## 2.5 Three factor classification (or) Latin Square Design (LSD)

We have seen data from a Latin square experiment results in a three way classification say (i) variety of seeds (ii) types of spacing (or plots) and (iii) different manure treatment.

The columns stand for different varities ($x_i$). The rows stand for different types of spaing ($y_i$) and the letters stand for different manurial treatments.

Let $x_{ij}$ be the value corresponding to the $i^{th}$ row, $j^{th}$ column and $k^{th}$ letter. Total number of values $\{x_{ij}\}$ is $N\,(= n^2)$ for $n \times$ LSD.

We have to test the null hypothesis.

$H_0$ : There is no difference between the rows, between the columns and between the letters.

Alternate hypothesis : $H_1$ : Not all are the same.

$T$ = Total of all $x_{ij}$ values, $N = n^2$

$$SST = \sum_{i=1}^{n} x_i^2 - \frac{T^2}{N}$$

$$SSC = \frac{1}{n}\left[(\Sigma x_1)^2 + (\Sigma x_2)^2 + \dots + (\Sigma x_n)^2\right] - \frac{T^2}{N}$$

$$SSR = \frac{1}{n}\left[(\Sigma y_1)^2 + (\Sigma y_2)^2 + \dots + (\Sigma_n)^2\right] - \frac{T^2}{N} \quad \text{and} \quad SSK = \frac{1}{n}\sum_{i=1}^{n} L_i^2 - \frac{T^2}{N},$$

where $L_i$ is the sum of all $x_{ij}$ receiving the $i^{th}$ treatment (or $i^{th}$ letter)

$SSE = SST - SSC - SSR - SSK$

We now from the ANOVA table as below

| Source of variation | Sum of squares | d.f | Mean squares | Variance ratio F |
|---|---|---|---|---|
| Between columns | SSC | $n-1$ | $MSC = \frac{SSC}{n-1}$ | $F_C = \frac{MSC}{MSE}$ |
| Between rows | SSR | $n-1$ | $MSR = \frac{SSR}{n-1}$ | $F_R = \frac{MSR}{MSE}$ |
| Between letters (treatments) | SSK | $n-1$ | $MSK = \frac{SSK}{n-1}$ | $F_K = \frac{MSK}{MSE}$ |
| Residual (error) | SSE | $(n-1)(n-2)$ | $MSE = \frac{SSE}{(n-1)(n-2)}$ | |
| | SST | $n^2-1$ | | |

$F_C$, $F_R$, $F_K$ values should be $> 1$

**Inference :** If the calculated value of $F <$ the table value of $F$, then $H_0$ is accepted.

If the calculated value of $F >$ the table value of $F$, then $H_0$ is rejected and we accept $H_1$.

## WORKED EXAMPLES

**Example 1 : An agricultural experiment on the latin square design gave the following results for the yield of wheat per acre, the letters corresponding to varieties, columns to treatments and rows to blocks. Discuss the variation of yield with each of these factors**

| | | | | |
|---|---|---|---|---|
| A 16 | B 10 | C 11 | D 9 | E 9 |
| E 10 | C 9 | A 14 | B 12 | D 11 |
| B 15 | D 8 | E 8 | C 10 | A 18 |
| D 12 | E 6 | B 13 | A 13 | C 12 |
| C 13 | A 11 | D 10 | E 7 | B 14 |

**Solution :** Given a Latin Square Design.

So, we use three-way classification of analysis of variance. So we assume

Null hypothesis : $H_0$ : there is no difference between rows, between columns and between varieties.

Alternative hypothesis : $H_1$ : there is difference in atleast one of these.

We shall form the coded data by subtracting 10 from each value.

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | Total | $x_1^2$ | $x_2^2$ | $x_3^2$ | $x_4^2$ | $x_5^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $y_1$ | A 6 | B 0 | C 1 | D $-1$ | E $-1$ | 5 | 36 | 0 | 1 | 1 | 1 |
| $y_2$ | E 0 | C $-1$ | A 4 | B 2 | D 1 | 6 | 0 | 1 | 16 | 4 | 1 |
| $y_3$ | B 5 | D $-2$ | E $-2$ | C 0 | A 8 | 9 | 25 | 4 | 4 | 0 | 64 |
| $y_4$ | D 2 | E $-4$ | B 3 | A 3 | C 2 | 6 | 4 | 16 | 9 | 9 | 4 |
| $y_5$ | C 3 | A 1 | D 0 | E $-3$ | B 4 | 5 | 9 | 1 | 0 | 9 | 16 |
| Total | 16 | $-6$ | 6 | 1 | 14 | $\boxed{31}$ | 74 | 22 | 30 | 23 | 86 |

$N = 5 \times 5 = 25 ; \qquad T = 31$

$$\frac{T^2}{N} = \frac{31^2}{25} = 38.44$$

$$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 + \Sigma x_4^2 + \Sigma x_5^2 - \frac{T^2}{N}$$

$$= 74 + 22 + 30 + 23 + 86 - 38.44 = 235 - 38.44 = 196.56$$

$$SSC = \frac{(\Sigma x_1)^2}{5} + \frac{(\Sigma x_2)^2}{5} + \frac{(\Sigma x_3)^2}{5} + \frac{(\Sigma x_4)^2}{5} + \frac{(\Sigma x_5)^2}{5} - \frac{T^2}{N}$$

$$= \frac{1}{5}(16^2 + 6^2 + 6^2 + 1^2 + 14^2) - 38.44$$

$$= \frac{525}{5} - 38.44 = 105 - 38.44 = 66.56$$

$$SSR = \frac{(\Sigma y_1)^2}{5} + \frac{(\Sigma y_2)^2}{5} + \frac{(\Sigma y_3)^2}{5} + \frac{(\Sigma y_4)^2}{5} + \frac{(\Sigma y_5)^2}{5} - \frac{T^2}{N}$$

$$= \frac{1}{5}(5^2 + 6^2 + 9^2 + 6^2 + 5^2) - 38.44$$

$$= \frac{203}{5} - 38.44 = 40.6 - 38.44 = 2.16$$

**To find SSK :**

We arrange the data according to the letters row wise

| | A | B | C | D | E |
|---|---|---|---|---|---|
| | 6 | 0 | 1 | −1 | −1 |
| | 4 | 2 | −1 | 1 | 0 |
| | 8 | 5 | 0 | −2 | −2 |
| | 3 | 3 | 2 | 2 | −4 |
| | 1 | 4 | 3 | 0 | −3 |
| **Total** | 22 | 14 | 5 | 0 | −10 |

$$SSK = \frac{22^2}{5} + \frac{14^2}{5} + \frac{5^2}{5} + 0 + \frac{(-10)^2}{5} - \frac{T^2}{N}$$

$$= \frac{805}{5} - 38.44 = 161 - 38.44 = 122.56$$

$$\therefore SSE = SST - SSC - SSR - SSK$$

$$= 196.56 - 66.56 - 2.16 - 122.56 = 5.28$$

We form the ANOVA Table

| **Source of variation** | **Sum of squares SS** | **d.f** | **Mean square MS** | **Variation F** |
|---|---|---|---|---|
| Between columns (i.e. treatments) | SSC = 66.56 | 4 | $MSC = \frac{66.56}{4} = 16.64$ | $F_C = \frac{MSC}{MSE} = \frac{16.64}{0.44} = 37.82$ |
| Between rows (i.e. blocks) | SSR = 2.16 | 4 | $MSR = \frac{2.16}{4} = 0.54$ | $F_R = \frac{MSR}{MSE} = \frac{0.54}{0.44} = 1.23$ |
| Between letters (varieties) | SSK = 122.56 | 4 | $MSK = \frac{122.56}{4} = 30.64$ | $F_K = \frac{MSK}{MSE} = \frac{30.64}{0.44} = 69.64$ |
| Residual (Error) | SSE = 5.28 | 12 | $MSE = \frac{5.28}{12} = 0.44$ | |
| Total | SST = 196.56 | 24 | | |

**Between columns :** $F_C = 38.82$

The table value of F (4, 12) at 5% level = 3.26

$\therefore$ the calculated value of $F_C >$ the table value of $F_C$

$\therefore$ the difference between treatments are significant.

so, the different treatments have effect on yield

**Between rows :** $F_R = 1.23$

The table value of F (4, 12) at 5% level = 3.26

$\therefore$ the calculated value of $F_R$ < the table value of $F_R$

$\therefore$ there is no difference between blocks or rows

**Between letters (or varieties) :**

$$F_K = 69.64$$

The table value of F (4, 12) at 5% level = 3.26

$\therefore$ the calculated value of $F_K$ > the table value of $F_K$

Hence the difference between letters or varieties is highly significant.

**Example 2 : A farmer wishes to test the effects of four different fertilizers A, B, C, D on the yield of wheat. In order to eliminate sources of error due to variability in soil fertility, he uses the fertilizers in a Latin square arrangement as indicated with following table, where the numbers indicate yields in bushels per unit area.**

| | | | |
|---|---|---|---|
| A 18 | C 21 | D 25 | B 11 |
| D 22 | B 12 | A 15 | C 19 |
| B 15 | A 20 | C 23 | D 24 |
| C 22 | D 21 | B 10 | A 17 |

**perform an analysis of variance to determine if there is a significant difference between the fertilizers at $\alpha = 0.05$ levels of significance.** **[A.U-2006, 2007, 2009]**

**Solution :**

Given a Latin Square Design.

So, we use three-way classification of analysis of variance.

Null hypothesis : $H_0$ : there is no difference between rows between columns and between treatments.

Alternative hypothesis : $H_1$ : not all equal.

We shall code the data by subtracting 15 from each value.

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | Total | $x_1^2$ | $x_2^2$ | $x_3^2$ | $x_4^2$ |
|---|---|---|---|---|---|---|---|---|---|
| $y_1$ | A 3 | C 6 | D 10 | B −4 | 15 | 9 | 36 | 100 | 16 |
| $y_2$ | D 7 | B −3 | A 0 | C 4 | 8 | 49 | 9 | 0 | 16 |
| $y_3$ | B 0 | A 5 | C 8 | D 9 | 22 | 0 | 25 | 64 | 81 |
| $y_4$ | C 7 | D 6 | B −5 | A 2 | 10 | 49 | 36 | 25 | 4 |
| Total | 17 | 14 | 13 | 11 | **55** | 107 | 106 | 189 | 117 |

$$N = 16,\ T = 55,\quad \frac{T^2}{N} = \frac{55^2}{16} = 189.06$$

$$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 + \Sigma x_4^2 - \frac{T^2}{N}$$

$$= 107 + 106 + 189 + 117 - 189.06 = 519 - 189.06 = 329.94$$

$$SSC = \frac{(\Sigma x_1)^2}{4} + \frac{(\Sigma x_2)^2}{4} + \frac{(\Sigma x_3)^2}{4} + \frac{(\Sigma x_4)^2}{4} - \frac{T^2}{N}$$

$$= \frac{17^2}{4} + \frac{14^2}{4} + \frac{13^2}{4} + \frac{11^2}{4} - 189.06 = \frac{1}{4}(775) - 189.06 = 193.75 - 189.06 = 4.69$$

$$SSR = \frac{(\Sigma y_1)^2}{4} + \frac{(\Sigma y_2)^2}{4} + \frac{(\Sigma y_3)^2}{4} + \frac{(\Sigma y_4)^2}{4} - \frac{T^2}{N}$$

$$= \frac{15^2}{4} + \frac{8^2}{4} + \frac{22^2}{4} + \frac{10^2}{4} - 189.06 = \frac{873}{4} - 189.06 = 218.25 - 189.06 = 29.19$$

To find SSK, we arrange the data according to the letters rowwise

| | A | B | C | D |
|---|---|---|---|---|
| | 3 | −4 | 6 | 10 |
| | 0 | −3 | 4 | 7 |
| | 5 | 0 | 8 | 9 |
| | 2 | −5 | 7 | 6 |
| Total | 10 | −12 | 25 | 32 |

$$SSK = \frac{10^2}{4} + \frac{(-12)^2}{4} + \frac{25^2}{4} + \frac{32^2}{4} - \frac{T^2}{N}$$

$$= \frac{1893}{4} - 189.06 = 473.25 - 189.06 = 284.19$$

$$\therefore\ SSE = SST - SSC - SSR - SSK$$

$$= 329.94 - 4.69 - 29.19 - 284.19 = 329.94 - 318.07 = 11.87$$

**ANOVA Table**

| Source of variation | Sum of squares | d.f | Mean square | Variance ratio F |
|---|---|---|---|---|
| Between columns | SSC = 4.69 | 3 | $MSC = \frac{SSC}{3} = \frac{4.69}{3} = 1.56$ | $F_C = \frac{1.978}{1.56} = 1.27$ |
| Between rows | SSR = 29.19 | 3 | $MSR = \frac{29.19}{3} = 9.73$ | $F_R = \frac{9.73}{1.978} = 4.92$ |
| Between Treatments | SSK = 284.19 | 3 | $MSK = \frac{284.19}{3} = 94.73$ | $F_K = \frac{94.73}{1.978} = 47.89$ |
| Residual (error) | SSE = 11.87 | 6 | $MSE = \frac{11.87}{6} = 1.978$ | |
| Total | SST = 329.94 | 15 | | |

**Between columns :** $F_C = 1.27$

The table value for F (6, 3) at 5% level = 8.94

$\therefore$ calculated value of F < the table value of F

So, there is no significant difference between columns so far as fertility is concerned.

**Between rows :** $F_R = 4.92$

The table value of F (3, 6) at 5% level = 4.76

$\therefore$ calculated value of F > the table value of F at 5% level

$\therefore$ there is significant difference in fertility from row to row.

**Between treatments :** $F_K = 47.89$

The table value of F (3, 6) at 5% level = 4.76

$\therefore$ the calculated value of F > the table value of F.

Hence there is significant difference between the fertilizers.

**Example 3 : Analyse the variance in the latin square of yields in (kgs) of paddy where P, Q, R, S denote the different methods of cultivation.**

| | | | |
|---|---|---|---|
| S 122 | P 121 | R 123 | Q 122 |
| Q 124 | R 123 | P 122 | S 125 |
| P 120 | Q 119 | S 120 | R 121 |
| R 122 | S 123 | Q 121 | P 122 |

**Examine whether the different methods of cultivation have given significantly different yields.**

**[A.U-2006, 2009, 2014, 2015]**

**Solution :** Given a Latin Square Design.

So, we use three-way classification of analysis of variance.

Null hypothesis : $H_0$ : There is no significant difference between rows, between columns and between the methods of cultivation.

Alternative hypothesis : $H_1$ : Not all equal.