Lecture 8: Balanced Incomplete Block Design

Montgomery Section 4.4

Catalyst Experiment

Four catalysts are being investigated in an experiment. The experimental procedure consists of selecting a batch raw material, loading the pilot plant, applying each catalyst in a separate run and observing the reaction time. The batches of raw material are considered as blocks, however each batch is only large enough to permit three catalysts to be run.

	Block(raw material)				
Catalyst	1	2	3	4	$y_{i.}$
1	73	74	-	71	218
2	-	75	67	72	214
3	73	75	68	-	216
4	75	-	72	75	222
$y_{.j}$	221	224	207	218	870=$y_{..}$

Balanced Incomplete Block Design (BIBD)

- There are a treatments and b blocks;
- Each block contains k (different) treatments;
- Each treatment appears in r blocks;
- Each pair of treatments appears together in λ blocks;

Example 1. $a = 3$, $b = 3$, $k = 2$, $r = 2$, $\lambda = 1$

	block			incidence		
treatment	1	2	3	matrix		
A	A	-	A	1	0	1
B	B	B	-	1	1	0
C	-	C	C	0	1	1

Incidence Matrix: $\mathcal{N} = (n_{ij})_{a\times b}$ where $n_{ij} = 1$, if treatment i is run in block j;
=0 otherwise.

Example 2.

	block						incidence					
treatment	1	2	3	4	5	6	matrix					
A	A	A	A	-	-	-	1	1	1	0	0	0
B	B	-	-	B	B	-	1	0	0	1	1	0
C	-	C	-	C	-	C	0	1	0	1	0	1
D	-	-	D	-	D	D	0	0	1	0	1	1

$a = 4, b = 6, k = 2, r = 3, \lambda = 1, \mathcal{N} = (n_{ij})_{4\times 6}$

BIBD: Design Properties

a, b, k, r, and λ are not independent

- $N = ar = bk$, where N is the total number of runs $\Longrightarrow r = bk/a$
- $\lambda(a-1) = r(k-1) \Longrightarrow \lambda = r(k-1)/(a-1)$

  1. for any fixed treatment i_0
  2. two different ways to count the total number of pairs including treatment i_0 in the experiment.

  I. $a-1$ possible pairs, each appears in λ blocks, so $\lambda(a-1)$;

  II. treatment i_0 appears in r blocks. Within each block, there are $k-1$ pairs including i_0, so $r(k-1)$
- $b \geq a$ (a brainteaser for math/stat students).
- Nonorthogonal design

Extensive list of BIBDs can be found in Fisher and Yates (1963) and Cochran and Cox (1957).

BIBD: Design in SAS

```
TITLE 'Balanced Incomplete Block Design';
DATA candidates;
     DO treatment = 1 to 4;  OUTPUT; END;
PROC OPTEX DATA=candidates SEED=5140514 CODING=ORTH;
     CLASS treatment;
     MODEL treatment;
     BLOCKS STRUCTURE=(6)2; /* (b)k: b=6, k=2 */
     OUTPUT OUT=bibd BLOCKNAME=block;
PROC PRINT DATA=bibd; RUN; QUIT;
```

Obs	BLOCK	treatment	Obs	BLOCK	treatment
1	1	1	7	4	3
2	1	2	8	4	1
3	2	3	9	5	1
4	2	4	10	5	4
5	3	2	11	6	3
6	3	4	12	6	2

BIBD: Statistical Model

- Statistical Model

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases}$$

- additive model (without interaction)
- Not all y_{ij} exist because of incompleteness
- Usual treatment and block restrictions : $\sum \tau_i = 0; \qquad \sum \beta_j = 0$
- Nonorthogonality of treatments and blocks

Use Type III Sums of Squares and `lsmeans`

Model Estimates

- Least squares estimates for μ, etc.

$$\hat{\mu} = \bar{y}_{..} = \frac{\sum_{i,j} n_{ij} y_{ij}}{N}; \qquad \hat{\tau}_i = \frac{kQ_i}{\lambda a}; \qquad \hat{\beta}_j = \frac{rQ'_j}{\lambda b}$$

where

$$Q_i = r\bar{y}_{i.} - \sum_{j=1}^{b} n_{ij}\bar{y}_{.j}; \qquad Q'_j = k\bar{y}_{.j} - \sum_{i=1}^{a} n_{ij}\bar{y}_{i.}$$

- Note that $\bar{y}_{i.} = \sum_{j=1}^{b} n_{ij} y_{ij}/r, \quad \bar{y}_{.j} = \sum_{i=1}^{a} n_{ij} y_{ij}/k$

$$\begin{aligned} \mathrm{Var}(Q_i) &= \mathrm{Var}\left(r\bar{y}_{i.}\right) + \mathrm{Var}\left(\sum_{j=1}^{b} n_{ij}\bar{y}_{.j}\right) - 2\mathrm{Cov}\left(r\bar{y}_{i.}, \sum_{j=1}^{b} n_{ij}\bar{y}_{.j}\right) \\ &= r^2\frac{\sigma^2}{r} + r\frac{\sigma^2}{k} - 2r\frac{\sigma^2}{k} = \frac{(k-1)r}{k}\sigma^2 \end{aligned}$$

- $\mathrm{Var}\left(\hat{\tau}_i\right) = \left(\frac{k}{\lambda a}\right)^2 \mathrm{Var}(Q_i) = \left(\frac{k}{\lambda a}\right)^2 \frac{(k-1)r}{k}\sigma^2 = \frac{k(a-1)}{\lambda a^2}\sigma^2; \;\; \mathrm{S.E.}_{\hat{\tau}_i} = ?$
- $\mathrm{Var}\left(\hat{\tau}_i - \hat{\tau}_j\right) = \frac{2k\sigma^2}{\lambda a}; \;\; \mathrm{S.E.}_{\hat{\tau}_i - \hat{\tau}_j} = ?$

Analysis of Variance (Focus: Treatment Effects)

- $\text{SS}_{\text{T}} = \sum\sum y_{ij}^2 - N\bar{y}_{..}^2$ may be partitioned into

$$SS_T = SS_{Blocks} + SS_{Treatments(adjusted)} + SS_E$$

- $\text{SS}_{\text{Block}} = k\sum_j \bar{y}_{.j}^2 - N\bar{y}_{..}^2$
- $\text{SS}_{\text{Treatment(adjusted)}} = k\sum Q_i^2/(\lambda a) = \frac{\lambda a}{k}\sum \hat{\tau}_i^2$ is adjusted to separate the treatment and the block effects as each treatment is represented in a different set of r blocks (incompleteness)
- Note that $Q_i = r\bar{y}_{i.} - \sum_{j=1}^b n_{ij}\bar{y}_{.j}$
  - treatment i's **total** minus treatment i's block averages
  - $\sum Q_i = 0$
- SS_{E} by subtraction

Analysis of Variance Table (Type I SS)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Blocks	SS_{Block}	$b-1$	MS_{Block}	
Treatment	$\text{SS}_{\text{Treatment(adjusted)}}$	$a-1$	$\text{MS}_{\text{Treatment(adjusted)}}$	F_0
Error	SS_{E}	$N-a-b+1$	MS_{E}	
Total	SS_{T}	$N-1$		

- This table can be followed for testing hypothesis on TREATMENTS
  - If $F_0 > F_{\alpha,a-1,N-a-b+1}$ then reject $H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0$

Analysis of Variance (Focus: Block Effects)

- $\text{SS}_\text{T} = \sum\sum y_{ij}^2 - N\bar{y}_{..}^2$ may be partitioned into

$$SS_T = SS_{Treatments} + SS_{Blocks(adjusted)} + SS_E$$

- $\text{SS}_\text{Treatment} = r\sum_i \bar{y}_{i.}^2 - N\bar{y}_{..}^2$
- $\text{SS}_\text{Blocks(adjusted)} = r\sum_j Q_j'^2/(\lambda b)$ is adjusted for non-orthogonal treatment effects
- Note that $Q_j' = k\bar{y}_{.j} - \sum_{i=1}^a n_{ij}\bar{y}_{i.}$
  - block j's **total** minus block j's treatment averages
  - $\sum_j Q_j' = 0$
- SS_E by subtraction

Analysis of Variance Table (Type I SS)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Treatments	$\text{SS}_{\text{Treatment}}$	$a-1$	$\text{MS}_{\text{Treatment}}$	
Blocks	$\text{SS}_{\text{Block(adjusted)}}$	$b-1$	$\text{MS}_{\text{Block(adjusted)}}$	F_0'
Error	SS_{E}	$N-a-b+1$	MS_{E}	
Total	SS_{T}	$N-1$		

- This table can be followed for testing hypothesis on BLOCKS.
  – If $F_0' > F_{\alpha,b-1,N-a-b+1}$ then reject $H_0' : \beta_1 = \beta_2 = \cdots = \beta_b = 0$

Analysis of Variance Table (Type III SS)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Blocks	SS$_{\mathrm{Block(adjusted)}}$	$b-1$	MS$_{\mathrm{Block(adjusted)}}$	F_0'
Treatment	SS$_{\mathrm{Treatment(adjusted)}}$	$a-1$	MS$_{\mathrm{Treatment(adjusted)}}$	F_0
Error	SS$_{\mathrm{E}}$	$N-a-b+1$	MS$_{\mathrm{E}}$	
Total	SS$_{\mathrm{T}}$	$N-1$		

- SS$_{\mathrm{T}} \neq$ SS$_{\mathrm{Block(adjusted)}}+$ SS$_{\mathrm{Treatment(adjusted)}}+$ SS$_{\mathrm{E}}$

$$
\begin{aligned}
SS_{\mathrm{E}} &= SS_{\mathrm{T}} - SS_{\mathrm{Block}} - SS_{Treatment(adjusted)} \\
&= SS_{\mathrm{T}} - SS_{\mathrm{Treatment}} - SS_{Block(adjusted)}
\end{aligned}
$$

- This table can be followed for tests on BLOCKS or TREATMENTS
  - If $F_0' > F_{\alpha,b-1,N-a-b+1}$ then reject $H_0' : \beta_1 = \beta_2 = \cdots = \beta_b = 0$
  - If $F_0 > F_{\alpha,a-1,N-a-b+1}$ then reject $H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0$

Mean Tests and Contrasts

- Must compute adjusted means (`lsmeans`)
- Adjusted mean is $\widehat{\mu} + \widehat{\tau}_i$
- Standard error of adjusted mean is $\sqrt{\mathrm{MS_E}(\frac{k(a-1)}{\lambda a^2} + \frac{1}{N})}$
- Contrasts based on adjusted treatment totals

  For a contrast: $\sum c_i \mu_i$

  Its estimate: $\sum c_i \hat{\tau}_i = \frac{k}{\lambda a} \sum c_i Q_i$

  Contrast sum of squares:

$$SS_C = \frac{k(\sum_{i=1}^a c_i Q_i)^2}{\lambda a \sum_{i=1}^a c_i^2}$$

Pairwise Comparison

- Pairwise comparison $\tau_i - \tau_j$:

1. Bonferroni:

$$CD = t_{\alpha/(2m), ar-a-b+1}\sqrt{\mathrm{MS_E}\frac{2k}{\lambda a}}.$$

2. Tukey:

$$CD = \frac{q_\alpha(a, ar-a-b+1)}{\sqrt{2}}\sqrt{\mathrm{MS_E}\frac{2k}{\lambda a}}$$

Using SAS

```
options nocenter ps=60 ls=75;
data example;
  input trt block resp @@;
  datalines;
1 1 73 1 2 74 1 4 71 2 2 75 2 3 67 2 4 72
3 1 73 3 2 75 3 3 68 4 1 75 4 3 72 4 4 75
;

proc glm;
  class block trt;
  model resp = block trt;
  lsmeans trt / tdiff pdiff adjust=bon stderr;
  lsmeans trt / pdiff adjust=tukey;
  contrast 'a' trt 1 -1 0 0;
  estimate 'b' trt 0 0 1 -1;
  run; quit;
```

SAS Output

```
Dependent Variable: resp
                                 Sum of
Source                  DF      Squares   Mean Square  F Value  Pr > F
Model                    6  77.75000000   12.95833333    19.94  0.0024
Error                    5   3.25000000    0.65000000
Corrected Total         11  81.00000000

Source                  DF    Type I SS   Mean Square  F Value  Pr > F
block                    3  55.00000000   18.33333333    28.21  0.0015
trt                      3  22.75000000    7.58333333    11.67  0.0107

Source                 DF  Type III SS   Mean Square  F Value  Pr > F
block                  3  66.08333333   22.02777778    33.89  0.0010
trt                    3  22.75000000    7.58333333    11.67  0.0107
```

Least Squares Means

Adjustment for Multiple Comparisons: Bonferroni

trt	resp LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
1	71.3750000	0.4868051	<.0001	1
2	71.6250000	0.4868051	<.0001	2
3	72.0000000	0.4868051	<.0001	3
4	75.0000000	0.4868051	<.0001	4

Bonferroni Method:

i/j	1	2	3	4
1		-0.35806	-0.89514	-5.19183
		1.0000	1.0000	0.0209
2	0.358057		-0.53709	-4.83378
	1.0000		1.0000	0.0284
3	0.895144	0.537086		-4.29669
	1.0000	1.0000		0.0464
4	5.191833	4.833775	4.296689	
	0.0209	0.0284	0.0464	

Tukey's Method:

i/j	1	2	3	4
1		0.9825	0.8085	0.0130
2	0.9825		0.9462	0.0175
3	0.8085	0.9462		0.0281
4	0.0130	0.0175	0.0281	

Dependent Variable: resp

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
a	1	0.08333333	0.08333333	0.13	0.7349

Parameter	Estimate	Standard Error	t Value	Pr > |t|
b	-3.00000000	0.69821200	-4.30	0.0077

Other Incomplete Designs

- Youden Square
- Partially Balanced Incomplete Block Design
- Cyclic Designs
- Square, Cubic, and Rectangular Lattices