

Unit-5

DESIGN OF EXPERIMENT

Always remember test statistics part for each topic solving questions. It is the main important thing.

The main objective of design of experiment is either make a high degree of precision or minimizing the error. In design of experiment we are going to design experiment by using different principles of design of experiment which we will discuss later, let first we discuss some terms (or terminology) used in design of experiment.

Terminology in Experimental Design:-

- i) Experiment → Experiment is means of getting an answer to a question that the experimenter has in mind. Experiment can be divided into two categories; absolute and comparative. Absolute experiment consists of determining the absolute value of some characteristic such as finding correlation coefficient between two variable. While the comparative experiment consists of comparing different types of fertilizers, different varieties of crops etc.
- ii) Treatment → These are the inputs whose outcomes are to be estimated and compared. These contain different procedures under comparison in experiment. For example in agricultural experiment different types of fertilizers, varieties of crops, cultivation methods etc.
- iii) Experimental Unit → The smallest division of the experimental area in which the treatments are applied and the effects of treatments are measured is known as experimental unit. For example, in agricultural experiment plot are the experimental unit.
- iv) Yields (effects) → The outcome of the experiment due to the application of treatments in experimental units are called yields. In agricultural experiment production of crops on using different fertilizers are yields.

v) Blocks → The experimental field is divided into relatively homogenous subgroups which are uniform (similar) among themselves than the field as a whole are called blocks.

vi) Experimental error → The variation in experimental measurements made on different experimental units even when they get same treatment is called experimental error. It can be estimated by replication and can be controlled by use of local control.

④ Principles of design of experiment:-

According to RA Fisher there are three main principles of design of experiment;

i) Replication → The execution of an experiment more than once is called replication. The repetition of treatment results in more reliable estimate than that of possible with single observation. The main objective of replication is to minimize the error i.e., to maximize the accuracy.

ii) Randomization → In randomization the treatments are allocated to various plots in a random manner. This means each treatment will have equal chance of being assigned to an experimental unit. By the use of randomization it reduce the bias of applying a particular treatment to a particular unit.

iii) Local control → Local control is used to convert heterogeneous field into homogenous by applying local control along row-wise or along column-wise. It reduces the experimental error by making relatively heterogeneous experimental materials into relatively homogeneous block. So, the experimental material is divided into a number of blocks row-wise or column-wise.

1) Completely Randomised Design (CRD):

In completely randomised design the two principles are only used which are randomization and replication. In this design, treatments are assigned completely at random manner so each and every experimental unit has equal chance of receiving treatment and it is appropriate for the homogenous experimental material.

row-wise & column-wise
कोर्डिनेट रेपेटेसन की

Mathematical model:

$$Y_{ij} = \mu + T_i + e_{ij}, \text{ where, } i=1, 2, \dots, t \\ j=1, 2, \dots, r$$

Here,

Y_{ij} = j^{th} unit receiving i^{th} treatment.

T_i = i^{th} treatment effect.

e_{ij} = error due to chance

μ μ = Overall mean (OR General mean).

2) Problem to test:

H_0 of treatment

A, B, C are three treatments

$$\text{Null Hypothesis } (H_0) \rightarrow \mu_A = \mu_B = \mu_C$$

i.e., There is no significance difference between treatment means.

$$\text{Alternate Hypothesis } (H_1) \rightarrow \mu_A \neq \mu_B \neq \mu_C$$

i.e., There is significance difference between treatment means.

OR At least one treatment mean is different than any other.

3) Test Statistics (Calculated value):

$$F_{\text{cal}} = \frac{\text{MST} \text{ (i.e, Mean square due to treatment)}}{\text{MSE} \text{ (i.e, Mean square due to error)}}$$

$$\text{where, } \text{MST} = \frac{\text{SST}}{\text{d.f}} \text{ (i.e, Sum of square due to treatment)}$$

d.f (i.e, degree of freedom).

$$\text{MSE} = \frac{\text{SSE}}{\text{d.f}} \text{ (i.e, Sum of square due to error)}$$

d.f (i.e, degree of freedom).

asked frequently in VIVA
full form asked.

ANOVA table: (i.e, Analysis of variance table)

Source of variation	d.f.	Sum of Square	Mean sum of square	Fcal
Treatment	(t-1)	SST	$MST = \frac{SST}{t-1}$	
Error	t(r-1)	SSE	$MSE = \frac{SSE}{t(r-1)}$	$F_{cal} = \frac{MST}{MSE}$
Total	$\frac{rt-1}{OR N-1}$	TSS		

Here,

G_1 = Grand Total

$$\text{Correction Factor (C.F)} = \frac{G_1^2}{N}$$

$$\text{Total sum of Square (TSS)} = \sum_{i=1}^t \sum_{j=1}^r y_{ij}^2 - C.F$$

Now,

$$SSE = TSS - SST$$

iii) Critical value (tabulated value):-

The tabulated value of F at α level of significance with $\{t-1, t(r-1)\}$ d.f. is $F_{\alpha} \{t-1, t(r-1)\}$.

iv) Decision:- If $F_{tab} > F_{cal}$ then, H_0 is accepted otherwise rejected.

v) Conclusions- On the basis of decision we write conclusion as we did before in chapter 2.

Q1. Carry out ANOVA of following output of wheat per field obtained as a result of 3 varieties of wheat A, B, C.

A	10	B	5	A	20	C	15
B	6	A	15	C	11	B	10
C	22	B	12	C	18	A	16

(R.D related question since row-wise and column-wise both repeated)

Solution:

Given,

A	10	B	5	A	20	C	15
B	6	A	15	C	11	B	10
C	22	B	12	C	18	A	16

Now, By using deviation (i.e Subtracting 15 from each value).

साले value परे को
बना deviation
नगदी से हैं। तो
value साले पानी
के square हैं।
जिसे हैला बना
deviation जाता

A	-5	B	-10	A	5	C	0
B	-9	A	0	C	-4	B	-5
C	7	B	-3	C	3	A	1

→ we can randomly choose
any number. By the subtraction
of which we get small numbers.

Problem to test:

Null Hypothesis (H_0): $\mu_A = \mu_B = \mu_C$

i.e., there is no significance difference between average production
on three varieties of wheat.

Alternate Hypothesis (H_1): $\mu_A \neq \mu_B \neq \mu_C$

i.e., there is significance difference between three varieties of wheat.

OR At least one variety of wheat is different than other.

Test Statistics:

We know that,

$$F_{cal} = \frac{MST}{MSE}$$

where, $MST = \frac{SST}{d.f}$

$$MSE = \frac{SSE}{d.f}$$

ANOVA table

					T_i	
A	-5	0	5	1	1	ΣA
B	-9	-10	-3	-5	-27	
C	7	-4	3	0	6	
					$G_T = -20$	Grand total

Now, correction factor (C.F) = $\frac{G_T^2}{N} = \frac{(-20)^2}{12} = \frac{400}{12} = 33.33$

Total Sum of Square (TSS) = $\sum_{q=1}^t \sum_{j=1}^{r_j} y_{ij}^2 - C.F$

$$= (-5)^2 + 5^2 + 0^2 + 1^2 + (-10)^2 + (-9)^2 + (-5)^2 + (-3)^2 + 0^2 + (-4)^2 + 7^2 + (3)^2 - 33.33$$

$$= 306.67$$

Sum of Square of treatment (SST) = $\sum t_i^2 - C.F$

$$= \frac{1^2 + (-27)^2 + 6^2}{4} - 33.33$$

$$= 158.17$$

Now, SSE = TSS - SST

$$= 306.67 - 158.17$$

$$= 148.50$$

Now, ANOVA table is;

Source of variation	d.f	SS	MSS	F_{cal}
Treatment	$(t-1) = 3-1 = 2$	SST = 158.17	$MST = \frac{158.17}{2} = 79.085$	$F_{cal} = \frac{MST}{MSE} = \frac{79.085}{16.5} = 4.79$
Error	$t(r-1) = 3(4-1) = 9$ OR $11-2=9$	SSE = 148.50	$MSE = \frac{148.5}{9} = 16.5$	
Total	$(rt-1) = 12-1 = 11$	TSS = 306.67		

divided by d.f

Critical value: The tabulated value of F at 0.05 level of significance with (2, 9) d.f is $F_{0.05}(2, 9) = 4.26$ table value
page no
of 321
at (2, 9)

Decision: Since $F_{cal} = 4.79 > F_{tab} = 4.26$. So, H_0 is rejected.
i.e., H_1 is accepted.

Conclusion: Hence, there is significance difference between three varieties of wheat.

Q2 The yield of treatments in different plots are as shown below. Carry out analysis.

t_4	1401	t_3	2536	t_3	2459	t_1	2537	t_3	2827	t_3	2069
t_2	2211	t_1	1797	t_4	1170	t_4	1516	t_4	2104	t_3	2385
t_2	3366	t_1	2104	t_2	2591	t_3	2460	t_4	1077	t_2	2544

Solve (Practice Yourself)

Solving method is same only difference is, 4 blocks were in example 6 blocks in this. Total treatments N will be $6 \times 3 = 18$.

2) Randomised Block - Design (RBD):-

Randomised Block - Design is used for heterogenous experimental materials which is better than CRD. In RBD the treatments are allocated in random manner but randomization (repetition) is restricted that each treatment must occur once in each row or once in each column. Hence this design is row-wise or column-wise so it is based upon all the principles of design namely replication, randomization and local control.

Mathematical Model:

$$Y_{ij} = \mu + T_i + \beta_j + e_{ij}$$

This β_j or Block (B) is additional in this method

where, $i = 1, 2, \dots, t$.

$j = 1, 2, \dots, r$.

Here,

Y_{ij} = j th block receiving i th treatment.

μ = Overall mean

T_i = treatment effect

β_j = Effect due to j th block

e_{ij} = Error due to chance.

Problem to test:-

दृष्टिकोण hypothesis कुनैले यसमा
रख्ना, treatment को लागि व
अधिक Block को लागि,

(a) For treatment

Null Hypothesis (H_0T): $\mu_1 = \mu_2 = \dots = \mu_t$

i.e., There is no significance difference between the treatment means.

Alternate Hypothesis (H_{1T}): $\mu_1 \neq \mu_2 \neq \dots \neq \mu_t$

i.e., There is significance difference between the treatment means.

OR At least there is one treatment mean different than other.

(b) For Block

Null Hypothesis (H_{0B}): $\beta_1 = \beta_2 = \dots = \beta_r$

i.e., " " " (same as above) " "

Alternate Hypothesis (H_{1B}): $\beta_1 \neq \beta_2 \neq \dots \neq \beta_r$

i.e., " " " (same as above) " "

Test Statistics:

ANOVA table:

Source of variation	d.f.	S.S	MSS	F _{cal}
Treatment	t-1	SST	$MST = \frac{SST}{d.f}$	$F_T = \frac{MST}{MSE}$
Block	r-1	SSB	$MSB = \frac{SSB}{d.f}$	$F_B = \frac{MSB}{MSE}$
Error	$\frac{(t-1)(r-1)}{\text{Total-treatment-block}}$	SSE	$MSE = \frac{SSE}{d.f}$	
Total	$\frac{rt-1}{=N-1}$	TSS		

Here,

$$TSS = SST + SSB + SSE$$

where,

$$TSS = \sum_{i=1}^t \sum_{j=1}^r y_{ij}^2 - C.F.$$

$$\text{here, } C.F = \frac{G_1^2}{N}$$

$$SST = \frac{\sum T_{i.}^2}{r} - C.F$$

$$\text{if } SSB = \frac{\sum T_j^2}{t} - C.F$$

$$\text{So, } SSE = TSS - SSB - SST$$

We will understand better through following example questions.

Q1: Carry out ANOVA for following design.

A	8	C	10	A	6	B	10
C	12	B	8	B	9	A	8
B	10	A	8	C	10	C	9

Also calculate the relative efficiency of the design with respect to CRD.

Solution: By using deviation (OR Subtracting 8 from each value)

A	0	C	2	A	-2	B	2
C	4	B	0	B	1	A	0
B	2	A	0	C	2	C	1

Problem to test:

① For treatment:

Null Hypothesis (H_0T): $\mu_A = \mu_B = \mu_C$

i.e., There is no significance difference between treatment means.

Alternate Hypothesis (H_1T): $\mu_A \neq \mu_B \neq \mu_C$

i.e., There is significance difference between treatment means.

OR At least there is one treatment mean which is different than others.

② For Block:

Null Hypothesis (H_0B): $\mu_1 = \mu_2 = \mu_3 = \mu_4$

i.e., " (same as above) " --

Alternate Hypothesis (H_1B): $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$

i.e., " (same as above) " --

Test Statistics:

$$F_T = \frac{MST}{MSE}$$

$$\text{or } F_B = \frac{MSB}{MSE}$$

Here, $TSS = SST + SSB + SSE$

For the calculation of TSS , SST , SSB and SSE ANOVA table is as below;

	T_i				
A	0	0	-2	0	-2
B	2	0	1	2	5
C	4	2	2	1	9
T_j	6	2	1	3	$G=12$

Here,

Grand total (G) = 12

we know that,
 $TSS = SST + SSB + SSE$

$$\text{Correction Factor (C.F)} = \frac{G^2}{N} = \frac{144}{12} = 12$$

$$\text{So, } TSS = \sum_{i=1}^t \sum_{j=1}^r y_{ij}^2 - C.F.$$

$$= 4 + 4 + 1 + 4 + 16 + 4 + 4 + 1 - 12$$

$$= 38 - 12$$

$$= 26$$

$$SST = \sum \frac{T_j^2}{r} - C.F = \frac{(-2)^2 + 5^2 + 9^2}{4} - 12 = 15.5$$

$$SSB = \sum \frac{T_i^2}{t} - C.F = \frac{6^2 + 2^2 + 1^2 + 3^2}{3} - 12 = 4.67$$

Now,

$$\begin{aligned} SSE &= TSS - SSB - SST \\ &= 26 - 4.67 - 15.5 \\ &= 5.83 \end{aligned}$$

ANOVA table:

Source of variation	d.f	S.S	MSS	F _{cal}
Treatment	3-1=2	SST=15.5	$MST = \frac{15.5}{2} = 7.75$	$F_T = \frac{7.75}{0.4722} = 7.97$
Block	4-1=3	SSB=4.67	$MSB = \frac{4.67}{3} = 1.567$	$F_B = \frac{1.567}{0.5722} = 1.61$
Error	6	SSE= 5.83	$MSE = \frac{5.83}{6} = 0.9722$	
Total	12-1=11	TSS=26		

Critical value

For treatment:

The tabulated value of F at 0.05 level of significance with (2,6) d.f is $F_{0.05}(2,6) = 5.14$

For block

The tabulated value of F at 0.05 level of significance with (3,6) d.f is $F_{0.05}(3,6) = 4.76$

Decision:

For treatment

Since $F_T = 7.97 > F_{tab} = 5.14$. So, H_0 is ~~accepted~~ rejected, i.e., H_1 is accepted, i.e., This means there is significance difference between means.

For block

Since $F_B = 1.61 < F_{tab} = 4.76$. So, H_0 is ~~rejected~~ accepted. i.e., H_1 is rejected.
i.e., There is no significance difference between means.

Now,

$$\text{Precision of RBD / Precision of CRD} = \frac{r(t-1)MSE + (r-1)MSB}{(rt-1)MSE}$$

$$= \frac{4 \times 2 \times 0.977 + 3 \times 1.56}{11 \times 0.977}$$

$$= 1.164$$

Since, Precision of RBD > 1 . So, RBD is more efficient than CRD.

~~In case~~ $\begin{cases} < 1 \Rightarrow \text{less effective} \\ = 1 \Rightarrow \text{equally effective} \end{cases}$

Missing Plot for RBD:

Let us consider an RBD involving t treatment with r replication each. Let one of the observation say x (i.e., missing value) receiving i^{th} treatment in the j^{th} block is missing.

Let $G^t = \text{total of all known } rt-1 \text{ values}$

$T^j = \text{total of all known values of } j^{th} \text{ block.}$

$B^j = \text{total of all known values of } j^{th} \text{ block.}$

Then,

$$\text{missing value } (x) = \frac{T^j t + B^j r - G^t}{(t-1)(r-1)}$$

Now we substitute the value of x in place of missing value and carry out analysis as usual.

Because of the change in level of degree of freedom we obtain an upward bias in SST. Hence to get better result subtract an adjustment factor from SST.

$$\text{i.e., Adjustment factor } (k) = \frac{(B^j + T^j t - G^t)^2}{t(t-1)(r-1)^2}$$

$$\text{Adjusted SST (SST}_A\text{)} = \text{SST} - k.$$

We will understand better with following example:-

Q. The table given below are yields of 3 varieties in a 4 replicate experiment for which one observation is missing. Estimate the missing value and then analyse the data.

P	19	R	29	P	23	Q	33
Q	26	P	?	Q	27	R	26
R	21	Q	28	R	22	P	26

Solution:

Let the missing value be x . Now we rearrange table as follows:-

Varieties	Blocks (replicate)				T_i
	1st	2nd	3rd	4th	
P	19	x	23	26	T^1+x
Q	26	28	27	33	$=68+x$
R	21	29	22	26	114
Total (G_i)	66	$B^1+x=57+x$	72	85	$G^1+x=280+x$

Here, $T^1 = 68$, $B^1 = 57$, $G^1 = 280$, $r = 4$, $t = 3$

Now, missing value (x) = $\frac{T^1t + B^1r - G^1}{(r-1)(t-1)} = \frac{68 \times 3 + 57 \times 4 - 280}{3 \times 2} = 25.33$

Problem to test:

a) For varieties:- (OR For treatment)

Null Hypothesis (H_{0T}): $\mu_P = \mu_Q = \mu_R$

i.e, There is no significance difference between three varieties.

Alternate Hypothesis (H_{1T}): $\mu_P \neq \mu_Q \neq \mu_R$

i.e, There is significance difference between three varieties.

OR At least one variety is different than other.

b) For blocks:- (OR For replicates)

Null Hypothesis (H_{0B}): $\mu_1 = \mu_2 = \mu_3 = \mu_4$

i.e, There is no significance difference between four blocks (replicates).

Alternate Hypothesis (H_{1B}): $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$

i.e, There is significance difference between four replicates.

OR At least one replicate is different than other.

Test Statistics:

We know that, $F_T = \frac{MST}{MSE}$

$$F_B = \frac{MSB}{MSE}$$

मानकों square को sum

$$\begin{aligned} TSS &= \sum_{q=1}^t \sum_{j=1}^{r_j} (Y_{qj})^2 - C.F \\ &= 7927.6 - 7768.86 \\ &= 158.74 \end{aligned}$$

$$\begin{aligned} \text{f1 Correction Factor (CF)} &= \frac{G_1^2}{N} \\ &= \frac{(G_1 + x)^2}{N} \\ &= \frac{(280 + 25.33)^2}{12} \\ &= 7768.86 \end{aligned}$$

$$\begin{aligned} SST &= \frac{\sum_{q=1}^t T_q^2}{r} - C.F \\ &= \frac{(68 + 25.33)^2 + 114^2 + 98^2}{4} - 7768.86 \\ &= 58.76 \end{aligned}$$

Adjusted treatment of SST (SST_A) = $SST - k$

Here,

$$k = \frac{(B^1 + T^1 t - G^1)^2}{t(t-1)(r-1)^2} = \frac{(57 + 204 - 208)^2}{3 \times 2 \times 9} = 6.68$$

$$\therefore (SST_A) = 58.76 - 6.68 = 52.08$$

$$SSB = \frac{\sum_{j=1}^r T_j^2}{t} - C.F = \frac{66^2 + 82.33^2 + 72^2 + 85^2}{3} - 7768.86 = 78.88$$

$$SSE = TSS - SSB - SST_A$$

Now ANOVA table is as follows:

Source of variation	d.f.	Sum of square	Mean sum of square	Fcal
Treatment	$3-1=2$	$SST = 52.08$	$MST = \frac{52.08}{2} = 26.04$	$F_T = 26.04$
Block	$4-1=3$	$SSB = 78.88$	$MSB = \frac{78.88}{3} = 26.29$	$\frac{5.55}{26.29} = 4.69$
Error	5	$SSE = 27.728$	$MSE = \frac{27.728}{5} = 5.55$	$F_B = \frac{26.29}{5.55} = 4.734$
Total	$12-2 = 10$	$TSS = 158.68$		

Critical value:

For treatment

Value of F at 0.05 level of significance with (2,5) d.f

$$\text{is } F_{0.05}(2,5) = 5.79$$

For Block

Value of F at 0.05 level of significance with (3,5) d.f is

$$F_{0.05}(3,5) = 5.41$$

Decision:

For treatment

Since $F_T = 4.69 < F_{tab} = 5.79$, So, H_0 is accepted i.e., H_1 is rejected.

For block

Since $F_B = 4.734 < F_{tab} = 5.41$. So, H_0 is accepted i.e., H_1 is rejected.

Conclusion:

There is no significance difference between three varieties.

There is no significance difference between four replicates.

3.) Latin Square Design (LSD):

(not repeat)

LSD is used for non-homogenous material and it convert the non-homogenous into homogenous. So, LSD is better than RBD that means LSD is more efficient than RBD. Local Control is used in LSD along rowwise and column-wise, which controls error simultaneously. The shape of LSD is square since it contains equal number of rows and columns. It is based upon all the principles of design namely randomization, replication and local control.

Mathematical Model:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + e_{ijk} \quad \text{where,}$$

$$i = 1, 2, 3, \dots, m.$$

$$j = 1, 2, 3, \dots, m.$$

$$k = 1, 2, 3, \dots, m.$$

Here, y_{ijk} = i^{th} row and j^{th} column receiving k^{th} treatment.

α_i = constant effect

β_j = effect due to j^{th} row.

τ_k = effect due to k^{th} column

e_{ijk} = error due to chance

Problem to test:

(a) For treatment:-

Null Hypothesis (H_0): $\mu_1 = \mu_2 = \dots = \mu_m$.

i.e., There is no significance difference between means.

Alternate Hypothesis (H_A): $\mu_1 \neq \mu_2 \neq \dots \neq \mu_m$

i.e., There is significance difference between means.

OR At least one mean is different than others.

(b) For rows:-

Similar as for treatment

(c) For columns:-

Similar as for treatment.

We will understand better with example

Test Statistics:

$$F_T = \frac{MST}{MSE} \quad (\text{for treatment})$$

$$M_C = \frac{MSC}{MSE} \quad (\text{for columns})$$

$$M_R = \frac{MSR}{MSE} \quad (\text{for rows})$$

where, $MST = \frac{SST}{d.f}$, $MSC = \frac{SSC}{d.f}$, $MSR = \frac{SSR}{d.f}$ & $MSE = \frac{SSE}{d.f}$

$$\boxed{TSS = SST + SSR + SSC + SSE}$$

Now, Correction Factor (C.F) = $\frac{G^2}{N}$ where, G_i = Grand Total.

$$TSS = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m y_{ijk}^2 - C.F$$

$$SST = \frac{\sum T_{..k}^2}{m} - C.F$$

$$SSR = \frac{\sum T_{i..}^2}{m} - C.F$$

$$SSC = \frac{\sum T_{.j.}^2}{m} - C.F$$

Now,

$$\boxed{SSE = TSS - SSR - SSC - SST}$$

ANOVA table is as follows:

Source of variation	d.f	SS	MSS	F
Treatment	$m-1$	SST	$MST = \frac{SST}{m-1}$	$F_T = \frac{MST}{MSE}$
Row	$m-1$	SSR	$MSR = \frac{SSR}{m-1}$	$F_R = \frac{MSR}{MSE}$
Column	$m-1$	SSC	$MSC = \frac{SSC}{m-1}$	$F_C = \frac{MSC}{MSE}$
Error	$(m-1)(m-2)$	SSE	$MSE = \frac{SSE}{(m-1)(m-2)}$	
Total	$m^2 - 1$	TSS		

Q. Set up the analysis of variance for the following results of a design.

A	10	B	15	C	20
B	25	C	10	A	15
C	25	A	20	B	15

Solution: Also calculate the efficiency of design over RBD or CRD.

Taking deviation (i.e., subtracting 15 from each value).

A	-5	B	0	C	5
B	10	C	-5	A	0
C	10	A	-5	B	0

Problem to test:

① For treatment:

Null Hypothesis (H_0): $\mu_A = \mu_B = \mu_C$.

i.e., There is no significance difference between treatment means.

Alternate Hypothesis (H_1): $\mu_A \neq \mu_B \neq \mu_C$.

i.e., There is significance difference between three treatment means.

OR At least one treatment mean is different than other.

(b) For rows:

Null Hypothesis (H_{0R}): $\mu_{1R} = \mu_{2R} = \mu_{3R}$

i.e., There is no significance difference between rows.

जो मात्र लेखे परिवहन करें।

Alternate Hypothesis (H_{1R}): $\mu_{1R} \neq \mu_{2R} \neq \mu_{3R}$

i.e., There is significant difference between rows.

(c) For columns:

Null Hypothesis (H_{0C}): $\mu_{1C} = \mu_{2C} = \mu_{3C}$

i.e., There is no significance difference between columns.

Alternate Hypothesis (H_{1C}): $\mu_{1C} \neq \mu_{2C} \neq \mu_{3C}$

i.e., There is significant difference between columns.

Test Statistics:

We know that,

$$F_T = \frac{MST}{MSE} \quad (\text{for treatment})$$

$$F_C = \frac{MSC}{MSE} \quad (\text{for columns})$$

$$F_R = \frac{MSR}{MSE} \quad (\text{for rows})$$

$$\text{and } TSS = SST + SSR + SSc + SSE$$

Row \ Column	1	2	3	T _j
Row 1	-5	0	5	0
Row 2	10	-5	0	5
Row 3	10	5	0	15
T _j	15	0	5	G _T = 20

Here,

$$\text{Correction Factor (C.F)} = \frac{G^2}{N} = \frac{G^2}{m^2} = \frac{20^2}{9} = 44.44$$

Now,

$$TSS = \sum_{ijk} y_{ijk}^2 - CF$$

$$= (-5)^2 + 0^2 + 5^2 + 10^2 + (-5)^2 + 0^2 + 10^2 + 5^2 + 0^2$$

$$= 255.56$$

$$SST = \frac{\sum T_{i..}^2}{m} - C.F$$

Here,

$$T_{1..A} = -5 + 5 + 0 = 0$$

$$T_{1..B} = 10 + 0 + 0 = 10$$

$$T_{1..C} = 10 - 5 + 5 = 10$$

$$\text{So, } SST = \frac{0^2 + 10^2 + 10^2}{3} - 44.44 = \cancel{38.89} \quad 22.226$$

$$SSR = \frac{\sum T_{i..}^2}{m} - C.F = \frac{0 + 5^2 + 15^2}{3} - 44.44 = 38.89$$

$$SSC = \frac{\sum T_{..j}^2}{m} - C.F = \frac{15^2 + 0 + 5^2}{3} - 44.44 = 38.89$$

$$\text{Now, } SSE = TSS - SSR - SSC - SST$$

$$= 256.56 - 38.89 - 38.89 - 22.226 \\ = 156.56$$

Now, ANOVA table :-

Source of Variation	d.f.	SS	MSS	F
Treatment	$3-1=2$	$SST = 22.22$	$MST = \frac{22.22}{2} = 11.11$	$F_T = \frac{11.11}{78.28} = 0.1414$
Row	$3-1=2$	$SSR = 38.89$	$MSR = \frac{38.89}{2} = 19.44$	$F_R = \frac{19.44}{78.24} = 0.2485$
Column	$3-1=2$	$SSC = 38.89$	$MSC = \frac{38.89}{2} = 19.44$	$F_C = \frac{19.44}{78.24} = 0.2485$
Error	2	$SSE = 156.56$	$MSE = \frac{156.56}{2} = 78.24$	
Total	$m^2-1=9-1=8$			

Critical value:

For treatment → The value of F at 0.05 level of significance with (2,2)

d.f is $F_{0.05}(2,2) = 19.0$

For row → The value of F at 0.05 level of significance with (2,2)

d.f is $F_{0.05}(2,2) = 19.0$

For column → The value of F at 0.05 level of significance with (2,2)

d.f is $F_{0.05}(2,2) = 19.0$

Decision:

For treatment → Since $F_T = 0.1414 < F_{tab} = 19.0$, So, H_0 is accepted.

For row → since $F_R = 0.2485 < F_{tab} = 19.0$. So, H_0 is accepted.

For column → since $F_C = 0.2485 < F_{tab} = 19.0$. So, H_0 is accepted.

Conclusion:

There is no significance difference between treatment means.

There is no significance difference between rows.

There is no significance difference between columns.

Again,

i) Efficiency of LSD with relative to RBD = $\frac{(m-1)MSE + MSC}{m MSE}$

$$= \frac{(3-1) \times 78.24 + 19.44}{3 \times 78.24}$$

$$= 0.749$$

which is < 1 , so LSD is less efficient than RBD.

ii) Efficiency of LSD with relative to CRD = $\frac{(m-1)MSE + MSR + MSC}{(m+1) MSE}$

$$= \frac{(3-1) \times 78.24 + 19.44 + 19.44}{(3+1) \times 78.24}$$

$$= 0.624$$

which is < 1 , so LSD is less efficient than CRD.

Missing plot for LSD:-

Let us consider a $m \times m$ (square) LSD. Let one of the observation x occurring in i^{th} row, j^{th} column and k^{th} treatment is missing.

Let, $G^i = \text{total of all known } m^2 - 1 \text{ values}$.

$R^i = \text{total of all known values of } i^{\text{th}} \text{ row}$.

$C^j = \text{total of all known values of } j^{\text{th}} \text{ column}$.

$T^k = \text{total of all known values of } k^{\text{th}} \text{ treatment}$.

$$\text{missing value } x = \frac{m(R^i + C^j + T^k) - 2G^i}{(m-1)(m-2)}$$

We substitute the value in place of missing value x and carry out analysis as usual.

$$\text{Adjustment factor (k)} = \frac{\{(m-1)T^k + R^i + C^j - G^i\}^2}{\{(m-1)(m-2)\}^2}$$

$$\text{Adjusted SST (SST}_A) = \text{SST} - k.$$

Example: The table given below represents the yields of 4 varieties in a 4 replicate experiment for which one observation is missing. Estimate the missing value and then carry out the ANOVA.

A	12	C	19	B	10	D	8
C	18	B	12	D	6	A	?
B	22	D	10	A	5	C	21
D	12	A	7	C	27	B	17

Solution:

Let missing value be x .

					Total			
A	12	C	19	B	10	D	8	49
C	18	B	12	D	6	A	x	$\frac{R^i+x}{=36+x}$
B	22	D	10	A	5	C	21	58
D	12	A	7	C	27	B	17	63
Total	64	48	48	$\frac{C^j+x}{=46+x}$	$\frac{G^i+x}{=206+x}$			

Since this is the case of LSD.

So, missing value (x) = $\frac{m(R^I + C^I + T^I) - 2G^I}{(m-1)(m-2)}$

where, $m=4$, $R^I=36$, $C^I=46$, and $T^I=12+5+7=24$

$$\therefore x = \frac{4(36+46+24)-2 \times 206}{(4-1)(4-2)} \\ = 2$$

Now,

$$G^I+x = 206+x = 206+2 = 208 = G_I$$

Problem to test:-

@ For treatment.

Null Hypothesis (H_0) :- There is no significance difference between treatments.

Alternate Hypothesis (H_1) :- There is significance difference between treatments.

Similarly we write for row and column.

Test Statistics:-

We know that, $F_T = \frac{MST}{MSE}$ (for treatment)

$F_R = \frac{MSR}{MSE}$ (for row)

$F_C = \frac{MSC}{MSE}$ (for column).

$$\text{& } TSS = SSE + SSC + SSR + SST_A$$

where, $SST_A = SST - \text{adjustment factor (k)}$.

$$\begin{aligned} \text{Here, adjustment factor (k)} &= \frac{[(m-1)T^I + R^I + C^I - G^I]^2}{\{(m-1)(m-2)\}^2} \\ &= \frac{[(4-1) \times 24 + 36 + 46 - 206]^2}{\{(4-1)(4-2)\}^2} \\ &= 75.11 \end{aligned}$$

Here, Grand Total (G_I) = 208

Sum of all treatments of A excluding x .

Correction Factor (C.F) = $\frac{G_i^2}{N} = \frac{(208)^2}{16} = 2704$.

Now

$$\begin{aligned} TSS &= \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m y_{ijk}^2 - C.F \\ &= 12^2 + 19^2 + 10^2 + 8^2 + 18^2 + 12^2 + 6^2 + 2^2 + 22^2 + 10^2 + 5^2 + 21^2 \\ &\quad + 12^2 + 7^2 + 27^2 + 17^2 - 2704 \\ &= 734 \end{aligned}$$

$$\begin{aligned} SSR &= \frac{\sum_{i=1}^m T_{i..}^2}{m} - C.F \\ &= \frac{(49)^2 + 38^2 + 58^2 + 63^2}{4} - 2704 \\ &= 90.5 \end{aligned}$$

$$SSC = \frac{\sum_{i=1}^m T_{..i}^2}{m} - C.F = \frac{64^2 + 48^2 + 48^2 + 48^2}{4} - 2704 = 48$$

$$SST = \frac{\sum_{i=1}^m T_{...k}^2}{m} - C.F$$

Here, $T_{..A} = 12 + 5 + 7 + 2 = 26$

$$T_{..B} = 22 + 10 + 12 + 17 = 61$$

$$T_{..C} = 19 + 18 + 21 + 27 = 85$$

$$T_{..D} = 8 + 6 + 10 + 12 = 36$$

$$\therefore SST = \frac{26^2 + 61^2 + 85^2 + 36^2}{4} - 2704$$

$$= 525.5$$

Now,

$$\begin{aligned} SST_A &= SST - \text{adjustment factor (k)} \\ &= 525.5 - 75.11 \\ &= 450.39 \end{aligned}$$

$$\begin{aligned} SSE &= TSS - SSC - SSR - SST_A \\ &= 734 - 48 - 90.5 - 450.39 \\ &= 145.11 \end{aligned}$$

Now, ANOVA table is;

Consider if any calculation mistakes understand process to solve

Source of variation.	d.f.	SS	MSS	F
Treatment	$4-1=3$	$SST_A = 450.39$	$MST = \frac{450.39}{3} = 150.13$	$F_T = \frac{150.13}{24.185} = 6.2$
Row	$4-1=3$	$SSR = 90.5$	$MSR = \frac{90.5}{3} = 30.16$	$F_R = \frac{30.16}{24.185} = 1.24$
Column	$4-1=3$	$SSC = 48$	$MSC = \frac{48}{3} = 16$	$F_C = \frac{16}{24.185} = 0.66$
Error	5	$SSE = 145.11$	$MSE = \frac{145.11}{6} = 24.185$	
Total	$m^2-1 = 16-1 = 15$	$TSS = 734$		

Critical value:

since same value 3,6 for all

For treatment, row and column the value of F at 0.05 level of significance and (3,6) degree of freedom is $F_{0.05}(3,6) = 4.76$.

Decision:

for treatment \rightarrow since $F_T = 6.2 > F_{tab} = 4.76$. So, H_0 is rejected.

for row \rightarrow since $F_R = 1.24 < F_{tab} = 4.76$. So, H_0 is accepted.

for column \rightarrow since $F_C = 0.66 < F_{tab} = 4.76$. So, H_0 is accepted.

Conclusion:

There is significance difference between treatments.

There is no significance difference between rows.

There is no significance difference between columns.

Short Questions:

The questions we did till now are for long [10 marks], there is possibility of short questions [5 marks] also from this unit which are filling up the ANOVA table and taking out analysis. We have did long questions already so it will be quite easier now to know short questions. The following things should be in mind before solving short questions:-

- Firstly observe the given table e.g. what—what source of variations are provided, generally the questions are of RBD and LSB only.
- If treatments, blocks and error are given in source of variation then we should know that it is RBD related question then; we use formulas of RBD in table as we used before for finding values.
- Similarly if treatments, columns, rows and error are given in source of variation then we should know that it is LSB related question and use formulas in ANOVA table as we used before.

Example 1: Complete the following table for the analysis of variance of a design.

S.V.	d.f.	S.S	MSS	F
Blocks	4 (i.e.)	26.8	?	?
Treatment	3 (i.e.)	?	?	?
Error	?	?	2.5	
Total	?	85.3		

Solution:-

The given question is of randomised block design (RBD).

So, Total is given by = $t \times r - 1$ (for degree of freedom)

$$\begin{aligned}
 &= 5 \times 4 - 1 \\
 &= 19
 \end{aligned}$$

If Error = $19 - 4 - 3$ (for d.f.)
 $= 12$

Compare this table with
ANOVA table for RBD for better
understanding

$\Rightarrow t-1 = 4$ ^{Rough}
 $\Rightarrow t = 5$
 $\Rightarrow r-1 = 3$
 $\Rightarrow r = 4$

Here,

$$MSE = \frac{SSE}{d.f}$$

$$\text{or, } 2.5 = \frac{SSE}{12}$$

$$\text{or, } SSE = 30.$$

$$\begin{aligned} SST &= TSS - SSB - SSE \\ &= 85.3 - 26.8 - 30 \\ &= 28.4 \end{aligned}$$

$$MSB = \frac{SSB}{d.f} = \frac{26.8}{4} = 6.7$$

$$MST = \frac{SST}{d.f} = \frac{28.4}{3} = 9.467$$

$$F_T = \frac{MST}{MSE} = \frac{9.467}{2.5} = 2.786$$

$$\& F_B = \frac{MSB}{MSE} = \frac{6.7}{2.5} = 2.68$$

Hence the table on filling values now becomes;

S.V.	d.f	SS	MSS	F
Blocks	4	26.8	6.7	2.786
Treatment	3	28.4	9.467	2.68.
Error	12	30	2.5	
Total	19	85.3		

{ Now, we get ANOVA table and we can now easily calculate critical value and make decision and conclusion just same as we did before. It is quite easy.

We will do this only if testing of hypothesis is also asked generally in exams

Similarly we can calculate if LSB table is given.

Note: The most important thing is hence to remember ANOVA table for RBD and LSB and some formulas related to them.