UNIT 7 TWO-WAY ANALYSIS OF VARIANCE

Structure

7.1 INTRODUCTION

In Unit 6 of this block, we have discussed the one-way analysis of variance. In one-way analysis of variance, we have considered one independent variable at different levels which affects the response variable. Analysis of variance is a technique which split up the total variation of data which may be attributed to various “sources” or “causes” of variation. There may be variation between variables and also within different levels of variables. In this way, analysis of variance is used to test the homogeneity of several population means by comparing the variances between the sample and within the sample. In this unit, we will discuss the two-way analysis of variance technique. In two-way analysis of variance technique, we will consider two variables at different levels which affect the response variables.

In Section 7.2 the two-way ANOVA model is explained whereas the basic assumptions in two-way ANOVA are described in Section 7.3. The estimates of each level mean of each factor are found in Section 7.4. Test of hypothesis in two-way ANOVA is explained in Section 7.5 and the degrees of freedom of various sum of squares in two-way ANOVA are described in Section 7.6. Expectations of various sum of squares are derived in Section 7.7.

Objectives

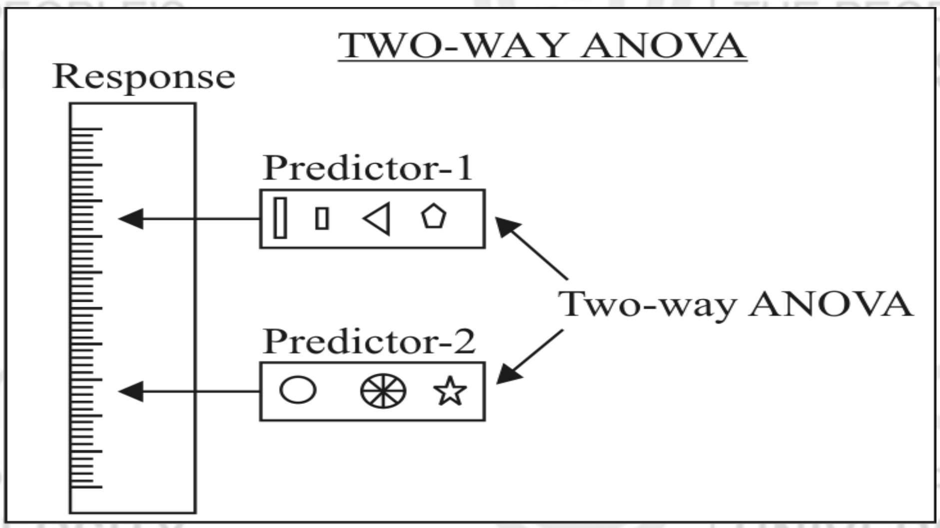
After reading this unit, you would be able to

- explain the two-way ANOVA Model;
- describe the basic assumptions in two-way ANOVA;
- find the estimate of each level mean of each factor;
- explain the degrees of freedom of various sum of squares in two-way ANOVA;

- derive the expectations of various sum of squares;
- construct the ANOVA Table for two-way classification and determine the critical differences; and
- perform the two-way ANOVA test.

7.2 TWO-WAY ANOVA MODEL WITH ONE OBSERVATION PER CELL

In the previous unit, we have considered the case where you had one categorized predictor/independent variable/explanatory at different levels. In this unit, consider a case with two categorical predictors. In general, some times you have more than one categorical predictor variables at a different levels and a continuous response variable then it is called two-way classification.



Two-way classified data can be treated in three ways

1. Analysis of two-way classified data with one observation per cell.
2. Analysis of two-way classified data with equal number of observations (m observation) per cell.
3. Analysis of two-way classified data with unequal number of observation per cell.

In this unit we shall discuss the first case i.e. Case I: Analysis of two-way classified data with one observation per cell (Fixed effect model) only.

We have an experiment in such a way as to study the effect of two factors in the same experiment. For each factor, there will be a number of classes/groups or levels. In the fixed effect model, there will be only fixed levels of the two factors. We shall first consider the case of one observation per cell. Let the factors be A and B and the respective levels be $A_1, A_2, \ldots, A_p$ and $B_1, B_2, \ldots, B_q$. Let y_{ij} be the observation/response/dependent variable under the i^{th} level of factor A and j^{th} level of factor B. The observations can be represented as follows:

Table of Two-Way Classified Data

A/B	B_1	B_2	…	B_j	…	B_q	Total	Mean
A_1	y_{11}	y_{12}	…	y_{1j}	…	y_{1q}	$y_{1.}$	$\bar{y}_{1.}$
A_2	y_{21}	y_{22}	…	y_{2j}	…	y_{2q}	$y_{2.}$	$\bar{y}_{2.}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_i	y_{i1}	y_{i2}	…	y_{ij}	…	y_{iq}	$y_{i.}$	$\bar{y}_{i.}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_p	y_{p1}	y_{p2}	…	y_{pj}	…	y_{pq}	$y_{p.}$	$\bar{y}_{p.}$
Total	$y_{.1}$	$y_{.2}$	…	$y_{.j}$	…	$y_{.q}$	$y_{..} = G$	
Mean	$\bar{y}_{.1}$	$\bar{y}_{.2}$	…	$\bar{y}_{.j}$	…	$\bar{y}_{.q}$		$\bar{y}_{..}$

Mathematical Model

Here, the mathematical model may be written as

$$y_{ij} = \mu_{ij} + e_{ij}$$

where, e_{ij} are independently normally distributed with common mean zero and variance σ_e^2. Corresponding to above table of observations following table is a table of expected values of observations:

Table of Expectations of Observations of y_{ij}

A/B	B_1	B_2	…	B_j	…	B_q	Mean	Difference
A_1	μ_{11}	μ_{12}	…	μ_{1j}	…	μ_{1q}	$\mu_{1.}$	$\mu_{1.} - \mu = \alpha_1$
A_2	μ_{21}	μ_{22}	…	μ_{2j}	…	μ_{2q}	$\mu_{2.}$	$\mu_{2.} - \mu = \alpha_2$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_i	μ_{i1}	μ_{i2}	…	μ_{ij}	…	μ_{iq}	$\mu_{i.}$	$\mu_i - \mu = \alpha_i$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_p	μ_{p1}	μ_{p2}	…	μ_{pj}	…	μ_{pq}	$\mu_{p.}$	$\mu_{p.} - \mu = \alpha_p$
Mean	$\mu_{.1}$	$\mu_{.2}$	…	$\mu_{.j}$	…	$\mu_{.q}$	$\mu_{..}$	
Difference	$\mu_{.1} - \mu = \beta_1$	$\mu_{.2} - \mu = \beta_2$	…	$\mu_{.j} - \mu = \beta_j$	…	$\mu_{.q} - \mu = \beta_q$		

Now μ_{ij} can be written as

$$\mu_{ij} = \mu + (\mu_{i.} - \mu) + (\mu_{.j} - \mu) + (\mu_{ij} - \mu_{i.} - \mu_{.j} + \mu)$$

$$= \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

or $\quad E(y_{ij}) = \mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$

or $\quad y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ij}$

where μ is a constant general mean effect, present in all the observations, $\alpha_i = \mu_{i.} - \mu$ is an effect due to the i^{th} level of the factor A, which is common to all the observations belonging to this level of A , $\beta_j = \mu_{.j} - \mu$ is an effect due to j^{th} level of the factor B, which is common to all the observations belonging of this levels of B and $(\alpha\beta)_{ij} = \mu_{ij} - \mu_{i.} + \mu_{.j} + \mu$ is called the interaction between the i^{th} level of A and j^{th} level of B. It is an effect peculiar to the combination $(A_i B_j)$. It is not present in the i^{th} level of A or in the j^{th} level of B if not taken together i.e.if the joint effect of A_i and B_j is different from the sum of the effects due to A_i and B_j taken individually, it means there is interaction and it is measured by $(\alpha\beta)_{ij}$.

In the case of two-way classified data with one observation per cell, the interaction $(\alpha\beta)$ cannot be estimated and, therefore, there is no interaction i.e. $(\alpha\beta)_{ij}$ for all i and j. So, the model becomes

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \qquad \ldots (1)$$

7.3 ASSUMPTIONS OF TWO-WAY ANOVA

For the validity of F-test, the following assumptions should be satisfied:

1. All the observations y_{ij} are independent;
2. Different effects (effects of levels of factor A, effects of levels of factor B and error effect) are additive in nature;
3. e_{ij} are independently and identically distributed normally with mean zero and variance σ_e^2, i.e. $e_{ij} \approx i\,i\,d\; N(0, \sigma_e^2)$; and
4. There is no interaction between levels of factor A and the levels of factor B.

7.4 ESTIMATION OF PARAMETERS IN TWO-WAY CLASSIFIED DATA

From the above model given in equation (1)

$$e_{ij} = y_{ij} - \mu - \alpha_i - \beta_j \qquad i = 1, 2, \ldots, p$$

Squaring both the sides $\qquad j = 1, 2, \ldots, q$

$$e_{ij}^2 = (y_{ij} - \mu - \alpha_i - \beta_j)^2$$

Summing over i & j both the sides, we get

$$\sum_{i=1}^{p}\sum_{j=1}^{q} e_{ij}^2 = \sum_{i=1}^{p}\sum_{j=1}^{q} (y_{ij} - \mu - \alpha_i - \beta_j)^2$$

Let us suppose that this can be written as

$$E = \sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \mu - \alpha_i - \beta_j\right)^2$$

For obtaining the least square estimates, we minimize E. For minimizing E, differentiating it with respect to μ we get

$$\frac{\partial E}{\partial \mu} = -2\sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \mu - \alpha_i - \beta_j\right)$$

and now equating this equal to zero, we have

$$\sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \mu - \alpha_i - \beta_j\right) = 0$$

or $$\sum_{i=1}^{p}\sum_{j=1}^{q} y_{ij} - pq\mu - q\sum_{i=1}^{p}\alpha_i - p\sum_{j=1}^{q}\beta_j = 0$$

It is clear from the above table that

$$\mu_{i.} = \mu - \alpha_i \quad \text{or} \quad \sum_{i=1}^{p}\left(\mu_{i.} - \mu\right) = \sum_{i=1}^{p}\alpha_i$$

or $$\sum_{i=1}^{p}\mu_{i.} - p\mu = \sum_{i=1}^{p}\alpha_i$$

or $$p\mu - p\mu = \sum_{i=1}^{p}\alpha_i$$

or $$\sum_{i=1}^{p}\alpha_i = 0$$

Similarly, $$\mu._j - \mu = \beta_j \Rightarrow \sum_{j=1}^{q}\beta_j = 0$$

Using these relations and substituting in the equation

$$\sum_{i=1}^{p}\sum_{j=1}^{q} y_{ij} - pq\mu - q\sum_{i=1}^{p}\alpha_i - p\sum_{j=1}^{q}\beta_j = 0$$

$$\sum_{i=1}^{p}\sum_{j=1}^{q} y_{ij} = pq\hat{\mu}$$

or $$\hat{\mu} = \frac{\sum_{i=1}^{p}\sum_{j=1}^{q} y_{ij}}{pq} = \bar{y}_{..}$$

Similarly, differentiating E with respect to $\alpha_1, \alpha_2, ..., \alpha_p$, we get

for i=1, i.e. for estimating α_1

$$E = \sum_{j=1}^{q}(y_{1j} - \mu - \alpha_1 - \beta_j)^2$$

$$\frac{\partial E}{\partial \alpha_1} = -2\sum_{j=1}^{q}\left(y_{1j} - \mu - \alpha_1 - \beta_j\right)$$

Now equating this equation to zero, we get

$$\sum_{j=1}^{q}\left(y_{1j} - \mu - \alpha_1 - \beta_j\right) = 0$$

$$\sum_{j=1}^{q} y_{1j} - q\hat{\mu} - q\hat{\alpha}_1 - \sum_{j=1}^{q}\beta_j = 0$$

or $\sum_{j=1}^{q} y_{1j} - q\hat{\mu} = \hat{\alpha}_1$ because $\sum_{j=1}^{q} \beta_j = 0$

or, $\hat{\alpha}_1 = \frac{\sum_{j=1}^{q} y_{1j}}{q} - \hat{\mu} = (\bar{y}_{1.} - \bar{y}_{..})$

or in general $\hat{\alpha}_i = (\bar{y}_{i.} - \bar{y}_{..})$ for all i =1, 2, ..., p.

Similarly for obtaining the least square estimate for β_j we have to differentiate E w.r.t. β_j and equating to zero, we have

$$\frac{\partial E}{\partial \beta_j} = -2\sum_{i=1}^{p}(y_{ij} - \mu - \alpha_i - \beta_j) = 0$$

or $\sum_{i=1}^{p} y_{ij} - \sum_{i=1}^{p} \mu - \sum_{i=1}^{p} \alpha_i - \sum_{i=1}^{p} \beta_j = 0$

or $\sum_{i=1}^{p} y_{ij} - p\hat{\mu} - p\hat{\beta}_j = 0$ because $\sum_{i=1}^{p} \alpha_i = 0$

or $\hat{\beta}_j = \frac{1}{p}\sum_{i=1}^{p} y_{ij} - \hat{\mu}$

$\hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..}$ for all j =1, 2, . . . , q

7.5 TEST OF HYPOTHESIS IN TWO-WAY ANOVA

It is clear that $\sum_{i=1}^{p} \alpha_i = 0$, $\sum_{j=1}^{q} \beta_j = 0$ and $\sum_{i=1}^{p}\sum_{j=1}^{q} (\alpha\beta)_{ij} = 0$

So, we have the model

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

The least square estimators, as already obtained in previous section by minimizing

$$E = \sum_{i=1}^{p}\sum_{j=1}^{q} (y_{ij} - \mu - \alpha_i - \beta_j)^2$$

and differentiating E with respect to μ, α_i, β_j, we get

$$\hat{\mu} = \frac{\sum_{i=1}^{p}\sum_{j=1}^{q} y_{ij}^2}{pq} = \bar{y}_{..}$$

$\hat{\alpha}_i = (\bar{y}_{i.} - \bar{y}_{..})$ for all i =1, 2 ,…, p.

$\hat{\beta}_j = (\bar{y}_{.j} - \bar{y}_{..})$ for all j =1, 2, … , q.

In this model each observation is the sum of four components and analysis of variance partitions $\sum_{i=1}^{p}\sum_{j=1}^{q} y_{ij}^2$ also in four components as follows:

$$y_{ij} = \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})$$

or
$$y_{ij} - \bar{y}_{..} = (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})$$

Squaring both sides and summing over i and j

$$\sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \bar{y}_{..}\right)^2 = \sum_{i=1}^{p}\sum_{j=1}^{q}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2 + \sum_{i=1}^{p}\sum_{j=1}^{q}\left(\bar{y}_{.j} - \bar{y}_{..}\right)^2$$

$$+ \sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right)^2 + 2\sum_{i=1}^{p}\sum_{j=1}^{q}\left(\bar{y}_{i.} - \bar{y}_{..}\right)\left(\bar{y}_{.j} - \bar{y}_{..}\right)$$

$$+ 2\sum_{i=1}^{p}\sum_{j=1}^{q}\left(\bar{y}_{i.} - \bar{y}_{..}\right)\left(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right)$$

$$+ 2\sum_{i=1}^{p}\sum_{j=1}^{q}\left(\bar{y}_{.j} - \bar{y}_{..}\right)\left(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right)$$

Therefore,

$$\sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \bar{y}_{..}\right)^2 = q\sum_{i=1}^{p}\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2 + p\sum_{j=1}^{q}\left(\bar{y}_{.j} - \bar{y}_{..}\right)^2$$

$$+ \sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right)^2$$

Other terms are zero because, sum of the deviation of observations from their mean is always zero.

So, Total Sum of Squares = Sum of Squares due to Factor A (SSA)

+ Sum of Squares due to Factor B (SSB)

+ Sum of Squares due to Error (SSE)

or in short TSS = SSA + SSB + SSE

The corresponding partition on the total degrees of freedom is as follows:

df of TSS = df of SSA + df of SSB + df of SSE

$pq - 1 = (p - 1) + (q - 1) + (p - 1)(q - 1)$

Dividing the sum of squares by their respective degrees of freedom (df), we get corresponding mean sum of squares.

In two-way classified data, we can test two hypotheses, one for levels of factor A that is equality of different levels of factor A

H_{0A}: $\alpha_1 = \alpha_2 = \ldots = \alpha_p = 0$

H_{1A}: $\alpha_1 \neq \alpha_2 \neq \ldots \neq \alpha_p \neq 0$

and for equality of different levels of factor B

H_{0B}: $\beta_1 = \beta_2 = \ldots = \beta_q = 0$

H_{1B}: $\beta_1 \neq \beta_2 \neq \ldots \neq \beta_q \neq 0$

To derive the appropriate tests for these hypotheses we have obtained the expected values of various sum of squares as explained in one-way classified data in Section 7.7.

So,

$$E(MSSA) = \sigma_e^2 + q\sum_{i=1}^{p}\alpha_i^2 /(p-1)$$

$$E(MSSB) = \sigma_e^2 + p\sum_{j=1}^{q}\beta_j^2 /(q-1)$$

$$E\ (MSSE) = \sigma_e^2$$

If H_{0A}: $\alpha_1 = \alpha_2 = \ldots = \alpha_p = 0$ is true, then E (MSSA) = E (MSSE) and hence

F = MSSA / MSSE will give the test to test H_{0A}.

So, a test for the hypothesis of equality of the effect of the different levels of factor A is provided by this F, which follows the Snedecor F- distribution with [(p - 1), (p - 1)(q - 1)] df.

Thus, the null hypothesis will be rejected at α level of significance if and only if F = MSSA/MSSE > F tabulated [(p - 1), (p - 1) (q - 1)] df at α level of significance.

Similarly if, H_{0B}: $\beta_1 = \beta_2 = \cdots = \beta_q = 0$ for the equality of the effects of the different levels of B is true then E(MSSB) = E(MSSE) and F = MSSB/MSSE will give the test to test H_{0B}. H_{0B} is rejected at the α level of significance if and only if F = MSSB/MSSE > F tabulated [(q - 1), (p - 1) (q - 1)]df at α level of significance. These calculations are shown in the following table.

ANOVA Table for Two-way Classified Data with One Observation per Cell

Source of Variation	DF	SS	MSS	F
Between the Levels of A	p - 1	$SSA = q\sum_{i=1}^{p}(\bar{y}_{i.} - \bar{y}_{..})^2$	MSSA = SSA/(p - 1)	$F_{(p-1),(p-1)(q-1)}$ = MSSA/MSSE
Between the Levels of B	q - 1	$SSB = p\sum_{j=1}^{q}(\bar{y}_{.j} - \bar{y}_{..})^2$	MSSB =SSA/(q - 1)	$F_{(q-1),(p-1)(q-1)}$ = MSSB/MSSE
Error	(p - 1) (q - 1)	SSE $= \sum_{i=1}^{p}\sum_{j=1}^{q}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$	MSSE $= \frac{SSE}{(p-1)(q-1)}$	
Total	pq - 1	$TSS = \sum_{i=1}^{p}\sum_{j=1}^{q}(y_{ij} - \bar{y}_{..})^2$		

If both the null hypotheses are rejected then the means of the different levels of factors A and B can be tested by the method of multiple comparisons given by Tukkey.

7.5.1 General Computation Procedure

For analysing the two-way classified data, with one observation per cell, one has to follow the following procedure:

1. Calculate G = Grand Total = Total of all observations = $\sum_{i=1}^{p}\sum_{j=1}^{q} y_{ij}$
2. Find the N = The number of observations
3. Find Correction Factor(CF) = G^2/N
4. Calculate Raw Sum Squares(RSS) = $\sum_{i=1}^{p}\sum_{j=1}^{q} y_{ij}^{2}$
5. Compute Total Sum of Squares (TSS) = RSS - CF
6. Calculate Sum of Squares due to Factor A (SSA)

   $= y_{1.}^2/q + y_{2.}^2/q + \ldots + y_{i.}^2/q + \ldots + y_{p.}^2/q - CF$
7. Compute Sum of Squares due to Factor B (SSB)

   $= y_{.1}^2/p + y_{.2}^2/p + \ldots + y_{.j}^2/p + \ldots + y_{.q}^2/p - CF$
8. Compute Sum of Squares due to Error (SSE) =TSS – SSA - SSB
9. Compute MSSA = SSA/df,

   MSSB = SSB/df

   MSSE = SSE/df
10. Find F_A = MSSA/MSSE

    and F_B = MSSB/MSSE
11. Compare the calculated value of F_A to tabulated value of F_A; if calculated value is greater than the tabulated value then reject the hypothesis H_{0A}, otherwise it may be accepted.
12. Compare the calculated value of F_B to tabulated value of F_B; if calculated value is greater than the tabulated value then reject the hypothesis H_{0B}, otherwise it may be accepted.

7.6 DEGREES OF FREEDOM OF VARIOUS SUM OF SQUARES

Total sum of squares (TSS) is calculated from pq observations of the form $\left(y_{ij} - \bar{y}_{..}\right)$ which carry $(pq - 1)$degrees of freedom (df), one degree of freedom being lost because of the linear constraints $\sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \bar{y}_{..}\right) = 0$.

Similarly, the sum of squares due to the levels of factor A $\sum_{i=1}^{p} q\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$ =SSA will have (p - 1) degrees of freedom,

since $\sum_{i=1}^{p} q\left(\bar{y}_{i.} - \bar{y}_{..}\right) = 0$.

The sum of squares due to the levels of factor B (SSB) = $\sum_{i=1}^{q} p\left(\bar{y}_{.j}-\bar{y}_{..}\right)^2$ will have (q - 1) degrees of freedom, since $\sum_{i=1}^{p} p\left(\bar{y}_{.j}-\bar{y}_{..}\right)=0$.

The error sum of squares (SSE) = $\sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..}\right)^2$ will have (p - 1) (q - 1) degrees of freedom (df). Since it is based on pq quantities which are subject to (p + q - 1) constraints or which can also be obtained by subtracting the degrees of freedom of sum of squares due to factor A and degrees of freedom of sum of squares due to factor B from the degrees of freedom for total sum of squares.

Hence, df of TSS = df (SSA) + df (SSB) + df (SSE)

(pq−1) = (p - 1) + (q - 1) + (p - 1) (q - 1)

7.7 EXPECTATIONS OF VARIOUS SUM OF SQUARES

For obtaining the valid test statistic, the expected value of various sum of squares should be obtained:

7.7.1 Expectation of Sum of Squares due to Factor A

We have

$$SSA = q\sum_{i=1}^{p}\left(\bar{y}_{i.}-\bar{y}_{..}\right)^2$$

$$E(SSA) = E\left[q\sum_{i=1}^{p}\left(\bar{y}_{i.}-\bar{y}_{..}\right)^2\right]$$

$$E(SSA) = q\,E\left[\sum_{i=1}^{p}\left(\bar{y}_{i.}-\bar{y}_{..}\right)^2\right] \quad \dots (2)$$

We know that $y_{ij}=\mu+\alpha_i+\beta_j+e_{ij}$

So

$$\bar{y}_{i.}=\frac{1}{q}\sum_{j=1}^{q}y_{ij}=\frac{1}{q}\left[\sum_{j=1}^{q}\left(\mu+\alpha_i+\beta_j+e_{ij}\right)\right]$$

or

$$\bar{y}_{i.}=\frac{1}{q}\left[q\mu+q\alpha_i+\sum_{j=1}^{q}\beta_j+\sum_{j=1}^{q}e_{ij}\right]$$

$$=\mu+\alpha_i+\bar{e}_{i.}$$

Similarly, the value of $\bar{y}_{..}$ will be

$$\bar{y}_{..}=\mu+\bar{e}_{..}$$

Substituting the values of $\bar{y}_{i.}$ and $\bar{y}_{..}$ in equation (2), we have

$$E(SSA) = q\,E\left[\sum_{i=1}^{p}\left(\mu+\alpha_i+\bar{e}_{i.}-\mu-\bar{e}_{..}\right)^2\right]$$

$$= q\,E\left[\sum_{i=1}^{p}(\alpha_i + \bar{e}_{i.} - \bar{e}_{..})^2\right]$$

$$= q\,E\left[\sum_{i=1}^{p}\alpha_i^2 + \sum_{i=1}^{p}(\bar{e}_{i.} - \bar{e}_{..})^2 + 2\sum_{i=1}^{p}\alpha_i(\bar{e}_{i.} - \bar{e}_{..})\right]$$

$$= q\left[\sum_{i=1}^{p}\alpha_i^2 + \sum_{i=1}^{p}E(\bar{e}_{i.} - \bar{e}_{..})^2 + 2\sum_{i=1}^{p}\alpha_i E(\bar{e}_{i.} - \bar{e}_{..})\right]$$

Since $e_{ij} \sim iidN(0, \sigma_e^2)$, so $E(e_{ij}) = 0 \quad Var(e_{ij}) = \sigma_e^2$

$$\bar{e}_{i.} \sim iidN\left(0, \frac{\sigma_e^2}{q}\right);\ E(\bar{e}_{i.}) = 0 \quad Var(\bar{e}_{i.}) = \frac{\sigma_e^2}{q}$$

and $$\bar{e}_{..} \sim iidN\left(0, \frac{\sigma_e^2}{pq}\right);\ E(\bar{e}_{..}) = 0 \quad Var(\bar{e}_{..}) = \frac{\sigma_e^2}{pq}$$

So $$E(SSA) = q\sum_{i=1}^{p}\alpha_i^2 + q\sum_{i=1}^{p}E(\bar{e}_{i.} - \bar{e}_{..})^2 + 0$$

Because $$E(\bar{e}_{i.}) = E(\bar{e}_{..}) = 0$$

or $$E(SSA) = q\sum_{i=1}^{p}\alpha_i^2 + qE\sum_{i=1}^{p}\left(\bar{e}_{i.}^2 + \bar{e}_{..}^2 - 2\bar{e}_{i.}\bar{e}_{..}\right)$$

$$= q\sum_{i=1}^{p}\alpha_i^2 + qE\left[\sum_{i=1}^{p}\bar{e}_{i.}^2 + p\,\bar{e}_{..}^2 - 2\,\bar{e}_{..}\sum_{i=1}^{p}\bar{e}_{i.}\right]$$

$$= q\sum_{i=1}^{p}\alpha_i^2 + qE\left[\sum_{i=1}^{p}\bar{e}_{i.}^2 - p\,\bar{e}_{..}^2\right] \quad \text{because } \bar{e}_{..} = \frac{1}{p}\sum_{i=1}^{p}\bar{e}_{i.}$$

$$= q\sum_{i=1}^{p}\alpha_i^2 + q\left\{\sum_{i=1}^{p}E(\bar{e}_{i.}^2) - pE(\bar{e}_{..}^2)\right\}$$

$$= q\sum_{i=1}^{p}\alpha_i^2 + q\left(p\frac{\sigma_e^2}{q} - p\,\frac{\sigma_e^2}{pq}\right)$$

$$E(SSA) = q\sum_{i=1}^{p}\alpha_i^2 + (p-1)\sigma_e^2$$

$$E\left[\frac{SSA}{p-1}\right] = \frac{q}{(p-1)}\sum_{i=1}^{p}\alpha_i^2 + \sigma_e^2$$

Under H_{0A}, MSSA provides an unbiased estimate of σ_e^2.

7.7.2 Expectation of Sum of Squares due to Factor B

Similarly, the E (SSB) can be obtained

$$SSB = p\sum_{j=1}^{q}(\bar{y}_{.j} - \bar{y}_{..})^2$$

$$E(SSB) = E\left[p\sum_{j=1}^{q}(\bar{y}_{.j} - \bar{y}_{..})^2\right]$$

$$= p\,E\left[\sum_{j=1}^{q}\left(\bar{y}_{.j}-\bar{y}_{..}\right)^2\right] \quad \ldots (3)$$

We know that $y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$

So $\bar{y}_{.j} = \frac{1}{p}\sum_{i=1}^{p} y_{ij} = \frac{1}{p}\left[\sum_{i=1}^{p}\left(\mu + \alpha_i + \beta_j + e_{ij}\right)\right]$

or $\bar{y}_{.j} = \frac{1}{p}\left[p\mu + \sum_{i=1}^{p}\alpha_i + p\beta_j + \sum_{i=1}^{p} e_{ij}\right]$

$$= \mu + \beta_j + \bar{e}_{.j}$$

Similarly, the value of $\bar{y}_{..}$ will be

$$\bar{y}_{..} = \mu + \bar{e}_{..}$$

Substituting the values of $\bar{y}_{.j}$ and $\bar{y}_{..}$ in equation (3), we have

$$E\,(SSB) = p\,E\left[\sum_{j=1}^{q}\left(\mu + \beta_j + \bar{e}_{.j} - \mu - \bar{e}_{..}\right)^2\right]$$

$$= p\,E\left[\sum_{j=1}^{q}\left(\beta_j + \bar{e}_{.j} - \bar{e}_{..}\right)^2\right]$$

$$= p\,E\left[\sum_{j=1}^{q}\beta_j^2 + \sum_{j=1}^{q}\left(\bar{e}_{.j} - \bar{e}_{..}\right)^2 + 2\sum_{j=1}^{q}\beta_j\left(\bar{e}_{.j} - \bar{e}_{..}\right)\right]$$

$$= p\left[\sum_{j=1}^{q}\beta_j^2 + E\sum_{j=1}^{q}\left(\bar{e}_{.j} - \bar{e}_{..}\right)^2 + 2E\sum_{j=1}^{q}\beta_j\left(\bar{e}_{.j} - \bar{e}_{..}\right)\right]$$

Since, $e_{ij} \sim \text{iid } N(0, \sigma_e^2)$, so $E\left(e_{ij}\right) = 0 \quad Var\left(e_{ij}\right) = \sigma_e^2$

$$\bar{e}_{.j} \sim \text{iid } N\left(0, \frac{\sigma_e^2}{p}\right) \quad E\left(\bar{e}_{.j}\right) = 0 \quad Var\left(\bar{e}_{.j}\right) = \frac{\sigma_e^2}{p}$$

and $\bar{e}_{..} \sim \text{iid } N\left(0, \frac{\sigma_e^2}{pq}\right) \quad E\left(\bar{e}_{..}\right) = 0 \quad Var\left(\bar{e}_{..}\right) = \frac{\sigma_e^2}{pq}$

Because $E\left(\bar{e}_{.j}\right) = E\left(\bar{e}_{..}\right) = 0$

Therefore, $E(SSB) = p\sum_{j=1}^{q}\beta_j^2 + pE\sum_{j=1}^{q}\left(\bar{e}_{.j}^2 + \bar{e}_{..}^2 - 2\bar{e}_{.j}\bar{e}_{..}\right)$

$$= p\sum_{j=1}^{q}\beta_j^2 + pE\left[\sum_{j=1}^{q}\bar{e}_{.j}^2 + q\bar{e}_{..}^2 - 2\bar{e}_{..}\sum_{j=1}^{q}\bar{e}_{.j}\right]$$

$$= p\sum_{j=1}^{q}\beta_j^2 + pE\left[\sum_{j=1}^{q}\bar{e}_{.j}^2 - q\bar{e}_{..}^2\right] \quad \text{because } \bar{e}_{..} = \frac{1}{q}\sum_{j=1}^{q}\bar{e}_{.j}$$

$$= p\sum_{j=1}^{q}\beta_j^2 + p\left[E\left(\sum_{j=1}^{q}\bar{e}_{.j}^2\right) - qE\left(\bar{e}_{..}^2\right)\right]$$

$$= p\sum_{j=1}^{q}\beta_j^2 + p\left(q\frac{\sigma_e^2}{p} - q\ \frac{\sigma_e^2}{pq}\right)$$

$$E(SSB) = p\sum_{j=1}^{q}\beta_j^2 + (q-1)\sigma_e^2$$

$$E\left[\frac{SSB}{q-1}\right] = \frac{p}{(q-1)}\ \sum_{j=1}^{q}\beta_j^2 + \sigma_e^2$$

Under H_{0B}, the MSSB is unbiased estimate of σ_e^2

7.7.3 Expectation of Sum of Squares Due to Error

$$E(SSE) = E\left(\sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right)^2\right)$$

$$= E\left(\sum_{i=1}^{p}\sum_{j=1}^{q}\left(\mu + \alpha_i + \beta_j + e_{ij} - \mu - \alpha_i - \bar{e}_{i.} - \mu - \beta_j - \bar{e}_{.j} + \mu + \bar{e}_{..}\right)^2\right)$$

$$= E\left(\sum_{i=1}^{p}\sum_{j=1}^{q}\left(e_{ij} - \bar{e}_{i.} - \bar{e}_{.j} + \bar{e}_{..}\right)^2\right)$$

$$= E\Bigg[\sum_{i=1}^{p}\sum_{j=1}^{q}e_{ij}^2 + \sum_{i=1}^{p}\sum_{j=1}^{q}\bar{e}_{i.}^2 + \sum_{i=1}^{p}\sum_{j=1}^{q}\bar{e}_{.j}^2 + \sum_{i=1}^{p}\sum_{j=1}^{q}\bar{e}_{..}^2$$

$$- 2\sum_{i=1}^{p}\sum_{j=1}^{q}e_{ij}\bar{e}_{i.} - 2\sum_{i=1}^{p}\sum_{j=1}^{q}e_{ij}\bar{e}_{.j} + 2\sum_{i=1}^{p}\sum_{j=1}^{q}e_{ij}\bar{e}_{..}$$

$$+ 2\sum_{i=1}^{p}\sum_{j=1}^{q}\bar{e}_{i.}\bar{e}_{.j} - 2\sum_{i=1}^{p}\sum_{j=1}^{q}\bar{e}_{i.}\bar{e}_{..} - 2\sum_{i=1}^{p}\sum_{j=1}^{q}\bar{e}_{.j}\bar{e}_{..}\Bigg]$$

$$= E\Bigg[\sum_{i=1}^{p}\sum_{j=1}^{q}e_{ij}^2 + q\sum_{i=1}^{p}\bar{e}_{i.}^2 + p\sum_{j=1}^{q}\bar{e}_{.j}^2 + pq\bar{e}_{..}^2 - 2q\sum_{i=1}^{p}\bar{e}_{i.}^2$$

$$- 2p\sum_{j=1}^{q}\bar{e}_{.j}^2 + 2pq\bar{e}_{..}^2 + 2pq\bar{e}_{..}^2 - 2pq\bar{e}_{..}^2 - 2pq\bar{e}_{..}^2\Bigg]$$

$$= E\left[\sum_{i=1}^{p}\sum_{j=1}^{q}e_{ij}^2 - q\sum_{i=1}^{p}\bar{e}_{i.}^2 - p\sum_{j=1}^{q}\bar{e}_{.j}^2 + pq\bar{e}_{..}^2\right]$$

$$E(SSE) = \sum_{i=1}^{p}\sum_{j=1}^{q}E(e_{ij}^2) - q\sum_{i=1}^{p}E(\bar{e}_{i.}^2) - p\sum_{j=1}^{q}E(\bar{e}_{.j}^2) + pqE(\bar{e}_{..}^2)$$

Since $e_{ij} \sim$ iid $N(0,\sigma_e^2)$, so $E(e_{ij}) = 0 \quad Var(e_{ij}) = \sigma_e^2$

so $\quad E(SSE) = pq\,\sigma_e^2 - qp.\frac{\sigma_e^2}{q} - pq\frac{\sigma_e^2}{p} + pq.\frac{\sigma_e^2}{pq}$

$$= pq\sigma_e^2 - p\ \sigma_e^2 - q\ \sigma_e^2 + \sigma_e^2$$

$$= (pq - p - q + 1)\ \sigma_e^2$$

$$E\ (SSE) = (p-1)\ (q-1)\ \sigma_e^2$$

$$E\left(\frac{SSE}{(p-1)(q-1)}\right) = \sigma_e^2$$

or $\quad E\,(MSSE) = \sigma_e^2$

Hence, mean sum of squares due to error is unbiased estimate of σ_e^2.

Example 1: Future group wishes to enter the frozen shrimp market. They contract a researcher to investigate various methods of groups of shrimp in large tanks. The researcher suspects that temperature and salinity are important factor influencing shrimp yield and conducts a two-way analysis of variance with their levels of temperature and salinity. That is each combination of yield for each (for identical gallon tanks) is measured.The recorded yields are given in the following chart:

Categorical variable Salinity (in pp)

Temperature	**700**	**1400**	**2100**	**Total**	**Mean**
60^0 F	3	5	4	12	4
70^0 F	11	10	12	33	11
80^0 F	16	21	17	54	18
Total	30	36	33	99	11

Compute the ANOVA table for the model.

Solution: Since in each call there is one observation. So we will use the model

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

where, y_{ij} is the yield corresponding to i^{th} temperature and j^{th} salinity, μ is the general mean, α_i is the effect due to i^{th} temperature, β_j is the effect due to j^{th} salinity and $e_{ij} \sim$ i i d N $(0, \sigma^2)$

Test of hypothesis that there is a significant difference in shrimp yield due to difference in the temprature or

$$H_{0A}: \alpha_1 = \alpha_2 = \alpha_3 \quad \text{or} \quad H_{0A}: \alpha_i = 0 \quad \text{for all } i = 1, 2, 3$$

$$H_{1A}: \alpha_1 \neq \alpha_2 \neq \alpha_3 \neq 0 \text{ or } \quad H_{1A}: \alpha_i \neq 0 \quad \text{at least one } i$$

Similarly, there is a significant difference in shrimp yield due to difference in the salinity or

$$H_{0B}: \beta_1 = \beta_2 = \beta_3 \quad \text{or} \quad H_{0B}: \beta_j = 0 \quad \text{for all } j = 1, 2, 3$$

$$H_{1B}: \beta_1 \neq \beta_2 \neq \beta_3 \neq 0 \quad \text{or} \quad H_{1B}: \beta_j \neq 0 \text{ at least one } j$$

The computations are as follows:

Grand Total = G = 99

No. of observations = N = 9

Correction Factor (CF) = $(99 \times 99) / 9 = 1089$

Raw Sum of Squares (RSS) = 1401

Total Sum of Squares (TSS) = RSS – CF = 1401 - 1089 = 312

Sum of Squares due to Temperature (SST)

$= (12)^2/3 + (33)^2/3 + (54)^2/3 - 1089$

$= (144 + 1089 + 2916)/3 - 1089$

$= (4149/3) - 1089$

$= 1383 - 1089 = 294$

Sum of Squares due to Salinity (SSS)

$= (30)^2/3 + (36)^2/3 + (33)^2/3 - 1089$

$= (900 + 1296 + 1089)/3 - 1089$

$= (3285/3) - 1089$

$= 1095 - 1089 = 6$

Sum of Squares due to Error = TSS – SST - SSS

$= 312 - 294 - 6 = 12$

ANOVA Table for Two-Way Classified Data

Sources of Variation	DF	SS	MSS	F-Test
Due to Temperature	2	294	147	F_T = MSST/MSSE =147/3 = 49
Due to Salinity	2	6	3	F_S = 3/3 = 1
Due to error	4	12	3	
Total	8	312		

Since tabulated value of $F_{2,4}$ at 5% level of significance is 10.649 which is less than the calculated $F_{T(2,4)} = 49$, for testing the significant difference in shrimp yield due to differences in levels of temperature. So, H_{0A} is rejected. Hence, there are differences in shrimp yield due to temperature at 5% level of significance.

Since tabulated value of $F_{2,4}$ at 5% level of significance is 10.649 which is greater than the calculated $F_{S(2,4)} = 1$, for testing the significant difference in shrimp yield due to difference in the level of salinity. So, H_{0B} is accepted. Hence there is no any difference in shrimp yield due the salinity level at 5% level of significance.

In summary we conclude that $\beta_j = 0$ for j =1, 2, 3 and that $\alpha_i \neq 0$ for at each one value of i. So this model implies that temperature is a more important factor than salanity in influencing shrimp yield. Difference in the level of salanity appears to have no effect on shrimp at all.

$H_{0A}: \alpha_i \neq 0$ i= 1, 2, 3, then we have to test pairwise. So, multiple comparison tests will be applied for different levels of temperature so the null hypothesis will be

$H_{0A}: \alpha_i = \alpha_j \quad i \neq j = 1, 2, 3$

$H_{1A}: \alpha_i \neq \alpha_j$

So, if $(\bar{y}_{1.} - \bar{y}_{2.}) > t_{(p-1)(q-1)}\sqrt{\frac{2MSSE}{3}}$

Then, H_{0A} is rejected

So $(\bar{y}_{1.} - \bar{y}_{2.}) = 7$, $(\bar{y}_{1.} - \bar{y}_{3.}) = 14$, $(\bar{y}_{2.} - \bar{y}_{3.}) = 7$

and the value $t_{(p-1)(q-1)}\sqrt{\frac{2\,MSSE}{3}} = 2.1776\times\sqrt{\frac{2\times 3}{3}} = 3.93$

Since, all $(\bar{y}_{i.} - \bar{y}_{j.}) > 3.93$. So, we conclude that there is a significance difference among the yield of shrimp due to different levels of temperature.

E 1) An experiment was conducted to determine the effect of different data of planting and different methods of planting on the field of sugar-cane. The data below show the fields of sugar-cane for four different data and the methods of planting:

Data of Planting

Method of Planting	October	November	February	March
I	7.10	3.69	4.70	1.90
II	10.29	4.79	4.50	2.64
III	8.30	3.58	4.90	1.80

Carry out an analysis of the above data.

E 2) A researcher wants to test four diets A, B, C, D on growth rate in mice. These animals are divided into 3 groups according to their weights. Heaviest 4, next 4 and lightest 4 are put in Block I, Block II and Block III respectively. Within each block, one of the diets is given at random to the animals. After 15 days, increase in weight is noted, which given in the following table:

Blocks	Treatments/Diets			
	A	B	C	D
I	12	8	6	5
II	15	12	9	6
III	14	10	8	5

Perform a two-way ANOVA to test whether the data indicate any significant difference between the four diets due to different blocks.

7.8 SUMMARY

In this unit, we have discussed:

1. The two-way analysis of variance model;
2. The basic assumptions in two-way analysis of variance;
3. Test of hypothesis using two-way analysis of variance; and
4. Expectations of various sum of squares.

7.9 SOLUTIONS /ANSWERS

E1) H_{0A} : There is no difference among the different method of planting.

$$H_{0A}: \alpha_i = \alpha_2 = \alpha_3$$

Against, $H_{1A}: \alpha_i \neq \alpha_2 \neq \alpha_3$

H_{0B} : There is no any difference among the different data of planting.

$$H_{0B}: \beta_i = \beta_2 = \beta_3 = \beta_4$$

$$H_{1B}: \beta_i \neq \beta_2 \neq \beta_3 \neq \beta_4$$

G = $\sum\sum y_{ij}$ = Grand Total = Total of all observations

= 7.10 + 3.69 + 4.70 + 1.90 + 10.29 + 4.79 + 4.58 + 2.64 + 8.30 + 3.58 + 4.90 + 1.80

= 58.28

N = No. of observations = 12

Correction Factor (CF) $= \dfrac{G^2}{N} = \dfrac{58.28 \times 58.28}{12} = 283.0465$

Raw Sum of Squares (RSS)

$$= (7.10)^2 + (3.69)^2 + (4.70)^2 + (1.90)^2 + (10.29)^2 + (4.79)^2 + (4.58)^2 + (2.64)^2 + (8.30)^2 + (3.58)^2 + (4.90)^2 + (1.80)^2$$

$= 355.5096$

Total Sum of Squares (TSS) = RSS − CF

= 355.5096 - 283.0465 = 72.4631

Sum of Squares due to Data of Planting (SSD)

$$= \frac{D_1^2}{3} + \frac{D_2^2}{3} + \frac{D_3^2}{3} + \frac{D_4^2}{3} - CF$$

$$= \frac{(25.69)^2}{3} + \frac{(12.06)^2}{3} + \frac{(14.18)^2}{3} + \frac{(6.35)^2}{3} - 283.0465$$

= 65.8917

Sum of Squares due to Method of Planting (SSM)

$$= \frac{M_1^2}{4} + \frac{M_2^2}{4} + \frac{M_3^2}{4} - CF$$

$$= \frac{(17.39)^2}{4} + \frac{(22.31)^2}{4} + \frac{(15.58)^2}{4} - 283.0465$$

= 286.3412 - 283.0465 = 3.2947

Sum of Squares due to Error (SSE) = TSS - SSD - SSM

= 72.4631 - 3.2947 - 65.8917

= 3.2767

ANOVA Table

Source of Variation	SS	DF	MSS	F
Between Method	3.2947	2	$\frac{3.2947}{2} = 1.6473$	$F_1 = \frac{1.6473}{0.5461} = 3.02$
Between Data	65.8917	3	$\frac{65.8917}{3} = 21.9639$	$F_2 = \frac{21.9639}{0.5461} = 40.22$
Due to Errors	3.2767	6	$\frac{3.2767}{6} = 0.5461$	
Total	72.4631	11		

The tabulated value of $F_{2,6}$ at 5% level of significance is 5.14 which is greater than the calculated value of F_M (3.02) so H_{0A} is accepted. So, we conclude that there is no significant difference among the different methods of planting.

The tabulated value of $F_{3,6}$ at 5% level of significance is 4.76 which is less than calculated value of F_D (40.22). So we reject the null hypothesis H_{0B} . Hence there is a significant difference among the data of planting. In all, we conclude that the different methods of planting affect the mean field of sugar-cane in the same manner. But the mean field differs with different data of planting.

If the four data of planting, included in the above experiment, be the only data in which we are interested. Then the next question that arises is: which one of the four data will give us the maximum mean field? To answer this question, we complete the critical difference (CD) at, say, the 5% level of significance,

$$CD = t_{\text{error df}} \sqrt{\frac{2MSSE}{3}} = 2.447 \times \sqrt{0.3641} \quad = 1.48$$

The mean field differences of the four data of planting are given as:

$$\left|\overline{D}_1 - \overline{D}_2\right| = \left|8.56 - 4.02\right| = 4.54$$

$$\left|\overline{D}_1 - \overline{D}_3\right| = \left|8.56 - 4.73\right| = 3.77$$

$$\left|\overline{D}_1 - \overline{D}_4\right| = \left|8.56 - 2.12\right| = 6.44$$

$$\left|\overline{D}_2 - \overline{D}_3\right| = \left|4.02 - 4.73\right| = 0.71$$

$$\left|\overline{D}_2 - \overline{D}_4\right| = \left|4.02 - 2.12\right| = 1.90$$

$$\left|\overline{D}_3 - \overline{D}_4\right| = \left|4.73 - 2.12\right| = 2.61$$

From above we conclude that there is no any significant difference between the data of planting November and February. But there is a significant difference between October, November, December and February. But the mean of October is the maximum. So, we can release the data of planting in the month of October.

E2) Null hypotheses are

H_{01}: There is no significant difference between mean effects of diets.

H_{02}: There is no significant difference between mean effects of different blocks.

Against the alternative hypothesis

H_{11}: There is significant difference between mean effects of diets

H_{12}: There is significant difference between mean effects of different blocks.

Blocks	**Treatments/Diets**				
	A	B	C	D	**Totals**
I	12	8	6	5	31 $T_{.1}$
II	15	12	9	6	42 $T_{.2}$
III	14	10	8	5	37 $T_{.3}$
Totals	41 $T_{1.}$	30 $T_{2.}$	23 $T_{3.}$	16 $T_{4.}$	110 Grand Total

Squares of observations

Blocks	**Treatments/Diets**				**Totals**
	A	B	C	D	
I	144	64	36	25	269
II	225	144	81	36	486
III	196	100	64	25	385
Totals	565	308	181	86	1140

Grand Total = G = $\sum\sum y_{ij}$ = 110

Correction Factor (CF) = $\frac{G^2}{n} = \frac{(110)^2}{12} = 1008.3333$

Raw Sum of Squares (RSS) = $\sum\sum y_{ij}{}^2 = 1140$

Total Sum Squares (TSS) = RSS - CF = 1140 - 1008.3333

= 131.6667

Sum of Squares due to Treatments/ Diets (SST)

$$= \frac{T_{1.}^2}{3} + \frac{T_{2.}^2}{3} + \frac{T_{3.}^2}{3} + \frac{T_{4.}^2}{3} - CF$$

$$= \frac{1}{3}\left[(41)^2 + (30)^2 + (23)^2 + (16)^2\right] - 1008.3333$$

$$= \frac{1}{3}\left[1681 + 900 + 529 + 256\right] - 1008.3333$$

= 1122 - 1008.3333

= 113.6667

Sum of Squares due to Block (SSB) = $\frac{1}{4}\left(T_{.1}^2 + T_{.2}^2 + T_{.3}^2\right) - CF$

$$= \frac{1}{4}\left[259 + 486 + 385\right] - 1008.3333$$

= 1023.5 – 1008.3333 = 15.1667

Sum of Squares due to Errors (SSE) = TSS – SST – SSB

$= 131.6667 - 113.6667 - 15.1667$

$= 2.8333$

Mean Sum of Squares due to Treatments (MSST)

$$= \frac{SST}{df} = \frac{113.6667}{3} = 37.8889$$

Mean Sum of Squares due to Blocks (MSSB)

$$= \frac{SSB}{df} = \frac{15.1667}{2} = 7.58335$$

Mean Sum of Squares due to Errors (MSSE)

$$= \frac{SSE}{df} = \frac{2.8333}{6} = 0.4722$$

ANOVA Table

Source of Variation	**SS**	**DF**	**MSS**	**F**
Between Treatments/ Diets	113.6667	3	$\frac{113.6667}{3} = 37.8889$	$F_1 = \frac{37.8889}{0.4722} = 80.2391$
Between Blocks	15.1667	2	$\frac{15.1667}{2} = 7.58335$	$F_2 = \frac{7.58335}{0.4722} = 160.596$
Due to Errors	2.8333	6	$\frac{2.8333}{6} = 0.4722$	
Total	131.6667	11		

Tabulated F at 5% level of significance with (3, 6) degrees of freedom is 4.76 & tabulated F at 5% level of significance with (2, 6) degrees of freedom is 5.14

Conclusion: Since calculated F_1 > tabulated F, so we reject H_{01} and conclude that there is significant difference between mean effect of diets.

Also calculated F_2 is greater than tabulated F, so we reject H_{02} and conclude that there is significant difference between mean effect of different blocks.