Practical : 1

A manufacturing Company has purchased 4 new machines of different makes and wishes to determine whether one of them is faster than the other in producing a certain output. Five hourly production figures are observed at random and the results are given below:

M_1:	24	30	36	38	31
M_2:	31	39	38	42	35
M_3:	25	31	28	25	28
M_4:	20	22	25	33	36

Carry out ANOVA to test whether the machines are significantly different in their avg speed at 5% LOS.

sol: AIM: To test whether the machines are significantly different in their average speed at 5% LOS.

CALCULATION :

i) Null hypothesis: The 4 machines donot differ significantly wrt avg speed. i.e, $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$.

ii) Alternative hypothesis: The 4 machines differ significantly wrt average speed i.e, $H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$.

To test the null hypothesis, the following calculations are made:

Raw sum of squares (RSS): $\sum_{i=1}^{t} \sum_{j=1}^{r_i} y_{ij}^2$

Correction factor (C.F) = $\frac{y_{..}^2}{N}$

Total sum of squares (TSS) = RSS - CF = $\sum_{i=1}^{t} \sum_{j=1}^{r_i} y_{ij}^2 - \frac{y_{..}^2}{N}$

sum of squares due to treatments (SST) = $\sum\sum \frac{y_{i.}^2}{r_i} - \frac{y_{..}^2}{N}$

SSE = TSS - SST.

M_1	M_2	M_3	M_4	
24	31	25	20	
30	39	30	22	
36	38	28	25	
38	42	25	33	
31	35	28	36	
159	185	136	136	$y_{..} = 616$
5177	6915	3718	3894	$= 19704$

$RSS = 19704$

$C.F = \frac{(616)^2}{20} = 18972.8$

$TSS = 19704 - 18972.8$
$= 731.2$

$SST = 19299.60 - 18972.8$
$= 326.8$

$SSE = TSS - SST$
$= 731.2 - 326.8$
$= 404.4$

The entire analysis can be presented in the following table known as ANOVA table for one-way classified data with equal no. of observations

source of variation	d.f	SST	M.S	F ratio	
				F_{cal}	$F_{critical}$
due to treatments	3	326.8	108.93	4.3098	3.24
due to error	16	404.4	25.2700		
total	19	731.2			

CONCLUSION:

since, $F_{cal} > F_{cri}$, we reject Null hypothesis & it is concluded that 4 machines differs significantly wrt their average speed.

In this case, we further proceed to find which pair of machines differ significantly.

for this, we can find c.d. $c.d = \sqrt{\frac{2 Se^2}{r}} \times t_{(0.05)}$

$= \sqrt{\frac{2(25.2750)^2}{5}} \times 2.12$

C.d = 6.740

Mean of each treatments is

$M_1 = 31.8$

$M_2 = 37$

$M_3 = M_4 = 27.2$.

In descending order of treatment means:

Treatment	mean	Δ		
M_2	37	5.2	9.8	9.8
M_1	31.8	4.6	4.6	
M_3	27.2	0		
M_4	27.2			

We have $M_2 - M_3 = 9.8 > 6.74$ & $M_2 - M_4 = 9.8 > 6.74$

$\Rightarrow$ Machines M_2 & M_3 as well as M_2 & M_4 differ significantly.

All the remaining differences are not significant.

Practical (3):

The following data relating to weekly sales record of 3 salesmen A, B & C during 13 sale calls are given below:

A	300	400	300	500	
B	600	300	301	400	
C	700	300	400	600	500

carryout ANOVA & test whether the sales of 3 salesman are different or not at 5% LOS.

Sol: AIM: To test whether the sales of 3 salesman are diff or not at 5%.

CALCULATION:

Null hypothesis: The sales of 3 salesman donot differ at 5%. LOS.

Alternative hypothesis: The sales of 3 salesman differ at 5% LOS.

To test Null Hypothesis, the following calculations are made:

i) $RSS = \sum_{i=1}^{t}\sum_{j=1}^{r_i} y_{ij}^2$ ii) $C.F = \frac{y..^2}{N}$ iii) $TSS = RSS - C.F$

iv) $SST = \frac{\sum_{i=1}^{k} y_{i.}^2}{r_i} - \frac{y..^2}{N}$ v) $S.S.E = TSS - SST$.

	A	B	C
	300	600	700
	400	300	300
	300	300	400
	500	400	600
			500
$y_{i.}$	1500	1600	2500
y_{ij}^2	590000	700000	1350000

$y.. = 5600$

$\Sigma = 2640000$

$RSS = 2640000$, $C.F = \frac{(5600)^2}{13} = 2412307.692$

$TSS = 2640000 - 2412307.692 = 227692.3080$

$SST = 2452500 - 2412307.692 = 40192.3080$

$SSE = 227692.3080 - 40192.3080 = 187500$

The entire analysis can be presented in the following table known as ANOVA table for one-way classified data with unequal no. of observat

Source of variation	d.f	S.S	M.S	F ratio F cal	F tab
due to treatments	2	40192.3080	20096.1540	1.0718	4.10
due to error	10	187500	18750		
total	12	227692.3080			

CONCLUSIONS

since, $F_{cal} < F_{tab}$. we accept Null hypothesis and it is concluded that the sales of 3 salesman doesnot differ at 5% LOS.

Practical: 3

5 doctors each applies treatments on the patients on a certain disease and records the no. of days each patient takes to recover the results are given below:

Doc	Treatments 1	2	3	4	5
A	10	14	23	19	20
B	11	15	24	17	21
C	9	12	20	16	19
D	8	13	17	17	20
E	12	15	19	15	22

Carryout ANOVA & discuss whether there is any significant difference between doctors and the treatments test at 5% LOS.

Sol: AIM: To carryout ANOVA and to examine whether there is any significant difference between the doctors and the treatments at 5%.

CALCULATIONS:

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

H_{02}: There is no significant difference between treatment effects.

Denote ANOVA model for 2 way classification, where y_{ij} = no. of days taken to recover the patient using j^{th} treatment of the doctor.

μ = General mean effect

α_i = effect of doctor

β_j = effect of j^{th} treatment

ϵ_{ij} = error effect due to chance

To test Null hypothesis, the following calculations are made:

$G = \sum_{i=1}^{5}\sum_{j=1}^{5} y_{ij} = y_{..}$

Correction factor $= \frac{G^2}{N} = 6658.56$

$RSS = \sum_{i=1}^{5}\sum_{j=1}^{5} y_{ij}^2 = 7130$, $TSS = RSS - C.F = 7130 - 6658.56 = 471.04$

$SS_{doc} = \frac{\sum_{i=1}^{5} y_{i.}^2}{5} - C.f = 27.44$

$SS_{tr} = \frac{\sum_{j=1}^{5} y_{.j}}{5} - C.f = 407.44$

$SSE = TSS - SS_{DOC} - SS_{tr} = 471.44 - 27.44 - 407.44 = 36.56$

$= -47$

Source of variation	d.f	SS	M.S	Fcal	Ftab
Doctor	4	27.44	6.86	3.0022	3.01
treatment	4	407.44	101.86	44.5777	3.01
Error	16	36.56	2.2850		
total	24	471.44			

CONCLUSION)

Since, $F_{cal} = 3.022 < F_{tab} = 3.01$ for the effects due to doctors, we accept H_{01} at 5% LOS & it is concluded that there is no significant difference between the effects of doctors.

Since, $F_{cal} > F_{tab}$ for the effects of due to treatments. we reject H_{02} at 5% LOS & it is concluded that there is significant difference between the effects of treatments

Practical (4):

4 varieties of potatoes are planted and each on 5 plots of grounds of the same size and each variety is treated with 5 different fertilizers. The yields in tons are given below:

Var \ (fertilizers)	1	2	3	4	5
A	1.9	2.2	2.6	1.8	2.1
B	2.5	1.9	2.3	2.6	2.2
C	1.7	1.9	2.2	2.0	2.1
D	2.1	1.8	2.5	2.3	2.4

Perform ANOVA & examine whether there is any significant difference between the different varieties and fertilizers at 5% LOS.

Sol:

AIM: To perform ANOVA & examine whether is any significant difference between different varieties and fertilizers at 5% LOS.

CALCULATIONS:

H_{01}: There is no significant difference between varieties. i.e, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$

H_{02}: There is no significant difference between fertilizer i.e, $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$ where $i = 1,2,3,4$ & $j = 1,2,3,4,5$

Denote ANOVA model for 2-way classification, where y_{ij} = no. of tons of i^{th} variety of potato treated with j^{th} fertilizers.

μ = General mean effect

α_i = effect of variety of potato

β_j = effect of j^{th} treatment

ϵ_{ij} = error effect due to chance.

To test null hypothesis, the following calculations are made.

var	fertilizers 1	2	3	4	5	$y_{i.}$
A	1.9	2.2	2.6	1.8	2.1	10.6
B	2.5	1.9	2.3	2.8	2.2	11.5
C	1.7	1.9	2.2	2.0	2.1	9.9
D	2.1	1.8	2.5	2.3	2.4	11.1
$y_{.j}$	8.2	7.8	9.6	8.7	8.8	43.1
$y_{.j}^2$	17.16	15.3	23.14	19.29	19.42	94.31

$G = 43.1$, $CF = \frac{G^2}{N} = 92.8805$, $RSS = \sum\sum y_{ij}^2 = 94.31$

$TSS = RSS - C.F = 94.31 - 92.8805 = 1.4295$

$SS_V = 0.2855$ $\quad SS_F = 0.4620$

$SSE = TSS - SS_V - SS_F = 1.4295 - 0.2855 - 0.4620 = 0.6820$

source of variation	d.f	s.s	M.S	fcal	ftab
varieties	3	0.2855	0.0714	1.2570	3.26
fertilizers	4	0.4620	0.1540	2.7113	3.49
error	12	0.6820	0.0568		
total	19				

CONCLUSION:

since, fcal < ftab, for the effects due to variations, we accept H_{01} at 5% LOS & it is concluded that there is no significant difference between the effects of potato variations.

since, fcal < ftab, for the effects of fertilizers, we accept H_{02} at 5% LOS & it is concluded that there is no significant difference between the effects of fertilizers.

(10)

Practical: Analysis of CRD:

To study the property of reflection, 4 colours of paint A, B, C & D are applied randomly on 20 homogenous metal unit each with 5 replications. The results are recorded on a CRD layout given below: Analyse the design and compare the property of reflection for various colours at 5% LOS.

A195	B45	D55	D50	C195
D120	A150	C230	B145	A160
B40	C115	B195	D80	C225
A205	C235	D135	A110	B65

sol: AIM: To analyse the given design and to compare the property of reflection of various colours at 5% LOS.

CALCULATION:

Ho: All 4 colours donot differ significantly wrt property of reflection

$H_0: T_A = T_B = T_C = T_D$.

Alternative hypothesis: All 4 colours differ significantly wrt property of reflection. i.e, $H_1: T_A \neq T_B \neq T_C \neq T_D$.

To test the null hypothesis, the following calculations are made:

	A	B	C	D	
	195	45	195	55	
	150	145	230	50	
	160	40	115	120	
	205	195	225	80	
	110	65	235	135	
$y_{i.}$	820	490	1000	440	$y_{..} = 2750$
y_{ij}^2	140250	66900	210000	44550	461700

RSS = 461700

$C.F = \frac{(2750)^2}{20} = 378125$, $TSS = 461700 - 378125 = 83575$

$$SST = \left[\frac{(820)^2 + (490)^2 + (1000)^2 + (440)^2}{5}\right] - \left[\frac{(2750)^2}{20}\right] = 43095$$

$SSE = TSS - SST = 83575 - 43095 = 40480$

The entire analysis can be presented in the following table known as ANOVA table for one-way classified data with equal no. of observati

source of variation	d.f	SS	M.S	F cal	F critical
due to treatments	3	43095	14365	2.3549	3.24
Error	16	40480	2530		
total	19	83575			

CONCLUSION:

since, $F_{cal} \not> F_{critical}$, We accept Null hypothesis and it is concluded that all 4 colours donot differ significantly wrt property of reflection at 5% LOS.

Practical (6):

The following data relating to the yields of potato crops for which seven kinds of fertilizers were used is given. Analyse the design and state your conclusion at 5% LOS.

f_3 90.12	O 123.5	S_6 182.5	F_{12} 104.5	S_6 242.1	S_{12} 175.6
O 107.3	S_3 75.82	F_{12} 48.56	F_6 102.9	S_3 215.1	O 249.3
f_3 98.95	S_{12} 73.77	F_6 181.3	O 306.2	F_6 185.7	S_{12} 165.3
f_3 99.99	O 182.6	S_{12} 192.9	S_6 192.3	O 325.9	f_{12} 55.52
S_6 189.5	O 341.2	S_6 201.8	F_{12} 181.9	F_{12} 42.35	f_3 100.4
f_6 291.4	f_6 305.2	S_{12} 225.4	O 241.8	f_3 99.91	f_3 78.52
S_6 214.8	S_6 235.4	F_3 111.6	F_6 55.89	O 120.0	O 171.3

Sol: AIM: To analyse the yields of potato crops when 7 kinds of fertilizers were used at 5% LOS.

CALCULATION:

Null hypothesis: Potato crops donot differ with 7 kinds of fertilizers.

$H_0: T_{F_3} = T_O = T_{S_6} = T_{F_{12}} = T_{S_{12}} = T_{S_3} = T_{F_6}$.

Alternative hypothesis: Potato crops differ with 7 kinds of fertilizers.

$H_1: T_{F_3} \neq T_O \neq T_{S_6} \neq T_{F_{12}} \neq T_{S_{12}} \neq T_{S_3} \neq T_{F_6}$

To test the null hypothesis, the following calculations are made

f_3	O	S_6	f_{12}	S_{12}	S_3	f_6
90.12	123.5	182.3	104.5	175.6	75.82	102.9
98.95	107.3	242.1	48.56	73.77	215.1	181.3
100.4	249.3	192.3	55.52	165.3	99.99	185.7
99.91	306.2	189.5	181.9	170.9	78.52	291.4
111.6	182.6	201.8	42.35	225.4		306.2

(3)

325.9	214.8	85.89
341.2	235.4	
241.8		
128.2		
171.3		

$y_{i.}$ 501.98 21773 1458.2 432.83 815.87 419.430 1152.39 7107.1

$$C.F = \frac{G^2}{N} = \frac{(10071.1)^2}{42} = 1169034.534.$$

$$RSS = \sum\sum y_{ij}^2 = 1428481.889$$

$$TSS = RSS - CF = 1428481.889 - 1169034.534 = 259447.3584$$

$$SST = 1275078.299 - C.F = 106043.7650$$

$$SSE = TSS - SST = 153403.5904.$$

Source of variation	d.f	S.S	M.S	F cal	F ratio
treatments	6	106043.7650	17673.9608	4.0324	2.34
error	35	153403.5904	4382.9597		
total	41	235973.6874			

CONCLUSION:

since, $F_{cal} > F_{tab}$, we reject H_0 i.e, ther is significant difference between the yields of potatoes wrt fertilizers.

Practical (7):

In agricultural field experimentation, 3 fertilizers are applied in 4 randomized blocks and the yields of wheat are given below:

	1	2	3	4
1	A8	C10	A6	B10
2	C12	B8	B9	A8
3	B10	A8	C10	C9

i) Analyse the above design & state your conclusions at 5% LOS

ii) find the efficiency of RBD over CRD.

Sol:

AIM:

i) To analyse the given RBD experiment & comment on the results at 5% LOS.

ii) To find efficiency of RBD over CRD.

CALCULATIONS.

let y_{ij} denote the yield of the i^{th} treatment in j^{th} block.

The ANOVA model for RBD is given by,

$y_{ij} = \mu + t_i + b_j + \epsilon_{ij}$ $i = 1,2,3$. & $j = 1, 2, 3, 4$.

$\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$ where μ = general mean effect, t_i = effect of i^{th} treatment, b_j = effect of j^{th} treatment & ϵ_{ij} = error random observation

Null hypothesis:

H_{01}: All the treatments are homogenous i.e, $H_{01} = t_1 = t_2 = t_3 = 0$

H_{02}: All the treatments are homogenous i.e, H_{02} : $b_1 = b_2 = b_3 = b_4 = 0$

The given design is RBD with columns taken as blocks since, there is no replication of treatments in any column.

To test the above two hypothesis, the following calculations are made:

Blocks	1	2	3	4	total
A	8	8	6	8	30
B	10	8	9	10	37
C	12	10	10	9	41
Block total	30	26	25	27	108

$G = 108$, $C.F = \frac{G^2}{N} = \frac{(108)^2}{12} = 972$, $RSS = \sum\sum y_{ij}^2 = 998$

$TSS = RSS - C.F = 998 - 972 = 26$

$SST = \frac{\sum y_{i.}^2}{r} - C.F = 15.5$

$SSB = \frac{\sum y_{.j}^2}{s_j} - C.F = 4.67$

$SSE = TSS - SST - SSB = 26 - 15.5 - 4.67 = 5.83$

source of variation	d.f	SS	M.S	fcal	fratio
treatments	2	15.5	7.75	7.9767	5.14
Blocks	3	4.67	1.5567	1.6020	4.76
Error	6	5.83	0.9717		
total	11	26			

Conclusion?

If $|f_{cal}| > f_{tab}$, we reject H_{01}

If $|f_{cal}| \le f_{tab}$, we accept H_{02}.

In the above case, we proceed to find which pair of fertilizers differ significantly.

Critical difference (c.d) = $\sqrt{\frac{2 se^2}{r}} \times t_{(r-1)(s-1)}$ at 0.05

$= \sqrt{\frac{2(0.9712)^2}{4}} \times 2.447 = 1.6813 \simeq 1.7$

fertilizer means arranged in descending order:

Treatment	means	Δ
A	7.5	$\Delta_{AB} = 1.75$
B	9.25	$\Delta_{BC} = 1$
C	10.25	$\Delta_{CA} = 2.75$

Among the fertilizers A, B & C, fertilizer C gives highest mean effect than A & B. As a critical difference is concerned, the fertilizer A & B, A & C tends to show a significant difference with $\Delta > 1.6813$.

ii) The efficiency of RBD over CRD is given by,

$$E_{RBD \text{ over } CRD} = \frac{(r-1) MSB + r(c-1)\sigma^2_{RBD}}{(rs-1)\sigma^2_{RBD}}$$

$$= \frac{(4-1)(1.5567) + 4(3-1)(0.9717)}{[4(3)-1](0.9717)} \times 100$$

$$= 116.4192$$

The gain by using RBD instead of CRD is 116.4192.

Practical (8):

An experiment was conducted in RBD layout with 4 insecticides and two methods of spraying. Comprising in total of 8 treatments combinations in 3 blocks. The treatment combinations are I_1S_1, I_1S_2, I_2S_2, I_1S_2, I_3S_1, I_3S_2, I_4S_1, I_4S_2 and the block are given as:

I_1S_1 15	I_2S_2 10	I_3S_2 15
I_3S_2 13	I_4S_1 12	I_1S_1 7
I_2S_2 10	I_3S_1 17	I_4S_2 18
I_2S_1 4	I_1S_2 10	I_4S_1 8
I_4S_2 16	I_4S_2 19	I_2S_1 15
I_3S_1 11	I_2S_1 16	?
I_1S_2 8	I_1S_1 8	I_2S_2 11
I_4S_1 9	I_3S_2 13	I_1S_2 6

estimate the missing value over RBD and then analyse the design at 5% LOS.

Sol: AIM: To estimate the missing value for the given RBD experiment and also to analyse the design & comment on the result at 5% LOS.

CALCULATIONS:

The given design is RBD with columns treated as blocks since there is no replication of treatments in any columns. Let the missing value be denoted by x. which is estimated during the formula:

$$\hat{x} = \frac{sy_{i.}' + ry_{.j}' - y_{..}'}{(r-1)(s-1)} \quad \text{where}$$

$y_{i.}'$ is the treatment total with missing yield involved.

$y_{.j}'$ is the block total when the missing yield is involved.

$y_{..}'$ is the grand total of all $(rs-1)$ available values s is no. of treatments & r is no of blocks.

The given layout is represented as follows:

Treatments	B_1	B_2	B_3	total
$I_1 S_1$	15	6	7	28
$I_1 S_2$	8	10	6	24
$I_2 S_1$	4	16	15	35
$I_2 S_2$	10	10	11	31
$I_3 S_1$	11	17	x	$28+x$
$I_3 S_2$	13	13	15	41
$I_4 S_1$	9	12	8	29
$I_4 S_2$	16	19	18	53
total	86	103	$80+x$	$269+x$

Here, $y_{i.}' = 28$, $y_{.j}' = 80$, $y_{..}' = 269$, $r = 3$, $s = 8$

$$\hat{x} = \frac{8(28) + 3(80) - 269}{(3-1)(8-1)} = 13.9286 \simeq 14.$$

After estimating the missing value, it is substituted back in the given data, the following calculations are made for testing null hypothesis

Null hypothesis:

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

i.e, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = 0$

H_{02}: There is no significant difference between effects of two methods of spraying

i.e, $S_1 = S_2 = S_3 = 0$

$G = 269 + 14 = 283$

$C.F = \frac{G^2}{N} = \frac{(283)^2}{24} = 3337.0417$

$RSS = 3727$

$TSS = 3727 - 3337.0417 = 389.9583$

$SSI = \left[\frac{28^2 + 24^2 + 35^2 + 31^2 + 42^2 + 41^2 + 29^2 + 53^2}{8}\right] - 3337.0417$

$= 209.9583$

$SSB = \frac{\sum y_{.j}^2}{s_j} - C.F = 3355.1250 - 3337.0417 = 18.0833$

S.V	d.o.f	S.S	M.S	Fcal	Ftab
treatments	7	209.9583	29.9940	2.4082	2.83
blocks	2	18.0833	9.0417	0.7259	3.81
errors	15	161.9167	12.4551		
total	22	389.9583			

Conclusion:

We accept H_{01} & H_{02}.

Practical (9):

4 manure treatment A, B, C & D are tested for their effect on the yield of rice in kilos - observed in a field experiment carried out in 4x4 LSD.

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

i) Analyse the above design & comment on the result at 5% LOS.

ii) Find the efficiency of design over CRD & RBD.

Sol: AIM: To analyse the given 4x4 LSD and to test whether rows, columns & treatments are homogenous at 5% LOS. Also to calculate the efficiency of LSD over CRD & RBD

CALCULATIONS:

The given design is LSD since, each treatments appear only once in each row & each column. let y_{ijk} denote the yield of the plot in ith row, jth column receiving the kth treatment then ANOVA model for LSD is $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijk}$.

Null hypothesis:

H_{01}: All rows are homogenous $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$

H_{02}: All columns are homogenous $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

H_{03}: All treatments are homogenous $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$.

To test the null hypothesis, the following calculations are made:

12	19	10	8	49	669
18	12	6	7	43	553
22	10	5	21	58	1050
12	7	27	17	63	1211
64	48	48	53	213	3483

treatment totals:

A = 12 + 7 + 5 + 7 = 31

B = 10 + 12 + 22 + 17 = 61

C = 19 + 18 + 21 + 27 = 85

D = 8 + 16 + 10 + 12 = 36

G = 213

$C.F = \frac{G^2}{N} = \frac{(213)^2}{16} = 2835.5625$, RSS = 3483

TSS = 3483 − 2835.5625 = 647.4375

$SSR = \left[\frac{49^2+43^2+58^2+63^2}{4}\right] - 2835.5625 = 60.1875$

$SSC = \left[\frac{64^2+48^2+48^2+53^2}{4}\right] - 2835.5626 = 42.6875$

$SST = \left[\frac{31^2+61^2+85^2+36^2}{4}\right] - 2835.5626 = 465.1875$

SSE = TSS − SSR − SSC − SST = 79.3750

The entire analysis can be presented in the following table known as ANOVA table.

source of variation	d.f	SS	M.S	F cal
Rows	3	60.1875	20.0625	F_R = 1.5165
Columns	3	42.6175	14.2292	F_C = 1.0756
treatments	3	465.1875	105.0626	F_T = 15.7212
Error	6	79.3750	13.2292	
total	15	647.4375		

CONCLUSION:

Since, $F_r < F_{tab}$ we accept H_{01} & $F_c < F_{tab}$ we accept H_{02}.
Hence, for rows & columns the effects are homogenous.
Since, $F_t > F_{tab}$, we reject H_{03} and it is concluded that the treatments effects are not homogenous.

To find which of the treatment pairs differs, we have to calculate critical difference:

$$C.D = \sqrt{\frac{2 S_E^2}{r}} \times t_{0.05} = \sqrt{\frac{2(13.2292)}{4}} \times 2.45 = 6.3011$$

The treatment means arranged in descending order:

A = 7.75, B = 15.25, C = 20.25, D = 9. ⇒ C, B, D, A.

		D
C	20.25	
		6
B	15.25	
		6.25
D	9	
		1.25
A	7.75	
		13.5

There is a higher difference between treatment C & treatment A.

The efficiency of LSD over CRD is $\frac{MSR + MSC + (r-1)\sigma^2_{LSD}}{(r+1)\sigma^2_{LSD}} \times 100$

$$= \frac{20.0625 + 14.2292 + (4-1)(13.2292)}{(4+1)(13.2292)} = 111.8424$$

The efficiency of LSD over RBD when rows are blocks is given by,

$$\frac{MSC + (r-1)\sigma^2_{LSD}}{r\sigma^2_{LSD}} \times 100 \Rightarrow \frac{14.2292 + (4-1)(13.2292)}{4 \times 13.2292} \times 100$$

$$= 101.8829$$

The efficiency of LSD over RBD where columns as blocks is given by,

$$E_{LSD \text{ to } RBD} = \frac{MSR + (r-1)\sigma^2_{LSD}}{r\sigma^2_{LSD}} \times 100$$

$$= \frac{21.0625 + (4-1)(13.2292)}{4 \times 13.2292} \approx 112.9133$$

CONCLUSION:

The efficiency of LSD over RBD is

(Rows as blocks) = 102

(Columns as blocks) = 113.

Practical (10):

To test the homogenity of average yield of 4 varieties A, B, C & D of wheat, LSD were selected to carry the experiment & the observed yields are given below & due to reasons, 1 datapoint is missing. estimate the missing observation & analyse the design. Comment on the result as 5% LOS.

D 29.1	B 18.9	C 29.4	A 5.7
C 16.4	A 10.2	D 21.2	B 19.1
—	D 35.8	B 24	C 37
B 24.9	C 41.7	A 9.5	D 28.9

estimate the efficiency of LSD over RBD.

Sol: AIM: To estimate the missing observation and to analyse the given LSD experiment to comment on the results at 5%. LOS. Also, to find the efficiency of design over RBD.

CALCULATION:

The given design is as LSD. Since, each treatment appears only once in each row & column.

let the missing observation be x. from the given design, it is easily understood that the missing observation belong to treatment A.

The missing observation can be estimated using the formula:

$$\hat{x} = \frac{r(R'+C'+T') - 2G'}{(r-1)(r-2)} = \frac{4(99.8+70.4+25.4) - 2(354.8)}{(4-1)(4-2)}$$

$$= 12.1333 \simeq 12.1$$

Therefore, $\hat{x} = 12.1$.

After estimating the missing value, it is substituted back in the the given data & data is analysed as usual with a little change of subtracting '1' d.o.f from the total sum of squares (TSS) and also from SSE.

D 29.1	B 18.9	C 29.4	A 5.7
C 16.4	A 10.2	D 21.2	B 19.1
A 12.1	D 38.8	B 24	C 37
B 24.9	C 41.7	A 9.5	D 28.9

Null hypothesis:

H_{01}: All rows are homogenous i.e, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$

H_{02}: All columns are homogenous i.e, $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

H_{03}: All treatments are homogenous i.e, $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$

To test the null hypothesis, the following calculations are made.

29.1	18.9	29.4	5.7	83.1	2100.87
16.4	10.2	21.2	19.1	66.9	1187.25
12.1	38.8	24	37	111.9	3596.88
24.9	41.7	9.5	28.9	105	3284.32
82.5	109.6	84.1	90.7	366.9	10169.33

Treatment totals:

$A = 5.7 + 10.2 + 12.1 + 9.5 = 37.5$

$B = 18.9 + 19.1 + 24 + 24.9 = 86.9$

$C = 29.4 + 16.4 + 37 + 41.7 = 124.5$

$D = 29.1 + 21.2 + 38.8 + 28.9 = 118$

$G = 366.9$

$C.F = \frac{(366.9)^2}{16} = 8413.4706$

$RSS = 10169.33$

$TSS = 10169.33 - 8413.4706 = 1755.8544$

$$SSR = \frac{(83.1)^2 + (66.9)^2 + (111.9)^2 + (108)^2}{4} - 8413.4706 = 318.4819$$

$$SSC = \frac{(82.5)^2 + (109.6)^2 + (84.1)^2 + (90.7)^2}{4} - 8413.4756 = 115.9019$$

$$SST = \frac{(37.5)^2 + 86.9^2 + (124.5)^2 + (118)^2}{4} - 8413.4706 = 1182.0519$$

$SSE = TSS - SSR - SSC - SST = 139.3687.$

The entire analysis can be presented in the following Anova.

Source of variation	d.f	S.S	M.S	F_{cal}	F_{tab}
Rows	3	318.4819	106.1606	4.5704	4.75
columns	3	115.9019	38.6506	1.6640	
treatments	3	1182.0519	394.0173	16.9630	
error	6	139.3617	23.2281		
total	15				

Conclusion:

Since, $F_r < F_{tab}$ we accept H_{01} & $F_C < F_{tab}$, we accept H_{02}.

Hence, for rows & columns the effects are homogenous.

$$C.D = \sqrt{\frac{2SE^2}{r}} \times t_{0.05} = \sqrt{\frac{2(23.2281)}{4}} \times 2.45 = 8.3490$$

$A = 37.5$, $B = 86.9$, $C = 124.5$, $D = 118.$

treatment means in descending order C D B A.

		Δ
C	124.5	
		6.5
D	118	
		31.1
B	86.9	
		49.4
A	37.5	
		85

There is a higher difference between treatment C & treatment A.

The efficiency of LSD over RBD is,

$$\frac{MSC + (r-1)\sigma^2_{LSD}}{r\sigma^2_{LSD}} \times 100$$

$$= \frac{38.6506 + (4-1)(23.2281)}{4(23.2281)} \times 100$$

$$= 116.5990$$

Practical (11):

Given below is the data regarding deaths in two districts. On the basis of given data, calculate CDR & STDR by considering District A as standard population.

Age group	District A population	no. of deaths	District B population	no. of deaths
0-10	2000	50	1000	20
10-55	7000	75	3000	30
55 & above	1000	25	2100	40

Sol: AIM: To calculate CDR & STDR considering district layer standard.

FORMULA:

$$CDR = \frac{D_x}{P_x} \times 100$$ when D_x = no. of deaths
P_x = population.

$$STDR = \frac{\Sigma m_x \cdot P_x^s}{\Sigma P_x^s}$$

when m_x = CDR & P_x^s = standard population.

CALCULATION:

Age group	District A P_x	D_x	m_x	$m_x P_x^s$	District B P_x	D_x	m_x	$m_x P_x^s$
0-10	2000	50	25	50,000	2000	20	50	50000
10-55	7000	75	10.7143	70,000	7000	30	100	300000
55 & above	1000	25	25	25000	1000	40	50	100000
	10,000	150	60.7143	1,50,000		90	200	4,50,000

$$m_x = \frac{D_x}{P_x} \times 1000 = \frac{50}{2000} \times 1000 = 25$$

$$StDR_A = \frac{1,50,000}{10000} = 15$$

$$CDR_A = \frac{150}{10,000} \times 1000 = 15$$

$$STDR_B = \frac{78000}{6000} \times \cancel{1000} = 13$$

$$CDR_B = \frac{90}{6000} \times 1000 = 15$$

Practical 12:

calculate CBR, GFR & CFR, TFR for the following data:

Age group	total population (1000's)	female population (1000's)	no. of births
15-19	30	15	600
20-24	32	14	1610
25-29	29	13	1716
30-34	25	10	1000
35-39	12	8	520
40-44	12	7	210
45-49	10	5	30.

sol) AIM: To calculate CBR, GFR, SFR, TFR

FORMULA:

$$CBR = \frac{B_x}{P_x} \times 1000$$

$$GFR = \frac{\Sigma B_x}{\Sigma P_x^f} \times 1000$$

$$SFR_1 = \frac{nB_x}{P_x^f} \times 1000$$

$$TFR = \left[\sum \frac{B_x}{P_x^f} \times 1000\right] \times 5$$

CALCULATION:

Age	P_x	P_x^f	B_x	CBR	SFR
15-19	30	15	600	2000	40
20-24	32	14	1610	50312.5	115
25-29	29	13	1716	56.172	132
30-34	25	10	1000	40	100
35-39	12	8	520	43.333	65
40-44	12	7	210	17.5	30
45-49	10	5	30	3	6
	150	72	5686	50492.5057	488

$SFR = \frac{5686}{72} = 78.9722$

$TFR = 488 \times 5 = 2440.$

$CBR = 50492.5057$

$SFR = 488.$

Practical (13):

fill in the blanks of the following table:

Age	l_x	d_x	p_x	q_x	L_x	T_x	e_x^0
25	5,62,324	?	?	?	?	2,52,76,840	?

Sol:

Aim: To find the missing observations in the given life table

FORMULA:

$d_x = l_x - l_{x+1}$

$q_x = \frac{d_x}{l_x}$

$p_x = 1 - q_x$

$L_x = \frac{l_x + l_{x+1}}{2}$

$e_x^0 = \frac{T_x}{l_x}$; $T_x = L_x + T_{x+1}$

CALCULATIONS:

Given, $x = 25$, $x+1 = 26$, $l_{25} = 5,62,324$, $l_{26} = 5,56,432$

$T_{25} = 2,52,76,840$.

$d_{25} = l_{25} - l_{26} = 5892$

$q_{25} = \frac{5892}{5,62,324} = 0.0105$; $P_{25} = 1 - 0.0105 = 0.9895$

$L_{25} = \frac{562324 + 556432}{2} = 559378$

$T_{26} = 25276840 - 559378 = 24717462$

$e^0_{25} = \frac{25276840}{562324} = 44.9507$

$$e^{o}_{26} = \frac{24717462}{556432} = 44.4214.$$

CONCLUSION:

$d_{25} = 5892$

$q_{25} = 0.0105$

$p_{25} = 0.9895$

$L_{25} = 559378$

$T_{26} = 24717462$

$e^{o}_{25} = 44.9587$

$e^{o}_{26} = 44.4214$

Practical (14)

Construct the Complete life table for the data given below:

Age (x)	l_x
0	1000
40	920
60	782
70	528
80	252
90	14
100	0

Sol: AIM: To construct the life table for the given table.

FORMULA:

$$d_x = l_x - l_{x+1}$$

$$q_x = \frac{d_x}{l_x}$$

$$p_x = 1 - q_x$$

$$L_x = \frac{l_x + l_{x+1}}{2}$$

$$T_x = L_x + T_{x+1}$$

$$= L_x + L_{x+1} + L_{x+2} + \cdots$$

$$\mathring{e}_x = \frac{T_x}{l_x}$$

CALCULATIONS:

Age(x)	l_x	d_x	q_x	p_x	L_x	T_x	e_x^0
0	1000	80	0.08	0.92	960	2996	2.996
40	920	138	0.15	0.85	851	2036	2.2130
60	782	264	0.3248	0.6752	655	1185	1.5153
70	528	276	0.5227	0.4773	390	530	1.0038
80	252	238	0.9444	0.0556	133	140	0.556
90	14	14	1	0	7	7	0.5
100	0						

Practical (15):

Compute GRR & NRR for the following data:

Age group	female population	female birth	survival rate
15-19	212100	1500	0.821
20-24	198000	8516	0.910
25-29	162000	6324	0.895
30-34	145000	4635	0.874
35-39	128000	2452	0.830
40-44	104000	856	0.814
45-49	82000	128	0.793.

Sol:

AIM: To calculate GRR & NRR for given data

FORMULA.

$$GRR = 5 \times \sum \frac{f_{B_x}}{f_{P_x}} \times 100$$

$$NRR = 5 \times \sum \left({}_n i_x \times {}_f {}_n \Pi_x \right)$$

CALCULATIONS:

Age group	${}^f{}_n P_x$	${}^f{}_n B_x$	${}^f{}_n i_x$	${}^f{}_n \Pi_x$	${}^f{}_n i_x \times {}^f{}_n \Pi_x$
15-19	212000	1500	7.0755	0.821	5.8089
20-24	198000	8516	43.0101	0.910	39.139
25-29	162000	6324	39.037	0.896	34.9772
30-34	145000	4635	31.965	0.874	27.9374
35-39	128000	2452	19.156	0.830	15.8997

40-44	104000	856	8.8308	6.6999	0.814
45-49	82000	128	1.5610	5.2379	0.793
			150.0354		131.6999.

$GRR = n \times \sum {}_{x}^{t}i_{n} = 5 \times 150.0334 = 750.1770$

$NRR = 5 \times \sum \left({}_{n}i_{x} \times {}_{n}^{t}r_{x} \right) = 5\ (131.6999) = 658.4995$

CONCLUSION)

GRR = 750.1770

NRR = 658.4995 .