3.11 Latin Square Designs

- The experimenter is concerned with a single factor having p levels. However, variability from two other sources can be controlled in the experiment.

- If we can control the effect of these other two variables by grouping experimental units into blocks having the same number of treatment levels as the factor of interest, then a *latin square design* may be appropriate.

- Consider a square with p rows and p columns corresponding to the p levels of each blocking variable. If we assign the p treatments to the rows and columns so that each treatment appears exactly once in each row and in each column, then we have a $p \times p$ **latin square design**.

- It is called a "latin" square because we assign "latin" letters A, B, C, ... to the treatments. Examples of a 4×4 and a 6×6 latin square designs are

		Column			
		1	2	3	4
Rows	1	A	C	B	D
	2	D	A	C	B
	3	B	D	A	C
	4	C	B	D	A

		Column					
		1	2	3	4	5	6
Rows	1	A	B	C	D	E	F
	2	B	C	D	E	F	A
	3	C	D	E	F	A	B
	4	D	E	F	A	B	C
	5	E	F	A	B	C	D
	6	F	A	B	C	D	E

- By blocking in two directions, the MSE will (in general) be reduced. This makes detection of significant results for the factor of interest more likely.

- The experimental units should be arranged so that differences among row and columns represent anticipated/potential sources of variability.

  - In industrial experiments, one blocking variable is often based on units of time. The other blocking variable may represent an effect such as machines or operators.

		Machines			
		1	2	3	4
Six-Hour Work Shift	1	A	C	B	D
	2	D	A	C	B
	3	B	D	A	C
	4	C	B	D	A

		Operator				
		1	2	3	4	5
Day	M	A	B	C	D	E
	Tu	B	C	D	E	A
	W	C	D	E	A	B
	Th	D	E	A	B	C
	F	E	A	B	C	D

  - In agricultural experiments, the experimental units are subplots of land. We would then have the subplots laid out so that soil fertility, moisture, and other sources of variation in two directions are controlled.

  - In greenhouse experiments, the subplots are often laid out in a continuous line. In this case, the rows may be blocks of p adjacent subplots and the columns specify the order within each row block.

Rep. 1	Rep. 2	Rep. 3	Rep. 4	Rep. 5	Rep. 6	Rep. 7
A D G B E C F	G F B C A E D	B C D G F A E	E G A D C F B	C B F E D G A	F E C A B D G	D A E F G B C

After converting Rep. 1 to Rep. 7 into row blocks, we get

		Column						
		1	2	3	4	5	6	7
	1	A	D	G	B	E	C	F
	2	G	F	B	C	A	E	D
	3	B	C	D	G	F	A	E
Reps	4	E	G	A	D	C	F	B
	5	C	B	F	E	D	G	A
	6	F	E	C	A	B	D	G
	7	D	A	E	F	G	B	C

- Like the RCBD, the latin square design is another design with restricted randomization. Randomization occurs with the initial selection of the latin square design from the set of all possible latin square designs of dimension p and then randomly assigning the treatments to the letters A, B, C, ... The following notation will be used:

  p = the number of treatment levels, row blocks and column blocks.

  y_{ijk} = the observation for the j^{th} treatment within the i^{th} row and k^{th} column.

  $N = p^2$ = total number of observations

  $y_{\cdot\cdot\cdot}$ = the sum of all p^2 observations

  R_i = the sum for row block i

  C_k = the sum for column block k

  T_j = the sum for treatment j

  $\overline{y}_{\cdot\cdot}$ = the grand mean of all observations = $y_{\cdot\cdot\cdot}/p^2$

  $\overline{y}_{i\cdot\cdot}$ = the i^{th} row block mean

  $\overline{y}_{\cdot\cdot k}$ = the k^{th} column block mean

  $\overline{y}_{\cdot j\cdot}$ = the j^{th} treatment mean

- There are three subscripts (i, j, k), but we only need to sum over two subscripts.

- The standard statistical model associated with a latin square design is:

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk} \tag{18}$$

  where μ is the baseline mean, α_i is the block effect associated with row i, β_k is the block effect associated with column k, τ_j is the j^{th} treatment effect, and ϵ_{ijk} is a random error which is assumed to be $IIDN(0, \sigma^2)$.

- To get estimates for the parameters in (18), we need to impose <u>three</u> constraints. If we assume $\sum_{i=1}^{p} \alpha_i = 0$, $\sum_{j=1}^{p} \tau_j = 0$, and $\sum_{k=1}^{p} \beta_k = 0$, then the least squares estimates are

$$\widehat{\mu} = \overline{y}_{\cdot\cdot\cdot} \qquad \widehat{\alpha}_i = \overline{y}_{i\cdot\cdot} - \overline{y}_{\cdot\cdot\cdot} \qquad \widehat{\beta}_k = \overline{y}_{\cdot\cdot k} - \overline{y}_{\cdot\cdot\cdot} \qquad \widehat{\tau}_j = \overline{y}_{\cdot j\cdot} - \overline{y}_{\cdot\cdot\cdot}$$

- Substitution of the estimates into (18) yields

$$\begin{aligned} y_{ijk} &= \widehat{\mu} + \widehat{\alpha}_i + \widehat{\tau}_j + \widehat{\beta}_k + e_{ijk} \\ &= \overline{y}_{\cdot\cdot\cdot} + (\overline{y}_{i\cdot\cdot} - \overline{y}_{\cdot\cdot\cdot}) + (\overline{y}_{\cdot j\cdot} - \overline{y}_{\cdot\cdot\cdot}) + (\overline{y}_{\cdot\cdot k} - \overline{y}_{\cdot\cdot\cdot}) + e_{ijk} \end{aligned} \tag{19}$$

  where e_{ijk} is the ijk^{th} residual from a latin square design and

$$e_{ijk} = y_{ijk} - \overline{y}_{i\cdot\cdot} - \overline{y}_{\cdot j\cdot} - \overline{y}_{\cdot\cdot k} + 2\,\overline{y}_{\cdot\cdot\cdot}$$

3.11.1 The ANOVA for a Latin Square Design

- Degrees of freedom (df): (Treatment df) = (Row df) = (Column df) = $p-1$
- $SStrt$ = the treatment sum of squares

  MS_{trt} = the treatment mean square = $SS_{trt}/(p-1)$
- SS_{row} = the sum of squares for rows

  MS_{row} = the mean square for rows = $SS_{row}/(p-1)$
- SS_{col} = the sum of squares for columns

  MS_{col} = the mean square for columns = $SS_{col}/(p-1)$
- SS_E = the error sum of squares

  The SS_E degrees of freedom = $(p-1)(p-2)$

  MS_E = the mean square error = $SS_E/[(p-1)(p-2)]$
- SS_{total} = the corrected total sum of squares

  The SS_{total} degrees of freedom = $N-1=p^2-1$

  The total sum of squares for the latin square design is partitioned into 4 components:

$$SS_{total} \quad = \quad SS_{row} \;+\; SS_{trt} \;+\; SS_{col} \;+\; SS_E$$

- Formulas to calculate SS_{total}, SS_{row}, SS_{trt} and SS_{col}:

$$SS_{total}=\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ijk}-\overline{y}_{..})^2=\sum_{i=1}^{p}\sum_{j=1}^{p}y_{ijk}^2-\frac{y_{...}^2}{p^2} \qquad SS_{row}=\sum_{i=1}^{p}p(\overline{y}_{i..}-\overline{y}_{..})^2=\sum_{i=1}^{p}\frac{R_i^2}{p}-\frac{y_{...}^2}{p^2}$$

$$SS_{trt}=\sum_{j=1}^{p}p(\overline{y}_{.j.}-\overline{y}_{..})^2=\sum_{j=1}^{p}\frac{T_j^2}{p}-\frac{y_{...}^2}{p^2} \qquad SS_{col}=\sum_{k=1}^{p}p(\overline{y}_{..k}-\overline{y}_{..})^2=\sum_{k=1}^{p}\frac{C_k^2}{p}-\frac{y_{...}^2}{p^2}$$

$$SS_E \;=\; SS_{total} \;-\; SS_{row} \;-\; SS_{trt} \;-\; SS_{col} \qquad \frac{y_{...}^2}{p^2}=\frac{y_{...}^2}{N}=\text{the correction factor.}$$

Latin Square Design ANOVA Table

Source of Variation	Sum of Squares	d.f.	Mean Square	F Ratio
Treatments	SS_{trt}	$p-1$	MS_{trt}	$\frac{MS_{trt}}{MS_E}$
Rows	SS_{row}	$p-1$	MS_{row}	
Columns	SS_{col}	$p-1$	MS_{col}	
Error	SS_E	$(p-1)(p-2)$	MS_E	
Total	SS_{total}	p^2-1		

- Note that there are two **restrictions on randomization** with latin square designs: (i) a row restriction that all treatments must appear in each row and (ii) a column restriction that all treatments must appear in each column.

- Because of these restrictions on randomization, there is no F-test for the equality of row block effects and no F-test for the equality of column block effects. This is not a problem, because the factor of interest is the treatment, and there is an F-test for treatments.

3.11.2 Latin Square Example (Peanut Varieties)

Example: A plant biologist conducted an experiment to compare the yields of 4 varieties of peanuts (A, B, C, D). A plot of land was divided into 16 subplots (4 rows and 4 columns) The following latin square design was run. The responses are given in the table to the right.

Treatment (peanut variety)

	Column			
Row	E	EC	WC	W
N	C	A	B	D
NC	A	B	D	C
SC	B	D	C	A
S	D	C	A	B

Response (yield)

	Column			
Row	E	EC	WC	W
N	26.7	19.7	29.0	29.8
NC	23.1	21.7	24.9	29.0
SC	29.3	20.1	29.0	27.3
S	25.1	17.4	28.7	35.1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	298.0056250	33.1117361	8.28	0.0091
Error	6	23.9837500	3.9972917		
Corrected Total	15	321.9893750			

R-Square	Coeff Var	Root MSE	yield Mean
0.925514	7.691552	1.999323	25.99375

Source	DF	Type III SS	Mean Square	F Value	Pr > F
columns	3	245.9118750	81.9706250	20.51	0.0015
rows	3	9.4268750	3.1422917	0.79	0.5439
peanut	3	42.6668750	14.2222917	3.56	0.0870

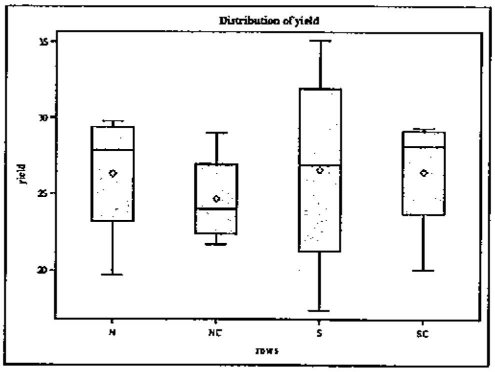
→ FAIL TO REJECT H_0 FOR $\alpha = .05$

REJECT H_0 FOR $\alpha = .10$ (NONCONCLUSIVE EVIDENCE)

$H_0: \tau_1 = \tau_2 = \tau_3 = \tau_4$

$H_1: \tau_i \neq \tau_{i'}$ FOR SOME $i \neq i'$

ROWS

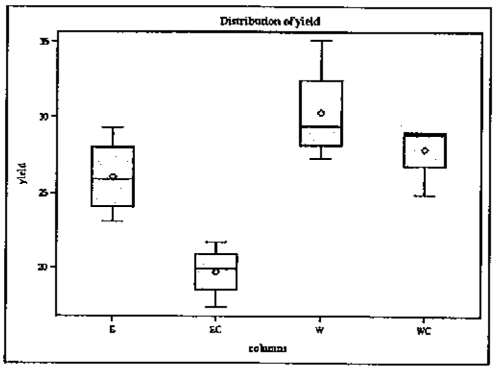


Level of rows	N	yield Mean	yield Std Dev
N	4	26.3000000	4.59202207
NC	4	24.6750000	3.16688596
S	4	26.5750000	7.38348382
SC	4	26.4250000	4.30764824

MS_{ROWS} IS SMALL

↑ LITTLE VARIABILITY

COLUMNS



Level of columns	N	yield Mean	yield Std Dev
E	4	26.0500000	2.61979643
EC	4	19.7250000	1.77458915
W	4	30.3000000	3.36551135
WC	4	27.9000000	2.00499377

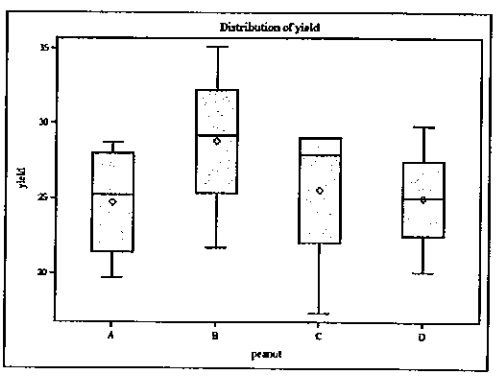
$MS_{COLUMNS}$ IS LARGE

↖ LARGE VARIABILITY

Tukey's Studentized Range (HSD) Test for reaction

Alpha	0.1
Error Degrees of Freedom	6
Error Mean Square	3.997292
Critical Value of Studentized Range	4.06509
Minimum Significant Difference	4.0637

PEANUT



$P = .0870$

Tukey Grouping		Mean	N	peanut
	A	28.775	4	B
	A			
B	A	25.525	4	C
B	A			
B	A	24.975	4	D
B				
B		24.700	4	A

Means with the same letter are not significantly different.

NOTE: WE WOULD REJECT $H_0: \bar{\mu}_{A.} = \bar{\mu}_{.B.}$ DESPITE THE F-TEST P-VALUE $= .0870$

$\bar{\mu}_{.j.}$ = MEAN FOR TREATMENT j AVERAGED OVER ROWS AND COLUMNS

diag
DATA SET

Obs	rows	columns	peanut	yield	pred	resid
1	N	E	C	26.7	25.8875	0.8125
2	N	EC	A	19.7	18.7375	0.9625
3	N	WC	B	29.0	30.9875	-1.9875
4	N	W	D	29.8	29.5875	0.2125
5	NC	E	A	23.1	23.4375	-0.3375
6	NC	EC	B	21.7	21.1875	0.5125
7	NC	WC	D	24.9	25.5625	-0.6625
8	NC	W	C	29.0	28.5125	0.4875
9	SC	E	B	29.3	29.2625	0.0375
10	SC	EC	D	20.1	19.1375	0.9625
11	SC	WC	C	29.0	27.8625	1.1375
12	SC	W	A	27.3	29.4375	-2.1375
13	S	E	D	25.1	25.6125	-0.5125
14	S	EC	C	17.4	19.8375	-2.4375
15	S	WC	A	28.7	27.1875	1.5125
16	S	W	B	35.1	33.6625	1.4375

$y_{ijk} - \hat{y}_{ijk}$

y_{ijk} $\hat{y}_{ijk} = \hat{\mu} + \hat{\alpha}_i + \hat{\tau}_j + \hat{\beta}_k$

SAS Code for Latin Square Design Example

```
DM 'LOG; CLEAR; OUT; CLEAR;';

ODS GRAPHICS ON;
ODS PRINTER PDF file='C:\COURSES\ST541\LATIN.PDF';
OPTIONS NODATE NONUMBER;

DATA in;
   DO rows    = 'N ', 'NC', 'SC', 'S ';
   DO columns = 'E ', 'EC', 'WC', 'W ';
      INPUT peanut $ yield @@; output;
   END; END;
LINES;
C 26.7  A 19.7  B 29.0  D 29.8
A 23.1  B 21.7  D 24.9  C 29.0
B 29.3  D 20.1  C 29.0  A 27.3
D 25.1  C 17.4  A 28.7  B 35.1
;
PROC GLM DATA=in PLOTS=(ALL);
     CLASS columns rows peanut;
     MODEL yield = columns rows peanut / SS3;
     MEANS rows columns;
     MEANS peanut / TUKEY ALPHA=.10;
     OUTPUT OUT=diag P=pred R=resid;
TITLE 'ANALYSIS FOR LATIN SQUARE DESIGN';

PROC PRINT DATA=diag;
RUN;
```

ORDER DOES NOT MATTER IF NO DESIGN OBSERVATION IS MISSING. (TYPE I VS = TYPE III SS)

GENERATE A DATA SET 'diag' THAT CONTAINS RESIDUALS AND PREDICTED VALUES.

3.12 Matrix Forms for the Latin Square Design

- Consider the 4×4 Peanut Variety Latin Square Design Example
- Model: $y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk}$ for $i, j, k = 1, 2, 3, 4$ $\quad \epsilon_{ijk} \sim N(0, \sigma^2)$
- Assume (i) $\sum_{i=1}^4 \alpha_i = 0$, (ii) $\sum_{j=1}^4 \tau_j = 0$ and (iii) $\sum_{k=1}^4 \beta_k = 0$.

Using the Alternate Approach: Keeping $3p + 1$ Columns

$$
X = \begin{array}{c}
\begin{array}{ccccccccccccc}
 & \multicolumn{4}{c}{\text{Row}} & & & & & \multicolumn{4}{c}{\text{Column}} \\
\mu & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \tau_1 & \tau_2 & \tau_3 & \tau_4 & \beta_1 & \beta_2 & \beta_3 & \beta_4
\end{array} \\
\left[\begin{array}{c|cccc|cccc|cccc}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 1^{C} & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1^{A} & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1^{B} & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1^{D} & 0 & 0 & 0 & 1 \\
\hline
1 & 0 & 1 & 0 & 0 & 1^{A} & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1^{B} & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1^{D} & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1^{C} & 0 & 0 & 0 & 0 & 1 \\
\hline
1 & 0 & 0 & 1 & 0 & 0 & 1^{B} & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1^{D} & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1^{C} & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1^{A} & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\hline
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1^{D} & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 1^{C} & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 1^{A} & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 1^{B} & 0 & 0 & 0 & 0 & 0 & 1 \\
\hline
0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1
\end{array}\right]
\end{array}
\begin{array}{l}
\\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\
\sum \alpha_i = 0 \\
\sum \tau_j = 0 \\
\sum \beta_k = 0
\end{array}
\qquad
y = \begin{bmatrix} 26.7 \\ 19.7 \\ 29.0 \\ 29.8 \\ 23.1 \\ 21.7 \\ 24.9 \\ 29.0 \\ 29.3 \\ 20.1 \\ 29.0 \\ 27.3 \\ 25.1 \\ 17.4 \\ 28.7 \\ 35.1 \\ 0 \\ 0 \\ 0 \end{bmatrix}
$$

$$
X'X = \begin{bmatrix}
16 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \\
4 & 5 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
4 & 1 & 5 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
4 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
4 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
4 & 1 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
4 & 1 & 1 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 1 & 1 \\
4 & 1 & 1 & 1 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 1 \\
4 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 1 \\
4 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 5 & 1 & 1 & 1 \\
4 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 5 & 1 & 1 \\
4 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 5 & 1 \\
4 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 5
\end{bmatrix}
$$

$$
(X'X)^{-1} = \frac{1}{16}\begin{bmatrix}
4 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
-1 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4
\end{bmatrix}
$$

$$
X'y = \begin{bmatrix} 415.9 \\ 105.2 \\ 98.7 \\ 105.7 \\ 106.3 \\ 98.8 \\ 115.1 \\ 102.1 \\ 99.9 \\ 104.2 \\ 78.9 \\ 111.6 \\ 121.2 \end{bmatrix}
\qquad
(X'X)^{-1}X'y = \begin{bmatrix} 25.9938 \\ 0.3062 \\ -1.3188 \\ 0.4312 \\ 0.5813 \\ -1.2938 \\ 2.7812 \\ -0.4688 \\ -1.0188 \\ 0.0562 \\ -6.2688 \\ 1.9063 \\ 4.3062 \end{bmatrix}
\begin{array}{l} \\ \left.\begin{matrix}\\ \\ \\ \\ \end{matrix}\right\} \sum\hat{\alpha}_i = 0 \\ \left.\begin{matrix}\\ \\ \\ \\ \end{matrix}\right\} \underline{\sum\hat{\tau}_j = 0} \\ \left.\begin{matrix}\\ \\ \\ \\ \end{matrix}\right\} \sum\hat{\beta}_k = 0 \end{array}
\begin{bmatrix} \widehat{\mu} \\ \widehat{\alpha}_1 \\ \widehat{\alpha}_2 \\ \widehat{\alpha}_3 \\ \widehat{\alpha}_4 \\ \widehat{\tau}_1 \\ \widehat{\tau}_2 \\ \widehat{\tau}_3 \\ \widehat{\tau}_4 \\ \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \widehat{\beta}_3 \\ \widehat{\beta}_4 \end{bmatrix}
\begin{array}{l} \bar{y}_{\cdot\cdot\cdot} \\ \left.\begin{matrix}\\ \\ \\ \\ \end{matrix}\right\} \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot} \\ \left.\begin{matrix}\\ \\ \\ \\ \end{matrix}\right\} \bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot} \\ \left.\begin{matrix}\\ \\ \\ \\ \end{matrix}\right\} \bar{y}_{\cdot\cdot k} - \bar{y}_{\cdot\cdot\cdot} \end{array}
$$

NOTE

3.12.1 Selection of a Latin Square Design

- Randomization with a latin square design occurs by (i) randomly selecting a generating design, (ii) randomly permuting the p columns and (iii) randomly permuting the last $p-1$ rows.

- For some values of p, the number of unique designs that can be generated from permutations of a generating design are not all equal. This will be true for $p=5$ and $p=6$.

- For $p=3$, there is only one generating design

A	B	C
B	C	A
C	A	B

  - Randomly permute the three columns.
  - Randomly permute rows 2 and 3.

- For $p=4$, there are 4 generating designs

(a)				(b)				(c)				(d)			
A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
B	A	D	C	B	C	D	A	B	D	A	C	B	A	D	C
C	D	B	A	C	D	A	B	C	A	D	B	C	D	A	B
D	C	A	B	D	A	B	C	D	C	B	A	D	C	B	A

  - Randomly select generating design (a), (b), (c), or (d).
  - Randomly permute the four columns.
  - Randomly permute rows 2 to 4.

- For $p=5$, there are 2 generating designs

(a) $(1-50)$					(b) $(51-56)$				
A	B	C	D	E	A	B	C	D	E
B	A	D	E	C	B	C	D	E	A
C	E	A	B	D	C	D	E	A	B
D	C	E	A	B	D	E	A	B	C
E	D	B	C	A	E	A	B	C	D

  - Randomly select a number D between 1 and 56.
    If $1 \le D \le 50$, then select generating design (a).
    If $51 \le D \le 56$, then select generating design (b).
  - Randomly permute the five columns.
  - Randomly permute rows 2 to 5.

- For $p=6$, there are 17 generating designs (I, II, ..., XVII). These generating designs are shown on the next page.
  - Randomly select a number D between 1 and 9408.
  - Find the generating design with interval that contains D.
  - Randomly permute the six columns.
  - Randomly permute rows 2 to 6.

I

A	B	C	D	E	F
B	C	F	A	D	E
C	F	B	E	A	D
D	E	A	B	F	C
E	A	D	F	C	B
F	D	E	C	B	A

0001–1080
1081–2160

II

A	B	C	D	E	F
B	C	F	E	A	D
C	F	B	A	D	E
D	E	A	B	F	C
E	A	D	F	C	B
F	D	E	C	B	A

2161–3240

III

A	B	C	D	E	F
B	A	F	E	C	D
C	F	B	A	D	E
D	C	E	B	F	A
E	D	A	F	B	C
F	E	D	C	A	B

3241–4320

IV

A	B	C	D	E	F
B	A	E	F	C	D
C	F	B	A	D	E
D	E	A	B	F	C
E	D	F	C	B	A
F	C	D	E	A	B

4321–5400

V

A	B	C	D	E	F
B	A	E	C	F	D
C	F	B	A	D	E
D	E	F	B	C	A
E	D	A	F	B	C
F	C	D	E	A	B

5401–5940
5941–6480

VI

A	B	C	D	E	F
B	A	F	E	C	D
C	F	B	A	D	E
D	E	A	B	F	C
E	C	D	F	B	A
F	D	E	C	A	B

6481–7020

VII

A	B	C	D	E	F
B	C	D	E	F	A
C	E	A	F	B	D
D	F	B	A	C	E
E	D	F	B	A	C
F	A	E	C	D	B

7021–7560

VIII

A	B	C	D	E	F
B	A	E	F	C	D
C	F	A	E	D	B
D	C	B	A	F	E
E	D	F	C	B	A
F	E	D	B	A	C

7561–7920
7921–8280

IX

A	B	C	D	E	F
B	A	E	F	C	D
C	F	A	B	D	E
D	E	B	A	F	C
E	D	F	C	B	A
F	C	D	E	A	B

8281–8640

X

A	B	C	D	E	F
B	C	F	A	D	E
C	F	B	E	A	D
D	A	E	B	F	C
E	D	A	F	C	B
F	E	D	C	B	A

8641–8820

XI

A	B	C	D	E	F
B	C	A	F	D	E
C	A	B	E	F	D
D	F	E	B	A	C
E	D	F	C	B	A
F	E	D	A	C	B

8821–8940
8941–9060

XII

A	B	C	D	E	F
B	C	A	E	F	D
C	A	B	F	D	E
D	E	F	B	A	C
E	F	D	A	C	B
F	D	E	C	B	A

9061–9180

XIII

A	B	C	D	E	F
B	C	A	F	D	E
C	A	B	E	F	D
D	F	E	B	A	C
E	D	F	A	C	B
F	E	D	C	B	A

9181–9240

XIV

A	B	C	D	E	F
B	C	A	E	F	D
C	A	B	F	D	E
D	F	E	B	A	C
E	D	F	C	B	A
F	E	D	A	C	B

9241–9280

XV

A	B	C	D	E	F
B	A	F	E	D	C
C	D	A	B	F	E
D	F	E	A	C	B
E	C	B	F	A	D
F	E	D	C	B	A

9281–9316
9317–9352

XVI

A	B	C	D	E	F
B	A	E	C	F	D
C	E	A	F	D	B
D	C	F	A	B	E
E	F	D	B	A	C
F	D	B	E	C	A

9353–9388

XVII

A	B	C	D	E	F
B	C	A	F	D	E
C	A	B	E	F	D
D	E	F	A	B	C
E	F	D	C	A	B
F	D	E	B	C	A

9389–9408

FIG. 6. The 6×6 Latin squares.

No enumeration has as yet been made of squares larger than 6×6. In Fig. 7 we give six squares, with sides from 7 to 12, from which any square of the transformation sets which contain them may be generated by the permutation of all rows, columns, and letters amongst themselves. These transformation sets, or even the smaller sets generated by the permutation of rows and columns, or either and letters, will give sets of squares amply large enough to serve all agricultural purposes.

A	B	C	D	E	F	G
B	D	E	F	A	G	C
C	G	F	E	B	A	D
D	E	A	B	G	C	F
E	C	B	G	F	D	A
F	A	G	C	D	E	B
G	F	D	A	C	B	E

7×7

A	B	C	D	E	F	G	H
B	C	A	E	F	D	H	G
C	A	D	G	H	E	F	B
D	F	G	C	A	H	B	E
E	H	B	F	G	C	A	D
F	D	H	A	B	G	E	C
G	E	F	H	C	B	D	A
H	G	E	B	D	A	C	F

8×8

A	B	C	D	E	F	G	H	I
B	C	E	G	D	I	F	A	H
C	D	F	A	H	G	I	E	B
D	H	A	B	F	E	C	I	G
E	G	B	I	C	H	D	F	A
F	I	H	E	B	D	A	G	C
G	F	I	C	A	B	H	D	E
H	E	G	F	I	A	B	C	D
I	A	D	H	G	C	E	B	F

9×9

A	B	C	D	E	F	G	H	I	J
B	G	A	E	H	C	F	I	J	D
C	H	J	G	F	B	E	A	D	I
D	A	G	I	J	E	C	B	F	H
E	F	H	J	I	G	A	D	B	C
F	E	B	C	D	I	J	G	H	A
G	I	F	B	A	D	H	J	C	E
H	C	I	F	G	J	D	E	A	B
I	J	D	A	C	H	B	F	E	G
J	D	E	H	B	A	I	C	G	F

10×10

A	B	C	D	E	F	G	H	I	J	K
B	A	J	I	D	C	F	K	H	G	E
C	K	H	A	B	I	J	F	D	E	G
D	C	G	J	I	K	E	B	F	A	H
E	J	B	G	K	H	D	C	A	I	F
F	E	I	C	G	A	K	J	B	H	D
G	F	D	B	H	J	A	I	E	K	C
H	I	K	F	A	D	B	E	G	C	J
I	D	E	H	J	B	C	G	K	F	A
J	G	A	K	F	E	H	D	C	B	I
K	H	F	E	C	G	I	A	J	D	B

11×11

A	B	C	D	E	F	G	H	I	J	K	L
B	L	G	C	D	J	K	E	H	A	F	I
C	K	A	B	F	L	I	D	G	H	J	E
D	F	I	A	L	E	C	G	J	B	H	K
E	D	F	G	J	K	A	L	C	I	B	H
F	H	K	E	G	C	D	B	A	L	I	J
G	I	D	F	K	H	J	A	L	C	E	B
H	E	L	J	C	A	B	I	K	D	G	F
I	J	B	L	H	G	F	K	D	E	A	C
J	C	E	K	A	I	H	F	B	G	L	D
K	G	J	H	I	B	L	C	E	F	D	A
L	A	H	I	B	D	E	J	F	K	C	G

12×12

Fig. 7.

3.13 Balanced Incomplete Block Designs (BIBD)

- In many experiments containing blocks, it is not possible to assign all a treatments within each block. The goal is to assign a subset of k ($k < a$) treatments within each block.
- If we assign k ($k < a$) treatments within each block then we have an **incomplete block design**. Several types of incomplete block designs are
  - Balanced Incomplete Block Designs (BIBD)
  - Partially Balanced Incomplete Block Designs (PBIBD)
  - Lattice Designs, Youden Squares

  The BIBD because it is the most commonly-used incomplete block design. In a BIBD:
  - Treatments are applied in each block in a balanced manner so that any two treatments appear together in the same number (λ) of blocks.
  - Treatments are randomized within each block.
- **Notation:**

  $a =$ the number of treatments in the experiment

  $b =$ the number of blocks in the experiment

  $k =$ the number of treatments per block

  $r =$ the number of blocks in which any treatment appears

  $\lambda =$ the number of blocks for which any each pair of treatments appears together
- For any BIBD, $\lambda = \dfrac{r(k-1)}{a-1}$ and $N = bk = ra$.
- The following tables contain a BIBD example with $a = 4$, $b = 4$,$k = 3$, $r = 3$, and $\lambda = 2$

  If a cell in the left table is blank, then that treatment is missing in that 'incomplete' block.

Block	Treatment A	B	C	D
1	X	X	X	
2	X		X	X
3	X	X		X
4		X	X	X

$\Longrightarrow$

Block	Treatments		
1	A	B	C
2	A	C	D
3	A	B	D
4	B	C	D

3.13.1 ANOVA for a BIBD

- Assume μ is the overall mean, τ_i is the i^{th} treatment effect, β_j is the j^{th} block effect, and ϵ_{ij} is the random error for y_{ij}. The statistical model for a BIBD is

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \qquad \text{and} \qquad \epsilon_{ij} \sim IIDN(0, \sigma^2). \tag{20}$$

- Because the blocks are incomplete, the Type I and Type III sums of squares will be different. That is, the missing treatments in each block represent missing observations (but not missing 'at random').
- For both Type I and Type III analyses: $RSS_1 = SS_{total} = \sum_i \sum_j y_{ij}^2 - \dfrac{y_{..}^2}{N}$.

- We can use the sequential **Type I (V2) sum of squares**. The steps are the same for the RCBD with a missing observation (see Section 3.6).

Step	V2 Source	Fit	*df*	Type I SS for V2
1	Total	μ	$N-1$	RSS_1
2′	Block	β_j	$b-1$	$R(\beta\|\mu) = RSS_1 - RSS_2^*$
3	Trt(adj)	τ_i	$a-1$	$R(\tau\|\beta,\mu) = RSS_2^* - RSS_3^*$
3	Error	ϵ_{ij}	$N-a-b+1$	RSS_3

- We say "treatment sum of squares is adjusted for blocks". The goal is to account for the block variability before we assess the variability among the treatments.

- For a BIBD, there are formulas for Type I (V2) sums of squares for blocks and (adjusted) treatments:

$$SS_{block} = \sum_{j=1}^{b} \frac{y_{\cdot j}^2}{k} - \frac{y_{\cdot\cdot}^2}{N} \qquad SS_{trt(adj)} = \frac{k\sum_{i=1}^{a} Q_i^2}{\lambda a} \qquad SS_E = SS_{total} - SS_{block} - SS_{trt(adj)}$$

where $Q_i = y_{i\cdot} - \dfrac{1}{k}\sum_{j=1}^{b} n_{ij}y_{\cdot j}$ and $n_{ij} = 1$ if treatment i appears in block j, and $n_{ij} = 0$ otherwise.

$\dfrac{1}{k}\sum_{j=1}^{b} n_{ij}y_{\cdot j}$ = average of the block totals across blocks that contain treatment i.

Type I (V2) Analysis of Variance (ANOVA) Table for a BIBD

Source of Variation	Sum of Squares	d.f.	Mean Square	F Ratio
Blocks	SS_{block}	$b-1$		—
Treatment	$SS_{trt(adj)}$	$a-1$	$MS_{trt(adj)}$	$F_0 = MS_{trt(adj)}/MS_E$
Error	SS_E	$N-a-b+1$	MS_E	—
Total	SS_{total}	$N-1$	—	—

where $N = bk = ar$.

- For the **Type III sum of squares**, the steps are:

Step	Source	Fit	*df*	Type III SS
1	Total	μ	$N-1$	RSS_1
2	Trt(adj)	τ_i	$a-1$	$R(\tau\|\beta,\mu) = RSS_2^* - RSS_3$
3	Block(adj)	β_j	$b-1$	$R(\beta\|\tau,\mu) = RSS_2 - RSS_3$
1	Error	ϵ_{ij}	$N-a-b+1$	RSS_3

- The Type III and the Type I (V2) sums of squares will always produce the same F-statistic for testing the equality of treatment effects. The sums of squares for blocks will be different.

BUT WHO CARES?

- To test $H_0 : \tau_1 = \tau_2 = \cdots = \tau_a$ and $H_1 : \tau_i \neq \tau_j$ *for some* $i \neq j$:

  Compare F_0 to the critical value $F_\alpha(a-1, N-a-b+1)$.

  If $F_0 > F_\alpha(a-1, N-a-b+1)$ then **reject** $H_0 : \tau_1 = \tau_2 = \cdots = \tau_a$.

  If $F_0 \leq F_\alpha(a-1, N-a-b+1)$ then **fail to reject** $H_0 : \tau_1 = \tau_2 = \cdots = \tau_a$.

- Or, more simply, compare the p-value to α.

  If p-value $\leq \alpha$ then **reject** H_0. If p-value $> \alpha$ then **fail to reject** H_0.

3.13.2 BIBD Example (Fabric Wearability)

This example can be found in the 1978 text by Box, Hunter, and Hunter. The experiment involves a Martindale wear tester which is a machine used to test the wearing quality of types of cloth or other fabrics. It possesses the feature that four pieces of cloth may be processed simultaneously in one run of the machine. The responses is weight loss in tenths of a milligram suffered by the test fabric when it is rubbed against a standard grade of emery paper for 1000 revolutions of the machine. The wearing quality of specimens of seven different types of cloth (treatments) A, B, C, D, E, F, G are to be compared. Four treatments are randomized and mounted in the four Martindale wear tester specimen holders. The layout and results are given below:

	Cloth Type							
Block	A	B	C	D	E	F	G	$y_{\cdot j}$
1		627		248		563	252	1690
2	344		233			442	226	1245
3			251	211	160		297	919
4	337	537			195		300	1369
5		520	278		199	595		1592
6	369			196	185	606		1356
7	396	602	240	273				1511
$y_{i\cdot}$	1446	2286	1002	928	739	2206	1075	9682 $= y_{\cdot\cdot}$

$$\sum\sum y_{ij}^2 = 3974162$$

$a = 7 \qquad b = 7 \qquad k = 4 \qquad r = 4 \qquad \lambda = \dfrac{r(k-1)}{a-1} = \dfrac{12}{6} = 2$

Estimation of Model Effects

- μ, τ_i $(i = 1, 2, \ldots, a)$, and β_j $(j = 1, 2, \ldots, b)$ are not uniquely estimable. Constraints must be imposed. To be able to calculate estimates $\widehat{\mu}$, $\widehat{\tau}_i$, and $\widehat{\beta}_j$, we need to impose two constraints.

- We will assume constraints $\displaystyle\sum_{i=1}^{a} \tau_i = 0$ and $\displaystyle\sum_{j=1}^{b} \beta_j = 0$

- Applying these constraints will yield

$$\widehat{\mu} = \overline{y}_{\cdot\cdot} \qquad \widehat{\tau}_i = \frac{k}{\lambda a} Q_i \tag{21}$$

where $Q_i = y_{i\cdot} - \dfrac{1}{k}\displaystyle\sum_{j=1}^{b} n_{ij} y_{\cdot j}$.

- The derivation of these formulas is presented in the Montgomery text.

Therefore, for the fabric wear example, we have

$$Q_1 = 1446 - \tfrac{1}{4}(1245 + 1369 + 1356 + 1511) = 75.75$$
$$Q_2 = 2286 - \tfrac{1}{4}(1690 + 1369 + 1592 + 1511) = 745.5$$
$$Q_3 = 1002 - \tfrac{1}{4}(1245 + 919 + 1592 + 1511) = -314.5$$
$$Q_4 = 928 - \tfrac{1}{4}(1690 + 919 + 1356 + 1511) = -441$$
$$Q_5 = 739 - \tfrac{1}{4}(919 + 1369 + 1592 + 1356) = -570$$
$$Q_6 = 2206 - \tfrac{1}{4}(1690 + 1245 + 1592 + 1356) = 735.25$$
$$Q_7 = 1075 - \tfrac{1}{4}(1690 + 1245 + 919 + 1369) = -230.75$$

$$SS_{total} = 3974162 - \frac{9682^2}{28} = 626264.7143$$

$$SS_{blocks} = \frac{1690^2 + 1245^2 + 919^2 + 1369^2 + 1592^2 + 1356^2 + 1511^2}{4} - \frac{9682^2}{28} = 97394.7143$$

$$SS_{cloth(adj)} = \frac{k \sum Q_i^2}{\lambda a}$$

$$= \frac{4[75.75^2 + 745.5^2 + (-314.5)^2 + (-441)^2 + (-570)^2 + 735.25^2 + (-230.75)^2]}{(2)(7)} = 506798.5714$$

$$SS_E = 626264.7143 - 97394.7143 - 506798.5714 = 22071.4286$$

- You can calculate the estimates using the formulas in (22), or you can use the **LSMEANS** statement in SAS. The LSMEANS statement calculates $\widehat{\mu} + \widehat{\tau}_i$. Thus,

$$\widehat{\tau}_i = \text{LSMEAN}_i - \widehat{\mu} = \text{LSMEAN}_i - \overline{y}_{..}$$

- For the fabric wearability example, $\widehat{\tau}_i = \text{LSMEAN}_i - 345.786$:

$$\widehat{\tau}_1 = 21.643 \quad \widehat{\tau}_2 = 213.000 \quad \widehat{\tau}_3 = -89.929 \quad \widehat{\tau}_4 = -126.000$$
$$\widehat{\tau}_5 = -162.857 \quad \widehat{\tau}_6 = 210.071 \quad \widehat{\tau}_7 = -65.929$$

- You can also calculate the estimates directly from ESTIMATE statements using the same structure used for the oneway ANOVA and the RCBD.

- The LSMEANS statement can also be used to perform a multiple comparison procedure (MCP) using the ADJ= option. For example, The / **ADJ=BON** options creates a Bonferroni MCP table with adjusted p-values for all pairwise comparisons . For any adjusted p-value $< .05$, we would reject $\text{H}_0 : \overline{\mu}_i = \overline{\mu}_j$, or equivalently, reject $\text{H}_0 : \tau_i = \tau_j$.

  Note: $\overline{\mu}_i$ is the mean for treatment i averaged over the blocks.

- Summarizing the table results indicates:

Treatment	E	D	C	G	A	F	B
LSMEAN	182.9	219.8	255.9	279.9	367.4	555.9	558.8

Therefore, for the fabric wear example, we have

$$\begin{aligned}
Q_1 &= -\tfrac{1}{4}\,(\qquad + \qquad + \qquad + \qquad) = \\
Q_2 &= -\tfrac{1}{4}\,(\qquad + \qquad + \qquad + \qquad) = \\
Q_3 &= -\tfrac{1}{4}\,(\qquad + \qquad + \qquad + \qquad) = \\
Q_4 &= -\tfrac{1}{4}\,(\qquad + \qquad + \qquad + \qquad) = \\
Q_5 &= -\tfrac{1}{4}\,(\qquad + \qquad + \qquad + \qquad) = \\
Q_6 &= -\tfrac{1}{4}\,(\qquad + \qquad + \qquad + \qquad) = \\
Q_7 &= -\tfrac{1}{4}\,(\qquad + \qquad + \qquad + \qquad) =
\end{aligned}$$

$$SS_{total} = 3974162 - \frac{9682^2}{28} =$$

$$SS_{blocks} = \frac{1690^2 + 1245^2 + 919^2 + 1369^2 + 1592^2 + 1356^2 + 1511^2}{4} - \frac{9682^2}{28} =$$

$$\begin{aligned}
SS_{cloth(adj)} &= \frac{k\sum Q_i^2}{\lambda a} \\
&= \frac{4[75.75^2 + 745.5^2 + (-314.5)^2 + (-441)^2 + (-570)^2 + 735.25^2 + (-230.75)^2]}{(2)(7)} =
\end{aligned}$$

$$SS_E = 626264.7143 \ - \ 97394.7143 \ - \ 506798.5714 \ =$$

- You can calculate the estimates using the formulas in (21), or you can use the **LSMEANS** statement in SAS. The LSMEANS statement calculates $\widehat{\mu} + \widehat{\tau}_i$. Thus,

  $$\widehat{\tau}_i = \text{LSMEAN}_i - \widehat{\mu} = \text{LSMEAN}_i - \overline{y}_{\cdot\cdot}.$$

- For the fabric wearability example, $\widehat{\tau}_i = \text{LSMEAN}_i - 345.786$:

  $$\widehat{\tau}_1 = 21.643 \qquad \widehat{\tau}_2 = 213.000 \qquad \widehat{\tau}_3 = -89.929 \qquad \widehat{\tau}_4 = -126.000$$
  $$\widehat{\tau}_5 = -162.857 \qquad \widehat{\tau}_6 = 210.071 \qquad \widehat{\tau}_7 = -65.929$$

- You can also calculate the estimates directly from ESTIMATE statements using the same structure used for the oneway ANOVA and the RCBD.

- The LSMEANS statement can also be used to perform a multiple comparison procedure (MCP) using the ADJ= option. For example, The / **ADJ=BON** options creates a Bonferroni MCP table with adjusted p-values for all pairwise comparisons . For any adjusted p-value $< .05$, we would reject $\text{H}_0 : \overline{\mu}_i = \overline{\mu}_j$, or equivalently, reject $\text{H}_0 : \tau_i = \tau_j$.

  Note: $\overline{\mu}_i$ is the mean for treatment i averaged over the blocks.

- Summarizing the table results indicates:

Treatment	E	D	C	G	A	F	B
LSMEAN	182.9	219.8	255.9	279.9	367.4	555.9	558.8

SAS Code for BIBD Example

```
DM 'LOG; CLEAR; OUT; CLEAR;';

ODS GRAPHICS ON;
ODS PRINTER PDF file='C:\COURSES\ST541\BIBD.PDF';
OPTIONS NODATE NONUMBER;

***************************************;
*** Balanced Incomplete Block Design ***;
*** From Box, Hunter, and Hunter     ***;
***************************************;

DATA bibd; INPUT cloth $ block wear @@; CARDS;
F 1 563   D 1 248   G 1 252   B 1 627
C 2 233   A 2 344   G 2 226   F 2 442
G 3 297   D 3 211   E 3 160   C 3 251
E 4 195   G 4 300   B 4 537   A 4 337
B 5 520   E 5 199   C 5 278   F 5 595
D 6 196   A 6 369   E 6 185   F 6 606
D 7 273   C 7 240   B 7 602   A 7 396
 ;
TITLE 'BALANCED INCOMPLETE BLOCK DESIGN (BIBD)';

PROC GLM DATA=bibd PLOTS=(ALL);
     CLASS cloth block ;                  INCLUDE /SOLUTION TO GET ESTIMATES FOR
     MODEL wear = block cloth ;           SAS CONSTRAINTS τa=0, βb=0
     MEANS cloth ;                        BONFERRONI MCP
'MIXED' LSMEANS cloth / ADJ=BON ;
EFFECTS RANDOM block / TEST ;                    a-1                          a
CLOTH ESTIMATE 'cloth A effect' cloth 6 -1 -1 -1 -1 -1 -1 / DIVISOR=7;
FIXED, ESTIMATE 'cloth B effect' cloth -1 6 -1 -1 -1 -1 -1 / DIVISOR=7;
     ESTIMATE 'cloth C effect' cloth -1 -1 6 -1 -1 -1 -1 / DIVISOR=7;
BLOCKS ESTIMATE 'cloth D effect' cloth -1 -1 -1 6 -1 -1 -1 / DIVISOR=7;
RANDOM ESTIMATE 'cloth E effect' cloth -1 -1 -1 -1 6 -1 -1 / DIVISOR=7;
     ESTIMATE 'cloth F effect' cloth -1 -1 -1 -1 -1 6 -1 / DIVISOR=7;
     ESTIMATE 'cloth G effect' cloth -1 -1 -1 -1 -1 -1 6 / DIVISOR=7;
TITLE2 'BLOCKS FIRST, TREATMENTS SECOND';
RUN;
```

BALANCED INCOMPLETE BLOCK DESIGN (BIBD)
BLOCKS FIRST, TREATMENTS SECOND

The GLM Procedure

Dependent Variable: wear

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	604193.2857	50349.4405	34.22	<.0001
Error	15	22071.4286	1471.4286		
Corrected Total	27	626264.7143			

R-Square	Coeff Var	Root MSE	wear Mean
0.964757	11.09335	38.35920	345.7857

$\bar{y}_{..} = \hat{\mu}$

TYPE I SS
V 2

Source	DF	Type I SS	Mean Square	F Value	Pr > F
block	6	97394.7143	16232.4524	11.03	<.0001
cloth	6	506798.5714	84466.4286	57.40	<.0001

REJECT
$H_0: \tau_1 = \tau_2 = \cdots = \tau_7$

TYPE III SS

Source	DF	Type III SS	Mean Square	F Value	Pr > F
block	6	14570.0714	2428.3452	1.65	0.2015
cloth	6	506798.5714	84466.4286	57.40	<.0001

$\hat{\tau}_i$ ESTIMATES ASSUMING $\sum_{i=1}^{7} \hat{\tau}_i = 0$

Parameter	Estimate	Standard Error	t Value	Pr > |t|
cloth A effect	21.642857	18.9828832	1.14	0.2721
cloth B effect	213.000000	18.9828832	11.22	<.0001
cloth C effect	-89.928571	18.9828832	-4.74	0.0003
cloth D effect	-126.000000	18.9828832	-6.64	<.0001
cloth E effect	-162.857143	18.9828832	-8.58	<.0001
cloth F effect	210.071429	18.9828832	11.07	<.0001
cloth G effect	-65.928571	18.9828832	-3.47	0.0034

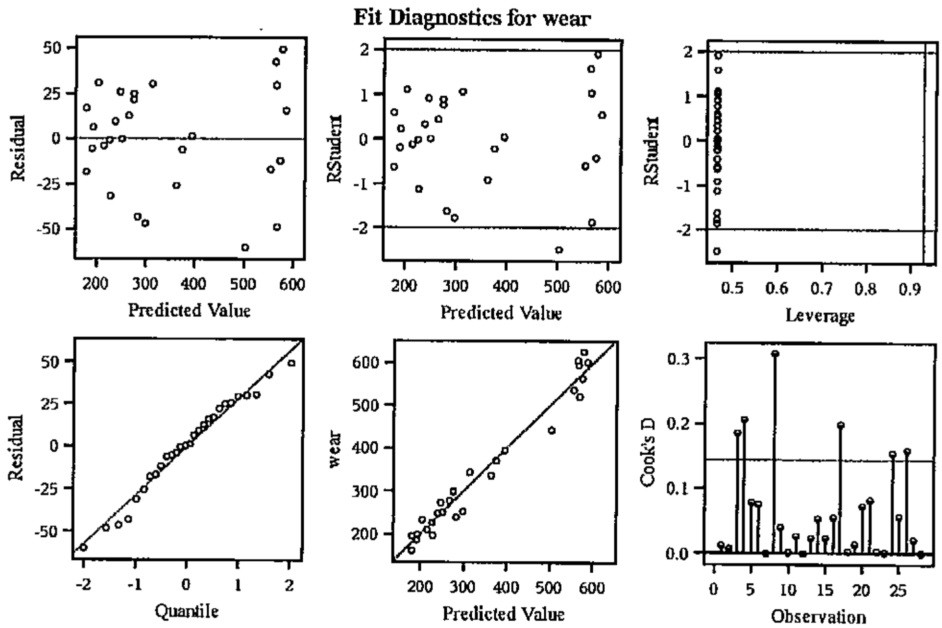
Tests of Hypotheses for Mixed Model Analysis of Variance

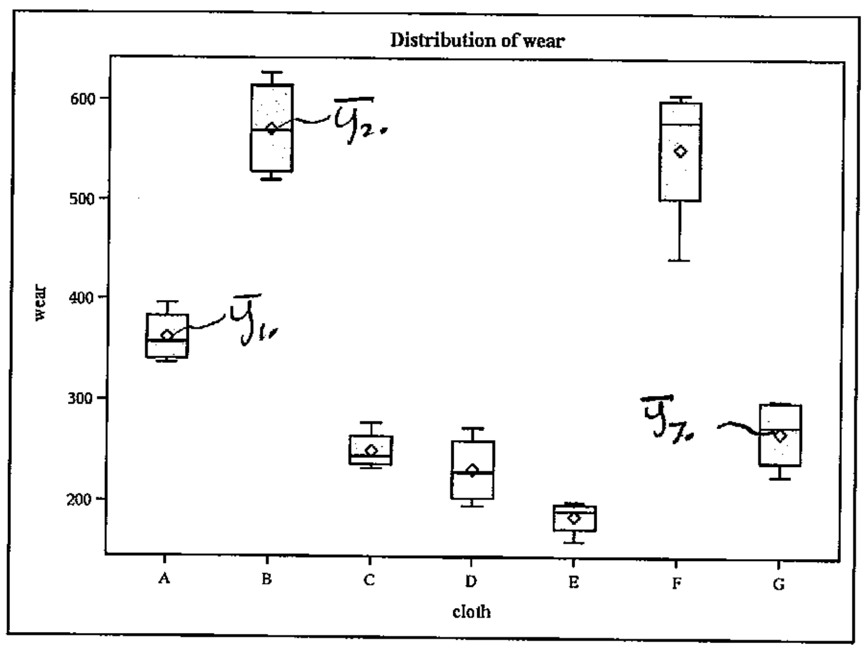
Variable: wear TREAT BLOCKS AS A RANDOM EFFECT

Source	DF	Type III SS	Mean Square	F Value	Pr > F
block	6	14570	2428.345238	1.65	0.2015
cloth	6	506799	84466	57.40	<.0001
Error: MS(Error)	15	22071	1471.428571		

SAME ANOVA RESULTS AS ABOVE BECAUSE MSE IS USED IN EACH F-TEST

DENOMINATOR OF F-TEST





RAW DATA BOXPLOTS:

WARNING: THESE MEANS ARE NOT ADJUSTED FOR BLOCK EFFECTS.

Level of cloth	N	wear Mean	wear Std Dev
A	4	361.500000	26.7893013
B	4	571.500000	51.1631378
C	4	250.500000	19.7737199
D	4	232.000000	34.9952378
E	4	184.750000	17.5190373
F	4	551.500000	75.2440474
G	4	268.750000	35.9756862

↖ RAW DATA SAMPLE MEANS (NOT ADJUSTED FOR BLOCK EFFECTS)

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Bonferroni

cloth	wear LSMEAN	LSMEAN Number
A	367.428571	1
B	558.785714	2
C	255.857143	3
D	219.785714	4
E	182.928571	5
F	555.857143	6
G	279.857143	7

LEAST SQUARES (LS) MEANS (ADJUSTED FOR BLOCK EFFECTS)
THESE ARE WHAT YOU WANT TO COMPARE

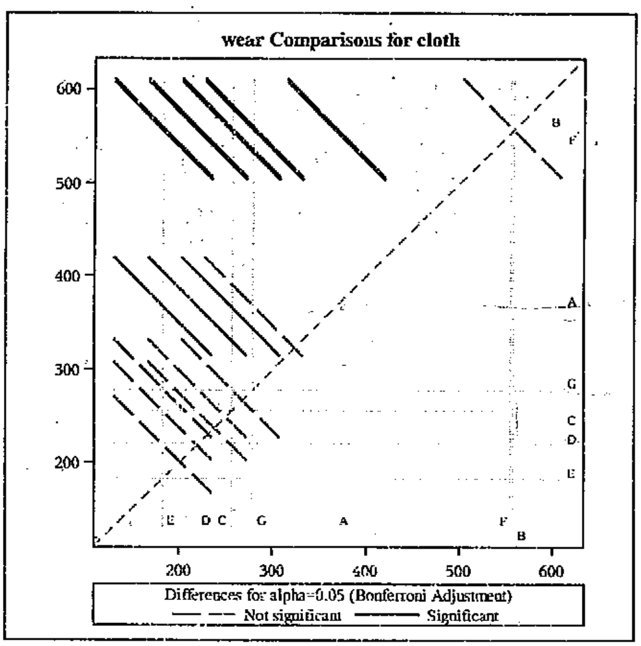
Least Squares Means for effect cloth
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: wear

	i/j	A 1	B 2	C 3	D 4	E 5	F 6	G 7
A	1		0.0002	0.0332	0.0028	0.0003	0.0002	0.1809
B	2	0.0002		<.0001	<.0001	<.0001	1.0000	<.0001
C	3	0.0332	<.0001		1.0000	0.4995	<.0001	1.0000
D	4	0.0028	<.0001	1.0000		1.0000	<.0001	1.0000
E	5	0.0003	<.0001	0.4995	1.0000		<.0001	0.0935
F	6	0.0002	1.0000	<.0001	<.0001	<.0001		<.0001
G	7	0.1809	<.0001	1.0000	1.0000	0.0935	<.0001	

BONFERRONI MCP

FIND CASES WITH ADJUSTED P-VALUES $\leq .05$

REJECT Ho:
$\tau_A=\tau_B$, $\tau_A=\tau_C$, $\tau_A=\tau_D$
$\tau_A=\tau_E$, $\tau_A=\tau_F$
$\tau_B=\tau_C$, $\tau_B=\tau_D$, $\tau_B=\tau_E$
$\tau_B=\tau_G$, $\tau_C=\tau_F$, $\tau_D=\tau_F$,
$\tau_E=\tau_F$, $\tau_F=\tau_G$



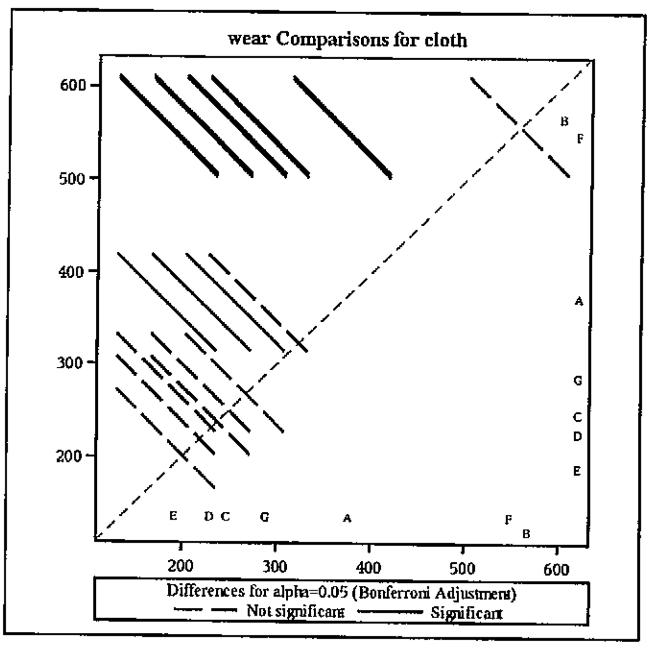
The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Bonferroni

cloth	wear LSMEAN	LSMEAN Number
A	367.428571	1
B	558.785714	2
C	255.857143	3
D	219.785714	4
E	182.928571	5
F	555.857143	6
G	279.857143	7

Least Squares Means for effect cloth
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: wear

i/j	1	2	3	4	5	6	7
1		0.0002	0.0332	0.0028	0.0003	0.0002	0.1809
2	0.0002		<.0001	<.0001	<.0001	1.0000	<.0001
3	0.0332	<.0001		1.0000	0.4995	<.0001	1.0000
4	0.0028	<.0001	1.0000		1.0000	<.0001	1.0000
5	0.0003	<.0001	0.4995	1.0000		<.0001	0.0935
6	0.0002	1.0000	<.0001	<.0001	<.0001		<.0001
7	0.1809	<.0001	1.0000	1.0000	0.0935	<.0001	



3.13.3 Matrix Forms for BIBD

- Model: $y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$ for $i = 1,2,3,4,5,6,7$ and $j = 1,2,3,4,5,6,7$

$$\epsilon_{ij} \sim N(0, \sigma^2) \qquad \beta_j \sim N(0, \sigma^2_\beta)$$

- Assume (i) $\sum_{i=1}^{7} \tau_i = 0$ and (ii) $\sum_{j=1}^{7} \beta_j = 0$.
- Goal: Estimate $[\, \mu, \tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6, \tau_7, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7 \,]$.

Alternate Approach: Keeping $a + b + 1$ Columns

$$X = \begin{array}{c}
\begin{array}{ccccccccccccccc} \mu & \tau_1 & \tau_2 & \tau_3 & \tau_4 & \tau_5 & \tau_6 & \tau_7 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 & \beta_7 \end{array} \\
\left[\begin{array}{c|ccccccc|ccccccc}
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\hline
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right]
\end{array}
\qquad
y = \left[\begin{array}{c}
627 \\ 248 \\ 563 \\ 252 \\ 344 \\ 233 \\ 442 \\ 226 \\ 251 \\ 211 \\ 160 \\ 297 \\ 337 \\ 537 \\ 195 \\ 300 \\ 520 \\ 278 \\ 199 \\ 595 \\ 369 \\ 196 \\ 185 \\ 606 \\ 396 \\ 602 \\ 240 \\ 273 \\ \hline 0 \\ 0
\end{array}\right]$$

CONSTRAINTS $\Sigma \hat{\tau}_i = 0$, $\Sigma \hat{\beta}_j = 0$

$$X'X = \begin{array}{cc} \begin{array}{ccc} \mu \qquad & \qquad \tau_i \qquad & \qquad \beta_j \end{array} & \\ \left[\begin{array}{c|ccccccc|ccccccc}
28 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ \hline
4 & 5 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\
4 & 1 & 5 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\
4 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\
4 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\
4 & 1 & 1 & 1 & 1 & 5 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\
4 & 1 & 1 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\
4 & 1 & 1 & 1 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ \hline
4 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 1 & 1 \\
4 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 1 & 1 \\
4 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 5 & 1 & 1 & 1 & 1 \\
4 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 5 & 1 & 1 & 1 \\
4 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 5 & 1 & 1 \\
4 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 5 & 1 \\
4 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 5
\end{array}\right] & \begin{array}{c} \mu \\ \\ \\ \tau_i \\ \\ \\ \\ \\ \\ \\ \beta_j \\ \\ \\ \\ \\ \end{array} \end{array}$$

$$(X'X)^{-1} = \frac{1}{196}\left[\begin{array}{c|ccccccc|ccccccc}
15 & -4 & -4 & -4 & -4 & -4 & -4 & -4 & -4 & -4 & -4 & -4 & -4 & -4 & -4 \\ \hline
-4 & 52 & -4 & -4 & -4 & -4 & -4 & -4 & 8 & -6 & 8 & -6 & 8 & -6 & -6 \\
-4 & -4 & 52 & -4 & -4 & -4 & -4 & -4 & -6 & 8 & 8 & -6 & -6 & 8 & -6 \\
-4 & -4 & -4 & 52 & -4 & -4 & -4 & -4 & 8 & -6 & -6 & 8 & -6 & 8 & -6 \\
-4 & -4 & -4 & -4 & 52 & -4 & -4 & -4 & -6 & 8 & -6 & 8 & 8 & -6 & -6 \\
-4 & -4 & -4 & -4 & -4 & 52 & -4 & -4 & 8 & 8 & -6 & -6 & -6 & -6 & 8 \\
-4 & -4 & -4 & -4 & -4 & -4 & 52 & -4 & -6 & -6 & 8 & 8 & -6 & -6 & 8 \\
-4 & -4 & -4 & -4 & -4 & -4 & -4 & 52 & -6 & -6 & -6 & -6 & 8 & 8 & 8 \\ \hline
-4 & 8 & -6 & 8 & -6 & 8 & -6 & -6 & 52 & -4 & -4 & -4 & -4 & -4 & -4 \\
-4 & -6 & 8 & -6 & 8 & 8 & -6 & -6 & -4 & 52 & -4 & -4 & -4 & -4 & -4 \\
-4 & 8 & 8 & -6 & -6 & -6 & 8 & -6 & -4 & -4 & 52 & -4 & -4 & -4 & -4 \\
-4 & -6 & -6 & 8 & 8 & -6 & 8 & -6 & -4 & -4 & -4 & 52 & -4 & -4 & -4 \\
-4 & 8 & -6 & -6 & 8 & -6 & -6 & 8 & -4 & -4 & -4 & -4 & 52 & -4 & -4 \\
-4 & -6 & 8 & 8 & -6 & -6 & -6 & 8 & -4 & -4 & -4 & -4 & -4 & 52 & -4 \\
-4 & -6 & -6 & -6 & -6 & 8 & 8 & 8 & -4 & -4 & -4 & -4 & -4 & -4 & 52
\end{array}\right]$$

$$X'y = \begin{bmatrix} 9682 \\ 1446 \\ 2286 \\ 1002 \\ 928 \\ 739 \\ 2206 \\ 1075 \\ 1690 \\ 1245 \\ 919 \\ 1369 \\ 1592 \\ 1356 \\ 1511 \end{bmatrix} \begin{array}{l} y_{..} \\ \left.\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array}\right\} y_{i.} \\ \left.\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array}\right\} y_{.j} \end{array}$$

$$(X'X)^{-1}X'y = \left[\begin{array}{r} 345.7857 \\ \hline 21.6429 \\ 213 \\ -89.9286 \\ -126 \\ -162.8571 \\ 210.0714 \\ -65.9286 \\ \hline 18.9286 \\ -53.5 \\ -4.8571 \\ -5 \\ 9.6429 \\ 7.5 \\ 27.2857 \end{array}\right] \begin{array}{c} \hat{\mu} \\ \\ \\ \\ \hat{\tau}_i \\ \\ \\ \\ \\ \\ \\ \hat{\beta}_j \\ \\ \\ \\ \end{array} = \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \hat{\tau}_3 \\ \hat{\tau}_4 \\ \hat{\tau}_5 \\ \hat{\tau}_6 \\ \hat{\tau}_7 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \\ \hat{\beta}_5 \\ \hat{\beta}_6 \\ \hat{\beta}_7 \end{bmatrix}$$

NOTE: $\hat{\tau}_i$ MATCHES SAS OUTPUT (p 120)