum_{j=1}^{n_i} \epsilon_{ij}^2 - \sum_{i=1}^k n_i \bar{\epsilon}_i.^2 \right) \right] \\
&= \left[ \sum_{i=1}^k \sum_{j=1}^{n_i} E \epsilon_{ij}^2 - \sum_{i=1}^k n_i E(\bar{\epsilon}_i.)^2 \right] = \sum_{i=1}^k \sum_{j=1}^{n_i} \sigma_e^2 - \sum_{i=1}^k n_i (\sigma_e^2 / n_i) \\
&= N \sigma_e^2 - \sum_{i=1}^k \sigma_e^2 = N \sigma_e^2 - K \sigma_e^2 = (N - K) \sigma_e^2 \Rightarrow E[ \text{ESS} / (N - K) ] = \sigma_e^2. \quad (9.2.11)
\end{aligned}$$

This means that the error mean sum of squares always gives an unbiased estimate of  $\sigma_e^2$ .

Thus under the null hypothesis  $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ , we have:

$$E[(\text{SST}/(K-1))] = E[\text{ESS} / (N-K)] = \sigma_e^2. \text{ or } E(S_t^2) = E(S_E^2). \text{ Otherwise, } E(S_t^2) > E(S_E^2). \quad (9.2.12)$$

Hence, to test the null hypothesis  $H_0$ , we use the variance ratio test namely, F test and the test statistic  $F = S_t^2 / S_E^2$ . If  $H_0$  is true  $F = 1$  otherwise,  $F > 1$ .

### 9.2.4. Application of Cochran's Theorem:

In order to find out if an observed value of F is significant or not we have to obtain the sampling distribution of the test statistic  $F = S_t^2 / S_E^2$ . Under  $H_0$  and using the Cochran's theorem,  $S_t^2 / \sigma_e^2$  and  $S_E^2 / \sigma_e^2$  are independently distributed Chi-square variables with (K-1) and (N-K) degrees of freedom respectively. Hence, the statistic:

$$F = [(S_t^2 / \sigma_e^2) / (K-1)] \div [(S_E^2 / \sigma_e^2) / (N - K)] = S_t^2 / S_E^2 \quad (9.2.13)$$

Follows Snedecor's F or Central F distribution with (K-1, N-K) degrees of freedom. Thus if an observed value of F obtained from (9.2.13) is less than the table value of F for (K-1, N-K) degrees of freedom at specified ( $\alpha$  %) level of significance, then,  $H_0$  is accepted otherwise, we reject  $H_0$ . Usually we consider  $\alpha = 5\%$  or  $1\%$  level of significance.

Now, we proceed to explain the Analysis of Variance (ANOVA) table for one way classification, which explains the above explained Statistical analysis in an elegant form. This is discussed in the following section.

### 9.3. ANOVA – Table of one way classification.

Analysis of variance (ANOVA) table of one way classification explains the Statistical analysis of the model discussed in section 9.2 in a nut shell, which is given as follows:

**Table (9.3.1): ANOVA table for one-way classified data.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F - cal)
Treatments (Columns)	K - 1	$S_t^2 = SST$	$V_1 = S_t^2 / K - 1$	$F = V_1 / V_2$
Error	N - K	$S_E^2 = ESS$	$V_2 = S_E^2 / N - K$	
Total	N - 1	$S_T^2 = TSS$		

**Remark:** 1. In the above table (9.3.1) one can observe that  $SST+ESS = TSS$ . Similarly the corresponding degrees of freedom (dof) are also additive. That is  $(K - 1) + (N - K) = N - 1$ . But it is important to note that the corresponding mean sum of squares do not add up to Total MSS. That is variances  $V_1$  and  $V_2$  are not equal to is not equal to  $TSS / (N - 1)$ .

2. Even though the technique is popularly known as Analysis of Variance (ANOVA), it is actually the analysis of the Sun of squares (SS).

#### 9.3.1: Simplified Procedure for ANOVA:

The procedure of ANOVA explained above is a very lengthy procedure and require more calculations. In practice, we adopt a short-cut procedure in analyzing many practical problems. The arithmetic calculations for obtaining various sum of squares (SS) are reduced to great extent if we follow the following simplified procedure. The simplified procedure is explained as follows:

Suppose that N observations  $x_{ij}$  ( $i = 1, 2, 3, \dots, k; j = 1, 2, \dots, n_i$ ) of random variable X grouped on some basis (usually difference among treatments) into k-classes of sizes  $n_1, n_2, \dots, n_k$  respectively such that  $\sum_{j=1}^k n_j = n$  as explained in the following table (9.3.1). First, calculate various treatment totals  $T_i$   $i = 1, 2, \dots, K$  and grand total  $G = \sum_{i=1}^k T_i$  and  $\bar{x}_i = T_i / n_i$   $i=1, 2, \dots, K$ .

Table (9.3.1) Anova one way classified data.

Treatments/rows	Data				Means	Total
Treatment – 1: $t_1$	$x_{11}$	$x_{12}$	$x_{13}$ . . . . .	$x_{1n_1}$	$\bar{x}_1$	$T_1$
Treatment – 2: $t_2$	$x_{21}$	$x_{22}$	$x_{23}$ . . . . .	$x_{2n_2}$	$\bar{x}_2$	$T_2$
Treatment – 3: $t_3$	$x_{31}$	$x_{32}$	$x_{33}$ . . . . .	$x_{3n_3}$	$\bar{x}_3$	$T_3$
.	.....				.	.
.	.....				.	.
.	.....				.	.
Treatment – k: $t_k$	$x_{k1}$	$x_{k2}$	$x_{k3}$ . . . . .	$x_{kn_k}$	$\bar{x}_k$	$T_k$
						<b>G = Grand total</b>

Step – 1: Calculate the **Correction factor (CF) =  $G^2 / N$** .

Step – 2: Calculate Treatment Sum of Squares (SST) =  $\sum_{i=1}^k (T_i)^2 / n_i - CF = A$ .

Step – 3: Calculate Total Sum of squares (TSS) =  $\sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij})^2 - CF = B$ .

Step – 4: Construct ANOVA table as follows:

Table (9.3.2). ANOVA table for one way classification.

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (Columns)	$K - 1$	$S_t^2 = SST = A$	$V_1 = S_t^2 / K - 1$	$F = V_1 / V_2$
Error	$N - K$	$S_E^2 = ESS = B - A$	$V_2 = S_E^2 / N - K$	
Total	$N - 1$	$S_T^2 = TSS = B$		

In the above ANOVA table (9.3.2), F- statistic follows F-distribution with  $(K - 1, N - K)$  degrees of freedom. Note down the table value at  $\alpha\%$  level of significance (los) at  $(K-1, N-K)$  degrees of freedom (dof) from the F-table given in appendix. Usually, we take  $\alpha = 0.05$  that is 5% los or 0.01 that is 1% los. This table of F is denoted by  $F_{tab}(\alpha)$  and the calculated value of F given in the 5<sup>th</sup> column of the table (9.3.1) as  $F_{cal}$ . If  $F_{cal} \leq F_{tab}$ , we accept the Null-hypothesis and conclude that  **$V(\mu_i) = 0$ . That is there is no significant variation present between the treat**

**means under consideration.** Otherwise, that is  $F_{cal} > F_{tab}$  we reject the null hypothesis or we accept the alternative hypothesis and conclude that  $V(\mu_i) > 0$ . **That is there exists significant evidence that the treatment means significantly differs from each other.  $V(\mu_i) = 0$ .**

**Remark:** 1. If we **reject the null-hypothesis** at 5% (0.05) level of significance, we represent the situation by (\*) and If we **reject the null-hypothesis** at 1% (0.01) level of significance, we represent the situation by (\*\*). Similarly if we **accept the null hypothesis** at 5% (0.05) level of significance, the situation is represented by no star ( ).

2. **The p-value of the test statistic:** The test statistic F is calculated is computed from the sample data collected and hence is a function of sample observations. Different samples will provide different test statistics. We are interested to know the chance of rejecting when it is really true. That is, if really  $H_0$  is true, what is the probability that it is rejected? This probability is known as “**p-value**” and this probability should be as small as possible. Thus p-value represent  $\Pr(F_{cal} > F_{crit})$ . In many computer outputs associated with the problem of testing will give this p-value. In order to calculate this p-value, we have to use the sampling distribution of the test statistic. That is for F-statistic, we have to use F-distribution. Similarly, for t-statistic, we have to use student’s t-distribution and Z-statistic, we have to use Z-distribution or Standard Normal distribution and so on. Hence, calculation of p-value manually is very difficult. But computers packages can calculate this p-value and is provided. This p-value represents the probability of wrongly rejecting the null-hypothesis.

**If  $p < 0.05$ , we say that the test result is significant at 5% (0.05) level of significance, we reject the null-hypothesis  $H_0$ . Otherwise, if  $p > 0.05$ , we accept the null-hypothesis  $H_0$ .**

3. We know that calculation of variance is independent of change of origin and scale. Hence the change of origin and scale reduces calculations of ANOVA to a very greater extent.

If we accept the null-hypothesis, implies that all the treatments under consideration are equally effective and one can use any one of the treatment, because, each one can give similar mean effect. But, if we reject the null-hypothesis, and conclude that the treatments under consideration are not equal, then next question arises is that “Which treatment is the Best one?” Or “which treatment is least preferable one?” and so on.

Thus the analysis is not complete, unless otherwise we answer the above type of questions. To do this we required to do “**Further Statistical Analysis**” which is explained in the following section 9.4.

#### **9.4. Further Statistical Analysis of ANOVA – one way classification**

**If the null-hypothesis  $H_0$  is rejected, simply stopping at that stage is not the end of the ANOVA analysis.** We must be in a position to say which treatment is more efficient or which treatment most preferable, which treatment is least preferable and so on. That is we want to sub-group the treatments such that treatments under the same group have no significant difference where as treatments under different sub-groups have significant difference. For example five treatments  $T_1, T_2, T_3, T_4$ , and  $T_5$  are sub-grouped as  $(T_1, T_4)$ ,  $(T_2, T_3, T_5)$  implies that there is no significant difference between  $(T_1$  and  $T_4)$ ,  $(T_2$  and  $T_3)$ ,  $(T_2$  and  $T_5)$  and  $(T_3$  and

$T_5$ ) because they are under the same group. Similarly, there exists significant difference between  $(T_1 \text{ and } T_2)$ ,  $(T_1 \text{ and } T_3)$   $(T_1 \text{ and } T_5)$   $(T_4 \text{ and } T_2)$ ,  $(T_3 \text{ and } T_4)$  or  $(T_5 \text{ and } T_4)$  because they are in different sub-groups. If there is no significant difference between any pair or treatments, they form one single sub-group in which all the treatments in the same sub-group. In that situation we accept the null-hypothesis  $H_0$ .

To do this type of exercise, we have to apply “**Further Statistical Analysis**”. There are four types of test procedures applicable in further statistical analysis. Namely:

- (1) Critical Difference (CD) Test.
- (2) Least Significant Difference (LSD) Test .
- (3) Duncan’s Multiple Range (DMR) Test and
- (4) Scheff’s Test.

Among the above mentioned tests, DMR or Scheff’s tests are based on DMR tables or Sheff’s tables respectively. These tables are specially prepared using order statistics and using their distributions. These tables are not available as popularly as Student’s t- tables. Hence we discuss CD test or LSD test which uses the most popular student t-tables, which are available in any fundamental book on statistics.

#### 9.4.1. Critical Difference (CD) or Least Significant Difference LSD Test.

The test procedure is explained as follows:

First, calculate all treatment means from the collected data. That is Calculate

$$\bar{x}_i = T_i / n_i \quad i=1,2,..,K.$$

If  $H_0$  is rejected, we conclude that there is significant difference between the treatment means under consideration. Then we are further interested to find out which pair(s) of treatments different significantly or which is the most useful treatment or which is the least preferable treatment and so on. To do this, usually we have to apply student’s t-test. Instead of calculating Student’s t-values for each pair of treatments, we calculate the “Least Significant Difference” at the given  $\alpha\%$  level of significance. This least significant difference is known as the “Critical Difference” denoted by  $CD_\alpha$  at  $\alpha\%$  los and is defined as follows:

$$CD_\alpha = \text{S.E. of the difference between two means} \times t_{\alpha\%} \text{ for error dof.} \quad (9.4.1)$$

The Critical Difference between two sample means  $(\bar{x}_i - \bar{x}_j)$  is given by:

$$CD_{\alpha\%} (\bar{x}_i - \bar{x}_j) = t_{\alpha\%} \text{ for error dof} \times S_E (1/n_i + 1/n_j)^{1/2}. \quad (9.4.2)$$

Where,  $S_E$  = is the square root of error Sum of square to be taken from the corresponding ANOVA table.

This is because of the fact that  $\text{Var} (\bar{x}_i - \bar{x}_j) = \sigma_e^2 / n_i - \sigma_e^2 / n_j = \sigma_e^2 (1 / n_i - 1/n_j)$ .

That is S.E.  $(\bar{x}_i - \bar{x}_j) = \sigma_e (1 / n_i - 1/n_j)^{1/2}$ , where,  $\sigma_e$  is the  $\sqrt{S_E^2}$  where  $S_E^2$  is the error sun of square calculated in ANOVA table.

Remark: In particular all the treatments have equal replications, that is  $n_i = n_j = n$  then the formula given in (9.4.2) will become:  $CD_{\alpha\%} (\bar{x}_i - \bar{x}_j) = t_{\alpha\%} \text{ for error dof} \times S_E (2/n)^{1/2} .(9.4.3)$ .

2. t 2.1 values can be adtainel form any reference book of ons fundamental book on ststiscal methods.

**Example (9.4.1):** The following data represents life time (in hrs) measured on four samples of electric bulbs manufactured from four different machines. Test whether these machines differ significantly with respect to the average life times of electric bulbs. Also comment whether there is any need to replace or repair any machine.

**Machine -1:** 1580 1640 1640 1700 1750.

**Machine-2:** 1600 1610 1650 1680 1700 1720 1800.

**Machine-3:** 1460 1550 1600 1620 1640 1660 1740 1820.

**Machine-4:** 1510 1520 1530 1570 1600 1680.

**Solution:** Since the variation in the data can be attributed to difference among the machines and random errors, one has to apply ANOVA one way classification. Tee procedure is explained as follows:

Step – 1: Frame the null-hypothesis  $H_0$  : There is no significant variation between the average means of different samples of electric bulbs produced from different machines.

Step – 2: Since the data items are large in sized, using the remark – 3 of Section 9.3.1, we can reduce calculations by considering origin as **1640** and Scale as dividing by **10**, we have the following table with the above change of origin and the scale.

Machines.	Transformed (Reduced) data.								Totals ( $T_i$ )	Squires ( $\sum_{j=1}^{n_i} (x_{ij})^2$ )
Machine -1:	-6	0	0	6	11	-	-	-	11	193
Machine -2:	-4	-3	1	4	6	8	16	28	28	398
Machine-3:	-18	-9	-4	-2	0	2	10	18	-3	853
Machine-4:	-13	-12	-11	-7	-4	4	-	-	-43	515
	Grand Totals								-7 = G	1,959 = RSS

Step – 3: Calculate row (machine) totals and squares and their total as shown above.

Step – 4: Calculate  $CF = G^2 / N = (-7)^2 / 26 = 49/26 = 1.884615$ .

Step – 5: Calculate Various Sum of Squares (SS) as follows:

Total SS = TSS = RSS – CF = 1959 – 1.884615 = **1,957.115385 = B.**

Machine (Treatment) SS =  $\sum_{i=1}^k (T_i)^2 / n_i - CF = [(11)^2 / 5 + (28)^2 / 7 + (-3)^2 / 8 + (-43)^2 / 6] - CF$   
**= 445.49167 – 1.884615 = 443.60706 = A.**

Step – 6: Calculate entries of ANOVA – one way table as follows:

Table (9.4.3). ANOVA table for one way classification for the example (9.4.1).

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F-cal)
Treatments (Columns)	$K - 1 = 4 - 1 = 3$	$S_t^2 = SST = A = 443.60706$	$V_1 = S_t^2 / K - 1 = 443.60706 / 3 = 147.86903$	$F = V_1 / V_2 = 147.86903 / 68.79583 = 2.1493$
Error	$N - K = 26 - 4 = 22$	$S_E^2 = ESS = B - A = 1,957.115385 - 443.60706 = 1513.50833$	$V_2 = S_E^2 / N - K = 1513.50833 / 22 = 68.79583$	
Total	$N - 1 = 26 - 1 = 25$	$S_T^2 = TSS = B = 1,957.115385$		

Step – 7: Conclusions: From the ANOVA table (9.4.3),  $F_{cal} = 2.1493$  and from F table (see annexure)  $F_{tab}$  at (3,22)degrees of freedom at 5% los  $F_{tab} = 3.491$  and at 1% los  $F_{tab} = 4.82$ . Since,  $F_{cal}$  is less than  $F_{tab}$  at 5% los, we can also represent this as  $p > 0.05$ , we accept the null-hypothesis  $H_0$  and conclude that the average life times **of electric bulbs produced from different machines does not show any statistical significance**. This means that, all machines are equal in the production of electrical bulbs with respect to the average lifetimes.

Since, we accept the null-hypothesis  $H_0$ , there is no need of further statistical analysis. Since all machines are equal in the production of electrical bulbs with respect to average life times, **there is no need for changing or repair of any machine**.

**Now we proceed to explain ANOVA – one way along with Further Statistical Analysis in the following example.**

**Example (9.4.2):** Following data represents the tensile strength of synthetic fiber measured in  $lb^2/m^2$  with various cotton percentages. Analyze the data whether the cotton percentage has any significant effect on the tensile strength of the fiber. Also determine which cotton percentage is best suitable for obtaining best tensile strength of the synthetic fiber.

Cotton Percentages	1	2	3	4	5
Treatment -1-20%	7	7	15	11	9
Treatment - 2-25%	12	17	12	18	18
Treatment-3-30%	14	18	18	19	19
Treatment-4-35%	19	25	22	19	23
Treatment-5-40%	7	10	11	15	11



Solution: Step – 1: Frame the null-hypothesis  $H_0$ : There is no significant variation between the average tensile strengths with different cotton percentages. Or  $V(\mu_i) = 0$ .

Or  $H_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ .

Step – 2: Calculate totals and squares as follows:

$G = 376$   $K = 5$  and  $N = 25$ ,  $T_1 = 49, T_2 = 77, T_3 = 88, T_4 = 108, T_5 = 54$  and  $CF = G^2/N = (376)^2 / 25 = 5,655.04$

$TSS = (7)^2 + (7)^2 + \dots + (11)^2 - CF = 636.96 = B$ .

$SST = (1/5)[(49)^2 + (77)^2 + \dots + (54)^2] - CF = 475.76 = A$ .

Step – 3: Calculate various entries of ANOVA – one way table as follows:

Table (9.4.4). ANOVA table for one way classification.

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (Columns)	$K - 1 = 4$	$S_t^2 = SST = A = 475.76$	$V_1 = S_t^2 / K - 1 = 118.94$	$F = V_1 / V_2 = 14.7568238$
Error	$N - K = 20$	$S_E^2 = ESS = B - A = 636.96 - 475.76 = 161.20$	$V_2 = S_E^2 / N - K = 8.06$	
Total	$N - 1 = 24$	$S_T^2 = TSS = B = 636.96$		

Step – 4: Conclusions:  $F_{cal} = 14.7568238$  and  $F_{tab}$  for (4,20) dos at 5% los is 2.8661 and at 1% los is 4.431. Since,  $F_{cal} > F_{tab}$  at 1% los or  $p < 0.01$ , we reject the null-hypothesis  $H_0$  at 1% los. We conclude that there exists statistically significant variance among various cotton percentage levels with respect to average tensile strengths of synthetic fiber. Hence, we require to apply further statistical analysis, which is explained as follows:

Step – 5: **Further Statistical Analysis- LSD or CD test:**

1. Calculate sample means as follows:

$$\bar{x}_1 = 49/5 = 9.8; \bar{x}_2 = 77/5 = 15.4; \bar{x}_3 = 88/5 = 17.6; \bar{x}_4 = 108/5 = 21.6 \text{ and } \bar{x}_5 = 54/5 = 10.8.$$

2. Calculate Least Significant Difference or Critical Difference CD at  $\sigma\%$  los which is given as follows:

Here  $S_E = \sqrt{(8.06)} = 2.839$  and  $t$  for 20 dos at 5% los is 2.09 and at 1% los is 2.85.

$$CD_{\alpha\%} (\bar{x}_i - \bar{x}_j) = t_{\alpha\%} \text{ for error dof } \times S_E (2/n)^{1/2}.$$

$$CD_{5\%} = 2.09 \times 2.839 (2/5)^{1/2} = 2.09 \times 2.839 \times 0.6324 = 3.7523.$$

Remark: For t values student can refer any fundamental book on statistical methods

$$CD_{1\%} = 2.85 \times 2.839(2/5)^{1/2} = 2.85 \times 2.839 \times 0.6324 = 5.1168.$$

#### Comparisons:

Treatments (t <sub>i</sub> ) or	value	5% Inference	1% Inference
<b>Cotton percentages</b>			
$ \bar{x}_1 - \bar{x}_2  =  9.8 - 15.4 $	5.6	Significant	Significant.
$ \bar{x}_2 - \bar{x}_3  =  15.4 - 17.6 $	2.2	No significant	No significant.
$ \bar{x}_1 - \bar{x}_3  =  9.8 - 17.6 $	7.8	Significant	Significant.
$ \bar{x}_1 - \bar{x}_5  =  9.8 - 10.8 $	1.0	No significant	No significant.
$ \bar{x}_4 - \bar{x}_5  =  21.6 - 10.8 $	10.8	Significant	Significant.
$ \bar{x}_3 - \bar{x}_5  =  17.6 - 10.8 $	6.8	Significant	Significant.
$ \bar{x}_3 - \bar{x}_4  =  17.6 - 21.6 $	4.0	Significant	No significant.

Like above we calculate the difference between treatment means and compare with respective CD values to determine significant effects. According to the above analysis Treatment – 4 is best which maximizes the tensile strength of the synthetic fiber. As per the effects we can arrange the treatments as: T- 4,T-3,T-2,T-5,T-1. Treat – 1 is least preferable and Treatment – 4 is best preferable.

Remark: Further Statistical Analysis is to be applied when we reject the null-hypothesis in ANOVA.

### 9.5. Summary

In this lesson we have learnt ANOVA one way classification, and discussed its analysis, assumptions and properties. Introduction to notations used in ANOVA when we accept or reject the null-hypothesis at 5% and 1% los along with the p-value are introduced. In computer packages or applications p-values are popularly used. Further Statistical Analysis used when the null-hypothesis  $H_0$  is rejected is discussed and the ANOVA – one way classification with unequal and equal sample sizes are discussed with examples. Application of LSD or CD test procedure is also explained with an example.

### 9.6. Self Assessment Questions

1. Explain various assumptions of ANOVA – one way classification.
2. Explain the method of estimating various sum of Squares.
3. Write down ANOVA table for one way classified data.
4. Explain the importance and meaning of p-value and its applications.
5. Explain Further Statistical Analysis and its application.

6. The following data represent the marks obtained in a GK test from four different universities. Analyse with a suitable analysis and apply Further Statistical Analysis if necessary.

University-1:	60	62	63	61	60	64	65	63
University-2:	59	61	60	62	63			
University-3:	61	60	65	64	62	63		
University-4 :	63	65	61	64	63	64	66.	

7. The following data represent the yield of a crop on three varieties measured in tones per acre.

Variety – 1:	10	13	12	15	14	13	13
Variety – 2:	12	12	11	10	12	10	
Variety – 3:	16	18	19	17	18.		

8. An experiment is conducted to measure moisture content of samples of a powder, each man carries a sample of six segments. The assessments are as follows:

Observer	Consignments					
	1	2	3	4	5	6
A	2	11	11	14	12	10
B	9	10	9	10	11	11
C	11	10	10	12	11	10
D	12	11	9	10	11	11.

9. Three varieties of coal were analyzed by four chemists in different laboratories. Ash content in the varieties was found as follows. Test whether the varieties different significantly, with respect to ash content.

Varieties	Ash content.			
V-1	3	6	5	4
V-2	8	5	5	7
V-3	7	6	4	4.

10. Explain ANOVA one way classification procedure applied in practice with an example.

### 9.7. Further Readings.

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons, New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 10.**

## **Analysis of Variance (ANOVA)-Two Way Classification**

**10.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the need of Analysis of variance – Two way classification.
- The concepts of ANOVA – Two way classification.
- Mathematical model of ANOVA – two way classification.
- Assumption of ANOVA – two way classification.
- Analysis of ANOVA two way classification with equal and un-equal number of observations per cell.
- Further models and topics of ANOVA .

**Structure**

- 10.1 Introduction to ANOVA – two way classification**
- 10.2 Mathematical model of ANOVA –two way classification**
- 10.3 Statistical Analysis of ANOVA – two way classification**
- 10.4. Models of ANOVA two way with multiple observations per cell**
- 10.5 Further aspects of ANOVA**
- 10.6 Summary**
- 10.7 Self Assessment Questions**
- 10.8 Further Readings**

**10.1. Introduction to ANOVA – two way classification**

In the last lesson we have discussed ANOVA-one way classification where, we assume that all the units considered for the experiment are uniform in all aspects except the difference in the treatments applied. Thus experimental results contains effects of two factors , namely (i) **Treatment Effects** and (ii) **Random Effects**. This may be possible in laboratory experiments but in many field experiments in agriculture, it is very difficult to get uniform plots. We may select uniform size plots, but they may differ with respect to fertility rate or type of soil or temperature of the soil and so on. Thus there is a need to consider the effects of third factor. Hence, there is a need to consider ANOVA – two way classification where we assume that all the experimental units are not homogeneous. Based on the heterogeneity character, we divide total units into 'Blocks' such that units within the block are assumes as homogeneous. For example patients are divided into various blocks as Children, Adults and Old age people.

Similarly, Male patients and female patients are considered as different blocks or patients suffering similar ailments can be considered as different blocks and so on. Similarly, in Agricultural experiments plots with similar fertility rates are considered as ‘blocks’. Thus heterogeneity may be present between blocks and homogeneity must be present within the block. For example, fertility rates changes in horizontal direction, we have to consider blocks in vertical direction such that plots within the block are having similar fertility rate as shown below:

Fertility rates

High	Medium	Low
Plot-1	Plot-2	Plot-3
Plot-4	Plot-5	Plot-6
Plot-7	Plot-8	Plot-9
Plot-10	Plot-11	Plot-12
<b>Block-I</b>	<b>Block – II</b>	<b>Block – III</b>

Similarly if the fertility rate changes vertically, blocks are to be considered as vertically as shown below.

High	Plot-1	Plot-4	Plot-7	Plot-10	<b>Block –I</b>
Medium	Plot-2	Plot-5	Plot-8	Plot-11	<b>Block – II</b>
Low	Plot-3	Plot-6	Plot-9	Plot-12	<b>Block – III</b>

Thus, units within the block are homogeneous and units in different blocks are not homogeneous. After dividing the heterogeneous experimental units into homogeneous blocks, treatments are to be applied randomly to the units within each block independently.

**10.2. Mathematical model of ANOVA –two way classification- with one observation per cell.**

Let us consider the experiments where there are two factors which may affect the experimental results  $x_{ij}$ ,  $i=1,2,\dots,K$  and  $j=1,2,\dots,n$ . Here,  $x_{ij}$  represents the yield of the plot in  $j^{th}$  block receiving the  $i^{th}$  treatment or yield of milk of the cow receiving  $j^{th}$  ration in  $i^{th}$  treatment and so on. The data under ANOVA two way classified data is arranged as follows:

Treatments/rows	Blocks				Means	Total
	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub> .....B <sub>j</sub> .....	B <sub>n</sub>		
Treatment – 1: t <sub>1</sub>	x <sub>11</sub>	x <sub>12</sub>	x <sub>13</sub> . . . . .	x <sub>1n</sub>	$\bar{x}_1$ .	<b>T<sub>1</sub>.</b>
Treatment – 2: t <sub>2</sub>	x <sub>21</sub>	x <sub>22</sub>	x <sub>23</sub> . . . . .	x <sub>2n</sub>	$\bar{x}_2$ .	<b>T<sub>2</sub>.</b>
Treatment – 3: t <sub>3</sub>	x <sub>31</sub>	x <sub>32</sub>	x <sub>33</sub> . . . . .	x <sub>3n</sub>	$\bar{x}_3$ .	<b>T<sub>3</sub>.</b>

Treatment – i: $t_i$	..... ..... $X_{ij}$ ..... .....				$\bar{x}_i$	$T_i$
Treatment – k: $t_k$	$X_{k1}$	$X_{k2}$	$X_{k3}$ . . . . .	$X_{kn}$	$\bar{x}_k$	$T_k$
Means	$\bar{x}_{.1}$	$\bar{x}_{.2}$	$\bar{x}_{.3}$	$\bar{x}_{.n}$		<b>G</b> = Grand total
Totals	$T_{.1}$	$T_{.2}$	$T_{.3}$	$T_{.n}$		

The mathematical model of ANOVA two way classified data is similar to the model of ANOVA one way where we consider block effects in addition to treatment effects as follows:

$$X_{ij} = \mu_{ij} + \epsilon_{ij} = \mu + \alpha_i + \beta_j + Y_{ij} + \epsilon_{ij} \tag{10.2.1}$$

Where,  $\sum_{i=1}^K \alpha_i = 0 = \sum_{j=1}^n \beta_j$  and  $\sum_{i=1}^K Y_{ij} = \sum_{j=1}^n Y_{ij} = 0$  for all  $i, j$ .

Here,  $Y_{ij}$  represent interaction effects between Factor A and Factor B.

Similarly  $\alpha_i$  represent  $i^{th}$  treatment effect  $i=1,2,\dots,K$  and  $\beta_j$  represent  $j^{th}$  block effect  $j=1,2,\dots,n$ .

As there is only one observation per cell and the observation corresponding to the  $j^{th}$  block receiving  $i^{th}$  treatment, it is not possible to estimate the interaction effect between factors A and B. It is important to note that we cannot estimate anything from single observation in Statistics. At least two observations must be there to get the estimate. More number of observations or replications are preferable, because more replications will reduce the experimental error. This is because of the fact that experimental error and the number of observations is inversely proportional. That is experimental error reduces as number observations increases. This type of estimations of interaction effects between the factors A and B can be done in **ANOVA – two way model with multiple observations per cell**. Thus in this model we consider  $Y_{ij} = 0$ . Thus the model given in (10.2.1) reduces to the following:

$$X_{ij} = \mu_{ij} + \epsilon_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} . \tag{10.2.2}$$

Now we proceed to discuss the Statistical Analysis of ANOVA – two way classified model with single observation per cell in the following section.

### 10.3. Statistical Analysis of ANOVA – two way model with one observation per cell

First we frame the null-hypotheses for ANOVA two way model as follows.

**10.3.1. Null-hypotheses:** We have to frame two null-hypotheses, one for treatment mean effects and another for the block means effects framed as follows.

For Treatments:

**H<sub>01</sub>: There is no significant variation between average treatment effects or**

$$H_{01}: \mu_{1.} = \mu_{2.} = \mu_{3.} = \dots \mu_{K.} = \mu \text{ or } : \alpha_1. = \alpha_2. = \alpha_3. = \dots \alpha_{K.} = 0. \tag{10.3.1}$$

For Blocks:

**H<sub>02</sub>: There is no significant variation between average block effects or**

$$\mathbf{H_{02}: \mu_{\cdot 1} = \mu_{\cdot 2} = \mu_{\cdot 3} = \dots \mu_{\cdot n} = \mu \text{ or } \beta_1 = \beta_2 = \beta_3 \dots = \beta_n = 0.} \quad (10.3.2)$$

### 10.3.2. Calculation and estimation of various parameters and sum of squares.

$$\text{Let us denote } \bar{x}_i = \text{Mean yield of } i^{\text{th}} \text{ treatment} = \sum_{j=1}^n x_{ij} / n; i = 1, 2, \dots, K. \quad (10.3.1)$$

$$\text{Similarly, } \bar{x}_j = \text{Mean yield of } j^{\text{th}} \text{ Block} = \sum_{i=1}^K x_{ij} / K; j = 1, 2, \dots, n. \quad (10.3.2)$$

$$\begin{aligned} \text{And } \bar{x}_{..} = \text{Over all mean} &= [\sum_{i=1}^K \sum_{j=1}^n x_{ij}] / nK = G / nK = \{ \sum_{i=1}^K [\sum_{j=1}^n x_{ij} / n] \} / K \\ &= [\sum_{i=1}^K \bar{x}_i] / K = \{ \sum_{j=1}^n [\sum_{i=1}^K x_{ij} / K] \} / n = [\sum_{j=1}^n \bar{x}_j] / n. \end{aligned} \quad (10.3.3)$$

Now we proceed to obtain the estimates of the parameters  $\mu$ ,  $\alpha_i$ ,  $i=1, 2, \dots, K$  and  $\beta_j$ ,  $j=1, 2, \dots, n$  of the model explained in section 10.3.1 by the method of least squares as follows:

$$\text{Consider } E = \sum_{i=1}^K \sum_{j=1}^n \epsilon_{ij}^2 = \sum_{i=1}^K \sum_{j=1}^n [x_{ij} - \mu - \alpha_i - \beta_j]^2. \quad (10.3.4)$$

The 'normal equations' for estimating the parameters  $\mu$ ,  $\alpha_i$ ,  $i=1, 2, \dots, K$  and  $\beta_j$ ,  $j=1, 2, \dots, n$  are:

$$\partial E / \partial \mu = 0 = -2 \sum_{i=1}^K \sum_{j=1}^n [x_{ij} - \mu - \alpha_i - \beta_j] \quad (10.3.5)$$

$$\partial E / \partial \alpha_i = 0 = -2 \sum_{j=1}^n [x_{ij} - \mu - \alpha_i - \beta_j] \quad i=1, 2, \dots, K. \quad (10.3.6)$$

$$\partial E / \partial \beta_j = 0 = -2 \sum_{i=1}^K [x_{ij} - \mu - \alpha_i - \beta_j] \quad j=1, 2, \dots, n. \quad (10.3.7)$$

$$\text{Under the null hypotheses we know that } \sum_{i=1}^K \alpha_i = \sum_{j=1}^n \beta_j = 0. \quad (10.3.8)$$

$$\hat{\mu} = (1 / nk) \sum_{i=1}^K \sum_{j=1}^n x_{ij} = \bar{x}_{..} \quad (10.3.9)$$

$$\hat{\alpha}_i = (1/n) [\sum_{j=1}^n x_{ij}] - \hat{\mu} = \bar{x}_i - \bar{x}_{..} \quad i=1, 2, \dots, K. \quad (10.3.10)$$

$$\hat{\beta}_j = (1/k) [\sum_{i=1}^K x_{ij}] - \hat{\mu} = \bar{x}_j - \bar{x}_{..} \quad j=1, 2, \dots, n. \quad (10.3.11)$$

Substituting above estimates in the model (10.2.2) we have:

$$x_{ij} = \bar{x}_{..} + (\bar{x}_i - \bar{x}_{..}) + (\bar{x}_j - \bar{x}_{..}) + (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}_{..}) \quad (10.3.12)$$

The error term is so chosen such that RHS is identically equal to LHS in equation (10.3.12).

Take the term  $\bar{x}_{..}$  to the LHS of (10.3.12) summing over  $i = 1, 2, \dots, K$  and  $j = 1, 2, \dots, n$  and squaring on both sides we have:

$$\begin{aligned} \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..})^2 &= \sum_{i=1}^K \sum_{j=1}^n [(x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}_{..}) + (\bar{x}_i - \bar{x}_{..}) + (\bar{x}_j - \bar{x}_{..})]^2 \\ &= \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}_{..})^2 + \sum_{i=1}^K \sum_{j=1}^n (\bar{x}_i - \bar{x}_{..})^2 + \sum_{i=1}^K \sum_{j=1}^n (\bar{x}_j - \bar{x}_{..})^2 \\ &\quad + 2 \sum_{i=1}^K \sum_{j=1}^n (\bar{x}_i - \bar{x}_{..})(x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}_{..}) \\ &\quad + 2 \sum_{i=1}^K \sum_{j=1}^n (\bar{x}_j - \bar{x}_{..})(x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}_{..}) \\ &\quad + 2 \sum_{i=1}^K \sum_{j=1}^n (\bar{x}_i - \bar{x}_{..})(\bar{x}_j - \bar{x}_{..}) \end{aligned} \quad (10.3.13)$$

$$\begin{aligned} \text{Now } \sum_{i=1}^K \sum_{j=1}^n (\bar{x}_i - \bar{x}_{..})(x_{ij} - \bar{x}_i - \bar{x}_j - \bar{x}_{..}) &= \sum_{i=1}^K [(\bar{x}_i - \bar{x}_{..}) \sum_{j=1}^n (x_{ij} - \bar{x}_i - \bar{x}_j - \bar{x}_{..})] \\ &= \sum_{i=1}^K [(\bar{x}_i - \bar{x}_{..}) \{ \sum_{j=1}^n (x_{ij} - \bar{x}_i) - \sum_{j=1}^n (\bar{x}_j - \bar{x}_{..}) \}] = 0 \end{aligned} \quad (10.3.14)$$

Since the algebraic sum of deviations of observations about their mean is always zero. Similarly, all product terms in (10.3.13) will become zero and hence we have:

$$\begin{aligned} \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..})^2 &= n \sum_{i=1}^K (\bar{x}_i - \bar{x}_{..})^2 + K \sum_{j=1}^n (\bar{x}_j - \bar{x}_{..})^2 \\ &+ \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}_{..})^2 \end{aligned} \quad (10.3.15)$$

$$\text{Above equation can also be written as : } S_T^2 = S_t^2 + S_b^2 + S_E^2 \quad (10.3.16)$$

$$\text{Where } S_T^2 = \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..})^2 = \text{Total Sum of Squares (TSS)} \quad (10.3.17)$$

$$S_t^2 = n \sum_{i=1}^K (\bar{x}_i - \bar{x}_{..})^2 = \text{Sum of Squares due to treatments (SST)} \quad (10.3.18)$$

$$S_b^2 = K \sum_{j=1}^n (\bar{x}_j - \bar{x}_{..})^2 = \text{Sum of Squares due to blocks (SSB)} \quad (10.3.19)$$

$$\begin{aligned} \text{And } S_E^2 &= \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}_{..})^2 = \text{Error Sum of Squares (ESS)} \\ &\text{or Residual sum of Squares (RSS)} \end{aligned} \quad (10.3.20)$$

We can also estimate various sum of squares as done in ANOVA one way classification. The difference is that here, we have to consider Block sum of squares additionally in ANOVA two way procedures.

**10.3.3 Degrees of freedom of various sums of squares and ANOVA table.**

Since the total sum of squares  $S_T^2$  is calculated from  $N=nK$  quantities which are subjected to one linear constraint,  $\sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..}) = 0$ , we have  $N - 1$  degrees of freedom for TSS.

Similarly, treatment sum of squares  $S_t^2$  has  $(K-1)$  degrees of freedom, block sum of squares has  $(n-1)$  degrees of freedom and Error sum of squares has  $(n-1)(K-1)$  degrees of freedom.

$$\text{Thus we have } N - 1 = (K-1) + (n-1) + (n-1)(K-1). \quad (10.3.21)$$

After calculating Various sum of squares, we proceed to form the ANOVA table for two way classification . It can be observed that the conditions of Cochran’s theorem are satisfied.

Table (10.3.22) ANOVA table of two factor model.

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F - cal)
Treatments (Rows)	$K - 1$	$S_t^2 = SST = A$	$V_1 = S_t^2 / K - 1$	$F_1 = V_1 / V_3$
Blocks (columns)	$n - 1$	$S_b^2 = SSB = B$	$V_2 = S_b^2 / n - 1$	$F_2 = V_2 / V_3$
Error	$(n-1)(K-1)$	$S_E^2 = ESS = C = D - (B + A)$	$V_3 = S_E^2 / (n-1)(K-1)$	
Total	$nK - 1$ or $N - 1$	$S_T^2 = TSS = D$		



Conclusions:

(1) **For treatments  $H_{01}$ :** Note  $F_{tab}$  at  $\alpha\%$  los at  $[(K-1),(n-1)(K-1)]$  from F – tables given in appendix.

If  $F_1$  calculated value is less than or equal to  $F_{tab}$  for treatments, we accept  $H_{01}$  and conclude that there is no significant variation among various treatment means under consideration. Otherwise, we reject the null hypothesis  $H_{01}$ .

(2)**For Blocks  $H_{02}$ :** Note  $F_{tab}$  at  $\alpha\%$  los at  $[(n-1),(n-1)(K-1)]$  from F – tables given in appendix.

If  $F_2$  calculated value is less than or equal to  $F_{tab}$  for blocks, we accept  $H_{02}$  and conclude that there is no significant variation among various block means under consideration. Otherwise, we reject the null hypothesis  $H_{02}$ .

Note: (1) Here also we can apply 'Further Statistical Analysis' when any hypothesis is rejected.

(2) We can also represent the conclusions using p-values.

(3) Above mentioned analysis is independent of change of origin and change of scale.

Now we proceed to explain various steps involved in practical problems, which reduces lengthy calculations.

#### 10.3.4: Procedure to be adopted in working out practical problems.

Let  $N = nK$  and  $T_i$  represent the  $i^{\text{th}}$  row (treatment) total;  $T_j$  represent  $j^{\text{th}}$  block (column) total and  $G$  represents Grand total as explained in section (10.2).

First frame the null-hypotheses as follows:

$H_{01}$ :  $\mu_1 = \mu_2 = \mu_3 = \dots = \mu_K = \mu$  or :  $\alpha_1 = \alpha_2 = \alpha_3 = \dots = \alpha_K = 0$ . Or there is no significant variation between the treatment means or  $V(\bar{x}_{i.}) = 0$ .

$H_{02}$ :  $\mu_{.1} = \mu_{.2} = \mu_{.3} = \dots = \mu_{.n} = \mu$  or  $\beta_1 = \beta_2 = \beta_3 = \dots = \beta_n = 0$ . Or there is no significant variation between various block means or  $V(\bar{x}_{.j}) = 0$ .

Step – 1: Calculate the **Correction factor (CF) =  $G^2 / N$** .

Step – 2: Calculate Treatment Sum of Squares (SST) =  $\sum_{i=1}^k (T_i)^2 / n - CF = A$ .

Step – 3 : calculate Block Sum of Squares (SSB) =  $\sum_{j=1}^n (T_{.j})^2 / K - CF = B$ .

Step – 4: Calculate Total Sum of squares (TSS) =  $\sum_{i=1}^k \sum_{j=1}^n (x_{ij})^2 - CF = D$ .

Step – 5: Calculate Error Sum of Square (ESS) =  $C = D - (A+B)$

Step – 6: Construct ANOVA table for two factor model as follows:

Table (10.3.23) ANOVA table of two factor model.

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F-cal)
Treatments (Rows)	$K - 1$	$S_t^2 = SST = A$	$V_1 = S_t^2 / K - 1$	$F_1 = V_1 / V_3$
Blocks (columns)	$n - 1$	$S_b^2 = SSB = B$	$V_2 = S_b^2 / n - 1$	$F_2 = V_2 / V_3$
Error	$(n-1)(K-1)$	$S_E^2 = ESS = C = D - (B + A)$	$V_3 = S_E^2 / (n-1)(K-1)$	
Total	$nK - 1$ or $N - 1$	$S_T^2 = TSS = D$		

**Conclusions: For treatments:**

(1) If  $F_1 \leq F_{tab}$  at  $[(K-1), (n-1)(K-1)]$  dof at  $\alpha\%$  los we accept  $H_{01}$  otherwise we reject  $H_{01}$ .

**For blocks:**

(2) If  $F_2 \leq F_{tab}$  at  $[(n-1), (n-1)(K-1)]$  dof at  $\alpha\%$  los we accept  $H_{02}$  otherwise we reject  $H_{02}$ .

Example (10.3.1): The following data represent the yield of paddy of five varieties in different experimental plots of size 10ftX10ft measured in Kgs. Analyze the data and determine whether there exists any variation in mean yields of different varieties or mean block yields?

	Blocks									
	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	B <sub>4</sub>	B <sub>5</sub>	B <sub>6</sub>	B <sub>7</sub>	B <sub>8</sub>	B <sub>9</sub>	B <sub>10</sub>
Variety – 1	97	65	62	99	60	80	76	92	90	89
Variety – 2	95	60	48	87	48	75	71	93	80	93
Variety – 3	99	70	90	99	65	85	75	70	85	92
Variety – 4	98	65	70	95	67	82	73	94	86	90
Variety – 5	96	65	80	95	70	88	70	51	84	91

Solution: First frame null-hypotheses:

$H_{01}$  : There is no significant variation between mean yields of different varieties.

$H_{02}$  : There is no significant variation between mean yields of different blocks.

Now calculate various totals, that is row totals, column totals and grand total G, the correction factor and various sum of squares as follows:

	Blocks										Totals
	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	B <sub>4</sub>	B <sub>5</sub>	B <sub>6</sub>	B <sub>7</sub>	B <sub>8</sub>	B <sub>9</sub>	B <sub>10</sub>	
Variety – 1	97	65	62	99	60	80	76	92	90	89	810

Variety – 2	95	60	48	87	48	75	71	93	80	93	750
Variety – 3	99	70	90	99	65	85	75	70	85	92	830
Variety – 4	98	65	70	95	67	82	73	94	86	90	820
Variety – 5	96	65	80	95	70	88	70	51	84	91	790
Totals	485	325	350	475	310	410	365	400	425	455	4,000=G

Here  $K=5, n=10$  and  $N = 50$  and Correction Factor =  $CF = G^2 / N = (4,000)^2 / 50 = 3,20,000$

Treatment SS =  $SST = (810)^2/10 + (750)^2/10 + (830)^2/10 + (820)^2/10 + (790)^2/10 - CF$   
 $= 400 = A.$

Block SS=  $SSB = (485)^2/5 + \dots + (455)^2/5 - CF = 6,810 = B.$

Total SS =  $TSS = (97)^2 + \dots + (91)^2 - CF = 3,29,948 - 3,20,000 = 9,948 = D.$

Error SS =  $ESS = C=D - (A+B) = 9,948 - (400+6,810) = 2,738.$

Then the ANOVA for two factors model is as follows:

Table (10.3.24) ANOVA table of two factors model.

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (Rows) Or Varieties	$K - 1 = 5 - 1 = 4$	$S_t^2 = SST = A = 400$	$V_1 = S_t^2 / K - 1$ $= 400 / 4 = 100$	$F_1 = V_1 / V_3$ $= 100 / 76.06 = 1.31$
Blocks (columns)	$n - 1 = 10 - 1 = 9$	$S_b^2 = SSB = B = 6,810$	$V_2 = S_b^2 / n - 1$ $= 6,810 / 9 = 756.67$	$F_2 = V_2 / V_3$ $= 756.67 / 76.06 = 9.95^{**}$
Error	$(n-1)(K-1)$ $= 9 \times 4 = 36$	$S_e^2 = ESS = C = D - (B + A)$ $= 2,738$	$V_3 = S_e^2 / (n-1)(K-1)$ $2,738 / 36 = 76.06$	
Total	$nK - 1$ or $N - 1 = 50 - 1 = 49$	$S_T^2 = TSS = D = 9,948$		

**Conclusions: (1) For varieties:** For  $F_1$ , table value of  $F_{tab}$  at [4,36] dof at 5% los is 2.63. Since  $F_1$  calculated value is less than  $F_{tab}$  at 5% los, we accept the null hypothesis  $H_{01}$  at 0.05 los and conclude that there is no statistical or significant variation between the varieties with respect to the average yield of paddy crop under consideration. Here we represent this situation in notation as  $P > 0.05$ .

**(2) For Blocks:** For  $F_2$ , table value of  $F_{tab}$  at [9,36] dof at 5% los is 2.15 and 1% los is 2.98. Since  $F_2$  calculated value is more than  $F_{tab}$  we reject the null hypothesis  $H_{02}$  at 0.01% los and conclude

that there is statistical or significant variation between the blocks with respect to the average yield of paddy crop under consideration. In this case we represent as  $P < 0.01$ .

Since  $H_{01}$  is accepted for treatments, Further Statistical analysis is not necessary for the treatments. Even though the null hypothesis  $H_{02}$  is rejected, blocking is not of that much important, and hence, we are not interested to apply further statistical analysis for blocks. If this is also important to determine, the we have to apply the analysis for blocks by calculating block means and the respective CD at 0.05 or 0.01 los. Now we proceed to explain ANOVA two way with multiple observations per cell in the following section.

**10.4. Models of ANOVA two way with multiple observations per cell**

In the ANOVA two factor model discussed contains only one observation per cell and hence, it is not possible to estimate interaction effects of factors A and B, namely  $Y_{ij}$  discussed in the model (10.2.1). To estimate these interaction effects, we have to consider multiple observations per cell. Here also we have two cases namely:

Case (1): ANOVA two way classified data with multiple but equal number of observations ( say r replications) in all the cells.

Case (2): ANOVA two way classified data with multiple but un-equal number of observations (say  $r_i, i=1,2,\dots,N$  observations) in  $i^{th}$  cell.

Case (1) is a particular case of Case (2) when  $r_1=r_2=\dots,r_N = r$ , that is when r replicates in each cell. The analysis of un-equal number of replicates is similar to that of equal replicates except the difference that in summing of observations we have to consider  $r_i$  replicates in the  $i^{th}$  cell instead of r replicates. In this section we outline the analysis of the ANOVA two way classified data with equal number of observations per cell. The model corresponding to this is explained in section 10.2 and equation (10.2.1) explains the model with interaction effects of factors A and B. The data collected under this set up is given as follows:

Treatments/rows	Blocks				Total
	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub> ..... B <sub>j</sub> .....	B <sub>n</sub>	
Treatment – 1: t <sub>1</sub>	X <sub>111</sub>	X <sub>121</sub>	X <sub>131</sub> . . . . .	X <sub>1n1</sub>	<b>T<sub>1..</sub></b>
	X <sub>112</sub>	X <sub>122</sub>	X <sub>132</sub> . . . . .	X <sub>1n2</sub>	
	.....	.....	.....	.....	
	X <sub>11r</sub>	X <sub>12r</sub>	X <sub>13r</sub>	X <sub>1nr</sub>	
	X <sub>211</sub>	X <sub>221</sub>	X <sub>231</sub> . . . . .	X <sub>2n1</sub>	
	X <sub>212</sub>	X <sub>222</sub>	X <sub>232</sub> . . . . .	X <sub>2n2</sub>	
Treatment – 2: t <sub>2</sub>	.....	.....	.....	.....	<b>T<sub>2..</sub></b>
	X <sub>21r</sub>	X <sub>22r</sub>	X <sub>23r</sub> . . . . .	X <sub>2nr</sub>	

Treatment – 3: $t_3$	$X_{311}$	$X_{321}$	$X_{331}$ . . . . .	$X_{3n1}$	$T_{3..}$
	$X_{312}$	$X_{322}$	$X_{332}$ . . . . .	$X_{3n2}$	
	.....	.....	.....	.....	
Treatment – i: $t_i$	$X_{31r}$	$X_{32r}$	$X_{33r}$ .....	$X_{3nr}$	$T_{i..}$
	.....	.....	.....	.....	
	..... $X_{ij}$ .....	.....	.....	.....	
Treatment – k: $t_k$	$X_{k11}$	$X_{k21}$	$X_{k31}$ . . . . .	$X_{kn1}$	$T_{k..}$
	$X_{k12}$	$X_{k22}$	$X_{k32}$ . . . . .	$X_{kn2}$	
	.....	.....	.....	.....	
Totals	$T_{.1.}$	$T_{.2.}$	$T_{.3.}$	$T_{.n.}$	$G = \text{Grand total} = T_{...}$

In the above table G is the total of all the (Knr) observations. K = number of treatments, n=number of blocks and r=number of replicates. From the above table, first totals all replicates in each cell is to be obtained to get the following table of replicate totals.

The analysis of ANOVA two classified data with multiple but equal number of observations per cell is briefly explained as follows:

**10.4.1. Brief analysis ANOVA two way classified data with ‘r’ number of observations per cell.**

**First obtain the table of cell total of each cell from the data by adding all replicate observations in each cell.**

**Table (10.4.1) Table of totals of all replicates in each cell**

Treatments/rows	Blocks				Total
	$B_1$	$B_2$	$B_3$ ..... $B_j$ .....	$B_n$	
Treatment – 1: $t_1$	$T_{11.}$	$T_{12.}$	$T_{13.}$ . . . . .	$T_{1n.}$	$T_{1..}$
Treatment – 2: $t_2$	$T_{21.}$	$T_{22.}$	$T_{23.}$ . . . . .	$T_{2n.}$	$T_{2..}$
Treatment – 3: $t_3$	$T_{31.}$	$T_{32.}$	$T_{33.}$ . . . . .	$T_{3n.}$	$T_{3..}$

Treatment – i: $t_i$	..... ..... $T_{ij}$ ..... .....				$T_{i..}$
Treatment – k: $t_k$	$T_{k1.}$	$T_{k2.}$	$T_{k3.}$ . . . . .	$T_{kn.}$	$T_{k..}$
Totals	$T_{.1.}$	$T_{.2.}$	$T_{.3.}$	$T_{.n.}$	$G = \text{Grand total} = T_{...}$

Step – 1: Frame the null hypotheses as follows:

**H<sub>01</sub>: There is no significant effect of Factor A or treatments.**

**H<sub>02</sub>: There is no significant effect of Factor B or blocks.**

**H<sub>03</sub>: There is no interaction effects between Factor A and Factor B.**

Step – 2: Calculate  $CF = G^2 / Knr = T_{...}^2 / Knr$

Step – 3: Calculate Treatment Sum of Squares (SST) =  $\sum_{i=1}^k (T_{i..})^2 / nr - CF = A$ .

Step – 4 : Calculate Block Sum of Squares (SSB) =  $\sum_{j=1}^n (T_{.j.})^2 / Kr - CF = B$ .

Step – 5 : Calculate Interaction sum of squares (SSI) =  $\sum_{i=1}^k \sum_{j=1}^n T_{ij}^2 - CF = C$

Step – 6: Calculate Total Sum of squares (TSS) =  $\sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^r (x_{ijl})^2 - CF = D$ .

Step – 7: Calculate Error Sum of Square (ESS) =  $E = D - (A + B + C)$

Step – 8: Construct ANOVA table for two factor model as follows:

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F-cal)
Treatments (Rows)	$K - 1$	$S_t^2 = SST = A$	$V_1 = S_t^2 / K - 1$	$F_1 = V_1 / V_4$
Blocks (columns)	$n - 1$	$S_b^2 = SSB = B$	$V_2 = S_b^2 / n - 1$	$F_2 = V_2 / V_4$
Interaction	$(K-1)(n-1)$	$S_{tb}^2 = SSI = C$	$V_3 = S_{tb}^2 / (K-1)(n-1)$	$F_3 = V_3 / V_4$
Error	$Kn(r-1)$	$S_e^2 = ESS = E = D - (A+B+C)$	$V_4 = S_e^2 / (n-1)(K-1)$	
Total	$Knr - 1$	$S_T^2 = TSS = D$		

**Conclusions:**

**(1) For treatments or Factor A:**

If  $F_1 \leq F_{tab}$  at  $[(K-1), Kn(r-1)]$ dof at  $\alpha\%$  los we accept  $H_{01}$  otherwise we reject  $H_{01}$ .

**(2) For blocks or Factor B:**

If  $F_2 \leq F_{tab}$  at  $[(n-1), Kn(r-1)]$ dof at  $\alpha\%$  los we accept  $H_{02}$  otherwise we reject  $H_{02}$ .

**(3) For interactions between Factor A and Factor B.**

If  $F_3 \leq F_{tab}$  at  $[(n-1), (k-D); Kn(r-1)]$ dof at  $\alpha\%$  los we accept  $H_{03}$  otherwise we reject  $H_{03}$ .

**Example (10.4.1):** The following data represent the effective Battery life measured in hours, manufactured with different material tested at different temperatures. In each category four randomly selected samples of size 4 are tested. Test the effects of material, temperature and also test whether there exists any interaction effects between material and temperature are present or not? The data obtained from the Battery experiment is given as follows:

**Table (10.4.2) Life times of Batteries measured in hours.**

		Temperatures ( <sup>0</sup> F)		
		20	80	120
Material Type	A	130,155,74,180	34,40,80,75	20,70,82,58
	B	150,188,159,126	136,122,106,115	25,70,58,45
	C	138,110,168,160	174,120,150,139	96,104,82,60

**Solution:** First frame the null-hypotheses and then calculate the table of totals of each cell, column and row totals and Grand total G as follows:

$H_{01}$ : There is no significant effect of Factor A or treatments or material types.

$H_{02}$ : There is no significant effect of Factor B or blocks or temperatures.

$H_{03}$ : There are no interaction effects between Factor A and Factor B.

**Table (10.4.3) Table of totals.**

		Temperatures ( <sup>0</sup> F)			Totals
		20	80	120	
Material Type	A	539	229	230	998
	B	623	479	198	1300
	C	576	583	342	1501
Totals		1738	1291	770	G = 3799

Here,  $K=3$ ,  $n=3$  and  $r = 4$ . Then calculate CF and various SS.

$$CF = G^2 / Knr = (3799)^2 / 36 = 77,646.97.$$

$$SST = [(998)^2 + (1300)^2 + (1501)^2] / 12 - CF = 10,683.72 = A$$

$$SSB = [(1738)^2 + (1291)^2 + (770)^2] / 12 - CF = 39,118.72 = B$$

$$SSI = [(539)^2 + (229)^2 + \dots + (342)^2] / 4 - CF = 10,683.72 = C$$

$$TSS = [(130)^2 + (155)^2 + \dots + (60)^2] - CF = 77,646.97 = D.$$

$$SSE = D - (A+B+C) = 18,230.75.$$

**Table (10.4.4) ANOVA table of two way classified data with multiple but equal number of observations per cell.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F-cal)
Treatments (Rows)	$K - 1 = 2$	$S_t^2 = SST = A = 10,683.72$	$V_1 = S_t^2 / K - 1$ $= 5,341.86$	$F_1 = V_1 / V_4$ $7.91^{**}$
Blocks (columns)	$n - 1 = 2$	$S_b^2 = SSB = B = 39,118.72$	$V_2 = S_b^2 / n - 1$ $= 19,559.36$	$F_2 = V_2 / V_4$ $28.97^{**}$
Interaction	$(K-1)(n-1) = 4$	$S_{tb}^2 = SSI = C = 10,683.72$	$V_3 = S_{tb}^2 / (K-1)(n-1)$ $= 2,403.44$	$F_3 = V_3 / V_4$ $3.56^*$
Error	$Kn(r-1) = 27$	$S_E^2 = ESS = E = D - (A+B+C)$ $= 18,230.75$	$V_4 = S_E^2 / kn (r-1)$ $= 675.21$	
Total	$Knr - 1 = 35$	$S_T^2 = TSS = D = 77,646.97$		

#### Conclusions:

**(1) For Materials or Factor A:** For  $F_1$ , table value of  $F_{tab}$  at [2,27] dof at 1% los is 5.49. Since  $F_1$  calculated value is greater than  $F_{tab}$  at 1% los, we reject the null hypothesis  $H_{01}$  at 0.01 % los and conclude that there is statistical or significant variation between the materials with respect to the average lifetimes of Batteries under consideration. Here we represent this situation in p-notation as  $P < 0.01$ .

**(2) For Blocks or Temperatures or Factor B:** For  $F_2$ , table value of  $F_{tab}$  at [2,27] dof at 1% los is 5.49. Since  $F_2$  calculated value is more than  $F_{tab}$  we reject the null hypothesis  $H_{02}$  at 0.01% los and conclude that there is statistical or significant variation between the blocks temperatures with respect to the average lifetimes of Batteries under consideration. In this case we represent as  $P < 0.01$ .

**(3) For Interactions between factors A and B Materials and Temperatures:** For  $F_3$ , table value of  $F_{tab}$  at [4,27] dof at 5% los is 2.73 and 1% los is 4.11. Since  $F_3$  calculated value is more than  $F_{tab}$  at 5% los and less than  $F_{tab}$  at 1% los hence we reject the null hypothesis  $H_{02}$  at 0.05 los and conclude that there is statistical or significant interaction effects between the Factors A and B or Materials and Temperatures with respect to the average lifetimes of Batteries under consideration. In this case we represent as  $P < 0.05$ .

Here also we can apply further statistical analysis in all the above three cases because we gave rejected all the three hypotheses by calculating CD test.



On similar lines we can analyze the two way model with multiple but un-equal number of observations. While calculating the Interaction sum of squares, we have to divide the cell total squares with the corresponding number of observation of that cell. Except this change, all other calculations runs on similar lines of ANOVA two way classified data, with equal number of observations per cell.

### 10.5 Further aspects of ANOVA

We can extend the ANOV two way classified data to **three way** and **higher way** on similar lines explained in this lesson. That is we can extend the analysis for three Factors A,B and C and their interactions. Interaction effects are many types and the detailed discussion of interaction effectors are discussed in the last unit while discussing “**Factorial Experiments**”. In factorial experiments, we discuss Main effects of different factors, various interaction effects like, Interaction effects of two factors and interaction effects of three factors and so on are discussed in detail in the last Unit.

### 10.6 Summary

In this lesson we have discussed ANOVA two way classification data, which is the extension of ANOVA one way classification. When all the experimental units are uniform, we can apply one way or one factor analysis. When the experimental units are heterogeneous, we have to divide the experimental units into blocks such that units within the Block are homogeneous. Then we have to consider the Block or Factor B in addition to Factor A or treatments. This analysis where we estimate the effects of Two Factors is known as ANOVA two factor analysis. First we have discussed ANOVA two factor analyses with single observation per cell. Then the analysis extended to the ANOVA two factors with multiple observations per cell. Here also we have two cases in multiple observations per cell. Each Cell contains multiple but equal number of observations per cell and the second case is that multiple but unequal number of observations per cell. Analysis all the above models are discussed in this lesson along with Further Statistical Analysis and further aspects like three factor analysis is discussed which is the extension of the analysis of two factor ANOVA.

### 10.7. Self Assessment Questions

1. Explain the need of ANOVA two factor analysis, with suitable examples.
2. Explain the method of estimation of various sum of squares of ANOVA two way classified data with one observation per cell.
3. Write down the ANOVA table for ANOVA two factor model with one observation per cell.
4. Write down the ANOVA table of ANOVA two factor model with multiple but equal number of observations per cell.
5. Write down the ANOVA table of ANOVA two factor model with multiple but un-equal number of observations per cell.

6. Following data represent the marks of students in 10<sup>th</sup> class in three different schools in a five different towns. Test whether any significant variation between Schools or significant variation in different towns. If necessary apply CD test and draw your conclusions.

		Towns				
		T1	T2	T3	T4	T5
Schools	A	999	997	996	995	998
	B	997	994	995	994	995
	C	996	995	996	997	996

7. The following data represent the yield of milk measured in lts. in three varieties of cows fed on four different rations.

		Rations			
		R1	R2	R3	R4
Varieties	V1	19	20	20	21
	V2	12	24	21	22
	V3	15	26	22	20

Apply suitable analysis and draw your conclusions.

8. A machine engineer is studying thrust force developed by a drill press. He suspects that the drilling speed and the feed rate of the material are most important factors. He selects four feed rates and uses high (200) and low (100) drill speed chosen to represent operating conditions. The results are tabulate as follows:

		Feed rates			
Drill Speed		0.015	0.045	0.060	0.080
High (200)		2.86,2.83	2.80,2.85	2.87,2.86	2.88,2.94
Low (100)		2.78,2.70	2.49,2.45	2.72,2.60	2.86,2.75

Test whether any significant effects speed rates and feed rates. Also test whether there exists any interaction effects between feed rates and the speed of the drilling machines.

8. The factors that influence the breaking strength of a synthetic fibber are being studied. Four production machines in three different shifts of workers are considered and data is collected as follows:

		Machines			
		M1	M2	M3	M4
Shifts					
I batch		109,110	112,115	106,109	110,109

II batch	114,116	115,118	117,120	115,119
III batch	110,116	118,121	110,111	110,109

Examine whether shifts and machines have any interaction effects along the effects of machines and effects of shifts.

9. An experiment is conducted to examine the influence of three operating temperatures and three types of face plate glass in the light output of oscilloscope tube. Test results are tabulated as follows:

Temperature ( <sup>0</sup> F)	Glass Type		
	A	B	C
100	546,575,599	570,568,580	579,530,550
120	1000,1035,1070	1090,1087,1085	1066,1053,1045
140	1386,1380,1392	1318,1328,1310	867,995,876.

Test whether glass type and temperature has any interaction effects. Similarly test whether temperature difference and glass type has any significant effects individually.

### 10.8 Further Readings

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons., New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 11.****Analysis of Covariance (ANCOVA)- One factor Model****11.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the need of Analysis of Co-variance (ANCOVA).
- The concepts of ANCOVA – one way classification.
- Mathematical model of ANCOVA – one way classification.
- Analysis of ANCOVA – one way classification.

**Structure****11.1. Introduction to ANCOVA****11.2. Mathematical model of ANCOVA –one way classification****11.3. Statistical Analysis of ANCOVA – one way classification****11.4. Example of ANCOVA – one factor model****11.5. Summary****11.6. Self Assessment Questions****11.7. Further Readings****11.1. Introduction to ANCOVA**

In the last lesson, we have introduced ANOVA two factors models which is to be applied when the experimental units are heterogeneous. The variation present in the experimental units may decrease the efficiency of the estimates. To avoid this, variation present in the experimental units are attributed as variation between blocks if we divide the heterogeneous units into blocks such that units within the block are homogeneous and any heterogeneity is classified as variation between blocks. Thus using this technique, we can reduce the experimental error or nuisance factor. Experimental errors are also known as Mean square errors or Random errors. Our objective is to reduce these errors so that the estimates will become more efficient. ANCOVA is one such tool which reduces further these Random errors.

Analysis of Co-variance (ANCOVA) is another technique that is occasionally useful to reduce the experimental and hence, improves the precision of the estimates of various factors average effects and their interactions. Suppose that in an experiment with the main response variable  $y$  there is another variable  $x$  which is linearly related with the main variable  $y$ . Further, it is important to note that the variable  $x$  cannot be controlled by the experimenter, but can be observed along with the main variable  $y$ . Such variable  $x$  is known as Concomitant or co-variate or independent variable on which the main variable  $y$  is depending. Using the information on  $y$  and  $x$  we can reduce further the experimental error in using ANCOVA. **Thus Analysis of co- variance (ANCOVA) is a statistical technique where, we combine the Analysis of variance (ANOVA) with the Regression Analysis.** Thus, ANCOVA technique involves adjusting

the observed response or main variable for the effect of concomitant variable. If such adjustment is not performed, the concomitant variable will inflate the error mean square or experimental error and make true differences in the response due to treatments will become difficult to detect or to estimate. Now we proceed to introduce the mathematical model of ANCOVA in the following section.

### 11.2. Mathematical model of ANCOVA –one way classification

Let  $y_{ij}$  represent  $j^{\text{th}}$  observation on the response variable belonging to the  $i^{\text{th}}$  treatment and  $x_{ij}$  is the observation made on the concomitant variate corresponding to  $y_{ij}$ . Let  $\bar{x}_{..}$  is the mean of the  $x_{ij}$  values and  $\mu$  is the overall mean of the main response variable  $Y$ . Let  $\tau_i$  is the effect of  $i^{\text{th}}$  treatment and  $\beta$  is the linear regression coefficient between the main response variable  $Y$  and the concomitant variable  $X$ . Let  $\epsilon_{ij}$  is the random error component. Then ANCOVA one factor model is given by:

$$y_{ij} = \mu + \tau_i + \beta (x_{ij} - \bar{x}_{..}) + \epsilon_{ij}, i=1,2,\dots,K \text{ and } j = 1,2,\dots,n \quad (11.2.1)$$

Assumptions: The ANCOVA one factor model has the following assumptions:

1. The regression coefficient or the slope  $\beta \neq 0$ .
2. The relation between  $X$  and  $Y$  is linear.
3. The errors  $\epsilon_{ij}$  independently and identically normally distributed with mean 0 and variance  $\sigma^2$ .
4. The regression coefficients for each treatment are identical that the treatment effects sum to zero. That is  $\sum_{i=1}^K \tau_i = 0$ .
5. The concomitant variable  $X$  is not affected by the treatments.

**It is important to note that the model given in (11.2.1) contains both ANOVA model of ANOVA one factor analysis and the linear model employed in the regression analysis.** That is, we have treatment effects  $\{\tau_i\}$  as in single factor ANOVA and a regression coefficient  $\beta$  as in linear regression line representing the 'slope' or the regression coefficient of  $Y$  on  $X$ . The concomitant variable  $X$  can be expressed as  $(x_{ij} - \bar{x}_{..})$  instead of  $x_{ij}$ , so that the parameter  $\mu$  is preserved as the overall mean of the main variable  $Y$ .

### 11.3. Statistical analysis of ANCOVA one way classification

The data for the ANCOVA for one factor model will be represented as follows:

Treatments.	Data	Totals
Treatment – 1:	$(y_{11}, x_{11}) (y_{12}, x_{12}) (y_{13}, x_{13}) \dots, \dots, (y_{1n}, x_{1n})$	$(y_{1.}, x_{1.})$
Treatment – 2:	$(y_{21}, x_{21}) (y_{22}, x_{22}) (y_{23}, x_{23}) \dots, \dots, (y_{2n}, x_{2n})$	$(y_{2.}, x_{2.})$
Treatment – i:	$(y_{i1}, x_{i1}) (y_{i2}, x_{i2}) (y_{i3}, x_{i3}) \dots, \dots, (y_{in}, x_{in})$	$(y_{i.}, x_{i.})$
Treatment – n:	$(y_{n1}, x_{n1}) (y_{n2}, x_{n2}) (y_{n3}, x_{n3}) \dots, \dots, (y_{nn}, x_{nn})$	$(y_{n.}, x_{n.})$
	Grand Total	$(y_{..}, x_{..})$

Now, we proceed to introduce various other notations used in the ANCOVA one factor analysis.

$$S_{yy} = \sum_{i=1}^K \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^K \sum_{j=1}^n y_{ij}^2 - y^2_{..} / Kn. \tag{11.3.1}$$

$$S_{xx} = \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..})^2 = \sum_{i=1}^K \sum_{j=1}^n x_{ij}^2 - x^2_{..} / Kn. \tag{11.3.2}$$

$$S_{xy} = \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..}) (y_{ij} - \bar{y}_{..}) = \sum_{i=1}^K \sum_{j=1}^n x_{ij}y_{ij} - (x_{..})(y_{..})/Kn. \tag{11.3.3}$$

$$T_{yy} = n \sum_{i=1}^K (\bar{y}_i - \bar{y}_{..})^2 = (1/n) \sum_{i=1}^K y_i^2 - y^2_{..} / Kn. \tag{11.3.4}$$

$$T_{xx} = n \sum_{i=1}^K (\bar{x}_i - \bar{x}_{..})^2 = (1/n) \sum_{i=1}^K x_i^2 - x^2_{..} / Kn. \tag{11.3.5}$$

$$T_{xy} = n \sum_{i=1}^K (\bar{x}_i - \bar{x}_{..}) (\bar{y}_i - \bar{y}_{..}) = (1/n) \sum_{i=1}^K x_i y_i - (x_{..})(y_{..})/Kn. \tag{11.3.6}$$

$$E_{yy} = \sum_{i=1}^K \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 = S_{yy} - T_{yy}. \tag{11.3.7}$$

$$E_{xx} = \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 = S_{xx} - T_{xx}. \tag{11.3.8}$$

$$E_{xy} = n \sum_{j=1}^n (x_{ij} - \bar{x}_i) (y_{ij} - \bar{y}_i) = S_{xy} - T_{xy}. \tag{11.3.9}$$

It is important to note that in general,  $S = T + E$ , where the symbols  $S, T$  and  $E$  are used to denote sum of squares and cross products for total, treatments and error, respectively. Further, the sum of squares of  $x$  and  $y$  must always be non-negative but, the sums of cross products ( $xy$ ) may be negative.

Now we proceed to estimate the parameters  $\mu$ ,  $\tau$  and  $\beta$  of the model given in (11.2.1) using method of least squares as done in earlier lessons to estimate parameters of those models. If we proceed on those lines we have:

$$\hat{\mu} = \bar{y}_{..}, \hat{\tau}_i = \bar{y}_i - \bar{y}_{..} - \hat{\beta} (\bar{x}_i - \bar{x}_{..}) \text{ and } \hat{\beta} = E_{xy} / E_{xx} \tag{11.3.10}$$

Further, the error sum of squares in this model is given by:

$$SS_E = E_{yy} - (E_{xy})^2 / E_{xx}, \text{ with } K(n-1) - 1 \text{ degrees of freedom.} \tag{11.3.11}$$

In addition to the testing of null-hypothesis  $H_0$ : There is no significant difference in the treatment effects, we frequently find it useful in interpreting the data to present the adjusted treatment means. These adjusted means computed as follows:

$$\text{Adjusted } \bar{y}_i = \bar{y}_i - \hat{\beta} (\bar{x}_i - \bar{x}_{..}), i=1,2,\dots,K.$$

This adjusted treatment mean is the least square estimator of  $\mu + \tau_i$   $i=1,2,\dots,K$  in the model given in (11.2.1). Various calculations of calculating adjusted regression are explained in the following ANCOVA table.

**Table (11.3.1) ANCOVA table for two factor model.**

SOV (1)	DOF (2)	SS and products			Regression Coefficient (4)	SS due to Reg. coeff. (5)
		X	XY	Y (3)		
Treatments	$K - 1$	$T_{xx}$	$T_{xy}$	$T_{yy}$	-----	-----
Error	$K(n-1)$	$E_{xx}$	$E_{xy}$	$E_{yy}$	$\hat{\beta} = E_{xy}/E_{xx}$	$\hat{\beta}(E_{xy})$
Total	$Kn - 1$	$S_{xx}$	$S_{xy}$	$S_{yy}$	-----	-----

Adjusted treatments	$K(n-1)+(K-1)$	$E_{xx}+T_{xx}$ $E_{xy}+T_{xy}$ $E_{yy}+T_{yy}$	$\hat{\beta}=(E_{xy}+T_{xy})/(E_{xx}+T_{xx})$	$\hat{\beta}(E_{xy}+T_{xy})$
---------------------	----------------	---	---	------------------------------

**Table (11.3.1) ANCOVA table for two factor model (cont.)**

Adjusted regression table.

SOV	Adjusted SS (6)	Adjusted DOF (7)	Adjusted MSS (8)	F-calculated (9)
Treatments	$T'_{yy}=C'_{yy} - E'_{yy}$	$K - 1$	$V_1=T'_{yy} / K-1$	$F_1 = V_1/V_2$
Error	$E'_{yy}= E_{yy} - \hat{\beta}E_{xy}$	$K(n-1)-1$	$V_2=E'_{yy} / [(K-1)(n-1)-1]$	
Total	-----	-----		
Adjusted treatments	$(E_{yy}+T_{yy}) - \hat{\beta}(E_{xy}+T_{xy})=C'_{yy}$	$K(n-1)+ (K-1)-1.$	-----	

$$\text{The required } F_{\text{cal}} = V_2 / V_1 = [SS_E / K(n-1)-1] / [(SS'_E - SS_E)/K-1] \quad (11.3.12)$$

Note  $F_{\text{Tab}}$  at  $\alpha\%$  los for  $[(K - 1),(K(n-1)-1)]$  dof. If  $F_{\text{cal}} \leq F_{\text{Tab}}$  We accept the  $H_0$  and conclude that there is no significant variation among treatment effects.

Further, we can also test the significance of regression coefficient  $\beta$ , by considering  $H_0 : \beta = 0$  and  $F_{\text{cal}} = [(Exy)^2/Exx] / V_1$ .

This calculated value of F is tested against  $F_{\text{Tab}}$  at  $\alpha\%$  los for  $[1,(K(n-1)-1)]$ dof. If  $F_{\text{cal}} > F_{\text{Tab}}$  we reject the null-hypothesis and conclude that the observed regression coefficient  $\beta$  is significant.

Now we proceed to explain the ANCOVA one factor analysis with an example.

#### 11.4. ANCOVA – one factor Model Example

Now we proceed to explain the method of analysis discussed above with an example as follows:

Example (11.4.1): The following data represent the breaking strength of yarn Y measured in pounds manufactured from 3 different machines. Let the concomitant variable X is the

diameter of the yarn measured  $10^{-3}$  inches. Apply the ANCOVA analysis to remove the effect of thickness X on the strength Y when testing for difference in strength between machines.

Machines					
M-1		M-2		M-3	
x	y	x	y	x	y
40	22	42	25	32	15
44	28	49	32	37	23
45	30	39	24	34	21
39	22	41	25	42	26
48	28	36	20	35	21

**Solution:** In this example,  $k = 3, n = 5$  and if we plot the scatter diagram between X and Y shows that there is strong linear relationship and hence, ANCOVA analysis is suggested to test the significant variation between machines with respect to the strength Y and the diameter X of the yarn. The procedure of ANCOVA is explained in the following discussion.

Step-1: Frame the null-hypotheses as follows:

$H_0$ : There is no significant variation in effects of machines with respect to the breaking strengths of the yarn or  $H_0: \tau_i = 0$ .

$H_0$ : The observed regression coefficient is insignificant. That is  $H_0: \beta = 0$ .

Step-2: Calculate various totals, squares and products as follows.

Table (11.4.1): Table of total.

Machines					
M-1		M-2		M-3	
y	x	y	x	y	x
40	22	42	25	32	15
44	28	49	32	37	23
45	30	39	24	34	21
39	22	41	25	42	26
48	28	36	20	35	21
Totals:	216 130	207 126		180 106	(603= $G_y$ , 362= $G_x$ )



Now we proceed to square each observation and the totals of squares useful in calculating total sum of squares:

Table (11.4.2): Table of squares and totals.

Machines

	M-1		M-2		M-3		
	y	x	y	x	y	x	
	1600	484	1764	625	1024	225	
	1936	784	2401	1024	1369	529	
	2025	900	1521	576	1156	441	
	1521	484	1681	625	1764	676	
	2304	784	1296	400	1225	441	
Totals:	9386	3436	8663	3250	6538	2312	(24587, 8998)

Table (11.4.3): Table of products of X and Y.

Machines

	M-1		M-2		M-3		
	y x		y x		y x		
	880		1050		480		
	1232		1568		851		
	1350		936		714		
	858		1025		1092		
	1344		720		735		
Totals:	5664		5299		3872	14835	

Step-3: calculate various SS as follows:

$$S_{yy} = (40)^2 + \dots + (35)^2 - (603)^2/15 = 24587 - 24240.6 = \mathbf{346.40};$$

$$S_{xx} = (22)^2 + \dots + (21)^2 - (362)^2/15 = 8998 - 8736.2667 = \mathbf{261.73};$$

$$S_{xy} = (40)(22) + \dots + (35)(21) - (603)(362)/15 = 14835 - 14552.4 = \mathbf{282.60};$$

$$T_{yy} = (1/5)[(216)^2 + (207)^2 + (180)^2] - (603)^2/15 = 140.40;$$

$$T_{xx} = (1/5)[(130)^2 + (126)^2 + (106)^2] - (362)^2/15 = 66.13;$$

$$T_{xy} = (1/5)[(130)(216) + (126)(207) + (106)(180)] - [(362)(603)] / 15 = 96.00;$$

$$E_{yy} = S_{yy} - T_{yy} = 346.40 - 140.40 = 206.00;$$

$$E_{xx} = S_{xx} - T_{yy} = 261.73 - 66.13 = 195.60 \text{ and } E_{xy} = S_{xy} - T_{xy} = 282.60 - 96.00 = 186.60.$$

Now we proceed to construct the ANCOVA table for one factor model as follows:

**Table (11.3.1) ANCOVA table for two factor model.**

SOV (1)	DOF (2)	SS and products X            XY            Y (3)	Regression Coefficient (4)	SS due to Reg. coeff. (5)
Treatments	$K - 1 = 3 - 1 = 2$	$T_{xx}$ $T_{xy}$ $T_{yy}$ 66.13    96.00    140.40	-----	-----
Error	$K(n-1) = 12$	$E_{xx}$ $E_{xy}$ $E_{yy}$ 195.60    186.60    206.00	$\hat{\beta} = E_{xy}/E_{xx}$ 186.6/195.6=0.954	$\hat{\beta}(E_{xy}) = 178.02$
Total	$Kn - 1 = 14$	$S_{xx}$ $S_{xy}$ $S_{yy}$ 261.73    282.60    346.40	-----	-----
Adjusted treatments	$K(n-1) + (K-1)$ $12 + 2 = 14$	$E_{xx} + T_{xx}$ $E_{xy} + T_{xy}$ $E_{yy} + T_{yy}$ 261.73    282.60    346.40	$\ddot{\beta} = (E_{xy} + T_{xy}) /$ $(E_{xx} + T_{xx})$ $= 282.6 / 261.73 = 1.08$	$\ddot{\beta}(E_{xy} + T_{xy})$ $= 305.208$

**Table (11.3.1) ANCOVA table for two factor model (cont.)**

Adjusted regression table.

SOV	Adjusted SS (6)	Adjusted DOF (7)	Adjusted MSS (8)	F-calculated (9)
Treatments	$T'_{yy} = C'_{yy} - E'_{yy}$ $= 41.192 - 27.98$	$K - 1 = 2$	$V_1 = T'_{yy} / K - 1$ $= 13.212 / 2 = 6.606$	$F_1 = V_1 / V_2$ $= 6.606 / 2.544$
Error	$= 13.212$	$K(n-1) - 1 = 11$	$V_2 = E'_{yy} / [(K-1)(n-1) - 1]$ $= 27.98 / 11 = 2.544$	$= 2.596698.$
Total	$E'_{yy} = E_{yy} - \hat{\beta}E_{xy}$ 206.00 - 178.02 = 27.98			

	-----	-----		
Adjusted treatments	$(E_{yy} + T_{yy}) - \hat{\beta}(E_{yy} + T_{yy}) = C'_{yy}$ $= 346.40 - 305.208 = 41.192$	$K(n-1) + (K-1) - 1$ $14 - 1 = 13$	-----	

Therefore,  $F_{cal} = V_2 / V_1 = [SS_E / K(n-1) - 1] / [(SS'_E - SS_E) / K - 1] = 6.606 / 2.544 = \mathbf{2.597}$ .

$F_{Tab}$  at 5% los for (2,11) dof we have 3.98. Since,  $F_{cal} < F_{Tab}$ , we accept the null-hypothesis and conclude that there is no significant variation in the production of yarn on the three different machines with respect to the strength of the yarn. In P notation, we can represent the situation as  $P > 0.05$ .

Similarly,  $\hat{\beta} = E_{xy} / E_{xx} = 186.60 / 195.60 = \mathbf{0.954}$  and to test the significance of the regression coefficient, that is  $H_0 : \beta = 0$  calculate

$$F_{cal} = [(186.60)^2 / 195.60] / 2.54 = \mathbf{70.08}$$

$F_{Tab}$  at 1% los for (1,11) dof is 9.65. Since,  $F_{cal} > F_{Tab}$  at 1% los, we Reject  $H_0$  and conclude that there is strong linear relation between the diameter of the yarn X and the Strength of the yarn Y. In P notation we can represent the situation as  $P < 0.01$ .

Now we proceed to explain in the following lesson ANCOVA two factors model, which can be extended on similar lines as discussed ANCOVA one factor model.

#### 11.4. Summary

In this lesson we have discussed the meaning of Analysis of covariance (ANCOVA) which is the combination of ANOVA one way classification analysis along with the linear regression analysis. After explaining the mathematical of ANCOVA model, and its assumptions were discussed. Further, we have discussed the statistical analysis of the model. The analysis is explained with an example.

#### 11.6. Self Assessment Questions.

1. Explain the need of combining ANOVA and Regression Analysis with an example.
2. Define Analysis of Co-variance technique along with its assumptions.
3. Explain ANCOVA one factor model along with its assumptions.
4. Explain the Statistical analysis of ANCOVA one way classification.
5. A soft drink distributor studying the effectiveness of delivery methods. Three different types of hand trucks have been developed and an experiment is performed in the companies laboratory. The variable of interest is the delivery time Y measured in minutes. The concomitant variable is the delivery items weight X measured in kgs. Each hand truck is used for four times and the delivery times along with item weights are recorded as follows:

Hand truck type					
Type-A		Type –B		Type-C.	
x	y	x	y	x	y
40	41	38	40	25	26
40	44	20	18	42	46
24	27	50	53	32	35
35	33	26	22	26	25

Analyze the data and draw your conclusions.

6. Compute the adjusted treatment means and the standard error for the adjusted treatment means using the following table.

SOV	Dof	SS OD x	Product of XY	SS of Y
Treatment	3	1500	1000	650
Error	12	6000	1200	550
Total	15	7500	2200	1200

### 11.7. Further Readings

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons, New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 12.****Analysis of Covariance (ANCOVA)-Two Factor Model****12.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the need of Analysis of Co-variance (ANCOVA) – two way classification.
- Mathematical model of ANCOVA – two way classification.
- Analysis of ANCOVA two way classifications with equal number of observations per cell.
- Example of Analysis of covariance (ANCOVA) two way classifications.

**Structure****12.1 Introduction to ANCOVA****12.2. Mathematical model of ANCOVA –one way classification****12.3. Statistical Analysis of ANCOVA – one way classification****12.4. Example of ANCOVA two-way classification****12.5. Summary****12.6. Self Assessment Questions****12.7. Further Readings****12.1. Introduction to ANCOVA two-way classification**

In the ANCOVA one way classification we have assumed that the units on which we have conducted, the experiment are uniform or homogeneous as far as possible. In fact, this is not a valid assumption particularly when sample are large in size. Naturally sample units are Heterogeneous which may affect the experimental results. Hence, there is a need to divide them into blocks such that units within the block are homogeneous and heterogeneity suspected is identified as variation between the blocks. In such situations, we have to apply ANOVA two way classification analyses. Similarly, if we can identify a concomitant variable  $X$  to the main response variable  $Y$ , we can apply regression analysis along with ANOVA two way analysis. Such combined analysis is known as Analysis of Covariance (ANCOVA) two way analysis and the model consider in such situation is known as **ANCOVA two factor model**.

**12.2. Mathematical model of ANCOVA – two factor Model**

In this model, we have to consider block effects along with the treatments effect and the regression coefficient  $\beta$ . The ANCOVA two factors model is given as follows:

$$y_{ij} = \mu + \tau_i + \alpha_j + \beta (x_{ij} - \bar{x}_{..}) + \epsilon_{ij} \quad i=1,2,\dots,K \text{ and } j = 1,2,\dots,n. \quad (12.2.1)$$

where:  $y_{ij}$  = Observation on the response or main variable of experimental unit receiving  $i^{\text{th}}$  treatment in  $j^{\text{th}}$  block.

$\mu$  = General mean effect of the main variable Y.

$\tau_i$  = Effect of  $i^{\text{th}}$  treatment,  $i = 1, 2, \dots, K$

$\alpha_j$  = Effect of  $j^{\text{th}}$  block,  $j = 1, 2, \dots, n$ .

$x_{ij}$  = Observation on the concomitant variable of the unit receiving  $i^{\text{th}}$  treatment in  $j^{\text{th}}$  block.

$\beta$  = the regression coefficient between the main variable X and concomitant variable Y.

$\epsilon_{ij}$  = Random effects or experimental error the  $i^{\text{th}}$  treatment unit belonging to  $j^{\text{th}}$  block.

The analysis of ANCOVA two factors runs on the similar lines of ANCOVA one factor model except the difference that calculations for blocks are to done additionally.

### 12.3: Statistical Analysis of ANCOVA two factor model.

First we proceed to frame the null hypotheses ANCOVA table for two factor model as follows:

$H_{01}$ : There are no significant effects of factor A or treatment effects or  $H_{01}: \tau_i = 0$ .

$H_{02}$ : There are no significant effects of Factor B or Block effects or  $H_{02}: \alpha_j = 0$ .

The data under this ANCOVA two factors model will be of the following form:

Treatments.	Blocks.					Totals.
	B-1	B-2	B-3	.....	B-n	
Treatment – 1:	$(y_{11}, x_{11})$	$(y_{12}, x_{12})$	$(y_{13}, x_{13})$	.....	$(y_{1n}, x_{1n})$	$(y_{1.}, x_{1.})$
Treatment – 2:	$(y_{21}, x_{21})$	$(y_{22}, x_{22})$	$(y_{23}, x_{23})$	.....	$(y_{2n}, x_{2n})$	$(y_{2.}, x_{2.})$
Treatment – i:	$(y_{i1}, x_{i1})$	$(y_{i2}, x_{i2})$	$(y_{i3}, x_{i3})$	.....	$(y_{in}, x_{in})$	$(y_{i.}, x_{i.})$
Treatment – n:	$(y_{n1}, x_{n1})$	$(y_{n2}, x_{n2})$	$(y_{n3}, x_{n3})$	.....	$(y_{nn}, x_{nn})$	$(y_{n.}, x_{n.})$
Totals	$(y_{.1}, x_{.1})$	$(y_{.2}, x_{.2})$	$(y_{.3}, x_{.3})$	.....	$(y_{.n}, x_{.n})$	$(y_{..}, x_{..}) = (G_{yy}, G_{xx})$

Now Calculate various sum of squares and cross products as follows:

$$S_{yy} = \sum_{i=1}^K \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^K \sum_{j=1}^n y_{ij}^2 - y^2_{..} / Kn. \quad (12.3.1)$$

$$S_{xx} = \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..})^2 = \sum_{i=1}^K \sum_{j=1}^n x_{ij}^2 - x^2_{..} / Kn. \quad (12.3.2)$$

$$S_{xy} = \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..}) (y_{ij} - \bar{y}_{..}) = \sum_{i=1}^K \sum_{j=1}^n x_{ij} y_{ij} - (x_{..})(y_{..}) / Kn. \quad (12.3.3)$$

$$T_{yy} = n \sum_{i=1}^K (\bar{y}_{i.} - \bar{y}_{..})^2 = (1/n) \sum_{i=1}^K y^2_{i.} - y^2_{..} / Kn. \quad (12.3.4)$$

$$T_{xx} = n \sum_{i=1}^K (\bar{x}_i - \bar{x}_{..})^2 = (1/n) \sum_{i=1}^K x_i^2 - x^2_{..} / Kn. \tag{12.3.5}$$

$$T_{xy} = n \sum_{i=1}^K (\bar{x}_i - \bar{x}_{..}) (\bar{y}_i - \bar{y}_{..}) = (1/n) \sum_{i=1}^K x_i y_i - (x_{..})(y_{..})/Kn. \tag{12.3.6}$$

$$B_{yy} = K \sum_{j=1}^n (\bar{y}_j - \bar{y}_{..})^2 = (1/K) \sum_{j=1}^n y^2_{.j} - y^2_{..} / Kn. \tag{12.3.7}$$

$$B_{xx} = K \sum_{j=1}^n (\bar{x}_j - \bar{x}_{..})^2 = (1/K) \sum_{j=1}^n x^2_{.j} - x^2_{..} / Kn. \tag{12.3.8}$$

$$B_{xy} = K \sum_{j=1}^n (\bar{x}_j - \bar{x}_{..}) (\bar{y}_j - \bar{y}_{..}) = (1/K) \sum_{j=1}^n x_j y_j - (x_{..})(y_{..})/Kn. \tag{13.3.9}$$

$$E_{yy} = \sum_{i=1}^K \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 = S_{yy} - T_{yy} - B_{yy}. \tag{12.3.10}$$

$$E_{xx} = \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 = S_{xx} - T_{xx} - B_{xx}. \tag{12.3.11}$$

$$E_{xy} = n \sum_{j=1}^n (x_{ij} - \bar{x}_i) (y_{ij} - \bar{y}_i) = S_{xy} - T_{xy} - B_{xy}. \tag{12.3.12}$$

After calculating all the above Sum of squares and product terms of the variable X and Y

Now we proceed to write the ANCOVA table for two factors model as follows:

Since the ANCOVA two factor model is very lengthy it is divided into two parts as follows:

**Table (12.3.1) ANCOVA table for two factor model.**

SOV (1)	DOF (2)	SS and products			Regression Coefficient (4)	SS due to Reg. coeff. (5)
		X	XY	Y (3)		
Treatments	K - 1	T <sub>xx</sub>	T <sub>xy</sub>	T <sub>yy</sub>	-----	-----
Blocks	n-1	B <sub>xx</sub>	B <sub>xy</sub>	B <sub>yy</sub>	-----	-----
Error	(K-1)(n-1)	E <sub>xx</sub>	E <sub>xy</sub>	E <sub>yy</sub>	$\hat{\beta} = E_{xy}/E_{xx}$	$\hat{\beta}(E_{xy})$
Total	Kn - 1	S <sub>xx</sub>	S <sub>xy</sub>	S <sub>yy</sub>	-----	-----
Adjusted treatments	(K-1)(n-1)+(K-1)	E <sub>xx</sub> +T <sub>xx</sub>	E <sub>xy</sub> +T <sub>xy</sub>	E <sub>yy</sub> +T <sub>yy</sub>	$\hat{\beta} = (E_{xy}+T_{xy}) / (E_{xx}+T_{xx})$	$\hat{\beta}(E_{xy}+T_{xy})$
Adjusted block effects	(K-1)(n-1)+(n-1)	E <sub>xx</sub> +B <sub>xx</sub>	E <sub>xy</sub> +B <sub>xy</sub>	E <sub>yy</sub> +B <sub>yy</sub>	$\hat{\beta} = (E_{xy}+B_{xy}) / (E_{xx}+B_{xx})$	$\hat{\beta}(E_{xy}+B_{xy})$

**Table (12.3.1) ANCOVA table for two factor model (cont.)**

Adjusted regression table.

SOV	Adjusted SS (6)	Adjusted DOF (7)	Adjusted MSS (8)	F-calculated (9)

Treatments	$T'_{yy} = C'_{yy} - E'_{yy}$	$K - 1$	$V_1 = T'_{yy} / K - 1$	$F_1 = V_1 / V_3$
Blocks	$B'_{yy} = D'_{yy} - E'_{yy}$	$n - 1$	$V_2 = B'_{yy} / n - 1$	$F_2 = V_2 / V_3$
Error	$E'_{yy} = E_{yy} - \hat{\beta} E_{xy}$	$(K-1)(n-1) - 1$	$V_3 = E'_{yy} / [(K-1)(n-1) - 1]$	
Total		-----		
Adjusted treatments	$(E_{yy} + T_{yy}) - \hat{\beta}(E_{yy} + T_{yy}) = C'_{yy}$	$(K-1)(n-1) + (K-1) - 1$	-----	
Adjusted block effects	$(E_{yy} + B_{yy}) - \hat{\beta}(E_{yy} + B_{yy}) = D'_{yy}$	$(K-1)(n-1) + (n-1) - 1$	-----	

Conclusions: If  $F_1$  calculated value is less than or equal to  $F_1$  table value at  $\alpha\%$  los for  $[(K-1), (K-1)(n-1)-1]$  dof, we accept the null hypothesis  $H_{01}$  and conclude that the treatment effects or factor A effects are insignificant. Otherwise we reject the null hypothesis  $H_{01}$ .

Similarly, if  $F_2$  calculated value is less than or equal to  $F_2$  table value at  $\alpha\%$  los for  $[(n-1), (K-1)(n-1)-1]$  dof, we accept the null hypothesis  $H_{02}$  and conclude that the block effects or factor B effects are insignificant. Otherwise we reject the null hypothesis  $H_{02}$ .

**12.4.Example of ANCOVA two factor model.**

Above explained ANCOVA two factors model is explained with the following example.

**Example (12.4.1):** The following data represent the yield of a crop Y measured in tones on three varieties of seeds in four different types of fields with different fertility rates. Rainfall measured in mm is the concomitant variable X. Analyze the data with a suitable ANCOVA model to test the effects of varieties of seeds and block effects.

Yiels of the crop (in tones)

	B1		B2		B3		B4	
	y	x	y	x	y	x	y	x
Variety - V1:	40	27	95	48	22	23	45	35
Variety - V2	22	24	67	38	35	28	25	15
Variety - V3	38	29	38	24	23	19	30	20.

**Solution:** In the given problem  $K=3, n=4$ , Factor –A is varieties , Factor – B blocks, Y = main variable yield and X= concomitant variable rainfall. Then, frame the null-hypotheses as follows:



$H_{01}$ : There is no significant variation among Seed varieties.

$H_{02}$ : There is no significant variation among blocks with respect to fertility rates.

Then calculate various totals, like row, columns and grand totals of Y and X as follows:

	Yields of the crop (in tones)									
	B1		B2		B3		B4		Totals	
	y	x	y	x	y	x	y	x	y	x
Variety - V1:	40	27	95	48	22	23	45	35	202	133
Variety – V2:	22	24	67	38	35	28	25	15	149	105
Variety – V3:	38	29	38	24	23	19	30	20	129	92
Totals:	100	80	200	110	80	70	100	70	480= $G_y$	330= $G_x$

Now we proceed to calculate various Sum of squares and products of Y and X as follows:

The following table will help us to calculate total sum of squares.

	Table of squares and their sums.									
	B1		B2		B3		B4		Totals	
	y	x	y	x	y	x	y	x		
Variety - V1:	1600	729	9025	2304	484	529	2025	1225	13134	4787
Variety – V2	484	576	4489	1444	1225	784	625	225	6823	3029
Variety – V3	1444	841	1444	576	529	361	900	400	4317	2178
Totals	3528	2146	14958	4324	2238	1674	3550	1850	24274	9994.

The following table will help us to calculate products of X and Y and their totals.

	Table of products of X and Y and their totals.				
	B1	B2	B3	B4	
	yx	yx	yx	yx	Total
Variety - V1:	1080	4560	506	1575	7721
Variety – V2	528	2546	980	375	4429
Variety – V3	1102	912	437	600	3051
Totals	2710	8018	1923	2550	15201

$$S_{yy} = \sum_{i=1}^K \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^K \sum_{j=1}^n y_{ij}^2 - y^2_{..} / Kn = (40)^2 + (95)^2 + \dots + (30)^2 - (480)^2 / 12$$

$$= 24274 - (480)^2 / 12 = 24274 - 19200 = 5074. \quad (12.4.1)$$

$$S_{xx} = \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..})^2 = \sum_{i=1}^K \sum_{j=1}^n x_{ij}^2 - x^2_{..} / Kn.$$

$$=9994 - (330)^2/12 = 9994 - 9075 = 919. \quad (12.4.2)$$

$$\begin{aligned} S_{xy} &= \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{..}) (y_{ij} - \bar{y}_{..}) = \sum_{i=1}^K \sum_{j=1}^n x_{ij}y_{ij} - (x_{..})(y_{..})/Kn. \\ &= 15201 - (480)(330)/12 = 15201 - 13200 = 2001. \end{aligned} \quad (12.4.3)$$

$$\begin{aligned} T_{yy} &= n \sum_{i=1}^K (\bar{y}_{i.} - \bar{y}_{..})^2 = (1/n) \sum_{i=1}^K y^2_{i.} - y^2_{..} / Kn. \\ &= (1/4)(202)^2 + (149)^2 + (129)^2 - (480)^2/12 = 1/4[40804 + 22201 + 16641] - 19200 \\ &= 19911.5 - 19200 = 711.5. \end{aligned} \quad (12.4.4)$$

$$\begin{aligned} T_{xx} &= n \sum_{i=1}^K (\bar{x}_{i.} - \bar{x}_{..})^2 = (1/n) \sum_{i=1}^K x^2_{i.} - x^2_{..} / Kn. \\ &= (1/4)[(133)^2 + (105)^2 + (92)^2] - (330)^2/12 = 1/4[17689 + 11025 + 8464] - 9075 \\ &= 9294.5 - 9075 = 219.5. \end{aligned} \quad (12.4.5)$$

$$\begin{aligned} T_{xy} &= n \sum_{i=1}^K (\bar{x}_{i.} - \bar{x}_{..}) (\bar{y}_{i.} - \bar{y}_{..}) = (1/n) \sum_{i=1}^K x_{i.} y_{i.} - (x_{..})(y_{..})/Kn. \\ &= (1/4)[(202)(133) + (149)(105) + (129)(92)] - (480)(330)/12 \\ &= (1/4)[26866 + 15645 + 11868] - 13200 = 13549.25 - 13200 = 349.25. \end{aligned} \quad (12.4.6)$$

$$\begin{aligned} B_{yy} &= K \sum_{j=1}^n (\bar{y}_{.j} - \bar{y}_{..})^2 = (1/K) \sum_{j=1}^n y^2_{.j} - y^2_{..} / Kn. \\ &= (1/3)[(100)^2 + (200)^2 + (80)^2 + (100)^2] - (480)^2/12 \\ &= 1/3[10000 + 40000 + 6400 + 10000] - 19200 \\ &= 22133.333 - 19200 = 2933.333. \end{aligned} \quad (12.4.7)$$

$$\begin{aligned} B_{xx} &= K \sum_{j=1}^n (\bar{x}_{.j} - \bar{x}_{..})^2 = (1/K) \sum_{j=1}^n x^2_{.j} - x^2_{..} / Kn. \\ &= (1/3)[(80)^2 + (110)^2 + (70)^2 + (70)^2] - (330)^2/12 = 1/3[6400 + 12100 + 4900 + 4900] - 9075 \\ &= 9433.333 - 9075 = 358.333. \end{aligned} \quad (12.4.8)$$

$$\begin{aligned} B_{xy} &= K \sum_{j=1}^n (\bar{x}_{.j} - \bar{x}_{..}) (\bar{y}_{.j} - \bar{y}_{..}) = (1/K) \sum_{j=1}^n x_{.j} y_{.j} - (x_{..})(y_{..})/Kn. \\ &= (1/3)[(100)(80) + (200)(110) + (80)(70) + (100)(70)] - (480)(330)/12 \\ &= 1/3[8000 + 22000 + 5600 + 7000] - 13200 = 14200 - 13200 = 1000. \end{aligned} \quad (12.4.9)$$

$$E_{yy} = \sum_{i=1}^K \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 = S_{yy} - T_{yy} - B_{yy} = 5074 - 711.5 - 2933.333 = 1429.167. \quad (12.4.10)$$

$$E_{xx} = \sum_{i=1}^K \sum_{j=1}^n (x_{ij} - \bar{x}_{i.})^2 = S_{xx} - T_{xx} - B_{xx} = 919 - 219.5 - 358.333 = 341.167. \quad (12.4.11)$$

$$E_{xy} = n \sum_{j=1}^n (x_{ij} - \bar{x}_{i.}) (y_{ij} - \bar{y}_{i.}) = S_{xy} - T_{xy} - B_{xy} = 2001 - 349.25 - 1000 = 651.75. \quad (12.4.12)$$

Now we proceed to form the ANCOVA table for two factors model as follows:

Table (12.4.2) ANCOVA table for two factor model.

SOV (1)	DOF (2)	SS and products X                      XY                      Y (3)	Regression Coefficient (4)	SS due to Reg. coeff. (5)
Treatments	$K - 1 = 3 - 1 = 2$	$T_{xx}$ $T_{xy}$ $T_{yy}$ <b>219.5    349.25    711.5</b>	-----	-----
Blocks	$n - 1 = 4 - 1 = 3$	$B_{xx}$ $B_{xy}$ $B_{yy}$ <b>358.333    1000    2933.333</b>	-----	-----
Error	$(K - 1)(n - 1) = 6$	$E_{xx}$ $E_{xy}$ $E_{yy}$	$\hat{\beta} = E_{xy}/E_{xx}$	$\hat{\beta}(E_{xy})$
Total	$Kn - 1 = 11$	<b>341.167    651.75    1429.167</b> $S_{xx}$ $S_{xy}$ $S_{yy}$ <b>919    2001    5074</b>	<b>=651.75/ 341.167 = 1.91</b>	=1245.074
Adjusted treatments	$(K - 1)(n - 1) + (K - 1)$ $6 + 2 = 8$	$E_{xx} + T_{xx}$ $E_{xy} + T_{xy}$ $E_{yy} + T_{yy}$ <b>560.667    1001    2140.667</b>	$\hat{\beta} = (E_{xy} + T_{xy}) /$ $(E_{xx} + T_{xx})$ $= 1001 /$ $560.667 = 1.785$	$\hat{\beta}(E_{xy} + T_{xy})$ <b>= 1787.159</b>
Adjusted block effects	$(K - 1)(n - 1) + (n - 1)$ $6 + 3 = 9$	$E_{xx} + B_{xx}$ $E_{xy} + B_{xy}$ $E_{yy} + B_{yy}$ <b>699.5    1651.75    4362.5</b>	$\hat{\beta} = (E_{xy} + B_{xy}) /$ $(E_{xx} + B_{xx})$ $= 1651.75 / 699.5 =$ <b>2.361</b>	$\hat{\beta}(E_{xy} + B_{xy})$ <b>3900.326</b>

Table (12.4.1) ANCOVA table for two factor model (cont.)

Adjusted regression table.

SOV	Adjusted SS (6)	Adjusted DOF (7)	Adjusted MSS (8)	F-calculated (9)
Treatments	$T'_{yy} = C'_{yy} - E'_{yy}$ <b>= 353.508 -</b> <b>184.093 = 169.415</b>	$K - 1 = 2$	$V_1 = T'_{yy} / K - 1$ <b>= 169.415 / 2 = 84.7075</b>	$F_1 = V_1 / V_3$ <b>= 84.7075 / 36.819</b> <b>= 2.3</b>

Blocks	$B'_{yy} = D'_{yy} - E'_{yy}$ <b>462.174</b> - <b>184.093 = 278.081</b>	$n - 1 = 3$	$V_2 = B'_{yy} / n - 1$ <b>= 278.081 / 3 = 92.694</b>	$F_2 = V_2 / V_3$ <b>= 92.694 / 36.819</b> <b>= 2.51</b>
Error	$E'_{yy} = E_{yy} - \hat{\beta} E_{xy}$ <b>= 1429.167</b> - <b>1245.074 = 184.093</b>	$(K-1)(n-1) - 1 = 5$	$V_3 = E'_{yy} / [(K-1)(n-1) - 1]$ <b>= 184.093 / 5 = 36.819</b>	
Total	-----	-----		
Adjusted treatments	$(E_{yy} + T_{yy})$ - $\hat{\beta}(E_{yy} + T_{yy}) = C'_{yy}$ <b>= 2140.667 -</b> <b>1787.159 = 353.508</b>	$(K-1)(n-1) +$ $(K-1) - 1 = 7$	-----	
Adjusted block effects	$(E_{yy} + B_{yy})$ - $\hat{\beta}(E_{yy} + B_{yy}) = D'_{yy}$ <b>= 4362.5</b> - <b>3900.326 =</b> <b>462.174</b>	$(K-1)(n-1) +$ $(n-1) - 1 = 8$	-----	

Conclusions: 1. For treatments: Calculated value of  $F_1 = 2.3$  and the corresponding table value of F at 5% los for (2,5) dof is 5.79. Since,  $F_1$  is less than F table value, we accept  $H_{01}$  at 0.05 los and conclude that all the treatments (varieties) are equally effective on the yield of the crop.

2. for blocks: Calculated value of  $F_2 = 2.51$  and the table value of F at 5% los for (3,5) dof is 5.41. Since,  $F_2$  calculated value is less than  $F_2$  table value, we accept the Null hypothesis  $H_{02}$  and conclude that all the blocks are equal with respect to the yield of the crop.

Since, we have accepted both null hypotheses, there is no need of further Statistical Analysis in this case.

We can also consider ANCOVA two-factor model with multiple observations per cell but the calculations will be large in number and hence, we required computer assistance for those calculations.

## 12.5.Summary

In this lesson we learnt the ANCOVA two factor model. To apply this analysis we have to divide the experimental units into blocks such that experimental error can further be reduced by adjusting the variation between blocks. It is important to note that units within must be homogeneous and any type of heterogeneity is attributed as the difference between units in different blocks. We have introduced the mathematical model of ANCOVA two factor model

and discussed its Statistical analysis. The method of analysis and estimation of various treatment effects, Adjusted treatment effects in two factors model is discussed with an example.

### 12.6. Self Assessment Questions

1. Explain the need of ANCOVA two factor model.
2. Explain the method of estimating the adjusted treatment means in ANCOVA two factor model.
3. Write the ANCOVA table for two factor model.
4. Explain various Sum of squares required in ANCOVA two factor model.
5. An Engineer is studying the effect of cutting speed on the rate of metal removal in a three machine operations. The amount of metal removed  $Y$  measured in grams and the hardness of the material  $X$  are recorded on four different machines and the data is recorded as follows given in the following table. Use appropriate analysis to determine whether there is any effect of cutting speeds and machine types on  $y$ .

	Cutting Speed (rpm)					
	1000		1200		1400	
	y	x	y	x	y	x
Machine – 1:	88	136	85	133	80	130
Machine – 2:	77	125	74	125	92	141
Machine – 3:	68	120	112	165	118	175
Machine – 4:	90	140	94	140	82	132

Also test whether there is any difference among the four machines.

### 12.7. Further Readings

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons, New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 13.****DESIGN OF EXPERIMENTS****13.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the need, concepts and definition of Design of Experiments.
- Basic principles of Experimental Designs.
- Efficiency of a design and uniformity trials.
- Determination of minimum number of replications.
- Determination of the size of the experimental units and its effect on the error.
- Types of Experimental Designs.

**Structure**

- 13.1 Introduction to Design of Experiments**
- 13.2 Basic Principles of Experimental Designs**
- 13.3 Efficiency of a design and Uniformity trials**
- 13.4 Determination of minimum number of replications**
- 13.5 Various types of Design of experiments**
- 13.6 Summary**
- 13.7. Self Assessment Questions**
- 13.8. Further Readings**

**13.1. Introduction to Design of Experiments**

Analysis of variance introduced in the unit – 3 is to be applied only when the assumptions discussed there must be true. Otherwise we cannot apply the analysis for that data. Thus, we need some rules and principles to conduct experiments, so that the statistical analysis discussed earlier is applicable. Such experiments are called as “**Statistical Experiments**” or “**Random Experiments**”. Usually in any experiment, variation in the experimental results is inevitable. This variability is only because of “**Random Causes**”, then it will be very small and can be neglected. Such ignorable variation, which arises because of nuisance factors, is negligible. Further it is important to note that these nuisance factors are unknown and uncontrolled and **is impossible to make the experiment free from such variations or errors.**

Since the random variation is very small we ignore such variation and consider the observations are approximately equal or we consider that there is no significant variation in the experimental results or the observations have statistically insignificant variation. Such ignorable variations are called “**Random Errors**” or “**Experimental Errors**”, which arises because of Random Factors or Nuisance Factors. If such random causes are alone present in the experiment, then we can estimate such effects statistically because random errors follow Statistical laws. Keeping these points in view we have to plan the experiment such that the results obtained from such experiment has Statistical Validity. This means that, we can test the required hypothesis on the experimental results, using Analysis of Variance techniques

discussed in unit – 3. Thus the plan of conducting the experiments, such that the experimental results have Statistical Validity is the subject matter of “**Design of Experiments**”.

Basic objectives of Design of Experiments are:

- (1) Planning of experiments which have Statistical Validity.
- (2) Obtaining experimental results with which we can test the Statistical hypothesis under study and
- (3) Making the Statistical Analysis applicable for the data obtained.

Experience has shown that proper care or consideration of the statistical analysis before the experiment is conducted. Thus it is necessary that the experimenter has to plan more carefully plan and well-designed experiment in advance such that we can draw statistically valid inferences from the results of the experiment. However, the certainty of the conclusion so drawn, regarding the acceptance or rejection of the null hypothesis, is given only in terms of probability. Thus, the design of Experiment is defined as follows:

**Definition (13.1.1): The logical construction of the experiment in which the degree of uncertainty with which the inference is drawn may be well defined.**

Experiments are basically two types, namely (i) Absolute and (ii) Comparative

Experiments. **Absolute experiments** consists in determining the absolute values of some characteristics like (a) Average Intelligent Quotient (I.Q) of a group of people (b) Finding the association or correlation coefficient  $r$  between two attributes or variables and (c) Effectiveness of medicine and the dosage and so on. On the other hand, **Comparative experiments** are designed to compare the effects of two or more treatments or objects on some population characteristic like yield of a crop. For example, we can study or compare two or more fertilizers or manures on the yield of crops. We can study different varieties of a crop or different cultivation methods or different irrigation methods different diets or different dosages of a medicine for a disease and so on. We can also study the effectiveness of training of employees and the industrial output or quality of the product and so on.

Frequently, we use the following terms in Design of Experiments, which require some Clarification as follows:

**13.1.1. Experimental unit:** In agricultural experiments, the smallest division of the experimental material to which we can apply the treatments and on which we make observations on the variable under study is termed as “**Experimental Unit**”. In other experiments like animal husbandry, a horse, a dog or a pig or a lactating animal or a fish will become experimental unit on which treatments are applied. In medical research, a patient suffering from a disease, organs affected in a patient will become the experimental unit. In business problems, a customer or a share of a company or a T.V or Computer of brand sold and so on.

**13.1.2. Treatment:** Various factors or objects of comparison in comparative experiments are known as “**Treatments**”. For example in agricultural experiments, different varieties seeds or different methods of cultivation or different irrigation methods, different fertilizers, different climatic conditions or different fertility rates can be considered as treatments. In industry,

items produced in different machines or from different raw material or items produced from different batches/workers, different designs, can be considered as treatments. Similarly, in medical research, different treatments or medicine or dosages can be considered as treatments. In animal husbandry different rations, different treatments or different fodders can be considered as treatments.

**13.1.3. Yield:** The measurement of the variable under study on different experimental units is known as “**yield**”. For example crop in a unit of land or milk produced from a cow or temperature or blood sugar or B.P of a patient or number of items produced from a machine per day or price of a commodity and so on.

**13.1.4: Block:** In agricultural experiments, most of the time, we divide the whole heterogeneous experimental unit (field) will be divided into relatively homogeneous sub-groups such that units within the sub-group are homogeneous. Such homogeneous sub-group of units is known as “**block**”. For example, patients suffering from a disease can be grouped into male and female patients. Similarly, they can be classified as Children, Adults and old patients and so on can be considered as blocks. Units produced from the same machine can be one blocks. Animals having same age group can be considered as a block. Customers belonging to a city or share holders of the same company can be considered as a block. Animals of the same breed in the same age can be considered as a block.

## 13.2. Basic Principles of Experimental Designs

According to “**Father of Statistics**”, Prof. Ronal A. Fisher, the following are the three basic principles of experimental designs.

- (1) Replication,
- (2) Randomization and
- (3) Local Conttrol.

**13.2.1: Replication:** Replication means, the repetition of treatments under investigation on different experimental units. An experimenter resorts to replication in order to average out the influence of the random factors or chance factors or experimental errors on different experimental units. Thus replication of treatments will provide more reliable estimate of the treatments effect, than with the single observation. Replication has the following advantages:

(a) Basically replication reduces the experimental error and hence, it enables the experimenter to obtain more precise estimates for the treatment effects. This can be justified as follows. We know from statistical theory that the Standard Error (SE) of the sample mean of a sample of size  $n$  is  $\sigma/\sqrt{n}$ , where  $\sigma$  is the standard deviation (per unit) of the population. Applying this result for ‘ $r$ ’ replicates for a treatment, S.E of its mean is  $\sigma/\sqrt{r}$ , where  $\sigma^2$  is the variance of the individual plot is estimated from the “**Error Variance**”. This implies that as replicates ‘ $r$ ’ increases, the S.E of the estimate reduces. This implies that **the precession of the experiment is inversely proportional to the square root of the replications**. Thus replications play an important role of increasing the efficiency of the experiment. **Thus replication increases the efficiency of the estimate of treatment effect.**



(b) Another important role of replication is to provide an estimate of the experimental error, without which we cannot test the significance difference between two treatment effects.

(c) it is desirable to have much uniformity or homogeneity as far as possible, within each replication but it is not important to have a great deal of uniformity between replications.

(d) A general rule to determine number of replications that the error degrees of freedom must be at least 12. Usually one should not use less than 4 replications.

**13.2.2. Randomization:** Application of treatments to the experimental units must be determined purely on random basis. In order to have statistical validity of experimental results, treatments are to be applied purely on random basis. That is allocation of treatments to experimental units so that each treatment gets an equal chance of showing its worth. This can be done only when the treatments are applied purely on random manner to the experimental material. **Thus randomization makes the estimated treatment effects as unbiased estimates.** The basic objectives of randomization are:

(a) To get the validity of the statistical tests of significance. For example, t- test for testing the significance difference between two treatment mean effects and F – test for testing the homogeneity of several means as discussed in ANOVA. Randomization provides logical basis to apply probability theory, statistical distributions and makes it possible to draw rigorous inductive inferences on the treatment effects. Randomizing treatments over the experimental units is an essential safeguard against the deviation of experimental results by un-anticipated factors like climatic conditions, instrumental or equipment errors, fertility of the soil, temperature conditions and so on. **Thus randomness makes the estimates obtained from experimental results as unbiased estimates.**

(b) The purpose of randomization is to assure that the sources of variation, not controlled in the experiment, operate randomly so that the average effect on any group of units is zero. In other words, randomization eliminates the bias present in any form in conducting the experiment and makes the estimates of the treatment effects as un-biased. It equalizes even factors of variation over which the experimenter has no control.

**(3) Local control:** If the experimental material, say agriculture field in agricultural experiments or different machines in industry or patients belonging to different age groups are heterogeneous, and different treatments are applied to different experimental units at random over entire experimental material, the heterogeneity present in the experimental material enters in the experimental results and thus increases the experimental error considerably. It is desirable to decrease the experimental error as smaller as possible. This can be done by dividing heterogeneous experimental units into sub-groups such that units within the sub-group are homogeneous as far as possible. Such homogeneous sub-groups are known as '**Blocks**'. The process of reducing the experimental error by dividing heterogeneous units into homogeneous sub-groups is known as '**Local Control**'. **Thus local control enables the experimenter to reduce the experimental error as small as possible.**

Above explained three principles are known as "**Basic Principles of Experimental Designs**" which are essential principles. Without these, we cannot apply statistical theory and we will not be in a position to apply statistical tests, like t-test or F-test, for testing the

statistical significance. Basically, designs of experiments are introduced by Prof. R.A. Fisher, to solve problems which arise in Agriculture. Recently, they are used in many other fields like medical, industrial production, market research, meteorology, biological studies, management and so on.

Now we proceed to explain some related concepts of experimental designs in the following section.

### 13.3. Efficiency of a design, Uniformity trials and size of the experimental units

Now we proceed to explain some basic concepts of experimental designs in the following sub-sections.

#### 13.3.1. Efficiency of a Design:

Let  $D_1$  and  $D_2$  be two designs, with error variances per unit  $\sigma_1^2$  and  $\sigma_2^2$  and replications  $r_1$  and  $r_2$  respectively. Then, the variance of the difference between two treatment means is given by  $2\sigma_1^2 / r_1$  and  $2\sigma_2^2 / r_2$  for designs  $D_1$  and  $D_2$  respectively.

The efficiency of design  $D_1$  with respect to the design  $D_2$  is denoted by  $E$  and is given by  $E = (r_1 / 2\sigma_1^2) / (r_2 / 2\sigma_2^2)$ . In other words, the Efficiency of design  $D_1$  with respect to the design  $D_2$  is defined as “the ratio of precisions of  $D_1$  and  $D_2$ ”.

If  $E = 1$ , then both designs  $D_1$  and  $D_2$  are said to be equally efficient.

If  $E > 1$ , then design  $D_1$  is more efficient than  $D_2$  and

If  $E < 1$ , design  $D_2$  is more efficient than  $D_1$ .

It is important to note that more efficient or sensitive design has greater ability to detect the differences of treatment effects than less efficient or sensitive design. Further, it is important to note that the efficiency of a design can be increased:

- (1) By controlling the experimental error per unit, that is  $\sigma^2$  of that design. This can be done by selecting experimental units as homogeneous as possible. If they are heterogeneous, by applying Local Control, divide the units into blocks such that units within the block are homogeneous.
- (2) By increasing number of replications. As number of replications are more, experimental error variance per unit will decrease and hence the precision of the estimates will be increased.

Now we proceed to explain the concept of “Uniformity Trials” which have special use in determining the efficiency of two designs.

#### 13.3.2. Uniformity Trials:

It is important to note that various designs of experiments are primarily introduced to solve agricultural problems. In agricultural problems, fertility rate of the experimental units may not be uniform. Even plots in the same field may not have same soil fertility rates. In a big field, fertility rates of the soil are distributed erratically over the entire field. **Uniformity trials** enable us to have an idea over the distribution of fertility rates variation

over the entire field. Thus, uniformity trials, we define as “a trial in which the experimental material or the field is divided into small units or plots and the same treatment or fertilizer is applied on each of the units and their yields are recorded.” From these yields, we can draw “Fertility Contour Map” which gives us graphical picture of the fertility rate variation of the soil and enables to form a good idea about the nature of the variation of the soil fertility variation. The fertility contour map is obtained by joining the points by joining the points with equality by straight lines.

It is already mentioned earlier that usually we expect that the entire field will vary with respect to the fertility rate of the soil. Hence, the field is to be divided into blocks such that units within the block are having homogeneous fertility rates. Thus blocks have relatively homogeneous plots, which is necessary to have control over the experimental error. Thus uniformity trials will help us to divide the entire heterogeneous field into homogeneous blocks and the size of the plots to be used. Further, we can observe that fertility maps that adjacent plots are more or less alike in fertility than those plots which are away from each other. Hence, blocks can be formed easily, by combining adjacent plots. Thus fertility maps helps in adopting ‘Local Control’.

**13.3.3. Size of the plot:** Determination of size of the plot is very important particularly, in agricultural experiments. Further, determination of the size of the plot depends on various factors like, total experimental area available, number of treatments to be tested, the type of the crop, soil fertility rate, climatic conditions and so on. If the total size of the experimental area remains fixed, then an increase in the size of the plot will result in decreasing the number plots and this will result in increasing the size of the block. Thus we obtain lesser number blocks. It is also important to note that deciding about the number of plots, we should keep in mind that increase in the number of plots will increase the ‘Non-experimental area’. This area is also known as guard-area, which means that the strips of land which are left out between consecutive plots and blocks, that is to distinguish different plots or boundaries of each plot or each block. Prof. Fairfield Smith, obtained an empirical relationship between plot size and plot variance, by conducting various uniformity trial experiments. The relation between plot variance and plot size is known as “Fairfield Smith’s Variance Law” and is given by:  $V_x = V_1 / x^b$  or  $\log V_x = \log V_1 - b \log x$ , where  $V_x$  is the variance of the yield per unit area from plots of size  $x$  units and  $V_1$  is the variance among the plots of size unity and  $b$  is the regression coefficient adjacent units with respect to some soil characteristic like temperature of the soil or fertility rate and so on.

We know that the limiting values of the regression coefficient  $b$  are 0 and 1. That is  $0 \leq b \leq 1$ . If  $b=1$  we mean that experimental unit composed of a random selection of  $x$  units. That is the units making the plots of size  $x$  and experimental units are un-correlated, or independent. Thus we have:  $V_x = V_1 / x$ . This implies that the precision of the experiment will increase along with the size of the plot. In agricultural experiments, the adjacent areas in a big plot are usually correlated. Hence, we obtain the  $b$  value less than unity. That is  $b < 1$ . If the regression coefficient  $b = 0$ , implies that the  $x$ -units are perfectly correlated and in this case, we have  $V_x = V_1$ . This implies that the increase in plot

size does not affect the variance among the experimental results. Thus precision of the estimate is independent of the size of the plot.

Now we proceed to explain the minimum number of experimental units required in the following section.

### 13.4. Determination of minimum number of replications

It is very important to determine number of replications required under each treatment.

This is because of the fact that number of replications is mainly concerned with experimental expenditure, experiment time and experimental error. Further, the adequate number of replications for various treatments in an experiment depends upon the knowledge of variability of the experimental material. On the other hand, the knowledge on the soil fertility rate is not known, in advance, unless otherwise we conduct the uniformity trials explained above. The following is the general rule one can follow in any experiment if the knowledge on the variability is not known.

#### General Thumb rules:

(1) Number of replications must be such that the error degrees of freedom must be at least 12. This is because of the fact that, the value of F-statistic do not decrease rapidly beyond for the denominator degrees of freedom is equal to 12.

(2) Minimum number of replications should not be less than 4.

Above explained rules to determine the minimum number of replications required in an experiment, can be explained with an example as follows:

**Example (13.4.1):** Determine the minimum number of replications required, so that an observed difference of 8% of the mean will be taken as significant at 5% level of significance and the Coefficient of variation of the plot values being 14%. Similarly, determine the minimum number of replications to get significance at 1% los in the above situation.

Solution: Let  $r$  be the minimum number of replications required. Let  $\mu$  be the common mean with  $\sigma^2$  as the process variance for the situation considered. Then, it is given that:

Coefficient of Variation = CV = 14% =  $100 \times (\sigma / \mu)$ . Then  $\sigma = 14\mu / 100 = 0.14\mu$ .

Further the difference between two sample means  $(\bar{x}_1 - \bar{x}_2) = 8\% \text{ of } \mu = 0.08\mu$ .

Then the t-statistic is given by:  $t_{\text{cal}} = |\bar{x}_1 - \bar{x}_2| / \sigma (\sqrt{2/r}) = 0.08\mu / 0.14\mu (\sqrt{2/r}) = (8/14)(\sqrt{2/r})$ .

To get significance  $t_{\text{cal}}$  at 5% los, minimum we have to get (for large degrees of freedom) is 1.96 and at 1% los the minimum value is 2.58. Thus we have:

Case (1): at 5%los:

We must get the calculated value of  $|t| > 1.96$ . This implies that  $(8/14)(\sqrt{2/r}) > 1.96$ . Thus we have:  $r > 2[1.96 \times (14/8)]^2 = 2(1.96 \times 1.75)^2 = 23.5298 \text{ or } \cong 24$ .

Thus minimum number replicates  $r$  is 24 to get significance at 5% los.

Case (2): at 1%los:

We must get the calculated value of  $|t| > 2.58$ . This implies that:  $(8/14)(\sqrt{2/r}) > 2.58$ . Thus we have:  $r > 2[2.58 \times (14/8)]^2 = 2(2.58 \times 1.75)^2 = 40.77045$  or  $\cong 41$ . Thus minimum number replicates  $r$  is 41 to get significance at 1% los. Thus we require minimum 24 replications to get significance at 5% los and 41 replications to get significance at 1% los.

Now we proceed to explain various types of Experimental Designs, we are going to discuss in the forthcoming lessons.

### 13.5. Various types of Design of experiments

Now we proceed to discuss various types of designs available in the literature using the three basic principles discussed in section (13.2). They are:

(A) General Designs.

- (1) Completely Randomized Designs (CRD) – Similar to ANOVA one factor model.
- (2) Randomized Block Design (RBD) – Similar to ANOVA – two factor model.
- (3) Latin Square Design (LSD) – Similar to three factor model.
- (4) Graeco-Latin Square Design (GLSD) – similar to four factor model.

(B) Factorial Designs.

- (1)  $2^k$  Factorial designs
- (2)  $3^k$  factorial designs.

(C) Balanced Designs.

(D) Balanced Incomplete Block Designs (BIBD) and so on.

There are many other designs like resolvable designs, rotatable designs, D-optimal and G-optimal designs, orthogonal designs and so on. These designs are not in the purview of this syllabus and hence are not discussed in this lesson series. Only above listed designs are explained in the forthcoming lessons.

### 13.6. Summary

In this lesson we have discussed the need and meaning of experimental designs. Basically, Design of experiments are introduced by Prof. R.A. Fisher to solve the problems arising in Agriculture. Now a days, they are popularly used many other fields like industry, business management, meteorology, medical research, biological studies and so on. We have discussed various categories of experiments namely absolute and comparative experiments. Latter, we have discussed some fundamental concepts relating to design of experiments and the three basic principles of experimental designs proposed by Prof. R.A. Fisher, who is popularly known as “Father of Statistics”.

Then we have introduced the concepts of determining efficiency of different designs, Uniformity trials, and thumb rules to determination of minimum number of replications in an experiment and determination of size of the plot or experimental unit and the relation between plot size and the precision of the estimate. Further we have introduced the concept of “Fertility Contour Map” and discussed its applications in determining various blocks. Finally,

the lesson is closed by explaining various types of experimental designs, which we are going to discuss in the forth coming lessons.

### 13.7. Self Assessment Questions.

1. Explain various objectives of Experimental Designs.
2. Define experimental designs and discuss their applications.
3. Explain the method of estimating the efficiency of a design.
4. Explain the method of determine the most efficient design among the given designs.
5. Explain Basic principles of experimental designs.
6. Explain the importance of randomization in obtaining estimates of treatment effects.
7. Explain the importance of replications in controlling the experimental error.
8. What is Local Control? Explain its need with suitable examples.
9. What are Uniformity trials? Explain their role in design of experiments.
10. Explain principles involved in determining minimum number replications in an experiment.
11. Determine the minimum number of replications required, so that an observed difference of 10% of the mean will be taken as significant at 5% level of significance and the Coefficient of variation of the plot values being 12%. Similarly, determine the minimum number of replications to get significance at 1% los in the above situation.
12. Determine the minimum number of replications required , so that an observed difference of 5% of the mean will be taken as significant at 5% level of significance and the Coefficient of variation of the plot values being 14%. Similarly, determine the minimum number of replications to get significance at 1% los in the above situation.
13. Determine the minimum number of replications required , so that an observed difference of 15% of the mean will be taken as significant at 5% level of significance and the Coefficient of variation of the plot values being 10%. Similarly, determine the minimum number of replications to get significance at 1% los in the above situation.
14. Explain various types of experimental designs.
15. What is meant by experimental error? Explain the method of controlling it.
16. What are Fertility Contour Maps? Explain the role in design of experiments.
17. Discuss the role of plot size with the precession of the estimate.
18. Explain Prof. Fairfield Smith's formula and discuss its applications.
19. Explain the effect of plot size on the number of replications.
20. Explain the concept of non-experimental area and it's effect.

**13.8. Further Readings.**

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons., New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 14.****Completely Randomized Design (CRD)****14.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the design of Completely Randomized Design (CRD).
- Assumptions, applications advantages and dis-advantages of CRD .
- Layout of CRD and its Statistical Analysis.
- CRD with equal number of replications and its analysis.
- CRD with unequal number of replications and its analysis.

**Structure**

- 14.1. Introduction to Completely Randomized Design (CRD)**
- 14.2. Important concepts of CRD**
- 14.3. Layout of CRD**
- 14.4. Statistical Analysis of CRD – with unequal number of replications**
- 14.5. Statistical Analysis of CRD – with equal number of replications**
- 14.6. Summary**
- 14.7. Self Assessment Questions**
- 14.8. Further Readings**

**14.1. Introduction to Completely Randomized Design (CRD)**

The completely Randomized Design (CRD) is the most fundamental and simplest design suitable for laboratory experiments. Hence this design is most popularly used by many scientists in many fields of research. This design is based on principles of Randomization and replication. This design is to be applied when all the experimental units are alike or uniform. Hence, we need not apply Local Control principle separately in this design. In this design, treatments are allocated at random to the experimental units over the entire experimental material. Number of replications under each treatment may be same or may not be same. Hence, CRD is of two types:

- (1) Completely Randomized Design – with equal number of replications and
- (2) Completely Randomized Design – with un-equal number of replications.

**14.2. Important concepts of CRD**

Completely Randomized Design (CRD) is most prominently used design in analyzing many real life problems, because of its simplicity and easiness in the analysis. CRD has the following advantages and disadvantages:



**Advantages:**

1. CRD is the design which uses maximum of the experimental units, because all the experimental material can be used without leaving any part of the plot.
2. The design is very flexible: Any number of treatments can be used and different treatments with different replications, without complicating the analysis.
3. Statistical analysis of CRD is very simple when compared to other designs like RBD, LSD or GLSD. If any observation is missing or lost due to natural calamities, we can carry out the analysis without estimating the missing value.
4. CRD provides maximum number of degrees of freedom for the error variance. Hence, CRD is more suitable design for smaller number treatments. Providing maximum number of degrees of freedom will increase the sensitivity or the precision of the design.
5. There is no restriction on the number of treatments or replication. We have freedom to choose number of treatments or replications depending up on the availability of experimental units.

Because of the above advantages, this design is most popularly used in many fields of research, where we want to test the effects of more than two treatments simultaneously. CRD is particularly useful for laboratory experiments and life testing experiments.

**Disadvantages:**

- (1) In certain situations, CRD suffers from the disadvantage of being inherently less informative, when compared to other sophisticated designs.
- (2) CRD is to be used only when the experimental material is uniform. Usually in agricultural experiments we may not get homogeneous experimental plots. Without our knowledge, some heterogeneity enters and makes the results less sensitive.
- (3) Presence of heterogeneity in the experimental material makes the design less efficient. Because of the above disadvantages, CRD is seldom used in agricultural experiments.

CRD is popularly used in scientific enquiries where laboratory experiments play a major role.

**Applications:**

(1) CRD is most popularly used in laboratory techniques and methodological studies. This design is most suited in Physics, Chemistry or Cookery, in Chemical and biological experiments, in green house studies, where, either the experimental material is relatively more homogeneous or the inherent variability between the experimental units is very small or ignorable.

(2) CRD is also has applications in life testing experiments in industry, where significant fraction of units is likely to be destroyed or to fail to respond or the test procedure is destructive in nature.

**14.3. Layout of CRD:**

Layout of the design means, the step by procedure of allocating treatments to the experimental units such that the experimental results are amicable for statistical treatment.

Let the number of experimental units are  $N$ , which are uniform in all respects and  $k$  treatments are to be applied on these  $N$  units randomly as follows:

Step – 1: Select  $r_1$  units randomly from  $N$  units and select one treatment randomly from  $K$  treatments. Apply this randomly selected treatment to the  $r_1$  randomly selected units.

Step – 2: Similarly, select randomly  $r_2$  units from the remaining  $(N-r_1)$  units and select randomly another treatment from the remaining  $(K-1)$  treatments. Apply the selected treatment to these  $r_2$  units selected randomly.

Step – 3: Repeat this procedure of selecting units and treatments randomly, until atleast all the treatments are completed.

Thus we obtain  $r_1$  replications under treatment-1,  $r_2$  replications under treatment -2,,,, and  $r_k$  replications under treatment-K. Such that  $r_1+r_2+.,.,.,+r_k = N$ .

Above layout is called **CRD with unequal number of replications under each treatment**. In particular if  $r_1=r_2=.,.,.,= r_k = r$  and  $Kr = N$ , then the layout is called **CRD with equal number of replicates (r replicates) under each treatment (with K treatments)**. Thus the Statistical Analysis of CRD is also of two types suitable to the above situations. Namely:

- (1) Statistical Analysis of CRD with unequal number of replications and
- (2) Statistical Analysis of CRD with equal number of replications.

#### 14.4. Statistical Analysis of CRD – with unequal number of replications

The statistical analysis of CRD – with unequal number of replications runs on the similar lines of ANOVA one way classification with unequal number of observations under each factor discussed in the lesson 9. Hence, we discuss only the brief outline of the analysis (shortcut method) for a ready reference.

The mathematical model suitable for CRD is given by:

$$X_{ij} = \mu_i + \epsilon_{ij} = \mu + (\mu_i - \mu) + \epsilon_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad (14.4.1)$$

$$i = 1, 2, \dots, K \text{ and } j = 1, 2, \dots, r_i$$

Where:  $x_{ij}$  represents the observation of  $j^{\text{th}}$  experimental unit receiving  $i^{\text{th}}$  treatment.

$$\mu \text{ is the general mean effect given by } \mu = \sum_{i=1}^k r_i \mu_i / N. \quad (14.4.2)$$

$$\alpha_i \text{ is the effect if } i^{\text{th}} \text{ treatment given by } (\mu_i - \mu), \text{ for } i=1, 2, \dots, K. \quad (14.4.3)$$

$\epsilon_{ij}$  is the error effect due to chance or random causes.

Here, we want to test the equality of the treatment means. That is the homogeneity of different treatment mean effects. Hence, we have the following Null Hypothesis  $H_0$ :

$$H_0 : \mu_1 = \mu_2 = \dots, = \mu_k = \mu \text{ or } V(\mu_i) = 0. \quad (14.4.4)$$

$$\text{Thus (14.3.4) implies that } H_0 : \alpha_1 = \alpha_2 = \dots, = \alpha_k = 0. \quad (14.4.5)$$

and the following Alternative Hypothesis  $H_1$

$$H_1: \mu_1 \neq \mu_2 \neq \dots, \neq \mu_k . = V(\mu_i) > 0 . \tag{14.4.6}$$

Is framed in CRD.

The experimental data collected under the CRD scheme is given as follows:

Suppose that there are N observations  $x_{ij}$  ( $i = 1, 2, 3, \dots, k; j = 1, 2, \dots, r_i$ ) of random variable X grouped on some basis (usually difference among treatments) into k-classes of sizes  $r_1, r_2, \dots, r_k$  respectively such that  $\sum_{i=1}^k r_i = N$  as explained in the following table:

Treatments/rows	Data				Means	Total
Treatment – 1: $t_1$	$x_{11}$	$x_{12}$	$x_{13} \dots \dots \dots$	$x_{1r_1}$	$\bar{x}_1$	$T_1.$
Treatment – 2: $t_2$	$x_{21}$	$x_{22}$	$x_{23} \dots \dots \dots$	$x_{2r_2}$	$\bar{x}_2$	$T_2.$
Treatment – 3: $t_3$	$x_{31}$	$x_{32}$	$x_{33} \dots \dots \dots$	$x_{3r_3}$	$\bar{x}_3$	$T_3.$
.	.....				.	.
.	.....				.	.
.	.....				.	.
Treatment – k: $t_k$	$x_{k1}$	$x_{k2}$	$x_{k3} \dots \dots \dots$	$x_{kr_k}$	$\bar{x}_k$	$T_k.$
						<b>G = Grand total</b>

Where,  $\bar{x}_i = T_i. / r_i \ i = 1, 2, \dots, K.$

Step – 1: Calculate Treatment total  $T_i.$  and the Grand total  $G = \sum_{i=1}^K T_i. = T..$

Step – 2: Calculate the Correction Factor=  $CF = G^2/N$ , where  $N = \sum_{i=1}^K r_i .$

Step – 3: Calculate the treatment Sum of square =  $SST = \sum_{i=1}^K T_i.^2 / r_i - CF = A$

Step – 4: Calculate Total sum of square =  $TSS = \sum_{i=1}^K \sum_{j=1}^{r_i} (x_{ij})^2 - CF = B.$

Step – 5: Construct the following ANOVA table for CRD.

**Table (14.4.1): ANOVA table for CRD with unequal number of replications.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (Columns)	$K - 1$	$S_t^2 = SST=A$	$V_1 = S_t^2 / K - 1$	$F = V_1 / V_2$
Error	$N - K$	$S_e^2 = ESS=B-A$	$V_2 = S_e^2 / N - K$	
Total	$N - 1$	$S_T^2 = TSS=B$		

Step – 6: Conclusions:

(1) If  $F_{cal}$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for  $[(K-1),(N-K)]$  dof, we accept the null hypothesis  $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$  and conclude that there is no significant effect of treatment means. In this case we need not apply Further Statistical Analysis because all the treatments are equally effective.

(2) If  $F_{cal}$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for  $[(K-1),(N-K)]$  dof, we reject the null hypothesis  $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$  and conclude that there is significant difference in the effect of treatment means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the treatment effects or to determine which treatment has maximum effect? Or which treatment has minimum effect? And so on.

Example (14.3.1): The effective life of insulating fluids at an accelerated load of 35 kv is being studied. Four types of fluids A,B,C and D are randomly applied on 20 units produced from the same machine and the data represents life of the unit measured in minutes and are recorded as follows:

Unit No:	12	18	06	01	05	13	10	19	04	20
Treatment:	C	A	A	B	D	C	B	A	D	A
Life in minutes:	176	189	153	---	174	169	216	236	185	193

Unit No:	02	17	07	15	09	14	03	11	08	16
Treatment:	B	A	C	A	D	D	B	C	C	D
Life in minutes:	223	246	182	---	223	203	232	182	---	218.

Analyze the data with suitable design and draw your conclusions. If necessary, apply further Statistical analysis.

Solution: The given data is to be arranged treatment wise as follows:

Treatment	life in minutes at 35 kv Load.						Totals.
A	189	153	236	193	246	---	1017
B	---	216	223	232			671
C	176	169	182	182	----		709
D	174	185	223	203	218		1003
							3400=G

For the given problem  $N = 17$ ,  $r_1=5, r_2=3, r_3=4$  and  $r_4 = 5$  and  $K = 4$ .

Treatment	Table of Squares of each observation.	Totals.
A	35721 23409 55696 37249 60516 ---	212591
B	--- 46656 49729 53824	150209
C	30976 28561 33124 33124 ----	125785
D	30276 34225 49729 41209 47524	202963
	Sum of Squares.	691548.

Step – 1:  $CF = G^2 / N = (3400)^2 / 17 = 680000$ .

Step – 2: Treatment SS= SST =  $(1017)^2 / 5 + (671)^2 / 3 + (709)^2 / 4 + (1003)^2 / 5 - 680000$   
 $= 206857.800 + 150080.3333 + 125670.25 + 201201.8 - 680000 = 683810.183 - 680000$   
 $= 3810.183 = A$ .

Step – 3: Total SS = TSS =  $691548 - 680000 = 11548 = B$ .

Step – 4: Construction ANOVA table.

**Table (14.4.2): ANOVA table for CRD with unequal number of replications.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (Columns)	$K - 1 = 4 - 1 = 3$	$S_t^2 = SST =$ $A = 3810.183$	$V_1 = S_t^2 / K - 1$ $= 3810.183 / 3$ $= 1270.061$	$F = V_1 / V_2$ $1270.061 / 595.2166923$ $= 2.133779204$
Error	$N - K = 17 - 4 = 13$	$S_E^2 = ESS = B - A$ $11548 - 3810.183$ $= 7737.817$	$V_2 = S_E^2 / N - K$ $= 7737.817 / 13$ $= 595.2166923$	
Total	$N - 1 = 17 - 1 = 16$	$S_T^2 =$ $TSS = B = 11548$		

Step-5: Conclusions:

$F_{tab}$  at 5% los for (3,13) is given by 3.41 (See F-table in Appendix). Since  $F_{cal} < F_{tab}$  we accept the null hypothesis  $H_0$  at 5% los. Hence we conclude that there is no significant difference between the treatment means or all the four types of fluids are equally effective with respect to the life of insulating fluids at an accelerated load of 35 kv. Since we have accepted the null-hypothesis, there is no need to apply further statistical analysis.

**Remark: (1)** It is interesting to note that in the analysis of CRD, it is very easy to deal the problem of missing values. If one or more values are missing we can still run the CRD analysis with the available observations. While calculating SST we have to divide with the number of observations available under the treatment 'i', namely  $r_i, i=1,2,...,K$ . this type of easiness will

not be there in other designs like RBD or LSD or GLISD. The missing values are to be estimated first and then carryout the analysis by reducing the degrees of freedom by number of missing values estimated. The procedure explained in the next lesson dealing with RBD.

(2) We can also reduce the data into smaller numbers by changing the origin. That is we can subtract a constant number  $K = 100$  or  $150$  or  $200$  from each observation and run the analysis. Whatever may be the value of  $k$   $F_{cal}$  will be same and conclusions will not change.

**14.5. Statistical Analysis of CRD – with equal number of replications**

The statistical analysis of CRD with equal number of replications runs on similar lines of analysis of CRD with unequal number of replications discussed in the last section. Only difference is that in the present case  $r_1 = r_2 = \dots = r_k = r$ . and  $N=rK$ . The data under this design is as follows:

Where,  $\bar{x}_i = T_i / r$ ;  $i = 1, 2, \dots, K$ .

Step – 1: Calculate Treatment total  $T_i$ . and the Grand total  $G = \sum_{i=1}^K T_i = T..$

Treatments/rows	Data				Means	Total
Treatment – 1: $t_1$	$X_{11}$	$X_{12}$	$X_{13} \dots \dots \dots$	$X_{1r}$	$\bar{x}_1$	$T_1.$
Treatment – 2: $t_2$	$X_{21}$	$X_{22}$	$X_{23} \dots \dots \dots$	$X_{2r}$	$\bar{x}_2$	$T_2.$
Treatment – 3: $t_3$	$X_{31}$	$X_{32}$	$X_{33} \dots \dots \dots$	$X_{3r}$	$\bar{x}_3$	$T_3.$
.	.....				.	.
.	.....				.	.
.	.....				.	.
Treatment – k: $t_k$	$X_{k1}$	$X_{k2}$	$X_{k3} \dots \dots \dots$	$X_{kr}$	$\bar{x}_k$	$T_k.$
						<b>G = Grand total</b>

Step – 2: Calculate the Correction Factor=  $CF = G^2/N$ , where  $N = rK$ .

Step – 3: Calculate the treatment Sum of square =  $SST = \sum_{i=1}^K T_i.^2 / r - CF = A$

Step – 4: Calculate Total sum of square =  $TSS = \sum_{i=1}^K \sum_{j=1}^r (x_{ij})^2 - CF = B$ .

Step – 5: Construct the following ANOVA table for CRD.

**Table (14.5.1): ANOVA table for CRD with equal number of replications.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (Columns)	$K - 1$	$S_t^2 = SST = A$	$V_1 = S_t^2 / K - 1$	$F = V_1 / V_2$



	Squires of the Yield						Totals.
Variety – 1(V <sub>1</sub> ):	3600	25600	23409	17161	32400	3364	105534
Variety – 2(V <sub>2</sub> ):	5184	24025	6889	4624	784	5625	47131
Variety – 3(V <sub>3</sub> ):	2500	13689	10000	5625	6084	5625	43523
Variety – 4(V <sub>4</sub> ):	1600	3600	3364	5625	9025	4900	28114
Variety – 5(V <sub>5</sub> ):	9604	7744	34225	22500	19600	14400	108073
							Sum of squares 332375

Step – 2: now frame the null hypothesis as follows:

H<sub>0</sub>: there is no significant variation between seed varieties.

Step -3:  $CF = (2897)^2/30 = 279753.63333$

Step – 4: Now calculate various sum of squares as follows:

$$SST = [(742)^2 + (481)^2 + (495)^2 + (398)^2 + (781)^2] / 6 - 279753.63333$$

$$= [550564 + 231361 + 245025 + 158404 + 609961] / 6 - 279753.63333 = 299219.1667 - 279753.63333$$

$$= 19465.53337 = A.$$

$$TSS = 332375 - 279753.63333 = 52621.36667 = B.$$

Step-5: Now construct ANOVA table for CRD with equal number of replicates as follows:

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (rows)	K – 1=4	$S_t^2 = SST=A$ =19465.53337	$V_1 = S_t^2 / K - 1$ =4866.38335	$F = V_1 / V_2$ =3.6693
Error	K(r-1)=25	$S_E^2 = ESS=B-A$ =33155.8333	$V_2 = S_E^2 / K(r-1)$ =1326.2333	
Total	Kr – 1=29	$S_T^2 = TSS=B$ =52621.36667		

Conclusion:

Since  $F_{cal}$  is greater than the  $F_{tab}$  at 5% los for (4,25) dof, that is 2.76, we reject the null hypothesis and conclude that the seeds variety have significant variation among them. Since we have rejected the null hypothesis, we have to apply further statistical analysis to sub-group these seed varieties. This problem is left as an exercise for the reader to determine which variety yields most? Or which variety yields less? Which variety of seeds have similar effect and so on. The procedure is explained in section 9.4.



Now we proceed to explain another important design when the experimental material is not homogeneous. Then we have to use Randomized Block Design (RBD), which is explained in the following lesson.

## 14.6 Summary

In this lesson, Completely Randomized Design (CRD) is introduced along with its advantages, disadvantages and applications. It is important to note that CRD is to be applied only when we have homogeneous experimental material. Otherwise we obtain meaningless results and we obtain experimental error more than the treatments effect. Statistical analysis CRD is of two types, Namely (1) CRD with unequal number of replications and (2) CRD with equal number of replications. Both methods are explained with suitable example.

## 14.7. Self Assessment Questions

1. Explain the conditions under which CRD can be applied.
2. Discuss various advantages and disadvantages of CRD.
3. Briefly explain the analysis of CRD with unequal number of replications.
4. Briefly explain the analysis of CRD with equal number of replications.
5. Explain the layout of the CRD.
6. Explain various applications of CRD.
7. Explain the easiness in dealing missing values in CRD.
8. To test the effect of small portion of coal in the sand used for manufacturing concrete, several batches were mixed under particularly identical conditions except the variation of the percentage of coal. From each batch, several bricks were made tested for their breaking strength. The results are recorded as follows:

0.00:	1690	1580	1745	1685
0.05:	1550	1445	1645	1545
0.10:	1625	1450	1510	-----
0.50:	1725	1550	-----	1530
1.00:	1530	1545	1665	1520

Analyze the above data with suitable design and draw your conclusions.  
(Ans:  $F_{cal} = 1.637$ ; Accept the null hypothesis).
9. A set of topical feeds A,B,C and D tried on 20 chicks and the increased is recorded in Grams.  
The data is recorded as follows:

Chick No:	02	17	07	15	09	14	03	11	08	16
Treatment:	B	A	C	A	D	D	B	C	C	D
Weight in gr.:	61	55	42	49	169	137	112	97	81	169
Chick No:	12	18	06	01	05	13	10	19	04	20
Treatment:	C	A	A	B	D	C	B	A	D	B
Weight in gr:	95	42	21	30	154	92	89	52	85	63.

Analyze the data with a suitable design and draw your conclusions.

10. Write down the ANOVA table CRD with equal and unequal number of replicates.

### 14.8. Further Readings

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons., New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 15****Randomized Block Design (RBD)****15.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the design of Randomized Block Design (RBD).
- Assumptions, applications advantages and dis-advantages of RBD .
- Layout of RBD.
- Statistical Analysis of RBD.
- Comparing the efficiency of RBD with CRD.
- Missing plot technique.

**Structure**

- 15.1. Introduction to Randomized Block Design (RBD)**
- 15.2. Important concepts of RBD**
- 15.3. Layout of RBD**
- 15.4. Statistical Analysis of RBD – with one observation per cell**
- 15.5. Important Concepts of RBD**
- 15.6. Summary**
- 15.7. Self Assessment Questions**
- 15.8. Further Readings**

**15.1. Introduction to Randomized Block Design (RBD)**

In the previous lesson, we discussed CRD, where Local control is not applied, because we have assumed that the experimental material is uniform or homogeneous. In agricultural related problems, this situation may not occur because plots or field differs significantly with each other with respect to many factors like climatic conditions or irrigation type or fertility rate or soil temperature and so on. Hence, there is a need to apply “**Local Control**” and make the heterogeneous fields into homogeneous blocks. Thus we have to apply all the three basic principles of experimental designs, namely Randomization, Replication and Local control. Such a design is known as Randomized Block Design (RBD).

In field experimentation if the whole of the experimental area is not homogeneous, and the fertility gradient is only in one direction, then a simple method of controlling the variability of the experimental material consists in stratifying or grouping the whole area into relatively homogeneous strata or sub-plots or blocks or replicates (usually they are called), perpendicular to the direction of the fertility gradient. Now all treatments are to be applied randomly to the

units in each block. Such a design is known as Randomized Block Design (RBD), because randomization is applied within each block independently.

## 15.2. Important concepts of RBD

Now we proceed to compare two designs, namely CRD and RBD before discussing various advantages and disadvantages of RBD.

### 15.2.1. Critical Comparison between CRD and RBD.

S.No.	Completely Randomized Design (CRD)	Randomized Block Design (RBD)
1	Local Control is not applied specifically because Experimental material is uniform.	Since experimental material or site is heterogeneous, we have to sub-group the entire site into relatively homogeneous Strata or blocks or replications.
2	Treatments are applied randomly applied throughout the experimental material.	Treatments are applied randomly within the block or replication. Hence Randomization is restricted to block here.
3	Variation among different replications is mixed with experimental error. Hence, Experimental error is more in CRD when compared with RBD.	The variation between the replications is separated out from experimental error and hence error is smaller in RBD when compared to CRD.
4	Error degree of freedom is maximum when compared to RBD or LSD.	Error degree of freedom is smaller when compared to CRD and larger when compared to LSD.
5.	This design is more suited laboratory experiments in scientific enquires, in industry, biological studies, medicine meteorology and animal husbandry.	The design is more suitable for field experiments in agricultural problems or where the experimental material is heterogeneous.

Now we proceed to discuss various advantages and disadvantages of Randomized Block Design (RBD) in the following section.

### 15.2.2. Advantages and disadvantages of RBD.

**Advantages:** Chief advantages of Randomized Block Design (RBD) may be outlined as follows:

1. Accuracy: The design has been shown to be more efficient or accurate than CRD for most types of experimental work. The elimination of between blocks sum of squares, from the error or residual sum of squares and hence, in general experimental error in RBD will decrease, when compared to CRD.
2. Flexibility: In RBD no restrictions are placed on the number of treatments or number of replications (blocks). Usually minimum two replicates are required to carry out the test of significance. (Factorial Design is an exceptional case

here). We can also add or delete any treatment or block if we wish, without complicating the analysis.

3. Ease of Analysis: The analysis is easy when compared to LSD and little lengthy when compared to CRD. We have freedom to choose any number of treatments and any number of blocks or replications, without complicating the analysis.
4. RBD is more suitable for solving many agronomical problems, when the fertility gradient is varying in one direction.
5. RBD is more efficient design than CRD and Less efficient than Latin Square Design (LSD).

**Disadvantages:**

**The design RBD is not a suitable design, when we want to test large number of treatments or for cases in which considerable variation is present within the blocks or replicates.**

**15.3. Layout of RBD**

Let us consider  $K$  treatments  $T_1, T_2, \dots, T_K$  treatments and the experimental material is divided into  $b$  blocks each of size  $K$ , thus we have  $K \cdot b$  experimental units, such that units within each block are homogeneous as far as possible. Any variation present in the experimental material is attributed to the variation between blocks. Now  $K$  treatments are to be applied for  $k$  units in a block at random in one block. Similarly apply again these  $k$  treatments randomly independent of the previous block order of application of treatments. This implies that randomization is to be applied independently in each block. This is explained with the following example:

Example (15.3.1): Let us consider five treatments A,B,C,D and E which are to be applied in six blocks. Thus we have 30 experimental units with five blocks such that five uniform plots are there in each block. Then an example of a layout of RBD is given as follows:

Block – 1	C	A	D	E	B
Block – 2	D	B	C	A	E
Block – 3	A	B	D	C	E
Block – 4	E	D	C	B	A
Block – 5	D	C	B	A	E
Block – 6	B	E	A	D	C

For randomization, we may use Tippet's random number of tables or calculator or computer to generate random digits, such that 1=A,2=B,3=C,4=D and 5=E or in any order liked by the experimenter. After determining the layout, same treatments are to be applied to the respective units in each block as given in the Layout and experimental results are to be recorded for the analysis of RBD.

Now we proceed to explain the statistical analysis of RBD in the following section.

**15.4. Statistical Analysis of RBD – with one observation per cell**

Consider the layout of the RBD explained in the last section with one observation is made on each of the experimental units, then, the statistical analysis is analogous to ANOVA two factor model data with observation per cell discussed in section 10.3. Hence, we briefly outline the step wise analysis as follows: Let  $x_{ij}$  observation or data collected from the unit receiving  $i$ th treatment in  $j$ th block.

$$x_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \quad i=1,2,\dots,K \text{ and } j=1,2,\dots,b. \tag{15.4.1}$$

where  $\mu$  = General mean effect.

$\alpha_i$  =  $i$ th treatment mean effect.

$\beta_j$  =  $j$ th block effect.

$\epsilon_{ij}$  = Error influencing the yield of  $(i,j)$ th unit.

The errors  $\epsilon_{ij}$ 's are assumed to be independent and identically distributed with mean 0 and variance  $\sigma^2$ .

15.4.1. **Null-hypotheses:** We have to frame two null-hypotheses, one for treatment mean effects and another for the block means effects framed as follows:

For Treatments:

**$H_{01}$ : There is no significant variation between average treatment effects or**

$$H_{01}: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_K = \mu \text{ or } \alpha_1 = \alpha_2 = \alpha_3 = \dots = \alpha_K = 0. \tag{15.4.2}$$

For Blocks:

**$H_{02}$ : There is no significant variation between average block effects or**

$$H_{02}: \mu_{\cdot 1} = \mu_{\cdot 2} = \mu_{\cdot 3} = \dots = \mu_{\cdot b} = \mu \text{ or } \beta_1 = \beta_2 = \beta_3 = \dots = \beta_b = 0. \tag{15.4.3}$$

To test the above hypotheses, we use the statistical analysis, which is explained in the following steps:

15.4.2. The statistical analysis of RBD with one observation per experimental unit, runs on similar lines of analysis of ANOVA two factor model with equal number of replications.

Treatments/rows	Blocks				Means	Total
	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub> ..... B <sub>j</sub> .....	B <sub>b</sub>		
Treatment – 1: $t_1$	$x_{11}$	$x_{12}$	$x_{13}$ . . . . .	$x_{1b}$	$\bar{x}_1$	<b>T<sub>1</sub></b>
Treatment – 2: $t_2$	$x_{21}$	$x_{22}$	$x_{23}$ . . . . .	$x_{2b}$	$\bar{x}_2$	<b>T<sub>2</sub></b>
Treatment – 3: $t_3$	$x_{31}$	$x_{32}$	$x_{33}$ . . . . .	$x_{3b}$	$\bar{x}_3$	<b>T<sub>3</sub></b>

Treatment – i: $t_i$	..... ..... $x_{ij}$ ..... .....				$\bar{x}_i$	$T_i$
Treatment – k: $t_k$	$x_{k1}$	$x_{k2}$	$x_{k3}$ . . . . .	$x_{kb}$	$\bar{x}_k$	$T_k$
Means	$\bar{x}_{.1}$	$\bar{x}_{.2}$	$\bar{x}_{.3}$	$\bar{x}_{.b}$		<b>G</b> = Grand total
Totals	$T_{.1}$	$T_{.2}$	$T_{.3}$	$T_{.b}$		

Where,  $\bar{x}_i = T_i / r$  ;  $i = 1, 2, \dots, K$  and  $\bar{x}_{.j} = T_{.j} / K$  ;  $j = 1, 2, \dots, b$ .

The short cut step wise procedure adopted in many practical problems in analyzing RBD with one observation per experimental unit is explained as follows:

Step – 1: Calculate Treatment total  $T_i$ . and the Grand total  $G = \sum_{i=1}^K T_i = T_{..} = G$

Step – 2: Calculate the Correction Factor=  $CF = G^2 / N$ , where  $N = bk$  .

Step – 3: Calculate the treatment Sum of squares =  $SST = \sum_{i=1}^K T_i^2 / b - CF = A$ .

Step – 4: Calculate the Block Sum of Squares =  $BSS = \sum_{j=1}^K T_{.j}^2 / K - CF = B$ .

Step – 5: Calculate Total sum of square =  $TSS = \sum_{j=1}^b \sum_{i=1}^K (x_{ij})^2 - CF = C$ .

Step – 6: Construct the following ANOVA table for RBD with one observation per cell.

**Table (15.4.1): ANOVA table of RBD with one observation per cell.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (rows)	$K - 1$	$S_t^2 = SST = A$	$V_1 = S_t^2 / K - 1$	$F_T = V_1 / V_3$
Blocks (columns)	$b - 1$	$S_b^2 = BSS = B$	$V_2 = S_b^2 / b - 1$	$F_B = V_2 / V_3$
Error	$(K-1)(b-1)$	$S_E^2 = ESS = C - B - A$	$V_3 = S_E^2 / (K-1)(b-1)$	
Total	$Kb - 1$	$S_T^2 = TSS = C$		

Step – 7: Conclusions:

A – For treatments:

- (1) If the calculated value  $F_T$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for  $[(K-1), (K-1)(b - 1)]$  dof, we accept the null hypothesis for treatments  $H_{01} : \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$  and conclude that there is no significant effect of treatment means. In this case we

need not apply Further Statistical Analysis because all the treatments are equally effective.

- (2) If the calculated value  $F_T$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for  $[(K-1),(K-1)(b-1)]$  dof, we reject the null hypothesis for treatments  $H_{01} : \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$  and conclude that there is significant difference in the effect of treatment means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the treatment effects or to determine which treatment has maximum effect? Or which treatment has minimum effect? And so on.

B- For blocks:

- (1) If the calculated value  $F_B$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for  $[(b-1),(K-1)(b-1)]$  dof, we accept the null hypothesis for treatments  $H_{02} : \beta_1 = \beta_2 = \beta_3 \dots = \beta_b = 0$  and conclude that there is no significant effect of block means.

In this case we need not apply Further Statistical Analysis because all the blocks are uniform with respect to the yield. This means that blocking effect is negligible.

- (2) If the calculated value  $F_B$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for  $[(b-1),(K-1)(b-1)]$  dof, we reject the null hypothesis for treatments  $H_0 : \beta_1 = \beta_2 = \beta_3 \dots = \beta_b = 0$  and conclude that there is significant difference in the effect of treatment means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the blocks or to determine which block has maximum fertility rate? Which block has minimum fertility rate? Which blocks have similar fertility rates? And so on.

Now we proceed to explain the above analysis with an example.

Example (15.4.1): To test the effect of five varieties of seeds  $V_1, V_2, V_3, V_4$  and  $V_5$  CRD is conducted and the yield is measured in kgs. The data is recorded as follows:

$V_1 ; 360$	$V_4 ; 340$	$V_5 ; 398$	$V_2 ; 360$	$V_3 ; 350$	Block – 1.
$V_4 ; 372$	$V_2 ; 455$	$V_3 ; 417$	$V_1 ; 460$	$V_5 ; 388$	Block - 2.
$V_3 ; 400$	$V_1 ; 453$	$V_5 ; 485$	$V_2 ; 431$	$V_4 ; 358$	Block – 3.
$V_2 ; 383$	$V_3 ; 375$	$V_4 ; 375$	$V_5 ; 368$	$V_1 ; 480$	Block – 4.
$V_5 ; 450$	$V_1 ; 358$	$V_3 ; 378$	$V_4 ; 395$	$V_2 ; 328$	Block – 5.
$V_4 ; 370$	$V_5 ; 440$	$V_2 ; 375$	$V_1 ; 420$	$V_3 ; 375$	Block – 6.

Analyze the though RBD and draw your conclusions.

Solution: Step – 1: To simply the calculations we subtract  $K = 300$  and arrange the data Treatment / variety wise as follows:

	Yield of the crop (in kgs)						Totals.
	B-1	B-2	B-3	B-4	B-5	B-6	
Variety – 1( $V_1$ ):	60	160	153	180	58	120	731



Variety – 2(V <sub>2</sub> ):	60	155	131	83	28	75	532
Variety – 3(V <sub>3</sub> ):	50	117	100	75	78	75	495
Variety – 4(V <sub>4</sub> ):	40	72	58	75	95	70	410
Variety – 5(V <sub>5</sub> ):	98	88	185	68	150	140	729
Totals:	308	592	627	481	409	480	2897= G

	Squires of the Yield						Totals.
Variety – 1(V <sub>1</sub> ):	3600	25600	23409	32400	3364	14400	102773
Variety – 2(V <sub>2</sub> ):	3600	24025	17161	6889	784	5625	58084
Variety – 3(V <sub>3</sub> ):	2500	13689	10000	5625	6084	5625	43523
Variety – 4(V <sub>4</sub> ):	1600	5184	3364	5625	9025	4900	29698
Variety – 5(V <sub>5</sub> ):	9604	7744	34225	4624	22500	19600	98297
Sum of squares	20904	76242	88159	55163	41757	50150	332375

Step – 2: Now frame the null hypothesis as follows:

H<sub>0</sub>: There is no significant variation between seed varieties.

Step -3:  $CF = (2897)^2/30 = 279753.63333$

Step – 4: Now calculate various sum of squares as follows:

$$SST = [(731)^2 + (532)^2 + (495)^2 + (410)^2 + (729)^2] / 6 - 279753.63333$$

$$= [534361 + 283024 + 245025 + 168100 + 531441] / 6 - 279753.63333 = 293658.5 - 279753.63333$$

$$= 13904.8667 = A.$$

$$BSS = [(308)^2 + (592)^2 + (627)^2 + (481)^2 + (409)^2 + (480)^2] / 5 - 279753.63333.$$

$$= [94864 + 350464 + 393129 + 231361 + 167281 + 230400] / 5 - 279753.63333$$

$$= 293499.8 - 279753.63333 = 13746.1667 = B$$

$$TSS = 417216 - 257613.33333 = 159602.6667 = C$$

Step-5: Now construct ANOVA table for CRD with equal number of replicates as follows:

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (rows)	K – 1 = 4	$S_t^2 = SST = A$ = 13904.8667	$V_1 = S_t^2 / K - 1$ = 3476.216675	$F_T = V_1 / V_3$ = 0.52689

Blocks (columns)	$b - 1 = 6 - 1 = 5$	$S_B^2 = BSS = B$ =13746.1667	$V_2 = S_B^2 / b - 1$ =2749.23334	$F_B = V_2 / V_3$ =0.4167
Error	$(K - 1)(b - 1) = 20$	$S_E^2 = ESS = C - A - B$ =131951.6333	$V_3 = S_E^2 / K(r - 1)$ =6597.581665	
Total	$Kb - 1 = 29$	$S_T^2 = TSS = C$ =159602.6667		

Conclusion:

Since  $F_{cal} < 1$ , in both the cases, this is because of the fact that experimental error or  $V_2$  is more than  $V_1$ ;  $V_2$  and hence, the experiment failed to measure the effects of treatments as well as blocks. Based on the data we cannot estimate treatment effects, or block effects, because F calculated values are less than unity. It is important to note that F calculated value must be greater than 1 always. Then only we can test the null hypothesis. Otherwise, we have to discard the experiment and suggest performing fresh experiment.

### 15.5. Important Concepts of RBD

Now we proceed to discuss some important concepts like (1) Comparing Efficiency of RBD with CRD (2) Missing plot Technique and (3) Further Statistical analysis applicable in RBD in the following sections:

#### 15.5.1. Comparing the efficiency of RBD with CRD.

Now we proceed to compare the efficiency of RBD with CRD. In order to do this Treatment effects are same irrespective of the design. In other words treatments effects will not be affected by the type of design we are using. Hence, treatment effects are to be removed from the ANOVA table, which means that we assume that uniform treatment is applied for all the experimental units. We have to assume that uniform trials are conducted which represents we have to consider only block variances and error variances in RBD and only Error variance in CRD. This is done as follows. First consider the ANOVA table of RBD as follows:

**Table (15.5.1): ANOVA table of RBD with one observation per cell.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	E(MSS)
Treatments (rows)	$K - 1$	$S_t^2 = SST = A$	$V_1 = S_t^2 / K - 1$	$V_3 + b V_1$
Blocks (columns)	$b - 1$	$S_b^2 = BSS = B$	$V_2 = S_b^2 / b - 1$	$V_3 + K V_2$
Error	$(K - 1)(b - 1)$	$S_E^2 = ESS = C - B - A$	$V_3 = S_E^2 / (K - 1)(b - 1)$	$V_3$
Total	$Kb - 1$	$S_T^2 = TSS = C$		

Now consider the ANOVA table for the same experiment without blocks, which means that for the same experiment conducted in CRD (without block effects). Then we have the following ANOVA table:

**Table (15.5.2): ANOVA table of RBD without block effect.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	E(MSS)
Treatments (rows)	$K - 1$	$S_t^2 = SST = A$	$V_1'$ (say)	$V_2' + b V_1$
Error	$K(b-1)$	$S_E^2 = ESS = C - B - A$	$V_2'$ (say)	$V_2'$ (say)
Total	$Kb - 1$	$S_T^2 = TSS = C$		

Since the same experiment is performed with CRD instead of RBD, The total remains the same in each case and also the treatment effects. From the above tables (15.5.1) and (15.5.2) we have:

$$(K - 1) (V_3 + b V_1) + (b-1) (V_3 + K V_2) + (K-1)(b-1) V_3 = (K-1)(V_2' + b V_1) + K(b-1)V_2' \quad (15.5.1)$$

$$\Rightarrow \{(K-1)+(b-1)+(K-1)(b-1)\} V_3 + K(b-1)V_3 = [(K-1)+K(b-1)] V_2' \quad (15.5.2)$$

$$\Rightarrow (Kb-1)V_3 + K(b-1) V_3 = (Kb-1)V_2' \quad (15.5.3)$$

The efficiency of RBD relative to CRD is given by the ratio of relative amounts of information or the precessions of both designs. Hence using the equation (15.5.3), the efficiency of RBD relative to CRD denoted by E is given by:

$$E = V_2' / V_3 = 1 + [K(b-1)/(kb-1)] V_3 / V_2' \quad (15.5.4)$$

But for an RBD  $E(V_3) = \sigma\epsilon^2$  and  $E(V_2) = V_3 + K V_2 = \sigma\epsilon^2 + K \sigma b^2$ .

$$\text{Therefore } E[(V_2 - V_3)/K] = \sigma b^2 \quad (15.5.5)$$

Hence an unbiased estimate of  $\sigma b^2$  is given by  $[(V_2 - V_3)/K]$ .

Using the unbiased estimates for  $V_3 = \sigma\epsilon^2$  and  $V_2 = \sigma b^2$  we have from equation (15.5.4):

$$E = 1 + [K(b-1)/(kb-1)] [(V_2 - V_3)/K] \cdot (1/V_3) = \{[(Kb-1) - (b-1)]V_2 + (b-1)V_3\} / [(kb-1)V_3]$$

$$\Rightarrow E = [b(k-1)V_3 + (b-1) V_2] / (Kb-1)V_3 \quad (15.5.6)$$

**Example (15.5.1):** From the following ANOVA table of RBD, calculate the efficiency of RBD with CRD.

**Table (15.5.3): ANOVA table of RBD with one observation per cell.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)
Treatments	$K - 1 = 5$	$S_t^2 = SST = A$	$V_1 = S_t^2 / K - 1$

(rows) Blocks (columns)	b – 1=3	=901.19 $S_b^2 = BSS = B$ =219.43	=180.24 $V_2 = S_b^2 / b - 1$ =73.14
Error	(K-1)(b-1)=15	$S_E^2 = ESS = C - B - A$ =229.63	$V_3 = S_E^2 / (K-1)(b-1)$ =15.31
Total	Kb – 1=23	$S_T^2 = TSS = C$ =1350.25	

Solution: The relative efficiency of RBD with CRD using the formula (15.5.6) we have:

$$E = \frac{[b(k-1)V_3 + (b-1)V_2]}{(Kb-1)V_3} = \frac{[4(5) 15.31 + 3 (73.14)]}{23(15.31)} = \frac{[306.2+219.42]}{352.13} = 525.62/352.13 = 1.492687.$$

We can express E in terms of percentage. Hence we have E = 149.26%. Hence the gain in efficiency is about 49% by conducting RBD than CRD.

Now we proceed to explain the concept of missing observation problem in the following section.

**15.5.2. Estimation of missing observation in RBD.**

Usually if one or more observations are missed due to some natural calamities like Floods or earthquakes or fire accidents, we cannot analyze the data in RBD, unless those missing values are to be estimated from the data available. But CRD can be carried out estimating the missing values just by removing the missing replicates, CRD with unequal number of replications can be carried out. This is the basic advantage of CRD when compared to RBD or LSD. If we want to carry out the RBD analysis, all the observations must be available. If one or more observations are missing, they are to be estimated from the available data and substitute this estimated values in the data and then carryout RBD analysis by reducing the number of missing observations from total degrees of freedom and hence consequently from error degrees of freedom. Thus, the number of observations we have estimated is to be subtracted from total and error dof. The procedure of estimating single observation is explained first and then we discuss the procedure with multiple missing observations.

Let  $y_{ij} = Y_{ij}$  is the observation missed in  $j^{th}$  block receiving  $i^{th}$  treatment. Then the data under RBD is of the following form:

Treatments/rows	Blocks				Means	Total
	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub> ..... B <sub>j</sub> .....	B <sub>b</sub>		
Treatment – 1: t <sub>1</sub>	x <sub>11</sub>	x <sub>12</sub>	x <sub>13</sub> . . . . .	x <sub>1b</sub>	$\bar{x}_1$	<b>T<sub>1</sub></b>
Treatment – 2: t <sub>2</sub>	x <sub>21</sub>	x <sub>22</sub>	x <sub>23</sub> . . . . .	x <sub>2b</sub>	$\bar{x}_2$	<b>T<sub>2</sub></b>
Treatment – 3: t <sub>3</sub>	x <sub>31</sub>	x <sub>32</sub>	x <sub>33</sub> . . . . .	x <sub>3b</sub>	$\bar{x}_3$	<b>T<sub>3</sub></b>

Treatment – i: t <sub>i</sub>	.....y <sub>ij</sub> =Y.....					$\bar{x}_i$	$T_{i\cdot}' + Y$
Treatment – k: t <sub>k</sub>	x <sub>k1</sub>	x <sub>k2</sub>	x <sub>k3</sub>	.....	x <sub>kb</sub>	$\bar{x}_k$	$T_{k\cdot}$
Means	$\bar{x}_{\cdot 1}$	$\bar{x}_{\cdot 2}$	$\bar{x}_{\cdot 3}$		$\bar{x}_{\cdot b}$		$G' + Y$
Totals	$T_{\cdot 1}$	$T_{\cdot 2}$	$T_{\cdot 3}$	$T_{\cdot j}' + Y$	$T_{\cdot b}$		

Where:  $T_{i\cdot}'$  = Sum of the known observations under i<sup>th</sup> treatment,

Similarly  $T_{\cdot j}'$  = Sum of the known observations in j<sup>th</sup> block,

$G'$  = Grand total of all known observations and  $Y$  is the unknown missing plot value.

Then  $CF = (G' + Y)^2 / Kb$ .

Thus we have,  $TSS = \sum_{j=1}^b \sum_{i=1}^K (x_{ij})^2 - CF = Y^2 + \text{terms not involving } Y - CF$ ,

On similar lines we can write other sum of squares as follows:

$SST = (1/K) [(T_{i\cdot}' + Y)^2 + \text{terms not involving } Y] - CF$ .

And  $BSS = (1/b) [(T_{\cdot j}' + Y)^2 + \text{terms not involving } Y] - CF$ . Then error sum of squares ESS is given by:

$ESS = E = \text{Residual SS} = TSS - SST - BSS$ .

$$= Y^2 - (T_{i\cdot}' + Y)^2 / K + (T_{\cdot j}' + Y)^2 / b + (G' + Y)^2 / Kb + \text{constant terms not involving } Y. \quad (15.5.7)$$

Differentiating (15.5.7) w.r.t.  $Y$  and equating it to zero and evaluating for  $y$  we have:

$$\partial E / \partial Y = 0 = 2Y - (2/K) (T_{i\cdot}' + Y) - (2/b) (T_{\cdot j}' + Y) + (2/Kb) (G' + Y)$$

$$= Y [1 - 1/K - 1/b + 1/Kb] = T_{i\cdot}' / K + T_{\cdot j}' / b - G' / Kb$$

$$\Rightarrow Y = [b T_{\cdot j}' + K T_{i\cdot}' - G'] / (b-1)(K-1). \quad (15.5.8)$$

Thus, we can estimate the missing observation from the known values using (15.5.8).

We can extend this analysis for several missing values. But several missing values is not there in the syllabus hence, we confine only to single missing value in these lecture series.

**15.5.3. Statistical Analysis of RBD with missing values:**

After estimating the missing value using the (15.5.8), ANOVA of RBD is performed in the usual way, after substituting the estimated value in the place of missing observation. For each missing observation, 1 dof is to be subtracted from the total degrees of freedom and consequently from error dof. The adjusted treatment SS is to be obtained by subtracting the so called adjustment factor a.f. which is given as follows:

$$a.f. = (T_{\cdot j}' + K T_{i\cdot}' - G')^2 / [K(K-1)(b-1)^2] \quad (15.5.9).$$

**Example (15.5.2):** In the following table data represents the yields of 6 varieties in a 4 replicate RBD experiment with one missing value. Estimate the missing value and carryout the Analysis.

		Treatments					
Blocks.		V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>
Block – 1	B <sub>1</sub>	185	157	162	141	130	136
Block – 2	B <sub>2</sub>	117	----	129	144	169	125
Block – 3	B <sub>3</sub>	154	166	155	203	184	215
Block – 4	B <sub>4</sub>	165	186	127	157	165	180

Solution: First calculate the totals with the available data as follows:

		Treatments						
Blocks.		V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>	Block Totals
Block – 1	B <sub>1</sub>	185	157	162	141	130	136	911
Block – 2	B <sub>2</sub>	117	----	129	144	169	125	684=T <sub>.j</sub> '
Block – 3	B <sub>3</sub>	154	166	155	203	184	215	1078
Block – 4	B <sub>4</sub>	165	186	127	157	165	180	980
Treatment totals:		621	509=T <sub>i.</sub> '	573	645	648	657	3653=G'

Let the missing value if denoted by Y. The estimated value of Y using (15.5.8) we have":

$$\text{Estimated value of } Y = [\mathbf{b} T_{.j}' + K T_{i.}' - G'] / (\mathbf{b}-1)(K-1) = [4(648) + 6(509) - 3653] / 15$$

= [2592 + 3054 – 3653]/15 = 1993/15 = 132.86667  $\cong$  133. Include this estimated value in the table and carryout the totals as usual as follows:

		Treatments						
Blocks.		V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>	Block Totals
Block – 1	B <sub>1</sub>	185	157	162	141	130	136	911
Block – 2	B <sub>2</sub>	117	<b>133</b>	129	144	169	125	817
Block – 3	B <sub>3</sub>	154	166	155	203	184	215	1078
Block – 4	B <sub>4</sub>	165	186	127	157	165	180	980
Treatment totals:		621	642	573	645	648	657	3786=G

Null Hypothesis: For treatments: H<sub>01</sub>: There is no significant variation between varieties.

For blocks: H<sub>02</sub>: There is no significant variation between blocks.

Now calculate each observation squares as follows:

### Squires

Blocks.		V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>	Totals
Block – 1	B <sub>1</sub>	34225	24649	26244	19881	16900	18496	140395
Block – 2	B <sub>2</sub>	13689	<b>17689</b>	16641	20736	28561	15625	112941
Block – 3	B <sub>3</sub>	23716	27556	24025	41209	33856	46225	196587
Block – 4	B <sub>4</sub>	27225	34596	16129	24649	27225	32400	162224
Totals:		98855	104490	83039	106475	106542	112746	612147

$$\text{Step – 1: CF} = G^2/N = (3786)^2/24 = 597241.5$$

$$\begin{aligned} \text{Step – 2: SST} &= [(621)^2 + (642)^2 + (573)^2 + (645)^2 + (648)^2 + (657)^2]/4 - 597241.5 \\ &= [385641 + 412164 + 328329 + 416025 + 419904 + 431649]/4 - 597241.5 \\ &= 598428 - 597241.5 = 1186.5 = A. \end{aligned}$$

$$\begin{aligned} \text{BSS} &= [(911)^2 + (817)^2 + (1078)^2 + (980)^2] / 6 - 597241.5 \\ &= [829921 + 667489 + 1162084 + 960400]/6 - 597241.5 = 603315.6 - 597241.5 = 6074.1 = B \end{aligned}$$

$$\text{TSS} = 612147 - 597241.5 = 14905.5.$$

Step – 3: Now calculate the adjustment factor a.f. using (15.5.9).

$$\text{a.f.} = (T'_{.j} + KT'_{i.} - G')^2 / [K(K-1)(b-1)^2] = [684 + 6(509) - 3653]^2 / (6 \times 5 \times 9) = (7225)/270 = 26.75926$$

Therefore, adjusted value of treatment SS = 1186.5 - a.f. = 1186.5 - 26.75926 = 1159.74074 = A'

Step – 4: Construct ANOVA table of RBD with one observation missing.

**Table (15.5.4): ANOVA table of RBD with one missing observation.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	F – Cal
Treatments (columns)	K – 1 = 5	$S_t^2 = \text{Adj.SST} = A'$ = 1159.74074	$V_1 = S_t^2 / K - 1$ = 231.948	$F_1 = 231.948 / 547.97567 = 0.42$
Blocks (rows)	b – 1 = 3	$S_b^2 = \text{BSS} = B$ = 6074.1	$V_2 = S_b^2 / b - 1$ = 2024.7	$F_2 = 2024.7 / 547.97567 = 3.6949$
Error	(K-1)(b-1)-1 = 14	$S_E^2 = \text{ESS} = C - B - A'$ = 7671.66	$V_3 = S_E^2 / (K-1)(b-1)$ = 547.97567	
Total	Kb – 1 - 1 = 22	$S_T^2 = \text{TSS} = C$ = 14905.5		

Step – 5: For treatments/varieties: Since, the calculated value of F is less than 1 we cannot take any decision on varieties.

**For blocks:  $F_{tab}$  at 5% los for (3,14) dof is 3.34, and  $F_2$  is greater than the table value we reject the  $H_{02}$  and conclude that blocks differs significantly with each other.**

#### 15.5.4. Further Statistical Analysis used in RBD.

As in CRD, when we reject any null hypothesis, either for treatments or blocks, we have to apply further statistical analysis which is explained as follows:

As in CRD the Standard Error(S.E) of any treatment mean is given by:

$$V(\bar{t}_i) = S_E^2 / b \Rightarrow \text{S.E.}(\bar{t}_i) = S_E / \sqrt{b} \quad i=1,2,\dots,K. \quad (15.5.10)$$

Similarly the S.E. of difference between any two treatment means is obtained as  $(2S_E^2 / b)^{1/2}$ .

If the treatments shows significant effect, we might be interested to test for the significant difference between any two treatment means say  $\bar{t}_i$  and  $\bar{t}_j$   $i \neq j$ . Under the null hypothesis  $H_0$ : there is no significant difference between two treatments  $\bar{t}_i$  and  $\bar{t}_j$   $i \neq j$ .

$$\text{Then the test statistic } t = | \bar{t}_i - \bar{t}_j | / (2S_E^2 / b)^{1/2} \quad i \neq j. \quad (15.5.11)$$

Follow, Student's t distribution with error dof. That is  $(b-1)(K-1)$  dof. Let  $t_{0.05}$  at  $(b-1)(K-1)$  dof.

$$\text{Then the Critical Difference (C.D)} = [t_{0.05}(b-1)(K-1)] (2S_E^2 / b)^{1/2}. \quad (15.5.12)$$

We can also calculate the CD at 1% los by taking  $t_{0.01}$  at  $(b-1)(K-1)$  dof.

Compare each pair of treatment means difference with respective C.D. and conclusions can be drawn as usual. That is if  $| \bar{t}_i - \bar{t}_j | > \text{C.D.}$  at  $\alpha\%$  los, we there is significant difference between two treatments  $\bar{t}_i$  and  $\bar{t}_j$  for  $i \neq j$ . Group all those treatments with no significance difference in one group and treatments with significance difference in different groups as explained in CRD.

**Remark: If replications under two different treatments are not same, let  $r_i$  are number of replications under treatment  $t_i$  and  $r_j$  replications under treatment  $t_j$  then the S.E. of the difference between two treatment means is given by  $S_E (1/r_i + 1/r_j)^{1/2}$  as discussed in CRD with unequal number of replications under each treatment.**

#### 15.6. Summary

In this lesson we have introduced Randomized Block design (RBD) which is to be used when the experimental material is heterogeneous. For example, in agricultural experiments, experimental plots may differ w.r.t. fertility rate or climatic conditions or with respect to any other characteristic. In such situations we have to apply the Local Control principle by dividing heterogeneous units into homogeneous Blocks, such that units within the block are homogeneous. After dividing the experimental material into blocks, treatments are to be applied at random in each block separately and obtain the experimental data or yields. Such design is called RBD. After introducing the design, advantages, disadvantages, its statistical analysis is discussed. We have also discussed some important problems like comparing the



relative efficiency of RBD with CRD, missing plot technique and further statistical analysis applicable in RBD is discussed.

### 15.7. Self Assessment Questions

1. Explain the need of RBD and explain the Layout of the design.
2. Explain various advantages and disadvantages of RBD.
3. Briefly explain the Statistical Analysis of RBD.
4. Compare the relative efficiency of RBD with CRD.
5. Explain the method of estimating single missing observation in RBD.
6. Explain further statistical analysis to be applied in RBD.
7. Following data represents number of computer chips produced in 5 different machines in three shifts. These whether there is any significant effect of machines and shifts of workers.

	Machines				
	M <sub>1</sub>	M <sub>2</sub>	M <sub>3</sub>	M <sub>4</sub>	M <sub>5</sub>
Shift – 1:	555	549	542	521	552
Shift – 2:	561	612	530	589	563
Shift – 3:	669	637	669	685	654

Also apply further statistical analysis if necessary. Also calculate the Efficiency of RBD with CRD.

8. Consider the following experiment, where we have 4 blocks such that 6 plots in each block. Treatments A,B,C,D, E and F are applied randomly in block as follows:

A-247	C-277	B-206	D-162	E-162	F-249
C-227	B-288	A-273	D-150	F-225	E-170
F-263	A-385	D-196	C-368	B-395	E-154
E-177	D-141	B-310	F-226	A-285	C-349

Analyze the data with a suitable design and draw your conclusions. If necessary calculate further statistical analysis.

9. Estimate the missing value in the following table and carryout the suitable analysis.

	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>	V <sub>6</sub>
B <sub>1</sub>	16.2	14.1	18.5	13.6	13.0	15.7
B <sub>2</sub>	15.5	20.3	15.4	21.5	18.4	16.6
B <sub>3</sub>	11.7	12.9	----	14.4	16.9	12.5
B <sub>4</sub>	16.5	18.6	12.7	15.7	16.5	18.0

10. Explain the need of developing the theory of estimating the missing observation.

**15.8. Further Readings**

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons., New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 16.****Latin Square Design (LSD)****16.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the Latin Square Design (LSD).
- Assumptions, applications advantages and dis-advantages of LSD .
- Layout of LSD.
- Statistical Analysis of LSD.
- Comparing the efficiency of LSD with RBD with CRD.
- Missing plot technique in LSD.

**Structure**

- 16.1. Introduction to Latin Square Design (LSD)**
- 16.2. Layout, Advantages and disadvantages of LSD**
- 16.3. Mathematical model and assumptions of LSD**
- 16.4. Statistical Analysis of LSD – with one observation per cell**
- 16.5. Important Concepts of LSD**
- 16.6. Summary**
- 16.7. Self Assessment Questions**
- 16.8. Further Readings**

**16.1. Introduction to Latin Square Design (LSD)**

In the last lesson, we have discussed the Randomized Block Design, where the entire Experimental material is divided into homogeneous blocks and treatments are applied at random within each block independently. Here we have assumed that the plots or units within the block are homogeneous. In some situations, this assumption may not be true, particularly, when we want to test many treatments, say  $K > 10$  or  $12$ . It is very difficult to get so many homogeneous experimental units / plots. Hence, there is a necessity to further divide plots in each block. Thus we have two types of blocks namely, Horizontal Blocks (Rows) and Vertical blocks (Columns). Thus in Latin Square Design we apply the local control twice, once for rows and once for columns. Thus in LSD, we eliminate the variation in experimental units is controlled in two perpendicular directions. This type of situation will occur in Agricultural Experiments, it may happen that experimental area (field) exhibits fertility in strips say High fertility, medium fertility and Low fertility strips. These strips may be Horizontal or Vertical. In each Strip, plots are again vary with respect to some other character like type of cultivation or type of irrigation or water type like well water, bore water or rain water and so on. In such

situations, we have to control the variation in two directions one Vertical and one Horizontal direction. The design used in such situation is known as Latin Square Design (LSD).

### 16.2. Layout, Advantages and disadvantages of LSD.

Now we proceed to discuss some important concepts related to Latin Square Design (LSD) in this section. First we concentrate on the Layout of the design as follows:

**16.2.1: Layout of the design:** In this design, the number of treatments must be equal to Number of replications, so that we obtain a square of experimental units. That is if we have K treatments, then we must have K replications under each treatment and hence, we must have  $K \times K = K^2$  experimental units / plots which are arranged in the form a square having K-rows and k-columns. Then , these K treatments are to be arranged in such a manner that, **each treatment should occur once and only once in each row and in each column.** This property of occurrence of each treatment once in each row and once in each column is known as “**Latin Property**”. For example, consider the following design with treatments K = 5; say A,B,C,D and E. The following is one layout of the design LSD.

Table (16.2.1) Example of Latin Square Design.

A	B	C	D	E
B	C	D	E	A
C	D	E	A	B
D	E	A	B	C
E	A	B	C	D

**Important Note:** It is important to note that in the above example, each treatment occurs once and only once in each row and in each column. Simple rule to construct such design is that “Circular Rule”. That is assume that the treatments are in the form of a Circle. Then start with first letter A,B,C,D and E. The next row starts with the second letter B,C,D,E and A. Continue this procedure until last row. The resultant design is called Latin Square Design. We can also construct similar design with integers 1,2,3,4 and 5.

1	2	3	4	5
2	3	4	5	1
3	4	5	1	2
4	5	1	2	3
5	1	2	3	4

Thus we obtain a square design with Latin Property, that is **occurrence of each treatment of symbol once and only once in each row and in each column.** Fisher and Yates have tabulated different Latin Square Designs up to size K = 12. Such designs have the following important properties:

### 16.2.2. Properties of Latin Square Design (LSD):

1. **Interchanging of any rows does not affect the Latin property.** If we interchange any two rows, still occurrence of each treatment once in each row and in each column will not be affected. For example, inter change first with fourth row in the above design, then we obtain the following design:

4	5	1	2	3
2	3	4	5	1
3	4	5	1	2
1	2	3	4	5
5	1	2	3	4

In the resultant design also the Latin Property holds good. The resultant design is also a Latin Square Design.

1. **Interchange of any two columns does not affect the Latin Property.** That we can interchange 3<sup>rd</sup> column with 5<sup>th</sup> column, the resultant design is also a Latin Square Design.
2. **Similarly interchange of any symbols for treatments will not affect the Latin property.** For example Instead of A or 1 substitute D or 4 and in the place of D or 4; substitute A or 1 in the above designs, still we obtain a Latin Design.

Thus the Latin Property, in the above designs play a vital role which ensures that interaction effects between the treatments are absent. Thus we can estimate the treatment effects independently.

**Selecting a Random LSD:** First select an LSD as explained above of the given size K.

**Step – 1 :** Randomly inter change rows.

**Step – 2:** Randomly inter change columns in the design obtained in the Step – 1.

**Step – 3:** Randomly inter change symbols for treatments in the design obtained in Step – 2.

The design obtained after Step – 3 is the random LSD selected for implementation. That is apply the treatments to the experimental units in the same order, as given in the random design selected after the Step – 3.

### 16.2.3. Advantages of LSD:

1. Since, local control is applied twice in LSD, or grouping is done twice in LSD, once for rows and once for columns, we have smaller experimental error or residual sum of squares in LSD than CRD or RBD.
2. It can be observed that LSD is an incomplete 3-way layout. Its basic advantage when compared to 3-way classification is that we require only  $m^2$  experimental in LSD than  $m^3$  experimental units required for 3-way classification. Thus if  $K=4$ , for 3-way classification, we require  $4 \times 4 \times 4 = 64$  experimental where as for LSD  $4 \times 4 = 16$  experimental units. Thus  $64 - 16 = 48$  experimental units in LSD than in three way

classification. Similarly, if  $K=5$ , we are saving 100 experimental units in LSD than 3-way classification.

3. The Statistical analysis is relatively simple when compared to Greco-Latin Square design (GLSD) and slightly completed when compared to CRD or RBD.
4. Missing plot or mixed up plot problems can easily be dealt in LDS as in RBD.
5. More than one factor can be investigated In LSD than other designs like CRD or RBD.
6. Error mean square or experimental is smallest when compared to CRD or RBD. Similarly, this design has smallest degrees of freedom when compared to CRD or RBD.

#### 16.2.4: Disadvantages of LSD:

1. LSD basically assumes that there are no interaction effects between the factors considered, for the experiment. This assumption may not be true in general, in many practical problems.
2. Number of treatments are fixed in LSD. We have to consider as many treatment as many number of rows and columns  $K$ . We have no freedom to add or remove any rows or columns or treatments as we can do in CRD or RBD.
3. The problem of several missing values, and its analysis is complicated and quite complex. We can easily eliminate the corresponding column or row in CRD or RBD without disturbing other treatment effects. This facility or easiness in the analysis is not possible in LSD.
4. The lay out of RBD or CRD is easier than LSD. Randomization is restricted only to plots. In CRD randomization is applied though out the experimental units and in RBD randomization is restricted to each block. Where is randomization in LSD is restricted to each plot, because, each treatment must occur once and only once in each row and in each column, because of **Latin property**.

### 16.3. Mathematical model and Assumptions of LSD

Let  $y_{ijk}$  represent the yield or response of the plot or unit in  $i^{\text{th}}$  row, belonging to  $j^{\text{th}}$  column receiving  $k^{\text{th}}$  treatment.  $i, j, k = 1, 2, \dots, m$ . The randomly selected scheme of LSD is applied and the data is obtained in the  $m \times m$  matrix form as randomly selected LSD.

Then the mathematical of the observation  $y_{ijk}$  is given as follows:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \epsilon_{ijk} ; i, j, k = 1, 2, \dots, m. \quad (16.3.1)$$

where:  $\mu$  = General mean effect,

$\alpha_i$  =  $i^{\text{th}}$  row mean effect,

$\beta_j$  =  $j^{\text{th}}$  column mean effect,

$\tau_k$  =  $k^{\text{th}}$  treatment mean effect and

$\epsilon_{ijk}$  = Random effects on  $(i, j, k)$  plot.

**16.3.1. Assumptions of LSD:**

1. The triplet (l,j,k) assumes only  $m^2$  values of the possible  $m^3$  values of the randomly selected Latin Square design.
2. In LSD we assume that there is no interaction effect between different factors under consideration.
3. That is we assume that factors considered are independent of each other.
4. Random errors  $\epsilon_{ijk}$ 's are independent and identically distributed.
5. Random errors  $\epsilon_{ijk}$  are distributed Normal with mean zero variance  $\sigma^2\epsilon$ .
6. It is assumed that m treatments are arranged in  $m^2$  plots arranged in m rows and m columns.

**16.4. Statistical Analysis of LSD – with one observation per cell.**

First frame various null hypotheses for LSD as follows:

A – For Treatments:  $H_{01} : \tau_1 = \tau_2 = \dots, \tau_m = 0$  or **All the treatments are equally effective.**

B – For Rows:  $H_{02} : \alpha_1 = \alpha_2 = \dots, \alpha_m = 0$  or **all the rows are equally effective.**

C – For Columns:  $H_{03} : \beta_1 = \beta_2 = \beta_3 \dots \dots \dots = \beta_m = 0$  or **all the columns are equally effective.**

The short cut step wise procedure adopted in many practical problems in analyzing LSD with one observation per experimental unit is explained as follows:

First construct the table of treatment effects treatment wise and get their totals  $T_i$ , Row totals  $R_i$  and Column totals  $C_j$  from the given data.

Step – 1: Calculate the Grand total  $G = \sum_{k=1}^m T_k = \sum_{i=1}^m R_i = \sum_{j=1}^m C_j$ .

Step – 2: Calculate the Correction Factor=  $CF = G^2/N$ , where  $N = m^2$ .

Step – 3: Calculate the treatment Sum of squares =  $SST = \sum_{k=1}^m T_k^2 / m - CF = A$ .

Step – 4: Calculate the Rows Sum of Squares =  $RSS = \sum_{i=1}^m R_i^2 / m - CF = B$ .

Step – 5: Calculate the Columns Sum of Squares =  $CSS = \sum_{j=1}^m C_j^2 / m - CF = C$ .

Step – 6: Calculate Total sum of square =  $TSS = \sum_{j=1}^m \sum_{i=1}^m (y_{ij})^2 - CF = D$ .

Step – 7: Construct the following ANOVA table for LSD with one observation per cell.

**Table (16.4.1): ANOVA table of LSD with one observation per cell.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments	$m - 1$	$S_t^2 = SST = A$	$V_1 = S_t^2 / m - 1$	$F_T = V_1 / V_4$
Rows	$m - 1$	$S_r^2 = RSS = B$	$V_2 = S_r^2 / m - 1$	$F_R = V_2 / V_4$
Columns	$m - 1$	$S_c^2 = CSS = C$	$V_3 = S_c^2 / m - 1$	$F_C = V_3 / V_4$

Error	$(m-1)(m-2)$	$S_E^2 = ESS$ $= D - C - B - A = E$	$V_4 = S_E^2 / (m-1)(m-2)$	
Total	$N - 1$	$S_T^2 = TSS = D$		

Step – 8: Conclusions:

A – For treatments:

(1) If the calculated value  $F_T$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for  $[(m-1), (m-1)(m-2)]$  dof, we accept the null hypothesis for treatments  $H_{01} : \tau_1 = \tau_2 = \dots, \tau_m = 0$  and conclude that there is no significant effect of treatment means. In this case we need not apply Further Statistical Analysis because all the treatments are equally effective.

(2) If the calculated value  $F_T$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for  $[(m-1), (m-1)(m-2)]$  dof, we reject the null hypothesis for treatments  $H_{01} : \tau_1 = \tau_2 = \dots, \tau_m = 0$  and conclude that there is significant difference in the effect of treatment means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the treatment effects or to determine which treatment has maximum effect? Or which treatment has minimum effect? And so on.

B- For Rows:

(1) If the calculated value  $F_R$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for  $[(m-1), (m-1)(m-2)]$  dof, we accept the null hypothesis for rows  $H_{02} : \alpha_1 = \alpha_2 = \dots, \alpha_m = 0$  and conclude that there is no significant effect of row means. In this case we need not apply Further Statistical Analysis because all the rows are uniform with respect to the yield. This means that rows effect is negligible.

(2) If the calculated value  $F_R$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for  $[(m-1), (m-1)(m-2)]$  dof, we reject the null hypothesis for rows  $H_{02} : \alpha_1 = \alpha_2 = \dots, \alpha_m = 0$  and conclude that there is significant difference in the effect of row means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the rows or to determine which row has maximum fertility rate? Which row has minimum fertility rate? Which rows have similar fertility rates? And so on.

C. For Columns:

(1) If the calculated value  $F_C$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for  $[(m-1), (m-1)(m-2)]$  dof, we accept the null hypothesis for columns  $H_{03} : \beta_1 = \beta_2 = \beta_3 \dots \dots \dots = \beta_m = 0$  and conclude that there is no significant effect of column means. In this case we need not apply Further Statistical Analysis because all the rows are uniform with respect to the yield. This means that columns effect is negligible.

(2) If the calculated value  $F_C$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for  $[(m-1), (m-1)(m-2)]$  dof, we reject the null hypothesis for columns  $H_{03} : \beta_1 = \beta_2 = \beta_3 \dots \dots \dots = \beta_m = 0$  and conclude that there is significant difference in the effect of column means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the rows or to determine which



column has maximum fertility rate? Which column has minimum fertility rate? Which columns have similar fertility rates? And so on.

The above explained procedure is explained with an example as follows:

Example (16.4.1): The following data represent LSD with 4 treatments. Analyze the data and if necessary apply further statistical analysis, where ever it is required. Data is measured in Kgs.

C : 16.4	A : 10.2	D : 21.2	B : 19.1
D : 29.1	B : 18.9	C : 29.4	A : 5.7
B : 24.9	C : 41.7	A : 9.5	D : 28.9
A : 5.4	D : 38.8	B : 24.0	C : 37.0

Solution: In the given problem  $m = 4$  and  $N = 16$ . Then frame the null hypothesis as follows:

A – For Treatments:  $H_{01} : \tau_1 = \tau_2 = \dots = \tau_m = 0$  or **All the treatments are equally effective.**

B – For Rows:  $H_{02} : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$  or **all the rows are equally effective.**

C – For Columns:  $H_{03} : \beta_1 = \beta_2 = \beta_3 \dots = \beta_m = 0$  or **all the columns are equally effective.**

Now calculate treatment totals, row totals, column totals and Grand total as follows.

	Column – 1	Column - 2	Column - 3	Column – 4	Row totals( $R_i$ )
Row – 1	C : 16.4	A : 10.2	D : 21.2	B : 19.1	66.9
Row – 2	D : 29.1	B : 18.9	C : 29.4	A : 5.7	83.1
Row – 3	B : 24.9	C : 41.7	A : 9.5	D : 28.9	105.0
Row – 4	A : 5.4	D : 38.8	B : 24.0	C : 37.0	105.2
Column Totals( $C_j$ )	75.8	109.6	84.1	90.7	360.2=G

Now Calculate treatment totals as follows:

	Treatment totals ( $T_k$ )				
A:	5.4	10.2	9.5	5.7	30.8
B:	24.9	18.9	24.0	19.1	86.9
C:	16.4	41.7	29.4	37.0	124.5
D:	29.1	38.8	21.2	28.9	118.0
	G=360.2.				

Now calculate various sum of squares as follows:

$$\text{Step – 1: } CF = G^2/N = (360.2)^2/16 = 8109.0025$$

$$\text{Step – 2: } SST = (1/4)[(30.8)^2 + (86.9)^2 + (124.5)^2 + (118.0)^2] - CF = 37924.50/4 - 8109.0025$$

$$= 1372.1225=A$$

$$\text{Step - 3: RSS} = (1/4)[(66.9)^2+(83.1)^2+(105.0)^2+(105.2)^2] - CF = 33473.26/4 - 8109.0025 \\ = 259.3125 = B.$$

$$\text{Step - 4: CSS} = (1/4)[(75.8)^2+(109.6)^2+(84.1)^2+(90.7)^2] - CF = 33057.10/4 - 8109.0025 \\ = 155.2725 = C.$$

$$\text{Step - 5: TSS} = [(16.4)^2 + \dots + (37.0)^2] - CF = 10052.08 - 8109.0025 = 1943.0775 = D.$$

Step - 6: Construct the ANOVA table for LSD as follows:

Table (16.4.2): ANOVA table of LSD.

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F - cal)
Treatments	$m - 1=3$	$S_t^2 = SST=A$ =1372.1225	$V_1 = S_t^2/m- 1$ =457.3742	$F_T = V_1 / V_4$ =17.5496
Rows	$m-1=3$	$S_r^2 = RSS=B$ =259.3125	$V_2 = S_r^2/m-1$ =86.4375	$F_R = V_2/ V_4$ =3.31664
Columns	$m- 1=3$	$S_c^2=CSS=C$ =155.2725	$V_3 = S_c^2/m-1$ =51.7575	$F_C = V_3 / V_4$ =1.98596
Error	$(m-1)(m-2)=6$	$S_E^2 = ESS$ =D- C-B-A =E =156.3700	$V_4 = S_E^2 /((m-1)(m-2))$ =26.0617	
Total	$N - 1=15$	$S_T^2 = TSS=D$ =1943.0775		

Step - 7: Conclusions:

A: For treatments: Since calculated value of  $F_T$  is greater than  $F_{tab}$  at 5% los at (3,6) dof is 4.76 and hence we reject  $H_{01}$  at 5% los and conclude that the treatments A,B,C and D differs significantly from each other. Hence, we have to apply further Statistical analysis to sub-group the treatments.

B: For rows: Since calculated value of  $F_R$  is less than  $F_{tab}$  at 5% los at (3,6) dof, that is 4.76, we accept the null hypothesis  $H_{02}$ . Hence, we conclude that there is no significant variation in the rows considered. Since we have accepted the null hypothesis for rows, we need not apply further statistical analysis for rows.

C: For columns: Since calculated value of  $F_C$  is less than  $F_{tab}$  at 5% los at (3,6) dof, that is 4.76, we accept the null hypothesis  $H_{03}$ . Hence, we conclude that there is no significant variation in

the columns considered. Since we have accepted the null hypothesis for columns, we need not apply further statistical analysis for columns.

#### Further Statistical Analysis for treatments:

Since, we have rejected for treatments, there is a need to apply further statistical analysis to sub-group the treatments, which is done as follows:

S.E. of difference between any two treatment means =  $S_E \sqrt{(2/m)} = \sqrt{(2(26.0616)/4)} = 3.609$ .

Now treatment means are calculated as follows:

For treatment – A:  $30.8/4 = 7.70 = \bar{y}_A$

For treatment – B:  $86.9/4 = 21.725 = \bar{y}_B$

For treatment – C:  $124.5/4 = 31.125 = \bar{y}_C$

For treatment – D:  $118.0/4 = 29.5 = \bar{y}_D$

Grouping is done using the following difference table:

Difference between treatments	Value	Significant or not.
1. $ \bar{y}_A - \bar{y}_B  =  7.70 - 21.725  = 14.025$	$> 3.609$	Significant difference.
2. $ \bar{y}_A - \bar{y}_C  =  7.70 - 31.125  = 23.425$	$> 3.609$	Significant difference.
3. $ \bar{y}_A - \bar{y}_D  =  7.70 - 29.5  = 21.8$	$> 3.609$	Significant difference.
4. $ \bar{y}_B - \bar{y}_C  =  21.725 - 31.125  = 9.4$	$> 3.609$	Significant difference.
5. $ \bar{y}_B - \bar{y}_D  =  21.725 - 29.5  = 7.775$	$> 3.609$	Significant difference.
6. $ \bar{y}_C - \bar{y}_D  =  31.125 - 29.5  = 1.625$	$< 3.609$	No significant difference.

Treatment – C has maximum yield and treatment – A is having least yield. Treatments C and D forms one group. Treatments A and B forms different groups. Thus we have to form three groups for treatments, namely **Group – 1: Treatment – A. Group – 2: Treatment – B and Group – 3: Treatments C and D.**

### 16.5. Important Concepts of LSD

Now we proceed to discuss some important concepts of LSD like (1) Estimating missing value(s) and (2) Comparing the efficiency of LSD with RBD and CRD.

#### 16.5.1: Estimation of missing value in LSD:

Let the value  $y_{ijk} = Y$  is missing value in  $i^{\text{th}}$  row,  $j^{\text{th}}$  column under  $k^{\text{th}}$  treatment, missed due to some natural calamity or un-avoidable circumstances like floods or fire accidents. Unless otherwise, we estimate the missing value, we cannot carry out the analysis of LSD. The method of estimation of missing value in LSD of size  $m$  is explained as follows:

Let  $R$  = Total of all observations in the row the missing value  $Y$  belongs.

Similarly, C = Total of all observations in the column in which the missing value Y belongs.

T = Total of all known observations under the treatment in which missing value Y belongs.

G' = Grand total of all known observations. Then we have various sum of squares as follows:

$$TSS = Y^2 + \text{Constant terms w.r.t. } Y - (G' + Y)^2/m^2.$$

$$RSS = (R+Y)^2/m + \text{Constant terms w.r.t. } Y - (G' + Y)^2/m^2.$$

$$CSS = (C+Y)^2/m + \text{Constant terms w.r.t. } Y - (G' + Y)^2/m^2.$$

$$SST = (T+Y)^2/m + \text{Constant terms w.r.t. } Y - (G' + Y)^2/m^2.$$

Then Experimental error = E = TSS – RSS – CSS – SST. Thus we have:

$$E = Y^2 - (1/m) [(R+Y)^2 + (C+Y)^2 + (T+Y)^2] + 2(G' + Y)^2/m^2. \quad (16.5.1)$$

Differentiating E in equation (16.5.1) w.r.t. Y and equating to zero we have:

$$\begin{aligned} \partial E / \partial Y = 0 &= 2Y - 2/m[R+C+T + 3Y] - 4(G' + Y)/m^2 \Rightarrow (m^2 - 3m + 2) Y = m(R+C+T) - 2G' \\ \Rightarrow Y &= [m(R+C+T) - 2G'] / (m-1)(m-2). \end{aligned} \quad (16.5.2)$$

By using equation (16.5.2) we can estimate the value of Y denoted by  $\hat{Y}$ .

**Statistical Analysis of LSD with one missing Value:** After substituting the estimated value  $\hat{Y}$ , in the place of Y we can carry out the analysis of LSD as usual except the difference that total degrees of freedom is to be reduced by 1 because we have estimated one missing value. The Adjustment Factor (AF) for treatment sum of squares is that:

$$AF = [(m-1)T+R+C - G']^2 / [(m-1)(m-2)]^2. \quad (16.5.3)$$

The Adjustment Factor is to be subtracted from treatment sum of squares (SST) to get Adjusted treatment sum of squares.

When further statistical analysis is required, we know that the S.E. of difference between two treatments is  $S_E \sqrt{(2/m)}$ , if no observation is missing. If the treatment involves missing observation then S.E. of such treatment difference is given by:

$$S_E \{ \{ (2/m) + [1 / (m-1)(m-2)] \}^{1/2} \} \quad (16.5.4)$$

Same procedure may be followed when we miss more than two, say 'M' missing values by differentiating E, m times we obtain M simultaneous equations. Solving these M simultaneous equations, we can estimate M missing values. After substituting these M estimated values in the data, analysis of LSD can be carried out as usual except the difference that total degrees of freedom is to be reduced by number of missing values estimated, namely M. That is degrees of freedom for total will become  $(m^2 - M - 1)$ . Here M represent number of missing values.

Example (16.5.1): Estimate the missing value in the following LSD.

B: 405	A : 525	E : 463	D : 441	C : 481
A : 552	C : 431	D : 425	E : 572	B : 451

E : 471	B : 492	A : 472	C : 381	D : 410
D : 430	E : 469	C : 432	B : -----	A : 460
C : 325	D : 445	B : 429	A : 513	E : 493

Solution: Let the missing value is denoted by Y. Now calculate the totals of known observations of each row, column and treatment as follows:

B: 405	A : 525	E : 463	D : 441	C : 481	2315
A : 552	C : 431	D : 425	E : 572	B : 451	2431
E : 471	B : 492	A : 472	C : 381	D : 410	2226
D : 430	E : 469	C : 432	B : Y	A : 460	1791+Y
C : 325	D : 445	B : 429	A : 513	E : 493	2205
2183	2362	2221	1907 +Y	2295	10968+Y

Treatment Totals:

A:	552	525	472	513	460	2522
B:	405	492	429	Y	451	1777+Y
C:	325	431	432	381	481	2050
D:	430	445	425	441	410	2151
E:	471	469	463	572	493	2468

From the above calculations we have R = 1791; C= 1907; T = 1777 m = 5; M = 1 and  $G' = 10968$ .

Using the formula (16.5.2), we have:

$$Y = [m(R+C+T) - 2G'] / (m-1)(m-2) = [5(1791+1907+1777) - 2(10968)] / 4(3)$$

$$= (27375 - 21936) / 12 = 5439 / 12 = 453.25.$$

The estimated missing value  $\hat{Y}$  is 453.25. Substitute this estimated value in the place of Y and carryout the analysis of LSD by reducing 1 dof from the total degrees of freedom. Thus we have total degrees of freedom is 23 and error degrees of freedom will become 11.

Now we proceed to explain the procedure to calculate the efficiency of LSD with RBD and CRD as follows.

### 16.5. Efficiency of LSD

In section 15.5.1, we have discussed the method of comparing two designs using Uniformity Trials and compared the efficiency of RBD with CRD. On similar lines, we discuss the method of comparing LSD with RBD and CRD. When we want to compare LSD with RBD, we have two cases (i) When rows are considered as blocks and (ii) When columns are considered as blocks. Now we discuss above two cases as follows:

#### Case – 1: Relative efficiency of LSD over RBD when rows as considered as blocks:

Proceeding on similar lines in section (15.5.1), we finally get the formula for obtaining relative efficiency of LSD with RBD when columns are considered as blocks is:

$$RE = [V_3 + (m - 1) V_4] / m(V_4) \quad (16.5.1)$$

Where,  $V_3$  and  $V_4$  are the values obtained in the ANOVA table of LSD given in table (16.4.1).

Example (16.5.1): Applying the formula (16.5.1) for the example (16.4.1), we have:

$$SE = [51.7575 + 3(26.0616)]/4(26.0616) = 1.2465.$$

Expressed SE in terms of percentage we have app. 125%. That is the efficiency of LSD is 125 and the efficiency of RBD is 100%. Thus a gain of 25% of the reduction of error is attained by conducting LSD instead of RBD with rows as blocks.

### Case – 2: Relative efficiency of LSD over RBD when columns are considered as blocks:

Proceeding on similar lines in section (15.5.1), we finally get the formula for obtaining relative efficiency of LSD with RBD when columns are considered as blocks is:

$$RE = [V_2 + (m - 1) V_4] / m(V_4) \quad (16.5.2)$$

Where,  $V_2$  and  $V_4$  are the values obtained in the ANOVA table of LSD given in table (16.4.1).

Example (16.5.2): Applying the formula (16.5.2) for the example (16.4.1), we have:

$$SE = [86.4375 + 3(26.0616)]/4(26.0616) = 1.5792.$$

Expressed SE in terms of percentage we have app. 158%. That is the efficiency of LSD is 158 and the efficiency of RBD is 100%. Thus a gain of 58% of the reduction of error is attained by considering LSD than RBD with columns as blocks.

Now we proceed to discuss the relative efficiency of LSD with CRD in the following sub-section:

#### 16.5.1: Relative efficiency of LSD over CRD:

Proceeding on similar lines in section (15.5.1), we finally get the formula for obtaining relative efficiency of LSD with CRD is:

$$RE = [V_2 + V_3 + (m - 1) V_4] / m(V_4) \quad (16.5.3)$$

Where,  $V_2$ ,  $V_3$  and  $V_4$  are the values obtained in the ANOVA table of LSD given in table (16.4.1).

Example (16.5.3): Applying the formula (16.5.3) for the example (16.4.1), we have:

$$SE = [86.4375 + 51.7575 + 3(26.0616)]/4(26.0616) = 1.6605.$$

Expressed SE in terms of percentage we have app. 166%. That is the efficiency of LSD is 166 and the efficiency of CRD is 100%. Thus a gain of 66% of the reduction of error is attained by considering LSD instead of CRD.

## 16.6. Summary

In this lesson, we have introduced the Latin Square design (LSD), where we apply the principle of local control twice, once for rows and once for columns. In this design, each treatment occurs once and only once in each row and in each column. This is an important property in LSD which is known as “Latin Property”. With this property, we ensure that interaction effects will be absent and we can estimate the treatment effects independently. Properties of LSD, assumptions, and the analysis of LSD is discussed. We have also discussed the procedure and the analysis of missing plots technique and compared the relative efficiency

of LSD with RBD and CRD. It is interesting to note that in LSD mean square error is reduced when compared to RBD or CRD. We have also discussed that a gain of 66% can be obtained in LSD when compared to CRD.

### 16.7. Self Assessment Questions

1. Define the Latin property and discuss its properties.
2. Define the method of construction of LSD with an example.
3. Discuss various assumptions and applications of LSD.
4. Discuss the mathematical model of LSD along with its assumptions.
5. Explain the statistical analysis of LSD.
6. Explain the method of estimation of missing observation and its analysis.
7. Explain the method of comparing relative efficiency of LSD with RBD and CRD.
8. Write down the ANOVA table of LSD with 2 missing observations.
9. Analyze the following data suitably to test the effect of 5 varieties of seeds of a crop. The yield of the crop is measured in tons.

E : 3.14	A : 3.08	D : 3.23	B : 3.09	C : 3.03
B : 3.12	E : 3.03	C : 3.25	A : 3.07	D : 3.02
A : 3.16	D : 3.06	E : 3.10	C : 3.04	B : 2.94
C : 3.10	B : 3.12	A : 3.20	D : 3.06	E : 3.00
D : 3.09	C : 3.10	B : 3.24	E : 3.00	A : 3.05

Also calculate the efficiency of this design w.r.t. RBD and CRD.

10. Estimate the missing value in the following data.

C : -----	B : 12	A : 20	D : 16	E : 10
E : 08	C : 18	B : 12	A : 15	D : 12
D : 09	A : 07	C : 16	E : 14	B : 24
B : 16	E : 13	D : 11	C : 22	A : 17
A : 23	D : 14	E : 18	B : 09	C : 22

### 16.8. Further Readings

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons, New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 17.****Graeco-Latin Square Design (GLSD)****17.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the concept of orthogonal Latin Squares.
- Definition of Graeco-Latin Square Design (GLSD)
- Assumptions, applications advantages and dis-advantages of GLSD .
- Layout of GLSD.
- Statistical Analysis of GLSD.
- Mutually Orthogonal Latin Square Designs.

**Structure**

**17.1. Introduction to Orthogonal Latin Square Designs**

**17.2. Introduction of Graeco-Latin Square Design (GLSD)**

**17.3. Mathematical model and assumptions of GLSD**

**17.4. Statistical Analysis of GLSD**

**17.5. Mutually Orthogonal Latin Square Designs**

**17.6. Summary**

**17.7. Self Assessment Questions**

**17.8. Further Readings**

**17.1. Introduction to Orthogonal Latin Square Designs**

In the last lesson, we have discussed Latin Square Designs (LSD). Now we proceed to discuss some advanced properties of LSD. That is the concept of “**Orthogonality**” and its applications in the construction of LSD’s. The concept of Orthogonality is related to the effect of treatment-A in the presence of another treatment-B. If treatments A and B are independent effects, we can estimate their effects easily, because they are independent. Otherwise, we have to consider combined effect or interaction effect between both the treatments A and B.

Such type of interaction effects are studied through “**Factorial Designs**” which are discussed in the next unit. At present we assume that the treatments have independent effects. These effects are known as “**Main Effects**”. These effects can be estimated independently, if the treatments are **orthogonal** to each other. The concept of Orthogonality is defined as follows:

**17.1.1: Definition of Orthogonality:**

Two sets of numbers  $\{a_i\}$  and  $\{b_i\}$   $i=1,2,\dots,n$  are said to be orthogonal if the sum of their products is zero. That  $\sum_{i=1}^n (a_i)(b_i) = 0$ . For example, consider  $a_i = -2, 1, 1$  and  $b_i = 0, -1, 1$ . Since  $\sum_{i=1}^3 (a_i)(b_i) = -2 \times 0 + 1 \times -1 + 1 \times 1 = 0$ . Hence  $\{a_i\}$  and  $\{b_i\}$  are orthogonal to each other.



Generally the concept of Orthogonality is useful for “**Preplanned Comparisons**”. That is the treatments are specified prior to running of the experiment and obtaining the experimental results or data. This is because of the fact that the experimenter wants to estimate the effects of treatments independently, without depending on other effects. Orthogonality ensures the independent estimation of treatment’s effects. Hence, we can estimate interaction effects or main effects of the treatment effects independently from the experimental data.

On similar lines, we can extend the property of Orthogonality for two experimental designs as follows:

**Orthogonal LSDs:** Consider two Latin Square Designs, namely, LSD-1 and LSD-2 of same size say  $K \times K$ . Two LSD are said to be orthogonal to each other if “ **we super impose LSD-1 on LSD-2, each treatment combination must appear once and only once in the super imposed design**”.

For example consider the following two LSDs of size  $K = 4$ .

<b>LSD – 1 :</b>	2	1	4	3	and	<b>LSD – 2:</b>	D	C	B	A
	1	2	3	4			A	B	C	D
	3	4	1	2			B	A	D	C
	4	3	2	1			C	D	A	B

If we super impose LSD – 1 over LSD – 2 we have the following design:

<b>LSD – 1 X LSD – 2 =</b>	2D	1C	4B	3A
	1A	2B	3C	4D
	3B	4A	1D	2C
	4C	3D	2A	1B

In the resultant design, each treat pair occur once and only once in the entire design. For example, 2B occurred once and only once in the entire design. Then, we say that both LSDs are orthogonal to each other. With this, we can estimate two sets of treatment effects represented by numbers and alphabets can be estimated independently from the experimental data.

Now we proceed to explain Graeco Latin Square Design (GLSD) in the following section.

## 17.2. Introduction of Graeco-Latin Square Design (GLSD).

**Now we proceed to introduce some concepts of Graeco Latin square Design (GLSD).**

**Definition of GLSD:** In GLSD, we use two letters namely Latin letters A,B,C,D,.. and Greek letters  $\alpha, \beta, \gamma, \delta, \dots$ , to represent two sets of treatments and the layout of the design done such the selected design has orthogonal property. The resultant design is called Graeco Latin square Design (GLSD).

For example the GLSD of size 4 is given by:

Example of GLSD:	$\beta$ D	$\alpha$ C	$\delta$ B	Y A
	$\alpha$ A	$\beta$ B	Y C	$\delta$ D

Y B	δ A	α D	β C
δ C	YD	β A	α B.

Above design allows investigation of four factors, namely Rows, Columns, Latin letters and Greek letters and requires only  $4 \times 4 = 16$  experimental units. Such type of designs we can construct for all values of  $k > 2$  except for  $k = 6$ .

#### Advantages of GLSD:

- (1) We can apply local control three times and hence we can reduce the experimental error much smaller than LSD.
- (2) We can test double the number of treatments that is  $2K$  treatments in GLSD when compared to LSD.
- (3) We require same number of experimental units in GLSD that is  $K \times K$  experimental units as we require in LSD. Thus we can reduce the number of experimental units required to conduct the experiment for testing  $2K$  treatments.
- (4) GLSD reduces experimental units when compared to four way analysis. We use only  $4 \times 4 = 16$  experimental units in GLSD than  $4 \times 4 \times 4 \times 4 = 256$  units required for four way analysis.
- (5) Experimental error or Error Variance is least in GLSD when compared to all the previously introduced designs, namely LSD, RBD and CRD.

Now we proceed to explain the mathematical model of GLSD in the following section.

### 17.3. Mathematical model and assumptions of GLSD.

Let  $y_{ijkl}$  represent the yield or response of the plot or unit in  $i^{\text{th}}$  row, belonging to  $j^{\text{th}}$  column receiving  $k^{\text{th}}$  Latin letter treatment and  $l^{\text{th}}$  Greek letter treatment.  $i, j, k, l = 1, 2, \dots, m$ . The randomly selected scheme of GLSD is applied and the data is obtained in the  $m \times m$  matrix form as randomly selected LSD.

Then the mathematical of the observation  $y_{ijk}$  is given as follows:

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \tau_k + \Upsilon_l + \varepsilon_{ijkl} ; i, j, k, l = 1, 2, \dots, m. \quad (17.3.1)$$

where:  $\mu$  = General mean effect,

$\alpha_i$  =  $i^{\text{th}}$  row mean effect,

$\beta_j$  =  $j^{\text{th}}$  column mean effect,

$\tau_k$  =  $k^{\text{th}}$  Latin letter treatment mean effect

$\Upsilon_l$  =  $l^{\text{th}}$  Greek letter treatment mean effect and

$\varepsilon_{ijkl}$  = Random effects on  $(i, j, k, l)$  plot.

#### 17.3.1. Assumptions of GLSD:

1. The sequence  $(i, j, k, l)$  assumes only  $m^2$  values of the possible  $m^4$  values of the randomly selected Graeco-Latin Square design.

2. In GLSD we assume that there is no interaction effect between different factors under consideration.
3. That is we assume that factors considered are independent of each other.
4. Random errors  $\epsilon_{ijkl}$ 's are independent and identically distributed.
5. Random errors  $\epsilon_{ijkl}$  are distributed Normal with mean zero variance  $\sigma^2\epsilon$ .
6. It is assumed that  $2m$  treatments are arranged in  $m^2$  plots arranged in  $m$  rows and  $m$  columns.

#### 17.4. Statistical Analysis of GLSD

First frame various null hypotheses for GLSD as follows:

A – For Latin letter Treatments:  $H_{01} : \tau_1 = \tau_2 = \dots = \tau_m = 0$  or **All the Latin letter treatments are equally effective.**

B – For Greek letter Treatments:  $H_{02} : \gamma_1 = \gamma_2 = \dots = \gamma_m = 0$  or **All the Greek letter treatments are equally effective.**

C – For Rows:  $H_{03} : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$  or **all the rows are equally effective.**

D – For Columns:  $H_{04} : \beta_1 = \beta_2 = \beta_3 \dots = \beta_m = 0$  or **all the columns are equally effective.**

The short cut step wise procedure adopted in many practical problems in analyzing LSD with one observation per experimental unit is explained as follows:

First construct the table of treatment effects treatment wise both Latin letters and Greek letters and get their totals  $LT_i$ ,  $GT_i$ , Row totals  $R_i$  and Column totals  $C_j$  from the given data.

Step – 1: Calculate the Grand total  $G = \sum_{k=1}^m Tk = \sum_{i=1}^m Ri = \sum_{j=1}^m Cj$ .

Step – 2: Calculate the Correction Factor=  $CF = G^2/N$ , where  $N = m^2$ .

Step – 3: Calculate the Latin letter treatment Sum of squares =  $SSLT = \sum_{k=1}^m LTK^2 / m - CF = A$ .

Step – 4: Calculate the Greek letter treatments sum of Squares =  $SSGT = \sum_{l=1}^m GTk^2 / m - CF = B$ .

Step – 5: Calculate the Rows Sum of Squares =  $RSS = \sum_{i=1}^m Ri^2 / m - CF = C$ .

Step – 6: Calculate the Columns Sum of Squares =  $CSS = \sum_{j=1}^m Cj^2 / m - CF = D$ .

Step – 7: Calculate Total sum of square =  $TSS = \sum_{j=1}^m \sum_{i=1}^m (y_{ij})^2 - CF = E$ .

Step – 9: Construct the following ANOVA table for GLSD with one observation per cell.

**Table (17.4.1): ANOVA table of GLSD.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments(Latin)	$m - 1$	$S_{Lt}^2 = SSLT = A$	$V_1 = S_{Lt}^2 / m - 1$	$F_{LT} = V_1 / V_5$
Treatments(Greek)	$m - 1$	$S_{Gt}^2 = SSGT = B$	$V_2 = S_{Gt}^2 / m - 1$	$F_{GT} = V_2 / V_5$

Rows	m-1	$S_r^2 = \text{RSS} = C$	$V_3 = S_r^2 / m - 1$	$F_R = V_3 / V_5$
Columns	m-1	$S_c^2 = \text{CSS} = D$	$V_4 = S_c^2 / m - 1$	$F_C = V_4 / V_5$
Error	(m-1)(m-3)	$S_E^2 = \text{ESS}$ = E-D- C-B-A = F	$V_5 = S_E^2 / (m-1)(m-3)$	
Total	N - 1	$S_T^2 = \text{TSS} = E$		

Step – 9: Conclusions:

A – For Latin letter treatments:

(1) If the calculated value  $F_{LT}$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for [(m-1),(m-1)(m-3)] dof, we accept the null hypothesis for treatments  $H_{01} : \tau_1 = \tau_2 = \dots, = \tau_m = 0$  and conclude that there is no significant effect of Latin letter treatment means. In this case we need not apply Further Statistical Analysis because all the Latin letter treatments are equally effective.

(2) If the calculated value  $F_{LT}$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for [(m-1),(m-1)(m-3)] dof, we reject the null hypothesis for Latin letter treatments  $H_{01} : \tau_1 = \tau_2 = \dots, = \tau_m = 0$  and conclude that there is significant difference in the effect of Latin letter treatment means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the Latin letter treatment effects or to determine which Latin letter treatment has maximum effect? Or which Latin letter treatment has minimum effect? And so on.

B – For Greek letter treatments:

(1) If the calculated value  $F_{GT}$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for [(m-1),(m-1)(m-3)] dof, we accept the null hypothesis for Greek letter treatments  $H_{02} : \gamma_1 = \gamma_2 = \dots, = \gamma_m = 0$  and conclude that there is no significant effect of Greek letter treatment means. In this case we need not apply Further Statistical Analysis because all the Greek letter treatments are equally effective.

(2) If the calculated value  $F_{GT}$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for [(m-1),(m-1)(m-3)] dof, we reject the null hypothesis for Greek letter treatments  $H_{02} : \gamma_1 = \gamma_2 = \dots, = \gamma_m = 0$  and conclude that there is significant difference in the effect of Greek letter treatment means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the Greek letter treatment effects or to determine which Greek letter treatment has maximum effect? Or which Greek letter treatment has minimum effect? And so on.

C- For Rows:

(1) If the calculated value  $F_R$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for [(m-1),(m-1)(m-2)] dof, we accept the null hypothesis for rows  $H_{03} : \alpha_1 = \alpha_2 = \dots, = \alpha_m = 0$  and conclude that there is no significant effect of row means. In this case we need not apply Further Statistical Analysis because all the rows are uniform with respect to the yield. This means that rows effect is negligible.

(2) If the calculated value  $F_R$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for  $[(m-1), (m-1)(m-2)]$  dof, we reject the null hypothesis for rows  $H_{03} : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$  and conclude that there is significant difference in the effect of row means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the rows or to determine which row has maximum fertility rate? Which row has minimum fertility rate? Which rows have similar fertility rates? And so on.

D. For Columns:

(1) If the calculated value  $F_C$  is less than or equal to  $F_{tab}$  at  $\alpha\%$  los, for  $[(m-1), (m-1)(m-2)]$  dof, we accept the null hypothesis for columns  $H_{04} : \beta_1 = \beta_2 = \beta_3 \dots = \beta_m = 0$  and conclude that there is no significant effect of column means. In this case we need not apply Further Statistical Analysis because all the columns are uniform with respect to the yield. This means that column effect is negligible.

(2) If the calculated value  $F_C$  is greater than  $F_{tab}$  at  $\alpha\%$  los, for  $[(m-1), (m-1)(m-2)]$  dof, we reject the null hypothesis for columns  $H_{04} : \beta_1 = \beta_2 = \beta_3 \dots = \beta_m = 0$  and conclude that there is significant difference in the effect of column means. In this case we required, applying Further Statistical Analysis explained in section (9.4), to sub-group the rows or to determine which column has maximum fertility rate? Which column has minimum fertility rate? Which columns have similar fertility rates? And so on.

Now we proceed to explain the analysis of GLSD with an example.

Example (17.4.1): The following data represents yield of a crop measured in tons from plots of same size arranged in 5 rows and 5 columns. Latin letter treatments A, B, C, D and E represent varieties of seeds and Greek letter treatments  $\alpha, \beta, \gamma, \delta$  and  $\epsilon$  represents five fertilizers used in the experiment. Analyze the data and draw your conclusions.

C Y : 11	D $\epsilon$ : 31	E $\beta$ : 19	A $\delta$ : 20	B $\alpha$ : 14
E $\epsilon$ : 21	A $\beta$ : 23	B $\delta$ : 13	C $\alpha$ : 22	D $\gamma$ : 24
D $\delta$ : 19	E $\alpha$ : 24	A $\gamma$ : 19	B $\epsilon$ : 16	C $\beta$ : 15
A $\alpha$ : 17	B $\gamma$ : 13	C $\epsilon$ : 12	D $\beta$ : 17	E $\delta$ : 17
B $\beta$ : 10	C $\delta$ : 17	D $\alpha$ : 23	E $\gamma$ : 20	A $\epsilon$ : 29

Solution: Step-1: Frame null hypotheses as follows:

A – For Latin letter Treatments:  $H_{01} : \tau_1 = \tau_2 = \dots = \tau_m = 0$  or **All the Latin letter treatments varieties of seeds are equally effective w.r.t mean yields.**

B – For Greek letter Treatments:  $H_{02} : \gamma_1 = \gamma_2 = \dots = \gamma_m = 0$  or **All the Greek letter treatments or Fertilizers are equally effective w.r.t mean yields.**

C – For Rows:  $H_{03} : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$  or **all the rows are equally effective w.r.t mean yields.**

D – For Columns:  $H_{04} : \beta_1 = \beta_2 = \beta_3 \dots = \beta_m = 0$  or **all the columns are equally effective w.r.t mean yields.**

Step – 2: Calculate various totals as follows:

C Y : 11	D ε: 31	E β:19	Aδ: 20	B α:14	R <sub>1</sub> =95
E ε:21	A β:23	Bδ:13	C α:22	D Y:24	R <sub>2</sub> =103
Dδ:19	E α:24	A Y: 19	B ε:16	C β:15	R <sub>3</sub> =93
A α:17	B Y:13	C ε:12	D β:17	Eδ :17	R <sub>4</sub> =76
B β:10	Cδ :17	D α :23	E Y:20	A ε:29	R <sub>5</sub> =99
C <sub>1</sub> =78	C <sub>2</sub> =108	C <sub>3</sub> =86	C <sub>4</sub> =95	C <sub>5</sub> =99	G=466

Latin letter treatments:

$$A:20+23+19+17+29=108.$$

$$B:14+13+16+13+10=66.$$

$$C:11+22+15+12+17=77.$$

$$D:31+24+19+17+23=114.$$

$$E:19+21+24+17+20=101.$$

Greek letter treatments:

$$\alpha:14+22+24+17+23=100.$$

$$\beta:19+23+15+17+10=84.$$

$$Y:11+24+19+13+20=87$$

$$\delta:20+13+19+17+17=86$$

$$\varepsilon:31+21+16+12+29=109$$

Now we proceed to calculate various squares as follows:

C Y : 121	D ε: 961	E β:361	Aδ: 400	B α:196	2039
E ε:441	A β:529	Bδ:169	C α:484	D Y:576	2199
Dδ:361	E α:576	A Y: 361	B ε:256	C β:225	1779
A α:289	B Y:169	C ε:144	D β:289	Eδ :289	1180
B β:100	Cδ :289	D α :539	E Y:400	A ε:841	2169
1312	2524	1574	1829	2127	9366

Step – 3: Now calculate CF and various SS as follows:

$$CF = (466)^2/25 = 8686.24.$$

$$S_{LT}^2 = SSSLT = [(108)^2 + (66)^2 + (77)^2 + (114)^2 + (101)^2]/5 - 8686.24$$

$$= [11664 + 4356 + 5929 + 12996 + 10201]/5 - 8686.24 = 9029.2 - 8686.24 = 342.96 = A$$

$$S_{GT}^2 = SSGT = [(100)^2 + (84)^2 + (87)^2 + (86)^2 + (109)^2]/5 - 8686.24$$

$$= [10000 + 7056 + 7569 + 7396 + 11881]/5 - 8686.24 = 8780.4 - 8686.24 = 94.16 = B.$$

$$S_R^2 = RSS = [(95)^2 + (103)^2 + (93)^2 + (76)^2 + (99)^2]/5 - 8686.24 =$$

$$= [9025 + 10609 + 8649 + 5776 + 9801]/5 - 8686.24 = 8772 - 8686.24 = 85.76 = C.$$

$$S_C^2 = CSS = [(78)^2 + (108)^2 + (86)^2 + (95)^2 + (99)^2]/5 - 8686.24$$

$$=[6084+11664+7396+9025+9801]/5 = 8686.24 = 8794 - 8686.24 = 107.76 = D.$$

$$S_T^2 = TSS = 9366 - 8686.24 = 679.76 = E.$$

Step – 4: Now we proceed to construct the ANOVA table of GLSD as follows:

**Table (17.4.2): ANOVA table of GLSD.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F - cal)
Treatments(Latin)	$m - 1 = 4$	$S_{Lt}^2 = SSLT = A$ =342.96	$V_1 = S_{Lt}^2 / m - 1$ =85.74	$F_{LT} = V_1 / V_5$ =13.96417**
Treatments(Greek)	$m - 1 = 4$	$S_{Gt}^2 = SSLT = B$ =94.16	$V_2 = S_{Gt}^2 / m - 1$ =23.54	$F_{GT} = V_2 / V_5$ =3.83388
Rows	$m - 1 = 4$	$S_r^2 = RSS = C$ =85.76	$V_3 = S_r^2 / m - 1$ =21.44	$F_R = V_3 / V_5$ =3.49186
Columns	$m - 1 = 4$	$S_c^2 = CSS = D$ =107.76	$V_4 = S_c^2 / m - 1$ =26.94	$F_C = V_4 / V_5$ =4.38762*
Error	$(m - 1)(m - 3) = 8$	$S_E^2 = ESS$ =E-D- C-B-A =F =49.12	$V_5 = S_E^2 / (m - 1)(m - 3)$ =6.14	
Total	$N - 1 = 24$	$S_T^2 = TSS = E$ =679.76		

$F_{Tab}$  at 5% los is 3.84 and at 1% los is 7.01 for (4,8) dof.

Step – 5: Conclusions:

A: For Latin letter treatments or seed Varieties:  $F_{LT} > F_{Tab}$  at 1% los, we reject the null hypothesis  $H_{01}$  and conclude that there exists significant variation in Varieties of seeds of the crop. To subgroup these varieties, we have to apply further statistical analysis.

B: For Greek letter treatments or Fertilizers: Since  $F_{GT} < F_{Tab}$  (it is approximately equal to table value) hence we accept  $H_{02}$  at 5% los. Even if the calculated value of F is exactly equal to the table value of F, we have to accept the null hypothesis, for safer side. Hence, we conclude that all the Greek letter treatments or fertilizers are equally effective. Further statistical analysis for fertilizers or Greek letter treatments is not necessary because all the fertilizers are equally effective.

C: For Rows: Since  $F_R < F_{Tab}$ , here also we accept the null hypothesis  $H_{03}$  at 5% los and hence we conclude that units or plots in each row are homogeneous. There is no need for further statistical analysis.

D: For columns: Since  $F_C > F_{Tab}$  at 5% los, we reject the null hypothesis  $H_{04}$  at 5% los and conclude that units or plots in columns are not homogeneous. Still there exists some heterogeneity among columns. Here further statistical analysis is required to further sub-group plots within each column.

### 17.5: Mutually orthogonal Latin Squares

In the last section, we have introduced the concept of orthogonal Latin Square designs and this concept can be extended to n number of Latin Squares. Let  $L_1, L_2, \dots, L_n$  are n Latin squares. This set of Latin Squares are said to be “**Mutually Orthogonal Latin Squares**” if all

pairs of Latin Squares must be orthogonal. That is if we super impose any Latin Square  $L_i$  over  $L_j$  the resultant Latin square must be orthogonal. This property must be true for all values of  $i \neq j, i, j = 1, 2, \dots, n$ . This implies that every pair of treatments must occur once and only once in the resultant Latin square.

Remark: If n is a prime number and power any prime number, then we obtain (n-1) pairs of Mutually Orthogonal Latin Squares. Such type of Latin Squares, have very important to play in Advance designs, like factorial Designs, or Balanced Incomplete Block Designs (BIBD), Partially Balanced Incomplete Block Designs (PBIBD) and so on. Such advanced designs are discussed in the next unit.

### 17.5. Summary

In this lesson we have introduced special type of LSD namely “Orthogonal Latin Square Designs”. Such type of designs are known as Graeco-Latin Square Designs (GLSD), where we can control the Total variation through four factors.

GLSD is the most efficient design when compared to LSD or RBD or CRD, because the Error mean square is smaller in GLSD than other designs. After discussing the Mathematical model, and its assumptions, we have explained the Statistical analysis of the design. The analysis is explained with an example. Finally we have introduced the concept of Mutually orthogonal Latin Squares and discussed the applications of such designs.

### 17.6. Self Assessment Questions.

1. Explain the Special Characters of GLSD.
2. Explain the Mathematical model and assumptions of GLSD.
3. Briefly explain various steps involved in the Statistical analysis of GLSD.
4. Construct the ANOVA table of GLSD of size 5x5.
5. Explain the orthogonal property and discuss its applications.
6. Define mutually orthogonal Latin Squares and their applications.
7. Analyze the following data suitably to test the effect of 5 varieties of seeds represented by Latin letters and five varieties of fertilizers denoted by Greek letters for a crop. The yield of the crop is measured in tons and the experiment is conducted in 25 plots divided as five blocks vertically and five blocks horizontally. Treatments are applied such that each pair of treatments occurs once and only once in the entire experimental area.



$A\alpha : 3.14$	$BY : 3.08$	$C\epsilon : 3.23$	$D\beta : 3.09$	$E\delta : 3.03$
$B\beta : 3.12$	$C\delta : 3.03$	$D\alpha : 3.25$	$EY : 3.07$	$A\epsilon : 3.02$
$CY : 3.16$	$D\epsilon : 3.06$	$E\beta : 3.10$	$A\delta : 3.04$	$B\alpha : 2.94$
$D\delta : 3.10$	$E\alpha : 3.12$	$AY : 3.20$	$B\epsilon : 3.06$	$C\beta : 3.00$
$E\epsilon : 3.09$	$A\beta : 3.10$	$B\delta : 3.24$	$C\alpha : 3.00$	$DY : 3.05$

Analyze the following data using GLSD analysis.

$CY : 14$	$D\epsilon : 20$	$E\beta : 16$	$A\delta : 10$	$B\alpha : 12$
$E\epsilon : 08$	$A\beta : 12$	$B\delta : 15$	$C\alpha : 12$	$DY : 18$
$A\alpha : 09$	$BY : 16$	$C\epsilon : 14$	$D\beta : 24$	$E\delta : 07$
$B\beta : 16$	$C\delta : 11$	$D\alpha : 22$	$EY : 17$	$A\epsilon : 13$
$D\delta : 23$	$E\alpha : 18$	$AY : 09$	$B\epsilon : 22$	$C\beta : 14$

8. Critically compare GLSD with LSD with an example.
9. Explain the layout of GLSD with an example.

### 17.8. Further Readings.

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons, New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 18.****FACTORIAL EXPERIMENTS****18.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the need, concepts and definition of Factorial Experiments.
- Basic concepts of Factorial experiments.
- Main Effects and Interaction effects and their calculations.
- Yates method of computing factorial effects totals.
- Types of Factorial Designs.

**Structure**

- 18.1 Introduction to Factorial Experiments**
- 18.2 Basic Concepts of  $2^n$  Factorial Experiments**
- 18.3 Main effects and Interaction effects**
- 18.4 Yates method of computing Factorial effects totals**
- 18.5 Various types of Factorial Experiments**
- 18.6 Summary**
- 18.7 Self Assessment Questions**
- 18.8. Further Readings**

**18.1: Introduction to Factorial Experiments**

In the previous lesson, we have discussed single factor, double and triple factor design of experiments namely completely randomized Design (CRD), Randomized Block Design and Latin Square designs (LSD) respectively, which are usually known as “**Simple Designs**”, because, treatments considered in these designs are “**single set of treatments**” like, Varieties of crops, or manures or different methods of irrigation or different methods of cultivation and so on.

In this lesson, we can concentrate on advanced designs known as “**Complex Designs**” or “**Complex Experiments**”, where the treatments are combination of different factors at different levels and their effects are studied for its significance. That is the variation in the effect of one factor as a result, to different levels of other factors. For example, consider two fertilizers, namely Potash (K) and Nitrogen (N). Here, treatments are combination of these two factors at different levels. Let p and q are the levels of these factors Potash and Nitrogen respectively. In order to study effectiveness of these factors, we have to conduct two different simple experiments, one for Potash and the other for Nitrogen. Like this, if factors are more, conducting several experiments different for each factor and studying their effects, is very lengthy, time consuming and are costly. Further, through such experiments, we can only study individual effects of factors and is difficult to study **combined effects** or **interaction effects** or **factorial effects** like (PotashxNitrogen) (KxN). The only alternative to study or to investigate the

variations in several factors simultaneously is by conducting above experiment as a  $p \times q$  “**Factorial Experiment**”, where  $p$  and  $q$  are the levels of various factors under consideration.

Factorial Experiments are of two types, namely, (1) Levels of different factors are equal and (2) Levels of different factors are unequal. We confine ourselves to only first type of factors in these lessons. Second type of designs is studied in advanced design of experiments.

Factorial Experiments of the first type are usually denoted by  $s^n$ , where  $s$  represents number of levels and  $n$  represents number of factors. For example:  $2^2$  –Factorial experiments means we have two factors each at two levels. Similarly  $2^n$  – factorial experiments means we have  $n$ -factors each having 2 levels.  $3^2$  – Factorial experiments means we have 2-factors each at 3-levels, and  $3^3$  – factorial experiments means we have 3-factors each at 3-levels, so on. Treatments for the experiment are various combinations of different factors at different levels. Thus in Factorial experiments, we specially consider the treatments as combinations of different factors at different levels. This is the basic difference between the design of experiments discussed in the previous unit and the present unit. This distinction is to understood clearly by the reader.

Now we proceed to explain some basic concepts and definitions of Factorial Experiments in the following section.

## 18.2. Basic Concepts of $2^n$ Factorial Experiment

Let us first consider the simple factorial series of designs namely  $2^n$ –series, where, we consider  $n$ -factors each having 2-levels say 0 and 1. Usually levels of factor means, two quantitative levels or concentrations. For example, 0-level represent 10% concentration and 1-level represents 15% concentration of the manure of the chemical. Sometimes, 0-level represent absence of the factor and 1- represent the presence of the factor. Sometimes, 0-level represent ‘control group’ as 1- level represent the ‘experimental group’. In some experiments, 0-level represent one spices of a plant and 1-level represent another spices of a plant. In animal husbandry, 0-level represent one breed of lactation animal and 1-level another breed of lactation animal and so on.

First we introduce the notations for the simplest design  $2^2$  and latter extend these notations for other designs easily on similar lines.

### 18.2.1: Notations of $2^2$ – factorial design:

Here, we consider 2-factors, say A and B each at two levels say 0 and 1. Let  $a$  and  $b$  represent the level of the factor A and B respectively. Then we have  $2^2 = 4$  treatment combinations as follows:

Treatment – 1:  $a_0b_0$  or ‘1’ : Represents both factors A and B at first level that is 0<sup>th</sup> level.

Treatment – 2:  $a_1b_0$  or ‘a’ : Represents Factor A at second level and Factor B at first level.

Treatment – 3:  $a_0b_1$  or ‘b’ : Represents Factor A at first level and factor B at second level.

Treatment – 4:  $a_1b_1$  or ‘ab’ : Represents Both the factors at second level that is at 1<sup>st</sup> level.

Thus we have 4 treatment combinations and these treatment combinations can be tested using CRD with  $r_i$ -replicates under  $i^{\text{th}}$  treatment or RBD-with b-blocks or LSD of size  $4 \times 4$ .

### 18.2.2: Notations for $2^n$ – factorial experiments:

On similar lines, we can extend treatment combinations for n factors A,B,C,...,N each with two levels say 0 and 1. Then we have  $2^n$  factorial effects, represented as follows:

1,a,b,ab,c,ac,bc,abc,d,ad,bd,abd,cd,acd,bcd,abcd,e,ae,be,abe,ce,ace,bce,abce,de,ade,bde,abde,cde,acde,abcd,...,.,(2<sup>n</sup>)- terms. Thus we have 2<sup>n</sup>- treatments.

Important note: while writing various treatment combinations of factors, we have followed one rule. That “every factor is to be introduced one after the other and after introducing, we have to completely write all previously written combination in the same order. After completing all previously introduced combinations, we have to introduce another factor”. This order of writing all treatment combinations are called “**Standard symmetric Order**” and in various combinations, absence of any letter represent the corresponding factor at 1<sup>st</sup> level that is level ‘0’ and presence of any letter represent the corresponding factor at 2<sup>nd</sup> level that is level ‘1’. After determining various treatments, we can conduct the experiment in CRD with  $r_i$  replicates or RBD with r blocks or LSD of size  $n \times n$ .

Now we proceed to explain main effects and interaction effects in the following section.

### 18.3. Main effects and Interaction effects

Suppose that  $2^2 = 4$  treatments are conducted in CRD with r-replicates. Let [1],[a],[b] and [ab] denote the total yields of r- replicates receiving the treatments 1,a,b,ab respectively.

Let (1),(a),(b) and (ab) are the respective mean yields of the treatments 1,a,b,ab respectively.

Thus we have:  $(1) = [1]/r; (a) = [a]/r; (b) = [b]/r$  and  $(ab) = [ab]/r$ . Let **A** represent the Main Effect of the Factor-A, **B** represent the Main Effect of the Factor- B and **AB** represent the interaction effect of Factor- A and Factor-B and these effects are defined as follows:

#### 18.3.1: Main effects and Interaction Effects of different factors:

Definition: **Main effect of Factor – A**: Consider the  $2^2$  factorial Experiment with two factors each with two levels. The main effect of Factor – A can be represented by the difference between mean yields obtained at each level at each level of other factor –B. We have to consider two effects, namely:

Effect of Factor – A at the first level  $b_0$  of Factor – B =  $(a_1 b_0) - (a_0 b_0)$  or  $(a) - (1)$ . (18.3.1)

Similarly, Effect of Factor – A at the second level  $b_1$  of Factor – B

=  $(a_1 b_1) - (a_0 b_1)$  or  $(ab) - (b)$  (18.3.2)

Now main effect of Factor is defined as the mean effect of above two effects. Namely;

**Main Effect of Factor – A = A** =  $(1/2)[\{(a_1 b_0) - (a_0 b_0)\} + \{(a_1 b_1) - (a_0 b_1)\}]$   
 =  $(1/2)[(ab) - (b) + (a) - (1)] = (1/2)[(a-1)(b+1)]$  (18.3.3)

Similarly, **Main effect of Factor – B = B** =  $(1/2)[\{(a_0 b_1) - (a_0 b_0)\} + \{(a_1 b_1) - (a_1 b_0)\}]$

$$= (1/2)[(ab) - (a) + (b) - (1)] = (1/2)[(a+1)(b-1)] \quad (18.3.4.)$$

**Interaction Effects Factors A and B:** The interaction effect of two factors is defined as the failure of the levels of a factor, say Factor-A to remain the same order and magnitude of performance though out all levels of the second factor say, Factor – B. If the two factors are independent of each other, we should expect the true effect of one factor must be same at either levels of other factor. In other words, the effects defined in (18.3.1) and (18.3.2) were really the estimates of the same thing and hence will be equal. If the two expressions are not

equal, implies that two factors are not independent and hence, the difference between two terms is called interaction effect of two Factors A and B. Thus the difference between (18.3.1) and (18.3.2) measures the interaction effect of two factors A and B. Thus, we define the interaction of effect of Two Factors A and B denoted by **AB** is defined as follows:

$$\begin{aligned} \text{Interaction of Factor – A and Factor – B} = \mathbf{AB} &= (1/2)[\{(a_1b_1) - (a_0 b_1)\} - \{(a_1b_0) - (a_0 b_0)\}] \\ &= (1/2)[(ab) - (a) - (b) + (1)] = (1/2)[(a-1)(b-1)] \end{aligned} \quad (18.3.5)$$

It is important to note that the interaction effect of BA = (1/2)[\{(a\_1b\_1) - (a\_0 b\_1)\} - \{(a\_1b\_0) - (a\_0 b\_0)\}] = (1/2)[(ab) - (b) - (a) + (1)] = (1/2)[(b-1)(a-1)] which is equal to the interaction effect of the factors A and B. That is **AB = BA**.

**Remark:** We can club all the above Main Effects and interaction effects of two factors in a single algebraic formula as: [(a±1)(b±1)] and we have to consider – sign if the factor is present in the effect and + sign if the factor is absent in the effect. That is for main effect of factor-A the algebraic expression is [(a-1)(b+1)] , for the main effect of Factor-B the expression is [(a+1)(b-1)] and for interaction effect of factors A and B the algebraic expression if [(a-1)(b-1)].

On similar lines, we can extend expressions for three factors or n-factors each at two levels 0 and 1 are as follows:

For Three factors A,B and C each with two levels various main effects and interaction effects are defined as:  $(1/2^2)[(a±1)(b±1)(c±1)]$  . (18.3.6)

That is:

$$\text{Main Effect of A} = \mathbf{A} = (1/2^2) [(a-1)(b+1)(c+1)] \quad (18.3.7)$$

$$\text{Main Effect of B} = \mathbf{B} = (1/2^2) [(a+1)(b-1)(c+1)] \quad (18.3.8)$$

$$\text{Main Effect of C} = \mathbf{C} = (1/2^2) [(a+1)(b+1)(c-1)] \quad (18.3.9)$$

$$\text{Interaction Effect of AB} = \mathbf{AB} = (1/2^2) [(a-1)(b-1)(c+1)] \quad (18.3.10)$$

$$\text{Interaction Effect of AC} = \mathbf{AC} = (1/2^2) [(a-1)(b+1)(c-1)] \quad (18.3.11)$$

$$\text{Interaction Effect of BC} = \mathbf{BC} = (1/2^2) [(a+1)(b-1)(c-1)] \quad (18.3.12)$$

$$\text{Interaction Effect of ABC} = \mathbf{ABC} = (1/2^2) [(a-1)(b-1)(c-1)] \quad (18.3.13)$$

Interaction effects **AB=BA; AC=CA;BC=CB**. Similarly interaction effects **ABC=ACB=BAC =BCA =CAB=CBA**. Further, above interaction effects **AB,AC** and **BC** are called “**first order interactions**” and the interaction effect **ABC=ACB=BAC =BCA =CAB=CBA** are called “**Second order interactions**”.

On similar lines we can extend various main effects and Interaction effects of  $n$  factors,  $A, B, \dots, N$  ( $n$ -factors) are clubbed in a single formula as :

$$(1 / 2^{n-1})[(a \pm 1)(b \pm 1)(c \pm 1) \dots (N \pm 1) \text{ n terms}] \quad (18.3.14)$$

On similar lines, we can extend these definitions, for  $3^2, 3^3$  and so on. These effects will be discussed in lesson-20, where we discuss  $3^n$  factorial experiments. With the above defined factorial effects we can discuss  $2^n$  – factorial experiments.

Now we give the definitions of “Contrast” and “orthogonal contrasts” as follows:

**Definition: Contrast:** A linear combination  $\sum_{i=1}^K c_i t_i$  for  $K$  treatment means  $t_i, i=1, 2, \dots, K$ , is called a “Contrast” or a “Comparison” of treatment means if  $\sum_{i=1}^K c_i = 0$ . In other words, Contrast is a linear combination of treatment means such that sum of coefficients is zero.

It is interesting to note that all the main effects and interaction effects are **contrasts**.

**Definition: Orthogonal Contrasts:** Two contrasts of  $K$ -treatment means  $t_i, i=1, 2, \dots, K$  say;

$\sum_{i=1}^K c_i t_i$ , such that  $\sum_{i=1}^K c_i = 0$  and  $\sum_{i=1}^K d_i t_i$  such that  $\sum_{i=1}^K d_i = 0$ ; are said to be orthogonal to each other if  $\sum_{i=1}^K c_i d_i = 0$ . In other words, two contrasts are orthogonal to each other if the sum of the product of the coefficients of the corresponding treatments is **zero**.

Further it can be observed that main effect of factor  $A$  given in (18.3.3) and Main effect of factor  $B$  given in (18.3.4) are orthogonal contrasts because sum of the products of the coefficients in each effect is:  $(1/2)(1/2) + (-1/2)(1/2) + (1/2)(-1/2) + (-1/2)(-1/2) = 0$ . On similar lines the main effects **A, B**, and interaction effects **AB** are orthogonal to each other. Hence, all the tree factorials **A, B** and **AB** are “**mutually orthogonal**”.

Now we proceed to discuss an important procedure useful for calculating various factorial effects totals from the given experimental data in the following section.

#### 18.4. Yates method or algorithm of computing Factorial effects totals

Now we proceed to explain a useful method of calculating various factorial effects

totals, which simplifies our calculations in analyzing  $2^n$  – Factorial Experiments. Prof. F. Yates developed a special computational rule which enables us to calculate various factorial effects without using the algebraic formula method discussed in the last section. This method is popularly known as “**Yates method of computing various factorial effects**” or simply “**Yates algorithm**” which consists of the following steps:

**Step – 1:** In the first column, write all the treatment combinations in a standard Symmetric Order (as explained in section(18.2.2)). It is an essential part of the procedure that the treatment combinations must be written in the ‘Standard Symmetric Order’ in the first column.

For example the Standard Symmetric order for three factors is: **1,a,b,ab,c,ac,bc,abc**.

**Step – 2:** In the second column, write down the corresponding total yields of treatments from all the replicates in the same order as written in the first column.

**Step – 3:** The entries in the third column can be split into two halves. The first half is obtained by writing down in order, the pair wise sums of the values in column – 2 and the second half is

obtained by writing in the same order the pair wise differences of the values in column – 2. It is to be remembered that while calculating the differences, the first member in the pair is to be subtracted from the second member of that pair.

**Step – 4:** Complete the 4<sup>th</sup> column by applying the procedure explained in Step – 3. Similarly obtain the 5<sup>th</sup> column from 4<sup>th</sup> column by applying the procedure explained in Step – 3.

This procedure is explained at appropriate places with an example in the coming lessons.

### **18.5. Various types of Factorial Experiments.**

**Factorial Experiments with equal number of level under each treatment are classified in to the following categories. Namely:**

- (1) Factorial Experiments of  $2^n$  – Series.**
- (2) Factorial Experiments of  $3^n$  – Series.**
- (3) Balanced Incomplete Designs.**
- (4) Partially Balanced Incomplete Block Designs.**
- (5) Orthogonal Designs.**
- (6) Robust designs.**
- (7) Response Surface designs.**
- (8) Split-Plot designs.**
- (9) Optimal Designs.**
- (10) V-Optimal designs and so on.**

Among the above designs, we have:  $2^2, 2^3, 3^2, 3^3$  Factorial Designs and Balanced Incomplete Block design only in the present syllabus. Hence, these designs along with their analysis are discussed in the remaining three lessons.

There are other factorial designs, where we consider treatments with un-equal number of levels. These designs are out of the scope of this book.

### **18.6 Summary**

In this lesson we have introduced the concept of Factorial Experiments, its applications and different types of factorial Experiments. After defining some fundamental concepts like Factorial Effects, Main Effects, Interaction effects, we have defined contrast and orthogonal contrasts and mutually orthogonal contrasts. Then we have explained the Yates method of algorithm useful to calculate various treatment totals. Finally we have discussed various Factorial experiments of n-factors with equal number levels.

### **18.7 Self Assessment Questions.**

1. Explain factorial experiments and their applications.
2. Define main Effects and Interaction effects with suitable examples.
3. Explain the application of algebraic expression useful in calculating various Factorial Effects.

4. Explain various steps involved in the Yates method of algorithm and discuss its uses.
5. Define (1) contrast (2) orthogonal Contrasts and (3) Mutually orthogonal contrasts.

### **18.8. Further Readings**

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons, New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.



**Lesson – 19.** **$2^n$  -FACTORIAL EXPERIMENTS****19.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the need, concepts and definition of  $2^n$  - Factorial Experiments.
- Basic concepts of  $2^2$  - Factorial experiments.
- Statistical analysis of  $2^2$  – Factorial experiments.
- $2^3$  Factorial Experiments.
- Statistical analysis of  $2^3$  - Factorial Experiments.

**Structure****19.1 Introduction to  $2^n$  Factorial Experiments****19.2 The  $2^2$  Factorial Experiment****19.3 Statistical analysis of  $2^2$  – Factorial Experiments****19.4 The  $2^3$  – Factorial Experiment****19.5 Statistical analysis of  $2^3$  – Factorial Experiments****19.6 Summary****19.7 Self Assessment Questions****19.8 Further Readings****19.1: Introduction to  $2^n$  - Factorial Experiments**

In the previous lesson, we have discussed some fundamental concepts of Factorial Designs. In this lesson, we consider  $2^n$ -series of Factorial experiments.  $2^n$ -factorial designs are widely used in experiments involving n-factors each with 2-levels. These experiments are useful to study the joint effect of n-factors each at two levels on a response like yield of a crop or quality of a product or breaking strength of a brick and so on. The two levels in these experiments may be considered as produced from two machines, or two operators or 'high' and 'low' levels of factors or 'absence' and 'presence' of a factor and so on. Each complete replicate of such design requires  $2 \times 2 \times 2 \times \dots \times 2$ , n-times =  $2^n$ - experimental units. If we replicate these  $2^n$ -treatment combinations in r-replicates, we must have  $r \times 2^n$  experimental units. In these design, we have  $\binom{n}{1}$ -main effects,  $\binom{n}{2}$ - two factor interactions,  $\dots, \binom{n}{n}$  - nfactor

interactions. Thus in total we have  $\binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n} = 2^n - 1$  factorial effects in  $2^n$ -Factorial Experiments. Such a design is called  $2^n$  – Factorial Design.  $2^n$ -factorial designs are particularly useful in the early stages of experimental work, when there are likely to be many factors to be investigated.

This design provides the smallest number of experimental units with n-factors because each factor is only at two levels. Thus these designs are widely used in 'Factor Screening Experiments'. In this lesson we consider and study the details of such design with n = 2 and 3.

Other designs are out of the scope of the syllabus.

## 19.2. The 2<sup>2</sup>- Factorial Experiment

In this design we consider two factors A and B each with two levels. Thus we have 2<sup>2</sup>=4 factorial combinations, namely:

(a<sub>0</sub>,b<sub>0</sub>) or 1; (a<sub>1</sub>,b<sub>0</sub>) or a; (a<sub>0</sub>,b<sub>1</sub>) or b and (a<sub>1</sub>,b<sub>1</sub>) or ab. After considering these factorial combinations as four treatments, we can conduct the experiment in CRD with r-replicates or RBD with b-blocks or LSD of size 4x4 layouts. Thus we have rx2<sup>2</sup>= 4r observations and 2<sup>2</sup> – 1 = 3 factorial effects namely Main Effect of Factor – A denoted by **A**; Main effect of Factor – B denoted by **B**; and two factor Interaction effect or first order interaction effect between factors – A and B denoted by **AB**. It is important to note that interaction effect **AB** = interaction effect **BA**.

### 19.2.1: Yates method of Algorithm for a 2<sup>2</sup> – factorial experiment.

Here we have four treatment combinations written the Standard symmetric order is 1,a,b and ab. Let the total yield from all the replicates in the selected design is denoted by [1],[a],[b] and [ab] respectively. The Yate's method of algorithm explained in the last lesson is applied here for 2<sup>2</sup>-factorial experiment as follows:

Table (19.2.1) Yates method of algorithm for 2<sup>2</sup>-factorial experiment.

Treatment Combinations	Total yield from all replicates	Column -3	Column – 4	Treatment totals
1	[1]	[a] + [1]	[1]+[a]+[b]+[ab]	Grand total= <b>G</b>
a	[a]	[ab] + [b]	[ab] –[b] +[a]-[1]	<b>[A]</b>
b	[b]	[a] – [1]	[ab]+[b]– [a] –[1]	<b>[B]</b>
ab or ba	[ab]	[ab] – [b]	[ab]-[b]-[a]-[1]	<b>[AB] or[BA]</b>

Now we proceed to explain the Statistical analysis if 2<sup>2</sup>-factorial experiments in the following section.

## 19.3. Statistical analysis of 2<sup>2</sup>-factorial experiments

It is already mentioned that factorial experiments are conducted either in CRD or RBD or LSD and hence, the statistical analysis discussed in the previous unit can be applied for the corresponding design applied, except the difference that Treatment SS is further splits in to three orthogonal contrasts, namely, Main effect – A, Main effect – B and Interaction effect - AB each with one degree of freedom. Usually these effects are computed in practice by using the entries of Yates method of algorithm in the last column, namely, **[A]**, **[B]** and **[AB]**. Various SS of these factorial effects are calculated as follows:

$$\text{SS due to main effect – A} = \frac{[A]^2}{4r}, \quad (19.3.1)$$

$$\text{SS due to main effect – B} = \frac{[B]^2}{4r} \text{ and} \quad (19.3.2)$$

$$\text{SS due to main effect – AB} = \frac{[AB]^2}{4r}, \quad (19.3.3)$$

each with 1 dof. The null hypotheses to be tested  $2^2$  Factorial experiment conducted in RBD are:

$H_{0T}$ : there is no significant variation between various treatments.

$H_{0A}$ : There is no Significant Main effect of Factor – A.

$H_{0B}$ : There is no significant main effect of the factor – B.

$H_{0AB}$ : There is no significant interaction effect between the factors A and B.

$H_{0\text{Blocks}}$ : There is no significant variation between block means.

Thus the ANOVA table of  $2^2$ -factorial experiment conducted in RBD with b- blocks or replicates is given by:

**Table (19.3.1): ANOVA table of  $2^2$  Factorial Experiment conducted in RBD.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (rows):	$2^2-1$	$S_T^2=A+B+AB$	$V_1=S_T^2/2^2-1$	$F_T=V_1/V_6$
Main effect-A	1	$S_A^2 = [A]^2/4b=A$	$V_2 = S_A^2$	$F_A = V_2 / V_6$
Main effect-B	1	$S_B^2 = [B]^2/4b=B$	$V_3 = S_B^2$	$F_B = V_3/V_6$
Interaction effect AxB	1	$S_{AB}^2 = [AB]^2/4b=AB$	$V_4 = S_{AB}^2$	$F_{AB}= V_4/V_6$
Blocks (columns)	$b - 1$	$S_b^2=BSS=C$	$V_5 = S_b^2/b-1$	$F_{\text{Blocks}} = V_5 / V_6$
Error	$(2^2-1)(b-1)$	$S_E^2 = ESS=D-A-B-AB$	$V_6 = S_E^2 / (2^2-1)(b-1)$	
Total	$2^2b - 1$	$S_T^2 = TSS=D$		

Conclusions: (1): **For treatments:**

If  $F_A < F_{\text{tab}}$  at  $\alpha\%$  los for  $[1,3(b-1)]$  dof, we accept the null hypothesis  $H_{0A}$ .

Otherwise, we reject the null hypothesis  $H_{0A}$ . if necessary, we can apply further statistical analysis. Similarly, we can accept or reject the null hypotheses for Main effect – B and Interaction Effect AB.

(2) **For blocks:**

If  $F_{\text{Blocks}} < F_{\text{tab}}$  at  $\alpha\%$  los for  $[(b-1),3((b-1))]$  dof, we accept the null hypothesis  $H_{0\text{Blocks}}$ . Otherwise we reject the null hypothesis  $H_{0\text{Blocks}}$ . If necessary we can apply further statistical analysis for sub-grouping blocks.

Example (19.3.1): An experiment was planned to study the effect of sulphate of potash and super phosphate on the yield of potatoes. All the combinations of two levels of super phosphate [ 0 cent ( $p_0$ ) and 5 cent ( $p_1$ ) /acre ] and two levels of sulphate of potash [ 0 cent ( $k_0$ ) and 5 cent ( $k_1$ )/acre] were studied in an RBD with four blocks. The data is measured in lbs representing the yield of potatoes per plot and is given as follows:

Blocks:

A:	(1): 23	k:25	p:22	kp:38
B:	p:40	(1):26	k:36	kp:38
C:	(1):29	k:20	kp:30	p:20
D:	kp:34	k:31	p:24	(1):28

Analyze the data suitably and draw your conclusions.

Solution: Frame the null hypotheses as follows:

$H_{01}$ : there is no significant variation between various treatments.

$H_{02}$ : There is no main effect of sulphate of potash (k) on the yield of potatoes.

$H_{03}$ : There is no main effect of super phosphate (p) on the yield of potatoes.

$H_{04}$ : There is no interaction effect of the two factors kp on the yield of potatoes.

$H_{05}$ : There is no significant variation in block mean yield of potatoes.

Now calculate various totals as follows:

Blocks.	(1)	k	p	kp	Totals	Yield Squares				Totals
A:	23	25	22	38	108	529	625	484	1444	3082
B:	26	36	40	38	140	676	1296	1600	1444	5016
C:	29	20	20	30	99	841	400	400	900	2541
D:	28	31	24	34	117	784	961	576	1156	3477
Treatment totals:	106	112	106	140	464 = G					14116

$$CF: G^2/16 = (464)^2/16 = 13456.$$

$$BSS = (1/4)[(108)^2 + (140)^2 + (99)^2 + (117)^2] - CF = (1/4)[11664 + 19600 + 9801 + 13689] - 13456$$

$$= 13688.5 - 13456 = 232.5 = C.$$

$$TSS = 14116 - 13456 = 660 = D.$$

Now we proceed to calculate various factorial effects totals by using Yates method as follows:

Table (19.3.2) Yates method of algorithm for 2<sup>2</sup>-factorial experiment.

Treatment Combinations	Total yield from all replicates	Column -3	Column - 4	Treatment totals
------------------------	---------------------------------	-----------	------------	------------------

1	[1]=106	[a] + [1]=218	[1]+[a]+[b]+[ab]=464	Grand total= <b>G</b>
a	[a]=112	[ab] + [b]=246	[ab] –[b] +[a]-[1]=40	<b>[A]</b>
b	[b]=106	[a] – [1]=6	[ab]+[b]– [a] –[1]=28	<b>[B]</b>
ab or ba	[ab]=140	[ab] – [b]=34	[ab]-[b]-[a]-[1]=28	<b>[AB] or[BA]</b>

Main effect of factor k =  $(40)^2/16 = 100 = A$

Main effect of Factor p =  $(28)^2/16 = 49 = B$

Interaction effect kp =  $(28)^2/16 = 49 = AB$

Now we proceed to construct the ANOVA table for  $2^2$  factorial experiment conducted in RBD as follows:**Table (19.3.3): ANOVA table of  $2^2$  Factorial Experiment conducted in RBD.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (rows):	3	$S_A^2 + S_B^2 + S_{AB}^2 = S_T^2 = 198$	$V_1 = S_T^2/3 = 198/3 = 66$	$F_T = 2.59$
Main effect-A	1	$S_A^2 = 100 = A$	$V_2 = S_A^2 = 100$	$F_A = 3.9215$
Main effect-B	1	$S_B^2 = 49 = B$	$V_3 = S_B^2 = 49$	$F_B = 1.9215$
Interaction effect AxB	1	$S_{AB}^2 = 49 = AB$	$V_4 = S_{AB}^2 = 49$	$F_{AB} = 1.9215$
Blocks (columns)	$b - 1 = 3$	$S_b^2 = BSS = C = 232.5$	$V_5 = S_b^2/b-1 = 77.5$	$F_{Blocks} = 3.039$
Error	$(2^2-1)(b-1) = 9$	$S_E^2 = ESS = D - A - B - AB = 229.5$	$V_6 = S_E^2 / (2^2-1)(b-1) = 25.5$	
Total	$2^2b - 1 = 15$	$S_T^2 = TSS = D = 660$		

Conclusions:  $F_{tab}$  at 5% los for (1,9) dof is 5.12 and for (3,9) dof is 3.86. Since all the calculated values are less than the table values we accept all the null hypotheses framed and conclude that there is no significant treatment effects, no significant main effects of factors k,p and no interaction effects kp and there are no block effects. There is no need for further statistical analysis either for treatments or blocks.

Note: It is interesting to note that  $S_A^2 + S_B^2 + S_{AB}^2 = S_T^2$  in the above ANOVA table. This implies that Sum of Squares of Main effect of Factor- A; Main effect of Factor – B and Interaction effect AB will add up to Treatment sum of squares. That is  $100+49+49 = 198$ .

#### 19.4. The $2^3$ – Factorial Experiment.

In this design, we have three factors say A,B and C each with two levels. Here we have  $2^3 = 8$  factorial effects, namely (1) a,b,ab,c,ac,bc and abc. Main effects and Interaction effects in  $2^3$ - factorial experiments are defined as follows:

**19.4.1: The Main effect of the factor – A** in 2<sup>3</sup>-factorial experiment is defined as the average of the following four simple effects, namely:

Simple effect of A is defined as the result of increasing the factor – A from the levels of a<sub>0</sub> to a<sub>1</sub> at other levels of the factors B and C. This is explained as follows:

Level of Factor B	Level of Factor C	Simple effects of Factor A.
b <sub>0</sub>	c <sub>0</sub>	(a <sub>1</sub> b <sub>0</sub> c <sub>0</sub> ) - (a <sub>0</sub> b <sub>0</sub> c <sub>0</sub> ) = (a) - (1)
b <sub>1</sub>	c <sub>0</sub>	(a <sub>1</sub> b <sub>1</sub> c <sub>0</sub> ) - (a <sub>0</sub> b <sub>1</sub> c <sub>0</sub> ) = (ab) - (b)
b <sub>0</sub>	c <sub>1</sub>	(a <sub>1</sub> b <sub>0</sub> c <sub>1</sub> ) - (a <sub>0</sub> b <sub>0</sub> c <sub>1</sub> ) = (ac) - (c)
b <sub>1</sub>	c <sub>1</sub>	(a <sub>1</sub> b <sub>1</sub> c <sub>1</sub> ) - (a <sub>0</sub> b <sub>1</sub> c <sub>1</sub> ) = (abc) - (bc)

The main effect of factor – A is denoted by **A** and is defined the average of above simple effects of Factor – A. That is:

$$\text{Main Effect of factor A} = \mathbf{A} = (1/2^{3-1}) [(abc) - (bc) + (ac) - (c) + (ab) - (b) + (a) - (1)]$$

$$= (1/4)[\{(abc) + (ac) + (ab) + (a)\} - \{(bc) + (c) + (b) + (1)\}] = (1/4)[(a-1)(b+1)(c+1)]$$

This is given in the algebraic expression (18.3.7). On similar lines, we can define other factorial effects like Main effect of factor- B, Main Effect of factor – C first order interaction effects and second order interaction effects given from equations (18.3.8) to (18.3.13) in the lesson - 18.

Now we proceed to explain the Yates method of algorithm for 2<sup>3</sup>-factorial Experiment in the following section.

**19.4.2: Yates method of algorithm for 2<sup>3</sup>-factorial experiment:**

The Yates method of algorithm for 2<sup>3</sup>-factorial experiment is similar to 2<sup>2</sup> – factorial experiment, which is as follows:

Table (19.4.1) Yates method of algorithm for 2<sup>3</sup>-factorial experiment.

Treatment Combinations	Total yield from all replicates	Column -3	Column - 4	Treatment totals
1	[1]	[a] + [1]=U1	U1+U2	Grand total= <b>G</b>
A	[a]	[ab] + [b]=U2	U3+U4	<b>[A]</b>
B	[b]	[ac]+[c]=U3	U5+U6	<b>[B]</b>
ab or ba	[ab]	[abc]+[bc]=U4	U7+U8	<b>[AB] or [BA]</b>
C	[c]	[a] - [1]=U5	U2 – U1	<b>[C]</b>
Ac	[ac]	[ab] – [b]=U6	U4 – U3	<b>[AC]</b>
Bc	[bc]	[ac] – [c]=U7	U6 – U5	<b>[BC]</b>
Abc	[abc]	[abc] – [bc]=U8	U8 – U7	<b>[ABC]</b>

Various interaction effects SS are calculated as follows:

$$\text{SS of main effect of factor A} = [\mathbf{A}]^2/2^3r \quad (19.4.1)$$

$$\text{SS of main effect of factor B} = [\mathbf{B}]^2/2^3r \quad (19.4.2)$$

$$\text{SS of main effect of factor C} = [\mathbf{C}]^2/2^3r \quad (19.4.3)$$

$$\text{SS of first order interaction effect of factor AB} = [\mathbf{AB}]^2/2^3r \quad (19.4.4)$$

$$\text{SS of first order interaction effect of factor AC} = [\mathbf{AC}]^2/2^3r \quad (19.4.5)$$

$$\text{SS of first order interaction effect of factor BC} = [\mathbf{BC}]^2/2^3r \quad (19.4.6)$$

$$\text{SS of second order interaction effect of factor ABC} = [\mathbf{ABC}]^2/2^3r \quad (19.4.7)$$

Now we proceed to explain the Statistical analysis of  $2^3$  – factorial experiment in the following section.

### 19.5. Statistical analysis of $2^3$ – factorial experiment

Here also we can conduct the  $2^3$ -factorial experiment with the above 8 treatments in

CRD or RBD or LSD of size 8x8. The statistical analysis is as usual for the corresponding design except the difference that the treatment SS is split into seven orthogonal components each with 1 dof. Let us assume that the 8 treatment combinations are applied in RBD with b-blocks.

Then the null hypotheses to be framed are:

$H_{0T}$ : There is no significant variation between the treatments.

$H_{0A}$ : There is no main effect of the Factor A.

$H_{0B}$ : There is no main effect of the Factor B.

$H_{0C}$ : There is no main effect of the Factor C.

$H_{0AB}$ : There is no interaction effect between the Factor A and B.

$H_{0AC}$ : There is no interaction effect between the Factor A and C.

$H_{0BC}$ : There is no Interaction effect between the Factor B and C.

$H_{0ABC}$ : There is no Interaction effect between the Factor A,B and C.

$H_{0Blocks}$ : There is no variation between blocks.

Then the ANOVA table of  $2^3$ - factorial experiment is given as follows:

**Table (19.5.1): ANOVA table of  $2^3$  Factorial Experiment conducted in RBD.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	Variance Ratio F (F – cal)
Treatments (rows):	$2^3-1$	$S_t^2=*$	$V_1=S_t^2/2^2-1$	$F_T=V_1/V_{10}$
Main effect-A	1	$S_A^2 = [A]^2/8b=A$	$V_2 = S_A^2$	$F_A = V_2 / V_{10}$
Main effect-B	1	$S_B^2 = [B]^2/8b=B$	$V_3 = S_B^2$	$F_B = V_3/V_{10}$

Main effect-C	1	$S_C^2 = [C]^2 / 8b = C$	$V_4 = S_C^2$	$F_C = V_4 / V_{10}$
Interaction effect AxB	1	$S_{AB}^2 = [AB]^2 / 8b = AB$	$V_5 = S_{AB}^2$	$F_{AB} = V_5 / V_{10}$
Interaction effect AxC	1	$S_{AC}^2 = [AC]^2 / 8b = AC$	$V_6 = S_{AC}^2$	$F_{AC} = V_6 / V_{10}$
Interaction effect BxC	1	$S_{BC}^2 = [BC]^2 / 8b = BC$	$V_7 = S_{BC}^2$	$F_{BC} = V_7 / V_{10}$
Interaction effect AxBxC	1	$S_{ABC}^2 = [ABC]^2 / 8b = ABC$	$V_8 = S_{ABC}^2$	$F_{ABC} = V_8 / V_{10}$
Blocks (columns)	$b - 1$	$S_b^2 = BSS$	$V_9 = S_b^2 / b - 1$	$F_{Blocks} = V_9 / V_{10}$
Error	$(2^3 - 1)(b - 1)$	$S_E^2 = ESS = **$	$V_{10} = S_E^2 / (2^2 - 1)(b - 1)$	
Total	$2^3 b - 1$	$S_T^2 = TSS$		

In the above table, \* = A+B+C+AB+AC+BC+ABC, and

\*\* = By subtraction = TSS – BSS – SST =  $S_T^2 - S_b^2 - S_T^2$ .

**Conclusions:** Compare various F-calculated values for various treatment effects at  $\alpha\%$  los for  $[(2^3 - 1)(2^3 - 1)(b - 1)]$  dof. Main effects are to be compared with table value of F for  $[1, (2^3 - 1)]$  dof and blocks calculated value of F is to be compared for  $[(b - 1)(2^3 - 1)(b - 1)]$  dof at  $\alpha\%$  los and conclusions can be drawn accordingly depending on the calculated and table values of F. If necessary, we can apply further Statistical analysis where ever it is required.

**Example (19.5.1):** The following table gives the layout and the results of a 2<sup>3</sup>-factorial design laid out in four replicates. The objective of the experiment is to determine the effect of different kinds of fertilizers Nitrogen N, Potash K and phosphate P on tomato crop yield.

Block - 1	p:323	1:87	kp:423	np:324	k:279	npk:471	n:128	nk:334
Block - 2	nk:306	np:324	kp:407	n:89	npk:449	p:324	1:106	k:272
Block - 3	1:101	k:265	n:106	np:373	p:312	kp:391	nk:291	npk:450
Block - 4	kp:435	npk:437	1:131	k:302	np:361	nk:272	p:324	n:103

Analyze the data and draw your conclusions.

Solution: Frame the null hypotheses as follows:

$H_{0t}$ : There is no significant variation between the treatments N, P and K.

$H_{0N}$ : There is no main effect of the Factor N.

$H_{0P}$ : There is no main effect of the Factor P.

$H_{0K}$ : There is no main effect of the Factor K.

$H_{0NP}$ : There is no interaction effect between the Factor N and P.

$H_{0NK}$ : There is no interaction effect between the Factor N and K.

$H_{0PK}$ : There is no Interaction effect between the Factor P and K.

$H_{0NPK}$ : There is no Interaction effect between the Factor N, P and K.

$H_{0Blocks}$ : There is no variation between blocks.



Now calculate various totals as follows:

Totals

Block - 1	p:323	1:87	kp:423	np:324	k:279	npk:471	n:128	nk:334	2369
Block - 2	nk:306	np:338	kp:407	n:89	npk:449	p:324	1:106	k:272	2291
Block - 3	1:101	k:265	n:106	np:373	p:312	kp:391	nk:291	npk:450	2289
Block - 4	kp:435	npk:447	1:131	k:302	np:361	nk:272	p:324	n:103	2375

Treatment totals:

$$[1]=87+106+101+131=425; [n]=128+89+106+103=426; [p]=323+324+312+324=1283;$$

$$[k]=279+272+265+302=1118; [np]=324+338+373+361=1396; [nk]=334+306+291+272=1203;$$

$$[kp]=423+407+391+435=1656; [npk]=471+449+450+447=1817; G=9324.$$

Here  $N=2^3 \cdot b=8 \times 32$ . Now we proceed to apply Yates method of algorithm for calculating various factorial combinations totals as follows:

Table(19.5.2): Yates algorithm for  $2^3$  factorial experiment.

Treatment Combinations	Total yield	(1)	(2)	(3)	Effects total	S.S.
1	425	851	3172	9324	G	$G^2/32=2716780.5$
N	426	2321	6152	360	[N]	$[N]^2/32=4050$
K	1118	2679	86	2264	[K]	$[K]^2/32=160178$
Nk	1203	3473	274	132	[NK]	$[NK]^2/32=544.5$
P	1283	1	1470	2980	[P]	$[P]^2/32=277512.5$
Np	1396	85	794	188	[NP]	$[NP]^2/32=1104.5$
Kp	1656	113	84	-676	[KP]	$[KP]^2/32=14280.5$
Npk	1817	161	48	-36	[NPK]	$[NPK]^2/32=40.5$

Now calculate other SS as follows:

$$CF = G^2/N = [9324]^2/32 = 2716780.5$$

$$\text{Block SS} = (1/8)[(2369)^2 + (2291)^2 + (2289)^2 + (2375)^2] - 2716780.5$$

$$= (1/8)[5612161 + 5248681 + 5239521 + 5640625] - 2716780.5 = 2717623.5 - 2716780.5 = 843.$$

$$\text{Treatment SS} = (1/8)[(425)^2 + (426)^2 + (1118)^2 + (1203)^2 + (1283)^2 + (1396)^2 + (1656)^2 + (1817)^2] - CF$$

$$= (1/8)[180625 + 181476 + 1249924 + 1447209 + 1646089 + 1948816 + 2742336 + 3301489] - CF$$

$$= 12697964/8 - 2716780.5 = 1587245.5 - 2716780.5 = -1129535.$$

Table of Squares of each observation is to be calculated as follows:

Block - 1	104329	7569	178929	104976	77841	221841	16384	111556	823425
Block - 2	93636	114244	165649	7921	201601	104976	11236	73984	773247
Block - 3	10201	70225	11236	139129	97344	152881	84681	202500	768197
Block - 4	189225	199809	17161	91204	130321	73984	104976	10609	817289

Sum of squares of each observation = 3182158.

Total SS = 3182158 – 2716780.5 = 465377.5.

Now we proceed to construct ANOVA table of 2<sup>3</sup> factorial experiment conducted in RBD with 4 replicates is as follows:

**Table (19.5.3): ANOVA table of 2<sup>3</sup> Factorial Experiment conducted in RBD.**

Source of Variation (SOV)	Degrees of freedom (DOF)	Sum of Squares (SS)	Mean Sum Squares (MSS)	Variance Ratio F (F – cal)
Treatments (rows):	2 <sup>3</sup> -1=7	S <sub>t</sub> <sup>2</sup> =#t=457710.5	V <sub>1</sub> =S <sub>t</sub> <sup>2</sup> /2 <sup>2</sup> -1 =65387.21429	F <sub>t</sub> =V <sub>1</sub> /V <sub>10</sub> =201.2209**
Main effect-N	1	S <sub>N</sub> <sup>2</sup> = [N] <sup>2</sup> /32 =4050	V <sub>2</sub> = S <sub>N</sub> <sup>2</sup> = 4050	F <sub>N</sub> = V <sub>2</sub> / V <sub>10</sub> 12.46336**
Main effect-P	1	S <sub>P</sub> <sup>2</sup> =[P] <sup>2</sup> /32=277512.5 S <sub>K</sub> <sup>2</sup> =[K] <sup>2</sup> /32 = 160178	V <sub>3</sub> = S <sub>P</sub> <sup>2</sup> =277512.5	F <sub>P</sub> = V <sub>3</sub> /V <sub>10</sub> =854.0097**
Main effect-K	1	S <sub>NP</sub> <sup>2</sup> = [NP] <sup>2</sup> /32 =1104.5	V <sub>4</sub> = S <sub>K</sub> <sup>2</sup> =160178	F <sub>K</sub> =V <sub>4</sub> /V <sub>10</sub> =492.6199**
Interaction effect NxP	1	S <sub>NK</sub> <sup>2</sup> =[NK] <sup>2</sup> /32 =544.5	V <sub>5</sub> = S <sub>NP</sub> <sup>2</sup> = 1104.5	F <sub>NP</sub> = V <sub>5</sub> /V <sub>10</sub> =3.3989
Interaction effect NxK	1	S <sub>KP</sub> <sup>2</sup> =[KP] <sup>2</sup> /32 =14280.5	V <sub>6</sub> = S <sub>NK</sub> <sup>2</sup> =544.5	F <sub>NK</sub> =V <sub>6</sub> /V <sub>10</sub> =1.6756
Interaction effect KxP	1	S <sub>NPK</sub> <sup>2</sup> =[NPK] <sup>2</sup> /32 =40.5	V <sub>7</sub> =S <sub>KP</sub> <sup>2</sup> =14280.5	F <sub>KP</sub> =V <sub>7</sub> /V <sub>10</sub> =43.9464**
Interaction effect NxPxK	1	S <sub>b</sub> <sup>2</sup> =BSS=843	V <sub>8</sub> =S <sub>NPK</sub> <sup>2</sup> =40.5	F <sub>NPK</sub> =V <sub>8</sub> /V <sub>10</sub> Not possible
Blocks (columns)	b – 1=3		V <sub>9</sub> = S <sub>b</sub> <sup>2</sup> /b-1=281	F <sub>Blocks</sub> = V <sub>9</sub> / V <sub>10</sub> Not possible

Error	$(2^3-1)(b-1)$ =21	$S_E^2 = ESS = \#\# = 6824$	$V_{10} = S_E^2 / (2^2-1)(b-1)$ =324.952381	
Total	$2^3b - 1 = 31$	$S_T^2 = TSS = 465377.5$		

$\# = 4050 + 277512.5 + 160178 + 1104.5 + 544.5 + 14280.5 + 40.5 = 457710.5$  and

$\#\# = TSS - S_t^2 - S_b^2 = 465377.5 - 457710.5 - 843 = 6824$ .

Conclusions: F-table values:

- (1)  $F_{tab}$  at 5% los for (7,21) dof is 2.58 and at 1% los for (7,21) dof is 3.64,
- (2)  $F_{tab}$  at 5% los for (3,21) dof is 3.07 and at 1% los for (3,21) dof is 4.87 and
- (3)  $F_{tab}$  at 5% los for (1,21) dof is 4.32 and at 1% los for (1,21) dof is 8.02.

**For treatments:** Since  $F_t$  is larger than  $F_{tab}$  at 1% los for (7,21) dof we reject the null hypothesis  $H_{0t}$  and conclude that there exists significant variation between various treatment combinations of factors Nitrogen -N, Phosphate - P and Potash- K. Hence there is a need for further Statistical analysis for sub-grouping various treatment effects.

On similar lines, we conclude that there exists significant effects of main effect of factors Nitrogen - N; Phosphate - P and Potash- K on the tomato yield. Similarly there is significant interaction effect of factors Phosphate - P and Potash - K that Interaction effect KxP. All other treatment combinations have insignificant effect.

For blocks and second order interaction effect NPK are not possible to calculate because numerator for F is smaller than the denominator. It is important to note that calculated value of F must always be greater than 1. That is numerator for F must be greater than the denominator. Hence, we cannot take any decision for Blocks and second interaction effect of factors N,P and K.

**Remark:** Above significance of the first interaction between factors KxP can be interpreted as the factors Phosphate - K and the Potash - P operates full joint effect rather than individual factors. Whereas, the factors Nitrogen - N and Potash - K is that, these factors do not act independently of one another, the presence of both enhances their individual effects.

#### Further Statistical analysis for treatments:

The significance of various factorial effects, namely, Main effects and interaction effects can be tested directly, from the factorial totals as follows:

From the ANOVA table (19.5.3) we can observe that S.E. of any factorial total =  $\sqrt{b^2 S_E^2} = \sqrt{4 \times 8 \times 324.952381} = \sqrt{10398.47619} = 101.97292$ .

- Hence the Critical Difference (CD) at 5% and 1% los are respectively given by:

CD at 5% =  $t_{5\%}$  for 21 dof  $\times \sqrt{b^2 S_E^2} = 2.08 \times 101.97292 = 212.10367$  and

CD at 1% =  $t_{1\%}$  for 21 dof  $\times \sqrt{b^2 S_E^2} = 2.83 \times 101.97292 = 288.58336$ .

Comparing above calculated values, with that of Modulus values of various factorial effects totals given in column 4 of table (19.5.2), we can observe that:

All main effects namely, Nitrogen – N Phosphate – k and Potash – P and the interaction effect between K and P are significant at 1% los. Other factorial effects are insignificant.

Before closing this lesson on 2<sup>n</sup> – factorial designs, it is important to note that these designs are specifically useful to represent where the effects are curvy nature. Regression lines are used to explain linear relations. If the relations are curvy nature, 2<sup>n</sup> factorial experiments prove their worth and explain in an excellent way such curvy linear effects. These experiments allow one to keep the size and complexity of the design low and simultaneously obtain some protection against curvy nature. Now we proceed to explain 3<sup>n</sup> – factorial experiments in the next lesson.

## 19.6 Summary

In this lesson, we have discussed 2<sup>n</sup> factorial Experiments in particular when n = 2 and n=3 designs. Some fundamental concepts, various main effects and Interaction effects of 2<sup>2</sup> and 2<sup>3</sup> factorial designs are discussed. Finally, statistical analysis of these designs, along with examples are discussed.

## 19.7 Self Assessment Questions

1. Explain various factorial effects studied in 2<sup>2</sup> factorial designs.
2. Explain the Yates method of calculating various factorial totals in 2<sup>2</sup> factorial design.
3. Explain the Statistical analysis of 2<sup>2</sup> factorial experiments.
4. Analyze the following data obtained from 2<sup>2</sup> factorial experiment conducted in RBD with four blocks. The data obtained is measured in kgs / plot and the treatment combinations are applied randomly as given in the following table.

Block – 1	AB:138	A:125	B:130	1:168
Block – 2	B: 150	1: 175	AB: 133	A:114
Block – 3	A: 112	B:141	AB:117	1:179
Block – 4	1: 178	AB:110	A: 133	B:125

Also obtain the CD values for 5% and 1% los.

5. Explain various factorial effects of 2<sup>n</sup> factorial experiment. Also explain the Standard symmetric order of these factorial effects.
6. Explain various factorial effects of 2<sup>3</sup> factorial experiment.
7. Explain the Yates method of algorithm for 2<sup>3</sup> factorial experiment.
8. Write down the ANOVA table of 2<sup>3</sup> experiment conducted in LSD.

9. Explain the Statistical analysis of  $2^3$  factorial experiments.
10. In an N,P and K trial, with two levels, of each fertilizer, and 5 replicates, various treatment totals are given as follows:

1= 194;n=208;p=197;k=198;np=214;nk=223;kp=211 and nkp=224.

Further, error mean square is known to be 19.86. Write down ANOVA table for the above experiment and draw your conclusions. Also apply further statistical analysis if necessary.

### 19.8 Further Readings

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons, New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

**Lesson – 20** **$3^n$  - FACTORIAL EXPERIMENTS****20.0. Objectives :**

After going through this lesson, you should be able to:

- Understand the need, concepts and definition of  $3^n$  - Factorial Experiments.
- Basic concepts of  $3^2$  - Factorial experiments.
- Statistical analysis of  $3^2$  – Factorial experiments.
- $3^3$  Factorial Experiments.
- Statistical analysis of  $3^3$  - Factorial Experiments.

**Structure**

- 20.1 Introduction to  $3^n$  Factorial Experiments**
- 20.2 The  $3^2$  Factorial Experiment**
- 20.3 Statistical analysis of  $3^2$  – Factorial Experiments**
- 20.4 The  $3^3$  – Factorial Experiment**
- 20.5 ANOVA table of  $3^3$  – Factorial Experiments**
- 20.6 Summary**
- 20.7 Self Assessment Questions**
- 20.8 Further Readings**

**20.1: Introduction to  $3^n$  - Factorial Experiments**

In the last lesson we have discussed  $2^n$  factorial experiments, which are used in many fields like Agriculture, pharmacy and chemical industry. There are some extensions and variations of these designs, which are occasionally used when the factors have three levels. Such designs are popularly known as  $3^n$  factorial experiments. In these experiments, we consider a factorial arrangement of  $n$  – factors each at three levels. These three levels may be considered as Low - 0, medium - 1 and high -2. Thus if we have two factors A and B then various factorial levels are 00,01,02,10,,11,12,20,21 and 22. In this notation first letter represent the level of first factor A and the second letter represent the level of second factor B.

Thus we have  $3^2 = 9$  treatment combinations. Similarly in  $3^3$  factorial experiment we have 27 factorial treatment combinations. We can also consider factors which are quantitative in nature. Then we can consider -1,0 and +1 levels respectively represent low, medium and high levels of different factors. Now we proceed to explain  $3^2$  factorial design in the following section.

**20.2: The  $3^2$  – Factorial Designs**

In this design, we consider two factors say A and B each at 3 levels say 0 – low, 1- medium and 2- high concentrations of factors A and B. For this design, we have  $3^2 = 9$  factorial

combinations of treatments Namely:  $A_0B_0, A_0B_1, A_0B_2, A_1B_0, A_1B_1, A_1B_2, A_2B_0, A_2B_1$  and  $A_2B_2$  respectively denoted as 00,01,02,10,11,12,20,21 and 22. This design can be conducted in CRD with the above nine treatment combinations with  $r$ -replicates, or RBD in 9 treatments and  $b$ -blocks or LSD of size  $9 \times 9$ . Here we consider three factorial effects, namely Main effect of Factor A; main effect of the factor B and interaction effect of factors  $A \times B$ . The eight degrees of freedom for treatments splits into three orthogonal components as 2 dof for Main effect of factor - A; 2 dof for main effect of factor - B and 4 dof for interaction effect of factors  $A \times B$ .

Main effects and interaction effects are explained in the following section.

### 20.3. Statistical Analysis of 3<sup>2</sup> Factorial Experiments

The Statistical analysis of 3<sup>2</sup> factorial experiment conducted in CRD with  $R$  replicates is explained as follows:

Null-hypotheses:

$H_{01}$ : There is no significant variation between all the replicate means.

$H_{02}$ : The main effect of the factor N is not significant.

$H_{03}$ : the main effect of the factor – P is not significant.

$H_{04}$ : The interaction effect of  $I_{NP}$  is not significant.

$H_{05}$ : the Interaction effect of  $J_{NP}$  is not significant.

Here, we have  $3^2 r = N$  experimental plots required to conduct 3<sup>2</sup> factorial experiment. Then various treatment combinations of these two factors each at three levels are  $3^2 = 9$  and are denoted as follows:

		Factor P			
		0	1	2	Totals
Factor N	0	(00)	(01)	(02)	$n_0$
	1	(10)	(11)	(12)	$n_1$
	2	(20)	(21)	(22)	$n_2$
Totals:		$p_0$	$p_1$	$p_2$	G = Grand Total.

In the above notation, ( ) represent totals of the respective treatments effects from all the replicates,  $n_i$ 's are factor N (row) totals and  $p_j$ 's are factor – P (column) totals and G is the grand total respectively. After calculating the above totals from the given data, main effects and interaction effects are calculated as follows:

$$\text{Correction Factor} = CF = G^2 / 3^2 r \quad (20.3.1)$$

$$\text{SS due to Main effect of factor N} = [n_0^2 + n_1^2 + n_2^2] / 3r - CF. = S_2 \quad (20.3.2)$$

$$\text{SS due to Main effect of factor P} = [p_0^2 + p_1^2 + p_2^2] / 3r - CF. = S_3 \quad (20.3.3)$$

Interaction effects between factors N and P splits into two components namely I-interactions denoted by  $I_{NP}$  and J-interactions which are denoted by  $J_{NP}$ , which are defined as follows:

$$I_{NP} = [I_0^2 + I_1^2 + I_2^2] / 3r - CF. = S_4 \quad (20.3.4)$$

$$J_{NP} = [J_0^2 + J_1^2 + J_2^2] / 3r - CF. = S_5 \quad (20.3.5)$$

$$\text{Where } I_0 = (00) + (11) + (22) \quad (20.3.6)$$

$$I_1 = (10) + (21) + (02) \quad (20.3.7)$$

$$I_2 = (20) + (01) + (12) \quad (20.3.8)$$

$$J_0 = (00) + (12) + (21) \quad (20.3.9)$$

$$J_1 = (01) + (10) + (22) \quad (20.3.10)$$

$$J_2 = (20) + (11) + (02) \quad (20.3.11)$$

$$\text{SS due to interaction between the two factors N and P} = [I_{NP} + J_{NP}] \quad (20.3.12)$$

It is important to note that these I and J interactions have no physical meaning or difficult to interpreting the meaning of these effects. For the convenience of the analysis these interaction effects are introduced.

$$\text{SS due to replicates} = R_1^2 + R_2^2 + \dots + R_r^2 / 3r - CF = S_1 \quad (20.3.13)$$

$R_1, R_2, \dots, R_r$  are replicate totals.

$$\text{Calculate total sum of Squares} = \sum_{i=1}^9 \sum_{j=1}^r [y_{ij}]^2 - CF. \quad (20.3.14)$$

Now we proceed to form the ANOVA table for  $3^2$  factorial experiment as follows:

Table (20.3.1) ANOVA table of  $3^2$  factorial experiment conducted in CRD with r replicates.

Source Of Variation (SOV)	DOF	SS	MSS	$F_{cal}$
Replicates	$r - 1$	$S_1$	$V_1 = S_1 / r - 1$	$F_1 = V_1 / V_6$
Main Effect of Factor N	2	$S_2$	$V_2 = S_2 / 2$	$F_2 = V_2 / V_6$
Main effect of factor P	2	$S_3$	$V_3 = S_3 / 2$	$F_3 = V_3 / V_6$
Interaction effect $I_{NP}$	2	$S_4$	$V_4 = S_4 / 2$	$F_4 = V_4 / V_6$
Interaction effect of $J_{NP}$	2	$S_5$	$V_5 = S_5 / 2$	$F_5 = V_5 / V_6$
Error	*	**	$V_6 = ** / *$	
Total	$3^2 r - 1$	TSS		

In the above table  $* = 3^2 r - 1 - [r - 1 + 2 + 2 + 2 + 2] = 3^2 r - r - 8 = 8(r - 1)$ .

Similarly  $** = TSS - [S_1 + S_2 + S_3 + S_4 + S_5]$ .

Conclusions:  $F_1$  is distributed as F [(r-1),\*] degrees of freedom and  $F_2$  to  $F_5$  are distributed as F [2,\*] degrees of freedom. If  $F_{cal} \leq F$  table value we accept the corresponding Null-hypothesis.

Otherwise, we reject the same.

Now we proceed to explain the procedure with an example as follows:



Example (20.3.1): Consider an agricultural experiment involving two factors Nitrogen-N and Phosphorous-P each at three levels namely 1,1 and 2 to test the effect of these factors at different levels for increasing the yield of wheat crop conducted by CRD with 3 replicates. Analyze the data and draw your conclusions.

Replicate – 1: (00)	(20)	(10)	(01)	(21)	(11)	(12)	(22)	(02)
85	85	73	95	85	89	74	71	83
Replicate – 2: (21)	(22)	(00)	(10)	(20)	(12)	(01)	(02)	(11)
82	79	74	78	91	86	92	85	77
Replicate – 3: (10)	(20)	(00)	(11)	(01)	(02)	(22)	(12)	(21)
71	82	79	91	93	79	74	82	77.

Solution: Frame the null hypotheses as follows:

H<sub>01</sub>: There is no significant variation between the replicate means of wheat yield.

H<sub>02</sub>: There is no significant main effect of factor N.

H<sub>03</sub>: There is no significant main effect of factor P.

H<sub>04</sub>: There is no significant I interaction between factors N and P.

H<sub>05</sub>: There is no significant J interaction effect between factors N and P.

Now we proceed to calculate various factorial effect totals in the following table.

		Factor P			
		0	1	2	Totals
	0 (00):238	(01):280	(02):247	765=n <sub>0</sub>	
Factor N	1 (10):222	(11):257	(12):242	721=n <sub>1</sub>	
	2 (20):258	(21):244	(22):224	726=n <sub>2</sub>	
Totals:	p <sub>0</sub> =718	p <sub>1</sub> =781	p <sub>2</sub> =713	2212=G=Grand total.	

Correction Factor:  $CF = G^2/3^2r = (2212)^2/27 = 181220.1481$ .

SS due to N =  $[n_0^2 + n_1^2 + n_2^2]/3r - CF = [(765)^2 + (721)^2 + (726)^2]/9 - 181220.1481$   
 $= [585225 + 519841 + 527076]/9 - 181220.1481 = 181349.1111 - 181220.1481$   
 $= 128.963011 = S_2$ .

SS due to P =  $[p_0^2 + p_1^2 + p_2^2]/3r - CF = [(718)^2 + (781)^2 + (713)^2]/9 - 181220.1481$   
 $= [515524 + 609961 + 508369]/9 - 181220.1481 = 181539.3333 - 181220.1481 = 319.185233 = S_3$ .

$$I_0 = (00)+(11)+(22)=238+257+224=719 \quad (20.3.15)$$

$$I_1 = (10) + (21) + (02)=222+244+247=713 \quad (20.3.16)$$

$$I_2 = (20) + (01) + (12)=258+280+242=780 \quad (20.3.17)$$

$$J_0 = (00) + (12) + (21)=238+242+244=724 \quad (20.3.18)$$

$$J_1 = (01) + (10) + (22)=280+222+224=726 \quad (20.3.19)$$

$$J_2 = (20) + (11) + (02)=258+257+247=762 \quad (20.3.20)$$

$$\text{SS due to interaction effect } I_{NP} = [I_0^2 + I_1^2 + I_2^2]/3r - CF = [(719)^2 + (713)^2 + (780)^2]/9 - 181220.1481 \\ = [516961 + 508369 + 608400]/9 - 181220.1481 = 181525.5556 - 181220.1481 = 305.407456 = S_4.$$

$$\text{SS due to interaction effect } J_{NP} = [J_0^2 + J_1^2 + J_2^2]/3r - CF = [(724)^2 + (726)^2 + (762)^2]/9 - 181220.1481 \\ = [524176 + 527076 + 580644]/9 - 181220.1481 = 181321.7778 - 181220.1481 = 101.629678 = S_5.$$

Replicate – 1:	(00)	(20)	(10)	(01)	(21)	(11)	(12)	(22)	(02)	Totals
	85	85	73	95	85	89	74	71	83	740=R <sub>1</sub>

Replicate – 2:	(21)	(22)	(00)	(10)	(20)	(12)	(01)	(02)	(11)	
	82	79	74	78	91	86	92	85	77	744=R <sub>2</sub>

Replicate – 3:	(10)	(20)	(00)	(11)	(01)	(02)	(22)	(12)	(21)	
	71	82	79	91	93	79	74	82	77.	728=R <sub>3</sub>

$$\text{Replicate SS} = [R_1^2 + R_2^2 + R_3^2]/3r - CF = [(740)^2 + (744)^2 + (728)^2]/9 - 181220.1481 \\ = [547600 + 553536 + 529984]/9 - 181220.1481 = 181235.5556 - 181220.1481 = 15.407456 = S_1$$

Replicate – 1:	(00)	(20)	(10)	(01)	(21)	(11)	(12)	(22)	(02)	SquaresTotals
	7225	7225	5329	9025	7225	7921	5476	5041	6889	61356

Replicate – 2:	(21)	(22)	(00)	(10)	(20)	(12)	(01)	(02)	(11)	
	6724	6241	5476	6084	8281	7396	8464	7225	5929	61820

Replicate – 3:	(10)	(20)	(00)	(11)	(01)	(02)	(22)	(12)	(21)	
	5041	6724	6241	8281	8649	6241	5476	6724	5929.	59306

$$\text{Total SS} = (85)^2 + (85)^2 + \dots + (77)^2 - 181220.1481 = 182482 - 181220.1481 = 1261.8519 = \text{TSS.}$$

Now we proceed to form ANOVA table of 3<sup>2</sup> factorial design as follows:

Source Of Variation (SOV)	DOF	SS	MSS	F <sub>cal</sub>
Replicates	$r - 1 = 2$	$S_1 = 15.4075$	$V_1 = S_1 / r - 1 = 7.7038$	$F_1 = V_1 / V_6 = \dots$
Main Effect of Factor N	2	$S_2 = 128.9630$	$V_2 = S_2 / 2 = 64.4815$	$F_2 = V_2 / V_6 = 2.637$
Main effect of factor P	2	$S_3 = 319.1852$	$V_3 = S_3 / 2 = 159.5926$	$F_3 = V_3 / V_6 = 6.526^{**}$
Interaction effect I <sub>NP</sub>	2	$S_4 = 305.4025$	$V_4 = S_4 / 2 = 152.7013$	$F_4 = V_4 / V_6 = 6.244^{**}$
Interaction effect of J <sub>NP</sub>	2	$S_5 = 101.6297$	$V_5 = S_5 / 2 = 50.8149$	$F_5 = V_5 / V_6 = 2.078$
Error	*=16	**=391.2590	$V_6 = ** / * = 24.4537$	
Total	3 <sup>2</sup> r-1=26	TSS=1261.8519		

Conclusions: F<sub>Tab</sub> for (2,16) dof at 5% los is 3.63 and at 1% los is 6.23.

Since, F<sub>1</sub> calculated value is less than 1 it is not possible to infer about replicate means.

We reject H<sub>03</sub> and H<sub>04</sub> at 1% los. We accept H<sub>02</sub> and H<sub>05</sub> at 5% los. Hence, we conclude that there is significant effect of the factor P, that is Phosphorous and interaction effect of I<sub>NP</sub> on the increase of the yield of wheat crop.

#### 20.4. The 3<sup>3</sup> – Factorial Experiment

On similar lines of 3<sup>2</sup> factorial experiments, we can extend the analysis for three factors each at three levels, namely 3<sup>3</sup> factorial experiments. In this design, we have three factors each at three levels and hence, we have 3<sup>3</sup> = 27 treatment combinations. Let the three factors be represented by Factor A, Factor-B and Factor-C and the corresponding levels are denoted by 0,1 and 2. Then we have the following 27 treatment combinations as follows:

(0,0,0) (0,0,1) (0,0,2) (0,1,0) (0,1,1) (0,1,2) (0,2,0) (0,2,1) (0,2,2)  
 (1,0,0) (1,0,1) (1,0,2) (1,1,0) (1,1,1) (1,1,2) (1,2,0) (1,2,1) (1,2,2)  
 (2,0,0) (2,0,1) (2,0,2) (2,1,0) (2,1,1) (2,1,2) (2,2,0) (2,2,1) (2,2,2)

Main effects, Interaction effects and the analysis of 3<sup>3</sup> factorial experiments are discussed in the following section. It is important to note that interpretation of interaction effects in 3<sup>3</sup> factorial designs requires lot of care and concentration.

#### 20.5. Statistical analysis of 3<sup>3</sup> – Factorial Experiments.

In 3<sup>3</sup> factorial experiments we have three factors each at three levels and hence we have 27 factorial combinations. If all these combinations are considered and conducted in CRD with r replicates, we require '27r' experimental units. Thus we have 27r-1 degrees of freedom for total SS. Further, each main effect has 2 degrees of freedom; each two factor interactions have 4 degrees of freedom and three factors interaction has 8 degrees of freedom. Further, error degrees of freedom will become 27(r-1).

The sum of squares (SS) may be calculated using the standard methods of Statistical Analysis applied to analyze factorial experiments explained in previous lessons. Further, if the factors are quantitative, the main effects may be further partitioned into linear and quadratic components, each with a single degree of freedom. Similarly, two factor interactions may be decomposed into linear x linear, linear x quadratic, quadratic x linear, and quadratic x quadratic effects each with one degree of freedom. It is also possible to partition the two factor interactions into I, J components as we have done in  $3^2$  factorial design. For example, two factor interactions are denoted by **AB, AB<sup>2</sup>, AC, AC<sup>2</sup>, BC and BC<sup>2</sup>**. Then, each of these two factor interaction effects has two degrees of freedom.

Similarly, three factor interaction effects are too difficult for interpretation. But for the purpose of analysis we partition them into four orthogonal components each with 2 degrees of freedom. Namely, W(ABC);X(ABC);Y(ABC) and Z(ABC) interaction effects. These are usually known as W,X,Y and Z components of the interaction. They are also referred as  $AB^2C^2, AB^2C, ABC^2$  and ABC components of three factor interaction ABC respectively. These two notations are interchangeable as follows:

$$\mathbf{W(ABC) = AB^2C^2 ; X(ABC) = AB^2C; Y(ABC) = ABC^2 \text{ and } Z(ABC) = ABC.} \quad (20.4.1)$$

It is important to note that in the above notations, no first letter can have exponent other than 1. Like I, J components, in  $3^2$  factorial experiments, the letters W,X,Y and Z components have no practical meaning or interpretation. They are introduced for the convenience of the analysis of  $3^3$  factorial experiments. These interaction effects are useful in the construction of more complicated designs, which are out of the scope of this book.

Now we proceed to explain the ANOVA table of  $3^3$  factorial experiment with an example.

Example (3.4.1): Write down the ANOVA table of  $3^3$  factorial design conducted in CRD with r-replicates.

Solution: The ANOVA table of  $3^3$  factorial Experiment, conducted in CRD with r- replicates is given as follows:

Source Of Variation (SOV)	DOF	SS	MSS	F <sub>cal</sub>
Replicates	r - 1	S <sub>1</sub>	V <sub>1</sub> = S <sub>1</sub> /r-1	F <sub>1</sub> = V <sub>1</sub> / V <sub>9</sub>
Main Effect of Factor A	2	S <sub>2</sub>	V <sub>2</sub> = S <sub>2</sub> / 2	F <sub>2</sub> = V <sub>2</sub> / V <sub>9</sub>
Main effect of factor B	2	S <sub>3</sub>	V <sub>3</sub> = S <sub>3</sub> / 2	F <sub>3</sub> = V <sub>3</sub> / V <sub>9</sub>
Main effect of factor C	2	S <sub>4</sub>	V <sub>4</sub> = S <sub>4</sub> / 2	F <sub>4</sub> = V <sub>4</sub> / V <sub>9</sub>
Interaction effect of AB	4	S <sub>5</sub>	V <sub>5</sub> = S <sub>5</sub> / 4	F <sub>5</sub> = V <sub>5</sub> / V <sub>9</sub>
Interaction effect of AC	4	S <sub>6</sub>	V <sub>6</sub> = S <sub>6</sub> / 4	F <sub>6</sub> = V <sub>6</sub> / V <sub>9</sub>
Interaction effect of BC	4	S <sub>7</sub>	V <sub>7</sub> = S <sub>7</sub> / 4	F <sub>7</sub> = V <sub>7</sub> / V <sub>9</sub>
Interaction effect of ABC	8	S <sub>8</sub>	V <sub>8</sub> = S <sub>8</sub> / 8	F <sub>8</sub> = V <sub>8</sub> / V <sub>9</sub>

Error	$3^3(r-1)$	**	$V_9 = **/3^3(r-1)$	
Total	$3^3r-1$	TSS		

In the above table  $** = TSS - (S_1 + S_2 + S_3 + S_4 + S_5 + S_6 + S_7 + S_8)$  and SS are to be calculated as usual.

Thus we have to frame eight null hypotheses and can be compared respectively calculated values of F with the respective degrees of freedom at  $\alpha\%$  los. Conclusions can be drawn accordingly. On similar lines we can extend the analysis for  $3^k$  factorial experiments.

Statistical analysis of these designs are complicated and above the scope of this book.

## 20.6. Summary

In this lesson, we have discussed experiments involving different factors say k factors each at three Levels. That is  $3^k$  series of factorial experiments. First we have discussed the  $3^2$  factorial experiments and discussed the method of calculation of Main effects and interaction effects of two factors each at three levels. Statistical analysis of  $3^2$  factorial experiments is discussed and explained the procedure with an example. Further,  $3^3$  factorial experiments, that is three factors each at three levels are discussed. In  $3^3$  factorial experiments we have 27 various treatment combinations and discussed main effects and Interaction effects. Since, calculations are lengthy, only ANOVA table of  $3^3$  factorial experiment is discussed.

## 20.7. Self Assessment Questions

1. Explain the need of  $3^k$  factorial experiments with suitable examples.
2. Explain the main effects and interaction effects of  $3^2$  factorial experiments.
3. Explain the statistical analysis of  $3^2$  factorial experiment.
4. Explain main and interaction effects of  $3^3$  factorial experiments.
5. Write down the ANOV A table of  $3^3$  factorial experiments.

## 20.8. Further Readings

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons, New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

## Lesson – 21

# Balanced Incomplete Block Design (BIBD)

### 21.0. Objectives :

After going through this lesson, you should be able to:

- Understand the need, concepts and definition of incomplete block designs.
- Basic concepts of Balanced Incomplete Block Design (BIBD).
- Statistical analysis of BIBD.
- Efficiency of BIBD over Randomized designs.
- Further Statistical analysis applicable in BIBD.

### Structure

- 21.1 Introduction to Incomplete Block Designs
- 21.2 The Balanced Incomplete Block Design (BIBD)
- 21.3 Parameters and properties of BIBD
- 21.4 Types of BIBD and its properties
- 21.5 Intra Block Analysis of BIBD
- 21.6 Summary
- 21.7 Self Assessment Questions
- 21.8 Further Readings

### 21.1: Introduction to Incomplete Block Designs

In the previous designs, we have considered each block size equal to the number of treatments. Thus we can apply all the treatment combinations completely in each block. Hence, these designs are known as “**Complete Designs**”. For example, in  $2^2$ - factorial design we have blocks of size 4 each. Similarly, in  $2^3$  factorial design we require blocks of size 8 each and  $3^3$  factorial design we require to have blocks of size 9 and in  $3^3$  factorial design, block sizes must be 27. Thus, as the factors – n or levels-s are increasing, we required to have large number of treatments and hence, same number of experimental units in each block, in “**complete block designs (CBD)**”. That is in complete block designs; each block will receive all the treatments completely once. Thus each block will become one complete replicate of all treatments.

Basic problem in such complete block designs is that, when experimental units increases, the heterogeneity between units in each block will enter automatically and hence, “**Experimental Error**” will increase automatically, which is not a desirable situation. To bring Homogeneity between the units in each block, we have to reduce the size of the block or further divide the block into sub-blocks such that units within each sub-block are homogeneous. But in such sub-blocks we cannot apply all the treatments say ‘v’, which are larger in size than the block sizes say ‘k’, that is  $v > k$ . then, we have to plan such that these

treatments are to be applied in such a way that treatment effects can be estimated with same accuracy without increasing the “**Experimental Error**”. Such experiments are called “**Incomplete Block Designs**” whose definition is given as follows:

**Definition: Incomplete Block Design (IBD):**

An Incomplete Block Design (IBD) is one, where  $v$ -treatments are arranged in  $b$ -blocks each of size  $k$  such that  $v > k$  and each of the treatments is replicated  $r$ -times. Further, each pair of treatments, must occur once and only once in the same block.

The quantities  $v, b, k$  and  $r$  are called parameters of IBD.

Remark: In the above design, if  $v=k$ , then Incomplete Block Design (**IBD**) will become Complete Block Design (**CBD**).

Incomplete Block designs are two types as follows:

- (1) Balanced Incomplete Block Designs (BIBD) and
- (2) Partially Balanced Incomplete Designs (PBIBD).

Now we proceed to explain “Balanced Incomplete Block Designs” only, in this lesson because, Partially Balanced Incomplete Block Designs are out of the scope of the syllabus.

### 21.2: The Balanced Incomplete Block Design (BIBD)

The Balanced Incomplete Block Designs are introduced by Prof. F. Yates in a paper “A new method of arranging variety trials involving large number of varieties” in the Journal Agriculture Science Vol. 26, pp 424 – 455 in the year 1936. These designs are designed specially for experiments in ‘Plant Breeding’ and ‘Agriculture’ experiments to compare all treatment comparisons, among pairs of treatments is made with equal precession.

**Definition of BIBD:**

An arrangement of  $v$  treatments in  $b$  blocks of  $k$  plots each such that  $k < v$  is known as Balanced Incomplete Block Design (BIBD) if:

- (1) Each treatment occurs once and only once in  $r$  blocks and
- (2) Each pair of treatments occurs together in  $\lambda$  blocks.

The following example of BIBD will explain above conditions clearly:

(21.2.1) Example of BIBD:

Block – 1	Block - 2	Block – 3	Block – 4	Block – 5	Block-6
Treatment -1	Treatment - 1	Treatment - 1	Treatment - 2	Treatment - 2	Treatment - 3
Treatment - 2	Treatment - 3	Treatment - 4	Treatment - 3	Treatment - 4	Treatment - 4

In the above example we have four treatments, namely, 1,2,3 and 4. To test these four treatments in RBD, we require blocks each of size 4. But planning in the above example we can conduct the experiment with blocks of size 2 each. Thus we can reduce the total number of

experimental units required for RBD, that is 24 Units. But in BIBD we require only 12 experimental units. In the above example we have  $v=4; b=6; k=2$  and  $\lambda = 1$ . Thus we have  $k < v$ .

### 21.3: Parameters and properties of BIBD

The integers  $v, r, b, k$  and  $\lambda$  are called Parameters of BIBD, where:

$v$ =number of treatments or factorial combinations,

$b$ =number of blocks,

$k$ =number of plots in each block or block size,

$r$ =number of replicates and

$\lambda$ =number of blocks in which any pair of treatments occurs together or number of times any two treatments occur together in a block.

These parameter of BIBD satisfies the following conditions which are known as "Parametric relations of BIBD" and serves as necessary conditions for the existence of a BIBD.

$$(1) vr=bk \quad (2) \lambda(v-1) = r(k-1) \quad \text{and} \quad (3) b \geq v \quad (\text{Fisher's Inequality}). \quad (21.3.1)$$

Theorem (21.3.1):  $vr=bk$ .

Proof: Since there are  $v$  treatments and each is replicated  $r$  number of times.

Hence total number of experimental units required for the design is  $vr$ .

Similarly, we have  $b$  blocks each of size  $k$ .

Hence, total number of experimental plots required for the design =  $bk$ .

Total number of experimental units must be equal to total number of experimental plots.

Hence,  $vr$  must be equal to  $bk$  or  $vr=bk$ . Hence the proof.

To prove the relation (2) in (21.3.1), we require some basic concepts in matrix theory which are explained as follows:

**21.3.1: Incidence Matrix:** A matrix associated with any design  $D$  is the "Incidence Matrix" denoted by  $N = [n_{ij}]$   $i=1,2,\dots,v$ ; and  $j=1,2,\dots,b$ , where  $n_{ij}$  denotes the number of times  $i^{\text{th}}$  treatment occurred in  $j^{\text{th}}$  block.

Thus for a BIBD, an incidence matrix  $N$  is given by:

$$N = \begin{bmatrix} n_{11} & n_{12} & \dots & \dots & \dots & \dots & n_{1b} \\ n_{21} & n_{22} & \dots & \dots & \dots & \dots & n_{2b} \\ & & & & \vdots & & \\ & & & & \vdots & & \\ & & & & \vdots & & \\ n_{v1} & n_{v2} & \dots & \dots & \dots & \dots & n_{vb} \end{bmatrix} \quad (21.3.2)$$

Where,  $n_{ij} = 1$ , if  $i^{\text{th}}$  treatment occurs in  $j^{\text{th}}$  block.

= 0, otherwise.



**(21.2.2): Example of an Incidence Matrix:**

Thus the incidence for the BIBD given in example (21.2.1) is given as follows:

$$N_{4 \times 6} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \quad (21.3.3)$$

**(21.3.2) Properties of Incidence Matrix:**

$$1. \sum_{j=1}^b n_{ij} = \sum_{j=1}^b n_{ij}^2 = r ; i = 1, 2, \dots, v. \quad (21.3.4)$$

$$2. \sum_{i=1}^v n_{ij} = \sum_{i=1}^v n_{ij}^2 = k ; j = 1, 2, \dots, b \text{ and} \quad (21.3.5)$$

$$3. \sum_{j=1}^b n_{ij} n_{lj} = \lambda \quad \text{for } i \neq l = 1, 2, \dots, v. \quad (21.3.6)$$

This is because of the fact that  $n_{ij}n_{lj} = 1$  if and only if  $i$ th treatment and  $l$ th treatment occur together in the  $j$ th block. Otherwise the product  $n_{ij}n_{lj} = 0$ . Thus they occur together in  $\lambda$  blocks. Using the above properties of incidence matrix we can show that:

$$[NN^T]_{v \times v} = \begin{bmatrix} r & \lambda & \lambda & \dots & \dots & \lambda \\ \lambda & r & \lambda & \dots & \dots & \lambda \\ & & \vdots & & & \\ & & \vdots & & & \\ & & \vdots & & & \\ \lambda & \lambda & \lambda & \dots & \dots & r \end{bmatrix}, \text{ where } N^T \text{ is the transpose of the matrix } N. \quad (21.3.7)$$

Now we proceed to prove properties of BIBD  $\lambda(v-1) = r(k-1)$  and Fisher's inequality as follows:

**Theorem (21.3.2):** In a BIBD  $\lambda(v-1) = r(k-1)$  is always true.

Proof: Let  $E_{m \times n}$  denote an  $m \times n$  matrix consisting of all the elements equal to 1. We know from (21.3.7) that:

$$[NN^T]_{v \times v} = \begin{bmatrix} r & \lambda & \lambda & \dots & \dots & \lambda \\ \lambda & r & \lambda & \dots & \dots & \lambda \\ & & \vdots & & & \\ & & \vdots & & & \\ & & \vdots & & & \\ \lambda & \lambda & \lambda & \dots & \dots & r \end{bmatrix}. \quad (21.3.8)$$

Multiply both sides of (21.3.8) with column matrix consisting of 1's. That is with  $E_{v \times 1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ .

Then we have:

$$[NN^T]_{v \times v} E_{vx1} = \begin{bmatrix} r & \lambda & \lambda & \dots & \lambda \\ \lambda & r & \lambda & \dots & \lambda \\ & & \vdots & & \\ & & \vdots & & \\ \lambda & \lambda & \lambda & \dots & r \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = [r+\lambda(v-1)] \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = [r+\lambda(v-1)] E_{vx1}. \tag{21.3.9}$$

$$\begin{aligned} \text{Consider } NN^T E_{vx1} &= N(N^T E_{vx1}) = N \begin{bmatrix} n11 & n21 & \dots & \dots & \dots & nv1 \\ n12 & n22 & \dots & \dots & \dots & nv2 \\ & & \vdots & & & \\ & & \vdots & & & \\ n1b & n2b & \dots & \dots & \dots & nvb \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = N \begin{bmatrix} \sum_{i=1}^v ni1 \\ \sum_{i=1}^v ni2 \\ \vdots \\ \sum_{i=1}^v nib \end{bmatrix} \\ &= N \begin{bmatrix} k \\ k \\ \vdots \\ k \end{bmatrix} = k \begin{bmatrix} n11 & n12 & \dots & \dots & \dots & n1b \\ n21 & n22 & \dots & \dots & \dots & n2b \\ & & \vdots & & & \\ & & \vdots & & & \\ nv1 & nv2 & \dots & \dots & \dots & nvb \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = k \begin{bmatrix} \sum_{j=1}^b n1j \\ \sum_{j=1}^b n2j \\ \vdots \\ \sum_{j=1}^b nbj \end{bmatrix} = k \begin{bmatrix} r \\ r \\ \vdots \\ r \end{bmatrix} = kr E_{vx1}. \tag{21.3.10} \end{aligned}$$

From equations (21.3.9) to (21.3.10) we have:

$$[r+\lambda(v-1)] E_{vx1} = kr E_{vx1} \Rightarrow [r+\lambda(v-1)] = kr \Rightarrow \lambda(v-1) = r(k-1). \text{ Hence the theorem. } \tag{21.3.11}$$

**Theorem (21.3.3): Fisher’s Inequality:  $b \geq v$ .**

**Proof: We know that:**

$$[NN^T]_{v \times v} = \begin{bmatrix} r & \lambda & \lambda & \dots & \lambda \\ \lambda & r & \lambda & \dots & \lambda \\ & & \vdots & & \\ & & \vdots & & \\ \lambda & \lambda & \lambda & \dots & r \end{bmatrix} \Rightarrow |NN^T| = \begin{vmatrix} r & \lambda & \lambda & \dots & \lambda \\ \lambda & r & \lambda & \dots & \lambda \\ & & \vdots & & \\ & & \vdots & & \\ \lambda & \lambda & \lambda & \dots & r \end{vmatrix}$$

Adding 2<sup>nd</sup>, 3<sup>rd</sup>, ..., v<sup>th</sup> columns, to the first column and taking  $[r+(v-1)\lambda]$

$$\begin{vmatrix} 1 & \lambda & \lambda & \dots & \lambda \\ 1 & r & \lambda & \dots & \lambda \\ & & \vdots & & \\ & & \vdots & & \\ 1 & \lambda & \lambda & \dots & r \end{vmatrix}$$

Further, subtracting first row from 2<sup>nd</sup>, 3<sup>rd</sup>, ..., v<sup>th</sup> rows we have:

$$[r+(v-1)\lambda] \begin{vmatrix} 1 & \lambda & \lambda & \dots & \lambda \\ 0 & (r-\lambda) & 0 & \dots & 0 \\ & & \vdots & & \\ & & \vdots & & \\ 0 & 0 & 0 & \dots & (r-\lambda) \end{vmatrix} = [r+(v-1)\lambda](r-\lambda)^{v-1}. \tag{21.3.12}$$

We know that  $\lambda(v-1) = r(k-1)$ . (21.3.13)

Hence substituting in equation (21.3.12) we have:

$$[r + k(r-1)](r-\lambda)^{v-1} = kr(r-\lambda)^{v-1}. \quad (21.3.14)$$

Thus we have proved that  $|NN^T| \neq 0$ , and if  $r = \lambda$  then from equation (21.3.13) we have:

$(v-1) = (k-1) \Rightarrow v = k$ . Thus if  $r = \lambda$ , the design will become RBD with  $k$  treatments and  $b - k$  blocks.

Since  $NN^T$  is a non-singular matrix we have Rank of  $(NN^T) = v$ . This is because the order of the matrix  $NN^T$  is  $v$ . We know that Rank of the matrix  $(NN^T) = \text{Rank of the matrix } (N)$ .

Since the matrix  $N$  is of size  $v \times b$  and hence, the rank of the matrix is at most  $b$ .

This implies that  $v = \text{rank of the matrix } N \leq b \Rightarrow v \leq b$  or  $b \geq v$ . Hence the theorem.

Using the above proved we can prove the following consequent results.

**Result - 1:  $r \geq k$ .**

Proof: we know that  $vr = bk \Rightarrow r = (b/v)k$ . We know that  $b \geq v$  and hence we have:

$r \geq k$  because  $(b/v) > 1$ .

**Result - 2:  $b \geq v + r - k \geq v$ .**

Proof: We know that number of treatments in BIBD is greater than the block size  $k$ . This implies that  $(v - k) \geq 0$ . Similarly we know that  $(r - k) \geq 0 \Rightarrow (v-k)(r - k) \geq 0$ .

$$\Rightarrow (v/k - 1)(r - k) \geq 0 \Rightarrow (v/k)(r - k) - (r - k) \geq 0 \Rightarrow (vr/k) - v \geq (r - k). \quad (21.3.15)$$

Since  $vr = bk$  we have from (21.3.15)  $b - v \geq (r - k) \Rightarrow b \geq v + r - k$ . (21.3.16)

## 21.4. Types of BIBD and its properties

Balanced Incomplete Block Designs (BIBD) are classified into two categories, namely:

- (1) Symmetric BIBD,
- (2) Resolvable BIBD and
- (3) Affine resolvable BIBD.

### 21.4.1: Symmetric BIBD and their properties:

**Definition: A BIBD is said to be symmetric BIBD if  $b=v$  and  $r=k$ .**

Necessary conditions for the existence of symmetrical BIBD are:

1. Necessary condition for a symmetrical BIBD with  $v$  as even is that  $(r - \lambda)$  must be perfect square.
2. If  $N$  is the incidence matrix of the symmetrical BIBD, then:
  - (a) Every row sum is equal to  $r$ .
  - (b) Every column sum is equal to  $k$ .
  - (c) The inner product of any two rows of  $N$  is equal to  $\lambda$ .

Example of a symmetrical BIBD:

(21.4.1) Example of a symmetrical BIBD:

Block – 1	Block – 2	Block – 3	Block – 4	Block – 5	Block-6
Treatment -1	Treatment –1	Treatment - 1	Treatment - 2	Treatment –2	Treatment - 3
Treatment - 2	Treatment –3	Treatment - 4	Treatment - 3	Treatment –4	Treatment - 4

### Properties of symmetric BIBD:

1. In symmetric BIBD the incidence matrix is a Square Matrix.
2. In a symmetric BIBD the number of treatments common between any two blocks is  $\lambda$ .
3. If  $N$  is the incidence matrix of a symmetric BIBD, then:

$$(NN^T)^{-1} = \frac{1}{INNTI} \text{Adj}(NNT) \quad (21.4.1)$$

4. If  $N$  is the incidence matrix of a symmetric BIBD, then:

$$(NN^T)(N^T N) = (r - \lambda) N^T N + k^2 \lambda E_{v \times v}. \quad (21.4.2)$$

(21.4.2): Resolvable BIBD Design:

Definition: A BIBD with parameters  $v, r, b, k$  and  $\lambda$  is said to be resolvable BIBD if the blocks can be divided into  $r$  groups or sets of  $b/r$  blocks each,  $b/r$  being an integer, such that  $b/r$  blocks forming any of these sets give a complete replication of all  $v$  treatments.

Example (21.4.2): Example for a resolvable BIBD:

Consider  $v=4, b=6, r=3, k=2$  and  $\lambda = 1$ . Then the following layout of the treatments makes the BIBD a Resolvable BIBD.

Block – 1	Treatment – 1	Treatment – 2	First Set
Block – 2	Treatment - 3	Treatment – 4	
Block – 3	Treatment – 1	Treatment – 3	Second Set
Block – 4	Treatment - 2	Treatment – 4	
Block – 5	Treatment – 1	Treatment – 4	Third Set
Block – 6	Treatment – 2	Treatment - 3	

Here,  $b=6$  are divided into  $r=3$  sets each of  $(b/r)=n$ ; that is  $6/3 = 2$ , which is an integer and blocks are integer multiple of  $r$ . That is  $3 \times 2 = 6$ . Further, each set contains each of the treatments occurring once and only once. Further, each pair of treatments also occurs once and only once, that is  $\lambda = 1$ .

Properties of a Resolvable BIBD:

1. In a resolvable BIBD parameters  $v, b, r, k$  and  $\lambda$  satisfies  $b \geq v+r-1$ .
2. In a resolvable BIBD number of blocks  $b = rn$  where  $n$  is an integer.

### 21.4.3: Affine Resolvable BIBD:

Definition: A resolvable BIBD is said to be Affine Resolvable BIBD if  $b = r+v-1$  and any two blocks from different sets have  $k^2/v$  treatments common, where  $k^2/v$  is an integer.

Example (21.4.3): Affine resolvable BIBD:

Consider the resolvable design with  $v=4, b=6, r=3, k=2$  and  $\lambda = 1$ . Then, the layout is:

Block – 1	Treatment – 1	Treatment – 2	First Set
Block – 2	Treatment – 3	Treatment – 4	
Block – 3	Treatment – 1	Treatment – 3	Second Set
Block – 4	Treatment – 2	Treatment – 4	
Block – 5	Treatment – 1	Treatment – 4	Third Set
Block – 6	Treatment – 2	Treatment – 3	

We can observe that in the above design the condition  $b=v+r-1$  is satisfied and also  $k^2/v = 1$  is an integer. Further any two blocks from different sets have only one treatment and  $\lambda = 1$ .

Now we proceed to discuss the statistical analysis of BIBD, which is popularly known as “**Intra Block Analysis of BIBD**” in the following section.

### 21.5. Statistical analysis of BIBD (Intra Block analysis of BIBD):

This method of Statistical analysis of BIBD was developed by Prof. F. Yates using the method of “Standard Least square” technique and is popularly known as “Intra Block Analysis”. In this method, we obtain the estimates of various treatment effects, without recovery of “Inter Block” effects. Recovery of Inter Block information is not possible, because, blocks are of smaller size than treatments and hence, we cannot get the data with all the treatment combinations replicated in each block. The method is explained as follows:

#### 21.5.1: Intra Block Analysis:

Let ‘ $a$ ’ units of the experimental material with ‘ $b$ ’ blocks in which ‘ $v$ ’ treatments ( $v < b$ ) are applied such that each treatment is replicated ‘ $r$ ’ times such that  $r < b$  satisfying the conditions of BIBD. Let  $N = [n_{ij}]$  be the incidence matrix of the BIBD. Then,

$n_{ij} = 1$  if  $i^{\text{th}}$  treatment occurs in the  $j^{\text{th}}$  block and  
 $= 0$ , otherwise.

By using the properties of incidence matrix given in (21.3.4), (21.3.50 and 21.3.6) we have:

$$1. \sum_{j=1}^b n_{ij} = \sum_{j=1}^b n_{ij}^2 = r ; i = 1, 2, \dots, v. \quad (21.5.1)$$

$$2. \sum_{i=1}^v nij = \sum_{i=1}^v nij^2 = k ; j = 1, 2, \dots, b \text{ and} \quad (21.5.2)$$

$$3. \sum_{j=1}^b nij nlj = \lambda \quad \text{for } i \neq l = 1, 2, \dots, v. \quad (21.5.3)$$

Let  $Y_{ij}$  be the yield or observation recorded on the experimental unit receiving  $i^{\text{th}}$  treatment belonging to the  $j^{\text{th}}$  block where the pairs  $(i, j)$  runs through the set  $D$  for which  $n_{ij} = 1$ . The mathematical of BIBD is given by:

$$Y_{ij} = \mu + t_i + b_j + \epsilon_{ij}, \text{ for } (i, j) \text{ belonging to } D; \text{ for all values of } i=1, 2, \dots, v ; j=1, 2, \dots, b, (21.5.4)$$

Where  $\mu$  is the general mean effect,  $t_i$  ( $i=1, 2, \dots, v$ ) is the effect of the  $i^{\text{th}}$  treatment,  $b_j$  ( $j=1, 2, \dots, b$ )

is the  $j^{\text{th}}$  block effect and  $\epsilon_{ij}$  are the  $v \times r$  intra block errors which are assumed to be normally distributed with mean zero and common variance  $\sigma^2$ . That is errors  $\epsilon_{ij}$  are independently and identically distributed (iid)  $N(0, \sigma_\epsilon^2)$ . The intra block analysis assumes that **the treatment effects and block effects are fixed but unknown.**

According to the principle of least squares, the normal equations for estimating  $(v+b+1)$  parameters of the model given in (21.5.4), namely:  $t_i$  ( $i=1, 2, \dots, v$ );  $b_j$  ( $j=1, 2, \dots, b$ ) and  $\mu$  are obtained on minimizing the error sum of squares denoted by  $E$  and is given by:

$$E = \sum_{i,j}^{v,b} \epsilon_{ij}^2 = \sum_{i,j}^{v,b} (y_{ij} - \hat{\mu} - t_i - b_j)^2 \text{ for } (i, j) \in D. \quad (21.5.5)$$

Differentiating  $E$  with respect to  $\mu$ ,  $t_i$  and  $b_j$  we obtain the following normal equations.

$$\partial E / \partial \mu = -2 \sum_{i,j}^{v,b} (y_{ij} - \hat{\mu} - t_i - b_j) = 0; \text{ for } (i, j) \in D. \quad (21.5.6)$$

$$\partial E / \partial t_i = -2 \sum_{i,j}^{v,b} (y_{ij} - \hat{\mu} - t_i - b_j) = 0; \text{ for } j \in D_i. \quad (21.5.7)$$

$$\partial E / \partial b_j = -2 \sum_{i,j}^{v,b} (y_{ij} - \hat{\mu} - t_i - b_j) = 0; \text{ for } i \in D_j. \quad (21.5.8)$$

In the above equations,  $\hat{\mu}$ ,  $t_i$  and  $b_j$  are least square estimates of the parameters of  $\mu$ ,  $t_i$  and  $b_j$  respectively. Similarly  $D_i$  is the set of  $r$  values for which  $(i, j) \in D$  for a given  $i$  and  $D_j$  is the set of  $k$  values for which  $(i, j) \in D$  for given  $j$ .

In order to have unique solution for the normal equations obtained in equations (21.5.6), (21.5.7) and (21.5.8) we must add the following two conditions, Namely:

$$\sum_{i=1}^v t_i = 0 \text{ and } \sum_{j=1}^k b_j = 0. \quad (21.5.9)$$

Solving the above set of equations from (21.5.6) to (21.5.9) we have:

$$\hat{\mu} = \sum_{i,j}^{v,b} y_{ij} = \bar{y}.. \quad (21.5.10)$$

From equation (21.5.7), we have:

$$\sum_j^b y_{ij} - r\hat{\mu} - r t_i - \sum_{j=1}^k \beta_j = 0 \text{ for } j \in D_i \Rightarrow T_i = r\bar{y}.. + r t_i - \sum_{j=1}^k \beta_j = 0 \text{ for } j \in D_i \quad (21.5.11)$$

Where,  $T_i = \sum_j^b y_{ij}$ ;  $j \in D_i$  is the total yield of the  $i^{\text{th}}$  treatment.

Further, we know that  $D_i$  is the set of  $r$  values for which  $(i, j) \in D$  for a given  $i$ . that is  $n_{ij} = 1$  and  $i$  has a fixed value. Thus we have:

$$T_i = r\bar{y}_{..} + r\tau_i + \sum_{j=1}^k n_{ij} \beta_j \quad (21.5.12)$$

From equation (21.5.12) we can observe that the estimate of the treatment effect is no longer of the form of observed treatment mean minus the grand mean, since, the block effects do not enter the same way into all the observed treatment means. For example, a treatment may be favored by occurring only in the blocks with high block effects.

Similarly, from the equation (21.5.8), we have:

$$B_j = K\bar{y}_{..} + K\beta_j + \sum_{i=1}^v n_{ij} \tau_i \quad (21.5.13)$$

Where,  $B_j = \sum_{i=1}^v y_{ij}$ ;  $i \in D_j$  is the total yield of  $j^{\text{th}}$  block.

$$\text{Thus the quantity } (B_j / K) - \bar{y}_{..} = \beta_j + (1/K) \sum_{i=1}^v n_{ij} \tau_i \quad (21.5.14)$$

The expression given in (21.5.14) is called as the estimate of  $j^{\text{th}}$  block effect ignoring treatments.

Substituting the value of  $\beta_j$  from (21.5.14) in equation (21.5.13), we have:

$$T_i = r\bar{y}_{..} + r\tau_i + (1/K) \sum_{j=1}^b [n_{ij} (B_j - K\bar{y}_{..} - \sum_{p=1}^v n_{pj} \tau_p)] \quad (21.5.15)$$

$$\Rightarrow T_i - (1/K) \sum_{j=1}^b n_{ij} B_j = r\bar{y}_{..} + r\tau_i - \bar{y}_{..} \sum_{j=1}^b n_{ij} - (1/K) \sum_{j=1}^b \sum_{p=1}^v n_{ij} n_{pj} \tau_p$$

$$= r\tau_i - (1/K) \sum_{j=1}^b \sum_{p=1}^v n_{ij} n_{pj} \tau_p. \quad (21.5.16)$$

$$\text{Let } Q_i = T_i - \sum_{j=1}^b (n_{ij} B_j) / K \quad (21.5.17)$$

The quantity  $Q_i$  is called the  $i^{\text{th}}$  'adjusted treatment total' or the 'adjusted total yield for the  $i^{\text{th}}$  treatment'. The adjustment consists in subtracting from the treatment total  $T_i$ , the sum of the  $j^{\text{th}}$  block average  $B_j/K$ , (which is the average yield per plot for the  $j^{\text{th}}$  block) for those blocks in which  $i^{\text{th}}$  treatment occurs.

Using equations (21.5.16) and (21.5.17), we have:

$$Q_i = r\tau_i - (1/K) \sum_{j=1}^b \sum_{p=1}^v n_{ij} n_{pj} \tau_p$$

$$= -(\tau_1/K) \sum_{j=1}^b n_{1j} n_{1j} - (\tau_2/K) \sum_{j=1}^b n_{2j} n_{2j} \dots - (\tau_i/K) \sum_{j=1}^b n_{ij} n_{ij} \dots - (\tau_v/K) \sum_{j=1}^b n_{vj} n_{vj}$$

Using the equation (21.5.3), we have:

$$Q_i = -(\lambda/K) (\tau_1 + \tau_2 + \tau_3 \dots + \tau_v) + (r - r/K + \lambda/K) \tau_i \quad (21.5.18)$$

Since  $\sum_{i=1}^v \tau_i = 0$  we have:

$$Q_i = (r - r/K + \lambda/K) \tau_i = \{ [r(K-1) + \lambda]/K \} \tau_i = \{ [\lambda(v-1) + \lambda]/K \} \tau_i = (\lambda v/K) \tau_i \quad (21.5.19)$$

This is because of the fact that  $\lambda(v-1) = r(k-1)$ .

$$\text{From (21.5.19) we have: } \tau_i = Q_i (K/\lambda v); i=1,2,3,\dots,v. \quad (21.5.20)$$

$$\text{If we put } (\lambda v)/kr = E, \text{ then we have } \tau_i = Q_i / rE. \quad (21.5.21)$$

Further, using  $\lambda(v-1) = r(k-1)$ , we have:

$$E = (\lambda/r)(v/K) = [v(k-1)]/[k(v-1)] = [vK - v]/[vK - K] < 1; \text{ (since } v > K). \quad (21.5.22)$$

We know that the residual sum of squares for the model given in (21.5.4), we have:

$$\sum_{i,j}^{v,b} (y_{ij} - \hat{\mu} - \tau_i - \beta_j)^2 = \sum_{i,j}^{v,b} (y_{ij})^2 - n \hat{\mu} \bar{y}_{..} - \sum_{i=1}^v T_i \tau_i - \sum_{j=1}^b B_j \beta_j \quad (21.5.23)$$

We know that  $n=vr=bk$  and hence, SS due to  $\hat{\mu}$ ,  $\tau_i$  and  $\beta_j$  are given by:

$$\begin{aligned} S(\mu, b, \tau) &= n \hat{\mu} \bar{y}_{..} + \sum_{i=1}^v T_i \tau_i + \sum_{j=1}^b B_j \beta_j \\ &= \hat{\mu} \sum_{i,j}^{v,b} y_{ij} + \sum_{i=1}^v T_i \tau_i + \sum_{j=1}^b B_j \beta_j \\ &= \hat{\mu} \sum_i^v \sum_{j=1}^b y_{ij} + \sum_{i=1}^v T_i \tau_i + \sum_{j=1}^b B_j \beta_j \quad (i, j) \in D \\ &= \hat{\mu} \sum_{j=1}^b B_j + \sum_{j=1}^b B_j \beta_j + \sum_{i=1}^v T_i \tau_i \\ &= \sum_{j=1}^b (\hat{\mu} + \beta_j) B_j + \sum_{i=1}^v T_i \tau_i \end{aligned} \quad (21.5.24)$$

Substituting for  $\hat{\mu} + \beta_j$ , from equation (21.5.13), we have:

$$\begin{aligned} S(\mu, b, \tau) &= \sum_{j=1}^b [(\mathbf{1}/K) (\mathbf{B}_j - \sum_{i=1}^v n_{ij} \tau_i) \mathbf{B}_j] + \sum_{i=1}^v T_i \tau_i \\ &= \sum_{j=1}^b (\mathbf{B}_j)^2 / K + \sum_{i=1}^v \tau_i (T_i - \sum_{j=1}^b n_{ij} \mathbf{B}_j / K) \end{aligned} \quad (21.5.25)$$

$$= \sum_{j=1}^b (\mathbf{B}_j)^2 / K + \sum_{i=1}^v \tau_i \mathbf{Q}_i = n \hat{\mu} \bar{y}_{..} + [\sum_{j=1}^b (\mathbf{B}_j)^2 / K - n \hat{\mu} \bar{y}_{..}] + \sum_{i=1}^v \tau_i \mathbf{Q}_i \quad (21.5.26)$$

We know that the quantity  $\sum_{i=1}^v \tau_i \mathbf{Q}_i$  is the Sum of Squares due to treatments after adjusting for block effects. Hence, Sum of Squares of treatments adjusted is given by:

$$\text{SS due to treatments (adjusted)} = \sum_{i=1}^v \tau_i \mathbf{Q}_i = \sum_{i=1}^v (\mathbf{Q}_i)^2 / rE = S_2. \quad (21.5.27)$$

$$\text{The quantity } \sum_{j=1}^b (\mathbf{B}_j)^2 / K - n \hat{\mu} \bar{y}_{..} = \sum_{j=1}^b (\mathbf{B}_j)^2 / K - (1/n) [\sum_{i,j=1}^{v,b} y_{ij}]^2$$

$$= \sum_{j=1}^b (\mathbf{B}_j)^2 / K - \text{CF} = S_1. \quad (21.5.28)$$

The expression given in 9 (21.5.28) is the sum of squares due to blocks ignoring treatments. This represents the unadjusted block sum of squares.

Now we proceed to construct the ANOVA table for Intra Block Analysis of BIBD as follows:

Source of variation(SOV)	DOF	SS	MSS	F calculated
Between Blocks (un-adjusted)	b-1	$S_1 = \sum_{j=1}^b (\mathbf{B}_j)^2 / K - \text{CF}$	$V_1 = S_1 / b - 1$	$F_1 = V_1 / V_3$
Between the treatments (adjusted)	v-1	$S_2 = \sum_{i=1}^v (\mathbf{Q}_i)^2 / rE$	$V_2 = S_2 / v - 1$	$F_2 = V_2 / V_3$
Intra Block Error	bk-b-v+1	$S_3 = \text{TSS} - S_1 - S_2$	$V_3 = S_3 / bk - b - v + 1$	
Total	bk - 1	$\text{TSS} = \sum_{i,j=1}^{v,b} (y_{ij})^2 - \text{CF}$		

Conclusions: The null hypothesis to be tested here is  $H_0$ : There is no significant variation between treatment means or  $t_1 = t_2 = \dots = t_v$ . If  $F_2$  calculated value is less than  $F$  table value at  $\alpha\%$  los for  $(v-1, bk-b-v+1)$  dof, we accept the null hypothesis at  $\alpha\%$  los. Otherwise, we reject



the null hypothesis. If necessary we can apply further statistical analysis, for which the reader can refer books suggested below.

### 21.6. Summary

In this lesson, we have discussed concepts related to Balanced Incomplete Block Designs (BIBD). Definitions, parameters and relation between various parameters of BIBD are discussed and properties we proved. Finally the concept of Intra block analysis which is the analysis to be applied for BIBD is discussed.

Finally it is important to remember that there are many advanced topics not covered in this book, because they are out of the scope of the syllabus. If any reader wants to know about them, they can refer books suggested in further readings.

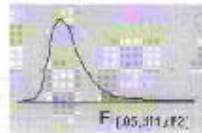
### 21.7. Self Assessment Questions

1. Define BIBD and parameters.
2. State various relations between parameters of BIBD prove any two of them.
3. Show that  $vr = bk$ .
4. State and prove Fisher's Inequality.
5. Define symmetric BIBD and give an example for the same.
6. Define an Incidence matrix and give an example for the same.
7. Show that  $\lambda(v-1) = r(k-1)$ .
8. Define affine resolvable BIBD and give an example of the same.
9. With usual notations, for a resolvable BIBD show that  $b \geq v+r-1$ .
10. Explain the Intra block analysis of BIBD.

### 21.8. Further Readings

1. C.R. Rao (1952) "Advanced Statistical Methods in Biometric research" John Wiley.
2. Fisher R.A. (1947) "The Design of experiments" Oliver and Boyd.
3. S.C. Gupta and V.K. Kapoor (2006) "Fundamentals of Applied Statistics" Sultan Chand & Sons, New Delhi.
4. Douglas C. Montgomery (2000) "Design and Analysis of Experiments" 5<sup>th</sup> edition. John Wiley & Sons, INC New York.

## APPENDIX

F Table for  $\alpha = 0.05$  (1/3)

$d.f_1/d.f_2$	1	2	3	4	5	6	7	8	9	10
1	161.4476	199.5000	215.7073	224.5832	230.1619	233.9860	236.7684	238.8827	240.5433	241.8817
2	18.5128	19.0000	19.1643	19.2468	19.2964	19.3298	19.3522	19.3710	19.3848	19.3959
3	10.1280	9.5521	9.2766	9.1172	9.0155	8.9406	8.8867	8.8452	8.8125	8.7855
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644
5	6.6279	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351
6	5.9274	5.1433	4.7571	4.5337	4.3874	4.2939	4.2067	4.1468	4.0990	4.0610
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7879	3.7257	3.6767	3.6365
8	5.3177	4.4500	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962	2.8536
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.6710
14	4.6001	3.7380	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6022
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943	2.4499
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117
19	4.3807	3.5219	3.1274	2.8951	2.7400	2.6283	2.5435	2.4768	2.4227	2.3779
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5141	2.4471	2.3928	2.3479
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660	2.3210
22	4.3009	3.4431	3.0491	2.8167	2.6611	2.5491	2.4638	2.3965	2.3419	2.2967

F Table for  $\alpha = 0.05$  (2/3)

df1/df2	12	15	20	24	30	40	60	120	$\infty$
1	249.0060	245.9499	248.0151	249.0513	250.0951	251.1432	252.1917	253.2529	254.3144
2	19.1625	19.1591	19.1458	19.1541	19.1624	19.1707	19.1791	19.1874	19.1957
3	8.7446	8.7429	8.7412	8.7385	8.7466	8.7544	8.7629	8.7719	8.7804
4	5.9117	5.9178	5.9225	5.9244	5.9250	5.9270	5.9287	5.9301	5.9311
5	4.6777	4.6788	4.6781	4.6772	4.6757	4.6738	4.6714	4.6685	4.6651
6	3.9999	3.9981	3.9742	3.9415	3.9082	3.8743	3.8398	3.8047	3.7689
7	3.5727	3.5107	3.4745	3.4395	3.3759	3.3404	3.3043	3.2674	3.2298
8	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	3.0729	3.0061	2.9369	2.9006	2.8637	2.8259	2.7872	2.7475	2.7067
10	2.9170	2.8450	2.7740	2.7373	2.6994	2.6609	2.6211	2.5801	2.5379
11	2.7876	2.7156	2.6464	2.6096	2.5705	2.5309	2.490	2.4486	2.4063
12	2.6866	2.6169	2.5426	2.5055	2.4663	2.4259	2.3842	2.3416	2.2982
13	2.6077	2.5331	2.4589	2.4202	2.3793	2.3382	2.2958	2.2524	2.2084
14	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2220	2.1778	2.1337
15	2.4753	2.4044	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0683
16	2.4247	2.3522	2.2756	2.2354	2.1938	2.1517	2.1058	2.0589	2.0126
17	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9601
18	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	2.3088	2.2341	2.1555	2.1141	2.0712	2.0261	1.9795	1.9302	1.8789
20	2.2776	2.2033	2.1242	2.0825	2.0391	1.9928	1.9464	1.8963	1.8442
21	2.2504	2.1757	2.0960	2.0543	2.0107	1.9645	1.9165	1.8657	1.8137
22	2.2258	2.1506	2.0707	2.0283	1.9842	1.9380	1.8894	1.8390	1.7861

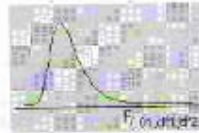
F Table for  $\alpha = 0.05$  (2/3)

df1/df2	12	15	20	24	30	40	60	120	$\infty$
1	249.0060	245.9499	248.0151	249.0513	250.0951	251.1432	252.1917	253.2529	254.3144
2	19.1625	19.1591	19.1458	19.1541	19.1624	19.1707	19.1791	19.1874	19.1957
3	8.7446	8.7029	8.6632	8.6385	8.6166	8.5944	8.5729	8.5519	8.5306
4	5.9117	5.8878	5.8625	5.7744	5.7450	5.7170	5.6907	5.6651	5.6381
5	4.6777	4.6388	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3651
6	3.9999	3.9811	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	3.5747	3.5107	3.4745	3.4305	3.3759	3.3404	3.3043	3.2674	3.2298
8	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	3.0729	3.0061	2.9368	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	2.9170	2.8450	2.7740	2.7373	2.6994	2.6609	2.6211	2.5801	2.5379
11	2.7876	2.7156	2.6464	2.6096	2.5705	2.5309	2.4900	2.4480	2.4048
12	2.6866	2.6169	2.5426	2.5055	2.4653	2.4259	2.3842	2.3410	2.2962
13	2.6077	2.5331	2.4589	2.4202	2.3793	2.3392	2.2966	2.2524	2.2069
14	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2220	2.1778	2.1307
15	2.4753	2.4044	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	2.4247	2.3522	2.2756	2.2354	2.1938	2.1517	2.1058	2.0589	2.0096
17	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0567	2.0077	1.9560
18	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	2.3086	2.2341	2.1555	2.1141	2.0712	2.0261	1.9795	1.9302	1.8780
20	2.2776	2.2033	2.1242	2.0825	2.0391	1.9928	1.9464	1.8963	1.8432
21	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	2.2258	2.1506	2.0707	2.0283	1.9842	1.9380	1.8894	1.8390	1.7831

F Table for  $\alpha = 0.05$  (3/3)

$d.f_2/d.f_1$	1	2	3	4	5	6	7	8	9
23	4.3797	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3211
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9917	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2282	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3401	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3344	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0411
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0184	1.9588
$\infty$	3.8415	2.9957	2.6049	2.3713	2.2141	2.0986	2.0096	1.9384	1.8790

$d.f_2/d.f_1$	10	12	15	20	24	30	40	60	120	$\infty$
23	2.2747	2.2035	2.1282	2.0476	2.0060	1.9605	1.9139	1.8698	1.8128	1.7570
24	2.2547	2.1834	2.1077	2.0267	1.9858	1.9390	1.8920	1.8424	1.7895	1.7330
25	2.2365	2.1649	2.0889	2.0075	1.9663	1.9192	1.8718	1.8217	1.7684	1.7110
26	2.2197	2.1479	2.0716	1.9898	1.9481	1.9010	1.8533	1.8027	1.7488	1.6906
27	2.2043	2.1323	2.0558	1.9736	1.9329	1.8842	1.8361	1.7851	1.7306	1.6717
28	2.1900	2.1179	2.0411	1.9586	1.9174	1.8687	1.8203	1.7689	1.7138	1.6541
29	2.1768	2.1045	2.0275	1.9446	1.9035	1.8540	1.8055	1.7537	1.6981	1.6376
30	2.1646	2.0921	2.0148	1.9317	1.8904	1.8405	1.7918	1.7396	1.6839	1.6223
40	2.0772	2.0035	1.9245	1.8389	1.7972	1.7471	1.6928	1.6373	1.5766	1.5089
60	1.9926	1.9174	1.8364	1.7480	1.7061	1.6561	1.5943	1.5343	1.4673	1.3893
120	1.9105	1.8337	1.7505	1.6587	1.6081	1.5543	1.4952	1.4290	1.3519	1.2549
$\infty$	1.8307	1.7522	1.6660	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0000

F Table for  $\alpha = 0.01$  (1/3)

df1 \ df2	1	2	3	4	5	6	7	8	9
1	4052.181	4999.500	5403.352	5624.583	5763.650	5858.986	5929.356	5981.070	6022.473
2	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388
3	34.116	30.817	29.457	28.710	28.237	27.911	27.673	27.489	27.345
4	21.198	18.000	16.654	15.977	15.522	15.207	14.976	14.799	14.659
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158
6	13.745	10.925	9.780	9.148	8.716	8.466	8.260	8.102	7.976
7	12.246	9.547	8.451	7.847	7.460	7.191	6.993	6.840	6.719
8	11.259	8.649	7.591	7.006	6.622	6.371	6.178	6.029	5.911
9	10.567	8.022	6.992	6.422	6.057	5.802	5.613	5.467	5.351
10	10.044	7.559	6.552	5.994	5.636	5.386	5.200	5.057	4.942
11	9.646	7.206	6.217	5.668	5.316	5.069	4.886	4.744	4.632
12	9.330	6.927	5.953	5.412	5.064	4.821	4.640	4.499	4.388
13	9.071	6.701	5.739	5.205	4.862	4.621	4.441	4.302	4.191
14	8.862	6.515	5.564	5.025	4.685	4.446	4.278	4.140	4.030
15	8.683	6.359	5.417	4.883	4.550	4.318	4.142	4.004	3.895
16	8.521	6.226	5.293	4.773	4.437	4.202	4.026	3.890	3.780
17	8.400	6.112	5.185	4.668	4.336	4.102	3.927	3.791	3.682
18	8.285	6.013	5.092	4.579	4.248	4.015	3.841	3.705	3.597
19	8.185	5.926	5.010	4.500	4.171	3.939	3.765	3.631	3.523
20	8.096	5.849	4.938	4.431	4.104	3.871	3.699	3.564	3.457
21	8.017	5.780	4.874	4.369	4.042	3.813	3.640	3.506	3.398
22	7.945	5.719	4.817	4.313	3.988	3.758	3.587	3.453	3.346

F Table for  $\alpha = 0.01$ (2/3)

d2/df1	10	12	15	20	24	30	40	60	120	$\infty$
1	6055.847	6106.321	6157.285	6208.730	6254.621	6260.649	6286.782	6313.030	6339.391	6365.864
2	99.399	99.416	99.433	99.449	99.458	99.466	99.474	99.482	99.491	99.499
3	27.229	27.052	26.872	26.690	26.558	26.505	26.411	26.316	26.231	26.135
4	14.546	14.374	14.198	14.020	13.929	13.838	13.745	13.652	13.559	13.463
5	10.051	9.888	9.722	9.553	9.466	9.379	9.291	9.202	9.112	9.020
6	7.874	7.718	7.559	7.396	7.313	7.229	7.143	7.057	6.969	6.880
7	6.620	6.469	6.314	6.155	6.074	5.992	5.908	5.824	5.737	5.650
8	5.814	5.667	5.515	5.359	5.279	5.198	5.116	5.032	4.946	4.859
9	5.237	5.111	4.962	4.808	4.729	4.649	4.567	4.483	4.398	4.311
10	4.849	4.706	4.558	4.405	4.327	4.247	4.165	4.082	3.996	3.909
11	4.539	4.397	4.251	4.099	4.021	3.941	3.860	3.776	3.690	3.602
12	4.296	4.155	4.010	3.858	3.780	3.701	3.619	3.535	3.449	3.361
13	4.100	3.960	3.815	3.663	3.587	3.507	3.425	3.341	3.255	3.168
14	3.939	3.800	3.656	3.505	3.427	3.348	3.266	3.181	3.094	3.007
15	3.805	3.666	3.522	3.372	3.294	3.214	3.132	3.047	2.959	2.873
16	3.691	3.553	3.409	3.259	3.181	3.101	3.018	2.933	2.845	2.758
17	3.593	3.455	3.312	3.162	3.084	3.003	2.920	2.835	2.746	2.659
18	3.508	3.371	3.227	3.077	2.999	2.919	2.835	2.749	2.660	2.573
19	3.434	3.297	3.153	3.003	2.925	2.844	2.761	2.674	2.584	2.497
20	3.368	3.231	3.088	2.938	2.859	2.778	2.695	2.608	2.517	2.430
21	3.310	3.173	3.030	2.880	2.801	2.720	2.636	2.548	2.457	2.370
22	3.258	3.121	2.978	2.827	2.749	2.667	2.583	2.495	2.403	2.315
23	3.211	3.074	2.931	2.780	2.701	2.620	2.536	2.447	2.355	2.267
24	3.168	3.031	2.888	2.737	2.658	2.576	2.492	2.403	2.311	2.223
25	3.128	2.991	2.848	2.697	2.618	2.536	2.452	2.362	2.270	2.182
26	3.090	2.953	2.810	2.659	2.580	2.498	2.414	2.324	2.232	2.144
27	3.054	2.917	2.774	2.623	2.544	2.462	2.378	2.287	2.195	2.107
28	3.020	2.883	2.740	2.589	2.510	2.428	2.344	2.253	2.161	2.073
29	2.987	2.850	2.707	2.556	2.477	2.395	2.311	2.220	2.128	2.040
30	2.955	2.818	2.675	2.524	2.445	2.363	2.279	2.188	2.096	2.008
32	2.923	2.786	2.643	2.492	2.413	2.331	2.247	2.156	2.064	1.976
34	2.892	2.755	2.612	2.461	2.382	2.300	2.216	2.125	2.033	1.945
36	2.862	2.725	2.582	2.431	2.352	2.270	2.186	2.095	2.003	1.915
38	2.833	2.696	2.553	2.402	2.323	2.241	2.157	2.066	1.974	1.886
40	2.805	2.668	2.525	2.374	2.295	2.213	2.129	2.038	1.946	1.858
45	2.750	2.613	2.470	2.319	2.240	2.158	2.074	1.983	1.891	1.803
50	2.700	2.563	2.420	2.269	2.190	2.108	2.024	1.933	1.841	1.753
60	2.630	2.493	2.350	2.199	2.120	2.038	1.954	1.863	1.771	1.683
70	2.570	2.433	2.290	2.139	2.060	1.978	1.894	1.803	1.711	1.623
80	2.520	2.383	2.240	2.089	2.010	1.928	1.844	1.753	1.661	1.573
90	2.475	2.338	2.195	2.044	1.965	1.883	1.799	1.708	1.616	1.528
100	2.435	2.298	2.155	2.004	1.925	1.843	1.759	1.668	1.576	1.488
120	2.375	2.238	2.095	1.944	1.865	1.783	1.699	1.608	1.516	1.428
150	2.305	2.168	2.025	1.874	1.795	1.713	1.629	1.538	1.446	1.358
200	2.195	2.058	1.915	1.764	1.685	1.603	1.519	1.428	1.336	1.248
300	2.095	1.958	1.815	1.664	1.585	1.503	1.419	1.328	1.236	1.148
400	2.015	1.878	1.735	1.584	1.505	1.423	1.339	1.248	1.156	1.068
500	1.955	1.818	1.675	1.524	1.445	1.363	1.279	1.188	1.096	1.008
600	1.905	1.768	1.625	1.474	1.395	1.313	1.229	1.138	1.046	0.958
700	1.865	1.728	1.585	1.434	1.355	1.273	1.189	1.098	1.006	0.918
800	1.830	1.693	1.550	1.399	1.320	1.238	1.154	1.063	0.971	0.883
900	1.800	1.663	1.520	1.369	1.290	1.208	1.124	1.033	0.941	0.853
1000	1.775	1.638	1.495	1.344	1.265	1.183	1.099	1.008	0.916	0.828
$\infty$	1.755	1.618	1.475	1.324	1.245	1.163	1.079	0.988	0.896	0.808

F Table for  $\alpha = 0.01$ (2/3)

df2/df1	10	12	15	20	24	30	40	60	120	$\infty$
1	6055.847	6106.321	6157.285	6208.730	6254.621	6260.649	6286.782	6413.030	6439.391	6465.864
2	99.399	99.416	99.433	99.449	99.458	99.466	99.474	99.482	99.491	99.499
3	27.229	27.052	26.872	26.690	26.558	26.505	26.411	26.316	26.231	26.135
4	14.546	14.374	14.198	14.020	13.929	13.838	13.745	13.652	13.558	13.463
5	10.051	9.888	9.722	9.553	9.466	9.379	9.291	9.202	9.112	9.020
6	7.874	7.718	7.559	7.396	7.313	7.229	7.143	7.057	6.969	6.880
7	6.620	6.469	6.314	6.155	6.074	5.992	5.908	5.824	5.737	5.650
8	5.814	5.667	5.515	5.359	5.279	5.198	5.116	5.032	4.946	4.859
9	5.237	5.111	4.962	4.808	4.729	4.649	4.567	4.483	4.398	4.311
10	4.849	4.706	4.558	4.405	4.327	4.247	4.165	4.082	3.996	3.909
11	4.538	4.397	4.251	4.099	4.021	3.941	3.860	3.776	3.690	3.602
12	4.296	4.155	4.010	3.858	3.780	3.701	3.619	3.535	3.448	3.361
13	4.100	3.960	3.815	3.665	3.587	3.507	3.425	3.341	3.255	3.168
14	3.938	3.800	3.656	3.505	3.427	3.348	3.266	3.181	3.094	3.007
15	3.805	3.666	3.522	3.372	3.294	3.214	3.132	3.047	2.959	2.872
16	3.691	3.553	3.409	3.259	3.181	3.101	3.018	2.933	2.845	2.758
17	3.593	3.455	3.312	3.162	3.084	3.003	2.920	2.835	2.746	2.659
18	3.508	3.371	3.227	3.077	2.999	2.919	2.835	2.749	2.660	2.573
19	3.434	3.297	3.153	3.003	2.925	2.844	2.761	2.674	2.584	2.497
20	3.368	3.231	3.088	2.938	2.859	2.778	2.695	2.608	2.517	2.431
21	3.310	3.173	3.030	2.880	2.801	2.720	2.636	2.548	2.457	2.370
22	3.258	3.121	2.978	2.827	2.749	2.667	2.583	2.495	2.403	2.315
23	3.211	3.074	2.931	2.780	2.701	2.620	2.536	2.447	2.355	2.267
24	3.168	3.031	2.888	2.737	2.658	2.576	2.492	2.403	2.311	2.223
25	3.128	2.991	2.848	2.697	2.618	2.536	2.452	2.362	2.270	2.182
26	3.090	2.953	2.810	2.659	2.580	2.498	2.414	2.324	2.232	2.144
27	3.054	2.917	2.774	2.623	2.544	2.462	2.378	2.287	2.195	2.107
28	3.020	2.883	2.740	2.589	2.510	2.428	2.344	2.253	2.161	2.073
29	2.987	2.850	2.707	2.556	2.477	2.395	2.311	2.220	2.128	2.040
30	2.955	2.818	2.675	2.524	2.445	2.363	2.279	2.188	2.096	2.008
32	2.900	2.763	2.620	2.469	2.390	2.308	2.224	2.133	2.041	1.953
34	2.857	2.720	2.577	2.426	2.347	2.265	2.181	2.090	1.998	1.910
36	2.816	2.679	2.536	2.385	2.306	2.224	2.140	2.049	1.957	1.869
38	2.777	2.640	2.497	2.346	2.267	2.185	2.101	2.010	1.918	1.830
40	2.740	2.603	2.460	2.309	2.230	2.148	2.064	1.973	1.881	1.793
45	2.665	2.528	2.385	2.234	2.155	2.073	1.989	1.898	1.806	1.718
50	2.603	2.466	2.323	2.172	2.093	2.011	1.927	1.836	1.744	1.656
60	2.520	2.383	2.240	2.089	2.010	1.928	1.844	1.753	1.661	1.573
70	2.457	2.320	2.177	2.026	1.947	1.865	1.781	1.690	1.598	1.510
80	2.403	2.266	2.123	1.972	1.893	1.811	1.727	1.636	1.544	1.456
90	2.357	2.220	2.077	1.926	1.847	1.765	1.681	1.590	1.498	1.410
100	2.317	2.180	2.037	1.886	1.807	1.725	1.641	1.550	1.458	1.370
120	2.250	2.113	1.970	1.819	1.740	1.658	1.574	1.483	1.391	1.303
140	2.196	2.059	1.916	1.765	1.686	1.604	1.520	1.429	1.337	1.249
160	2.151	1.994	1.851	1.700	1.621	1.539	1.455	1.364	1.272	1.184
180	2.113	1.956	1.813	1.662	1.583	1.501	1.417	1.326	1.234	1.146
200	2.080	1.923	1.780	1.629	1.550	1.468	1.384	1.293	1.201	1.113
250	2.017	1.860	1.717	1.566	1.487	1.405	1.321	1.230	1.138	1.050
300	1.966	1.809	1.666	1.515	1.436	1.354	1.270	1.179	1.087	1.000
400	1.907	1.750	1.607	1.456	1.377	1.295	1.211	1.120	1.028	0.940
500	1.861	1.704	1.561	1.410	1.331	1.249	1.165	1.074	0.982	0.894
600	1.823	1.666	1.523	1.372	1.293	1.211	1.127	1.036	0.944	0.856
700	1.790	1.633	1.490	1.339	1.260	1.178	1.094	1.003	0.911	0.823
800	1.760	1.603	1.460	1.309	1.230	1.148	1.064	0.973	0.881	0.793
900	1.733	1.576	1.433	1.282	1.203	1.121	1.037	0.946	0.854	0.766
1000	1.709	1.552	1.409	1.258	1.179	1.097	1.013	0.922	0.830	0.742
$\infty$	1.688	1.531	1.388	1.237	1.158	1.076	0.992	0.901	0.809	0.721



F Table for  $\alpha = 0.01$  (3/3)

df2/df1	1	2	3	4	5	6	7	8	9
23	7.881	5.664	4.765	4.264	3.939	3.710	3.539	3.405	3.299
24	7.823	5.617	4.718	4.218	3.895	3.667	3.496	3.363	3.256
25	7.770	5.568	4.675	4.177	3.855	3.627	3.457	3.324	3.217
26	7.721	5.526	4.637	4.140	3.818	3.591	3.421	3.288	3.182
27	7.677	5.488	4.601	4.106	3.785	3.558	3.388	3.256	3.149
28	7.636	5.453	4.568	4.074	3.754	3.528	3.358	3.226	3.120
29	7.598	5.420	4.538	4.045	3.725	3.499	3.330	3.198	3.092
30	7.562	5.390	4.510	4.018	3.699	3.473	3.304	3.173	3.067
40	7.314	5.179	4.313	3.828	3.514	3.291	3.124	2.993	2.888
60	7.077	4.977	4.126	3.649	3.339	3.119	2.953	2.823	2.718
120	6.851	4.787	3.949	3.480	3.174	2.956	2.792	2.663	2.559
$\infty$	6.635	4.605	3.782	3.319	3.017	2.802	2.639	2.511	2.407

df2/df1	10	12	15	20	24	30	40	60	120	$\infty$
23	5.211	3.074	2.931	2.781	2.702	2.620	2.535	2.447	2.354	2.256
24	3.168	3.042	2.889	2.736	2.659	2.577	2.492	2.403	2.310	2.211
25	3.129	2.993	2.850	2.699	2.620	2.538	2.453	2.364	2.270	2.169
26	3.094	2.958	2.815	2.664	2.585	2.503	2.417	2.327	2.233	2.131
27	3.062	2.926	2.783	2.632	2.552	2.470	2.384	2.294	2.198	2.097
28	3.032	2.896	2.753	2.602	2.522	2.440	2.354	2.263	2.167	2.064
29	3.005	2.868	2.726	2.574	2.495	2.412	2.325	2.234	2.138	2.034
30	2.979	2.843	2.700	2.549	2.469	2.386	2.299	2.208	2.111	2.006
40	2.801	2.665	2.522	2.369	2.288	2.203	2.114	2.019	1.917	1.805
60	2.632	2.496	2.352	2.198	2.115	2.028	1.936	1.836	1.726	1.601
120	2.472	2.336	2.192	2.035	1.950	1.860	1.763	1.658	1.533	1.381
$\infty$	2.321	2.185	2.039	1.878	1.791	1.696	1.592	1.473	1.325	1.000