

VU LEARNING HUB

vulearninghub.xyz

HIGHLIGHTED HANDOUT

STA408

Design and Analysis of Experiments

STA408 by VU Learning HUB

Prepared for VU students • Free study resources

<https://vulearninghub.xyz>

COURSE: STA 408

DESIGN & ANALYSIS OF EXPERIMENT

LECTURE DELIVERED BY

SARAH AHMAD

DEPARTMENT OF STATISTICS

VIRTUAL UNIVERSITY OF PAKISTAN



COURSE: STA 408
DESIGN & ANALYSIS OF EXPERIMENT

THE LATIN SQUARE DESIGN

LECTURE NO.5

DESCRIPTION

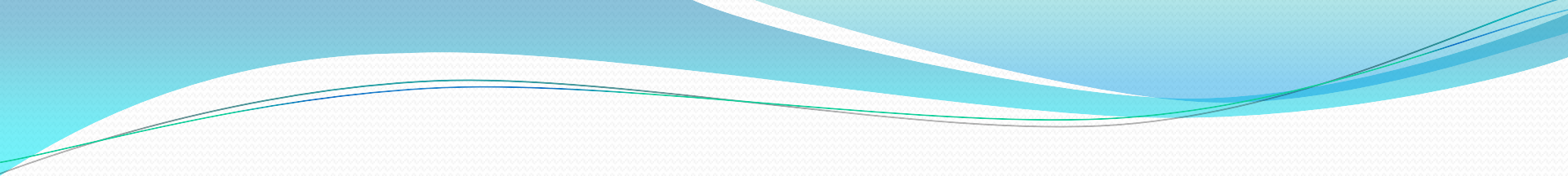
When the variation is found in two directions, it becomes necessary to remove these two sources of variation simultaneously. This is done by simultaneous blocking of experimental units in two mutually perpendicular directions called rows and columns. Since each row and each column is a complete block, the grouping for a balanced arrangement is performed by imposing the restriction that each treatment must appear “once and only once” in each row and in each column.

If there are “k” treatments then the experimental area will be divided into “k” rows and “k” columns in $k \times k$ plots or experimental units. As the experiment is laid out in a square pattern and the treatments are then assigned at random to plots or experimental units. Such a double blocking of experimental units and a corresponding double random assignment of treatments is called **“Latin Square Design”**.

LAYOUTS

Latin squares are always constructed by rotations, e.g. in the case of 4 treatments A, B, C and D, we get

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C



A large number of distinct Latin squares can be derived by interchanging rows and columns of a standard Latin square, where a **standard Latin square** is a Latin square in which the treatments in the first row and first column are arranged in alphabetical or numerical order.

A standard Latin square is also called a reduced Latin square.

- There is only one standard square for a 2 x 2 Latin square as,

A	B
B	A

- And for 3 x 3 Latin square we have also one standard square such as,

A	B	C
B	C	A
C	A	B

- But for a 4 x 4 Latin square, there are four possible standard squares as,

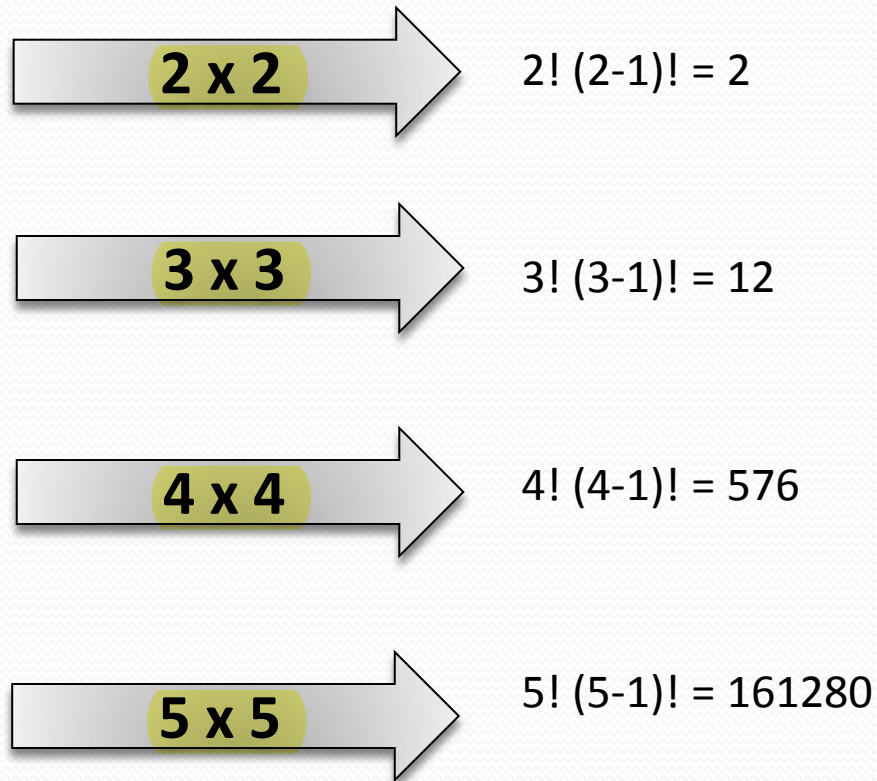
A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

A	B	C	D
B	A	D	C
C	D	B	A
D	C	A	B

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

A	B	C	D
B	D	A	C
C	A	D	B
D	C	B	A

Each Latin square of $k \times k$ yields $k! (k-1)!$ Different squares by permuting the rows except first and all the k columns. According we may obtain from a square of design.



MATHEMATICAL MODEL

In Latin square design, the observation $Y_{ij(h)}$,

$$Y_{ij(h)} = \mu + R_i + C_j + \tau_h + e_{ij(h)}$$

where $i, j, h = 1, 2, 3, \dots, k$.

The μ denote the general mean

R_i is the effect of the i th row

C_j the effect of the j th column

τ_h the effect of the h th treatment

$e_{ij(h)}$ is the experimental error, that is assumed to be normally and independently distributed with mean zero and common variance σ^2 .

- ① formulation of Hypothesis
- ② level of significance
- ③ test Statistic
- ④ Calculation
- ⑤ Critical Region
- ⑥ Conclusion

STATISTICAL ANALYSIS

1-Hypothesis

$$H_0: \tau_h = 0$$

$$H_1: \neq$$

1

$$H_0: \mu_{..1} = \mu_{..2} = \mu_{..3} \dots$$

H_1 : Not all treatment means are equal

3

$$H'_0: \mu_{.1} = \mu_{.2} = \mu_{.3} \dots \quad H_0: \tau_j = 0$$

H'_1 : Not all column means are equal $H_1: \neq$

2

$$H''_0: \mu_{1..} = \mu_{2..} = \dots \quad H_0: \tau_i = 0$$

H''_1 : Not all row means are equal $H_1: \neq$

2-Level of Significance

$$\alpha = ?$$

3-Test Statistics

$$F_1 = \frac{s_t^2}{s_e^2}$$

$$F_2 = \frac{s_r^2}{s_e^2}$$

$$F_3 = \frac{s_c^2}{s_e^2}$$

4-Calculation

The data obtained from a $k \times k$ Latin square design may be arranged in tabular form as,

		COLUMNS						R_i	
		1	2	3	4		K	Total	Means
ROWS	1	Y_{11A}	Y_{12B}	Y_{13C}	Y_{14D}			$R_{1..}$	$\bar{Y}_{1..}$
	2	:	:	:	:			$R_{2..}$	$\bar{Y}_{2..}$
	3	:	:	:	:			$R_{3..}$	$\bar{Y}_{3..}$
	4	Y_{41D}	Y_{42A}	Y_{43B}	Y_{44C}				
	:	:	:	:	:				
	:	:	:	:	:				
	K	:	:	:	:				
Total C_j		$C_{.1.}$	$C_{.2.}$					$G = T_{...}$	
Means		$\bar{Y}_{.1.}$	$\bar{Y}_{.2.}$	$\bar{Y}_{.3.}$	$\bar{Y}_{.4.}$				$\bar{Y}_{...}$

COMPUTATIONS FOR THE TREATMENT MEANS

	A	B	C	...	
	-	-	-	-	
	-	-	-	-	
	-	-	-	-	
	-	-	-	-	
$T_{kth}(T_h)$	$T_{..h}$	$T_{..h}$	$T_{..h}$	$T_{..h}$	
Means	$\bar{y}_{..h}$	$\bar{y}_{..h}$	...	..	

PARTITIONING OF TOTAL SUM OF SQUARES

$$\begin{array}{ccccccc}
 \nearrow \text{ISS} & & \nearrow \text{RSS} & & \nearrow \text{CSS} & & \nearrow \text{TreatSS} & & \nearrow \text{ESS} \\
 \text{Total SS} = & \text{Row SS} & + & \text{Column SS} & + & \text{Treatment SS} & + & \text{Error SS} \\
 \nwarrow & \nwarrow & & \nwarrow & & \nwarrow & & \nwarrow & \nwarrow \\
 n = K \times K & & & & & & & & \\
 = K^2 & & & & & & & & \\
 & K^2 - 1 & & K - 1 & & K - 1 & & K - 1 & (K - 1)(K - 2)
 \end{array}$$

$$K^2 - 1 - 3(K - 1)$$

$$K^2 - 1 - 3K + 3$$

$$K^2 - 3K + 2$$

$$K^2 - 2K - K + 2$$

$$K(K - 2) - 1(K - 2)$$

$$(K - 1)(K - 2)$$

$$\underline{\underline{ESS = ISS - RSS - CSS - \text{TreatSS}}}$$

$$C.F = \frac{(\bar{T} \dots)^2}{n} = \frac{G^2}{K^2}$$

$$TSS = \sum_i \sum_j (Y_{ij(h)})^2 - C.F$$

$$RSS = \frac{\sum_i R_i^2}{K} - C.F$$

$$CSS = \frac{\sum_j C_j^2}{K} - C.F$$

$$TreatSS = \frac{\sum_h T_h^2}{n} - C.F$$

ANOVA-TABLE for k x k Latin Square Design

Source of Variation	Degrees of freedom	Sum of Squares	Mean Square	F-Test
Between Treatment	$\nu_1 = k-1$	$Treat SS = \frac{\sum T_h^2}{k} - C.F$	$\delta_t^2 = \frac{Treat SS}{\nu_1}$	$F_1 = \delta_t^2 / \delta_e^2$
B/w Rows	$\nu_2 = k-1$	$RSS_2 = \frac{\sum R_i^2}{k} - C.F$	$\delta_r^2 = \frac{RSS}{\nu_2}$	$F_2 = \delta_r^2 / \delta_e^2$
B/w columns	$\nu_3 = k-1$	$CSS = \frac{\sum C_j^2}{k} - C.F$	$\delta_c^2 = \frac{CSS}{\nu_3}$	$F_3 = \delta_c^2 / \delta_e^2$
Within Errors	$\nu_4 = (k-1)(k-2)$	$TSS - Treat SS - CSS - RSS = ESS$	$\delta_e^2 = \frac{ESS}{\nu_4}$	
Total	$k^2 - 1$	$\sum \sum (y_{ij})^2 - C.F$	—	

5-Critical region

Reject the null hypotheses when

$$F_1 \geq F_{\alpha, (v_1, v_4)}$$

$$F_2 \geq F_{\alpha, (v_2, v_4)}$$

$$F_3 \geq F_{\alpha, (v_3, v_4)}$$

6-Conclusion

EXAMPLE

The atmosphere in 4 different districts of a large town was sampled, the samples being taken at 4 different heights. Four different tests for the presence of a certain chemical were made on the samples. The arrangement is shown in the following table with the percentage by weight of the chemical as determined by the tests. Letter denote the difference tests.

		DISTRICTS			
		1	2	3	4
HEIGHTS	1	A 8	B 5.3	C 4.1	D 5
	2	D 6.8	A 4.9	B 4.1	C 3.2
	3	B 6.3	C 4.7	D 4.0	A 5
	4	C 5.7	D 3.3	A 4.0	B 4.2

Is there any evidence of significant variation from district to district and between heights in the percentage of the chemical present in the atmosphere..

1-Hypothesis

$$H_0: \mu_{..1} = \mu_{..2} = \mu_{..3} = \mu_{..4}$$

H_1 : Not all treatments means are equal.

$$H'_0: \mu_{1..} = \mu_{2..} = \mu_{3..} = \mu_{4..}$$

H'_1 : Not all row means are equal

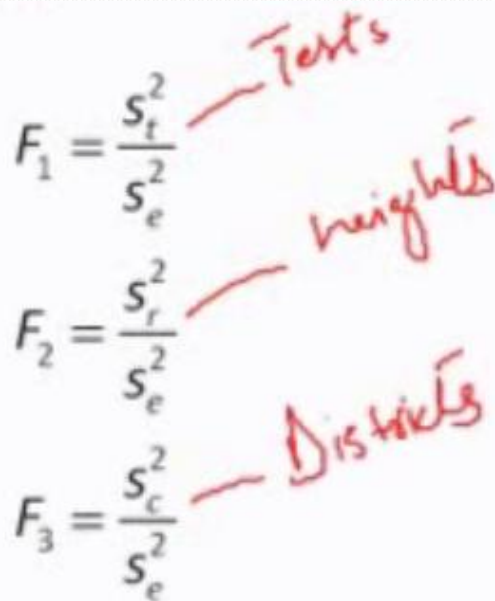
$$H''_0: \mu_{.1} = \mu_{.2} = \mu_{.3} = \mu_{.4}$$

H''_1 : Not all column means are equal.

2-Level of Significance

$$\alpha = 0.05$$

3-Test Statistic



Handwritten formulas for F-statistics with red annotations:

- $F_1 = \frac{s_t^2}{s_e^2}$ — Tests
- $F_2 = \frac{s_r^2}{s_e^2}$ — Heights
- $F_3 = \frac{s_c^2}{s_e^2}$ — Districts

4-Calculations

HEIGHTS	DISTRICTS				R_i	R_i^2
	1	2	3	4		
1	A 8	B 5.3	C 4.1	D 5	22.4	501.76
2	D 6.8	A 4.9	B 4.1	C 3.2	19.0	361.00
3	B 6.3	C 4.7	D 4.0	A 5	20.0	400
4	C 5.7	D 3.3	A 4.0	B 4.2	17.2	295.84
C_j	26.8	18.2	16.2	17.4	78.6	1558.60
C_j^2	718.4	331.24	262.44	302.76	1614.68	—
$\sum \sum f_{ij}(h)$	182.42	85.08	65.62	77.88	411.00	

	TESTS			
	A	B	C	D
	8	5.3	4.1	5
	4.9	4.1	3.2	6.8
	5	6.3	4.7	4.0
	4.0	4.2	5.7	3.3
T_n	81.9	19.9	17.7	19.1
T_n^2	479.61	396.01	313.29	364.81

Total

78.6

1553.72

$$TSS = \sum_i \sum_j y_{ij}^2 - C.F$$

$$= 411.00 - 386.1225$$

$$= 24.8775$$

$$RSS = \frac{\sum R_i^2}{K} - C.F = \frac{1558.60}{4} - 386.1225$$

$$RSS = 3.5275$$

$$C.F = \frac{G^2}{K^2} = \frac{(18.6)^2}{16}$$

$$C.F = 386.1225$$

$$CSS = \frac{\sum C_j^2}{K} - C.F = \frac{1614.68}{4} - 386.1225$$

$$CSS = 17.5475$$

$$TreatSS = \frac{\sum T_h^2}{K} - C.F = \frac{1553.72}{4} - 386.1225$$

$$TreatSS = 2.3075$$

$$ESS = TSS - RSS - CSS - TreatSS$$

$$ESS = 1.4950$$

ANOVA-TABLE for 4 x 4 Latin Square Design

Source of Variation	Degrees of freedom	Sum of Squares	Mean Square	F-Test
B/w Treatments	$v_1 = 4 - 1 = 3$	Treat SS = 2.3075	$b_1^2 = 5.8492$	$F_1 = 23.47$
B/w Rows	$v_2 = 3$	RSS = 3.5275	$b_2^2 = 1.1758$	$F_2 = 4.72$
B/w Columns	$v_3 = 3$	CSS = 17.5475	$b_3^2 = 0.7692$	$F_3 = 3.09$
Error	$v_4 = 6$	1.4950	$b_4^2 = 0.2490$	
Total	15	24.8775		

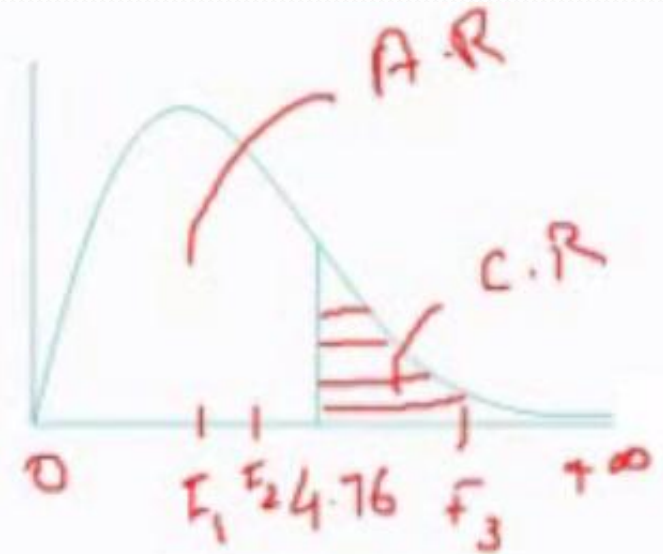
5-Critical Region

Reject Null hypothesis if

$$\begin{matrix} 3.09 \\ \leftarrow F_1 > F_{0.05}(3,6) = 4.76 \end{matrix}$$

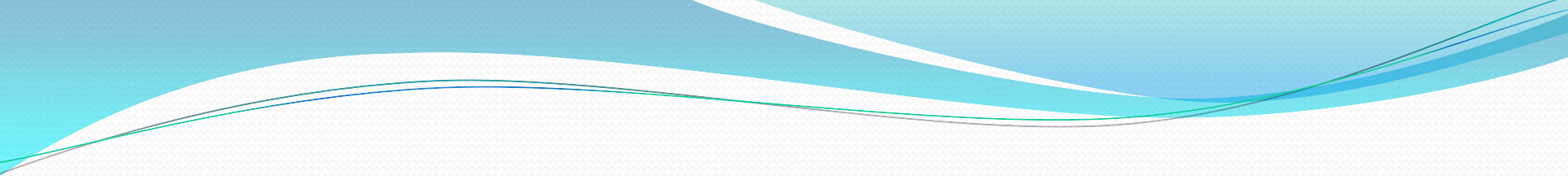
$$\begin{matrix} 4.72 \\ \leftarrow F_2 > F_{0.05}(3,6) = 4.76 \end{matrix}$$

$$\begin{matrix} 23.47 \\ \leftarrow F_3 > F_{0.05}(3,6) = 4.76 \end{matrix}$$



6-Conclusion

By comparing the computed values of F with the corresponding critical F , we conclude that the variation among the districts is significant and the variation among the heights is insignificant. Also, there is no difference between the sensitivity of the test.



THANK YOU