Experimental Design
Note 7-2
Blocking and confounding 2^K factorial design

Keunbaik Lee

Sungkyunkwan University

Recall: Blocking in Factorial Design-Example I

Battery Life Experiment:
An engineer is studying the effective lifetime of some battery. Two factors, plate material and temperature, are involved. There are three types of plate materials (1, 2, 3) and three temperature levels (15, 70, 125). Four batteries are tested at each combination of plate material and temperature, and all 36 tests are run in random order. The experiment and the resulting observed battery life data are given below in Table 5.1. block effect (fixed)
If we assume further that four operators (1, 2, 3, 4) were hired to conduct the experiment. It is known that different operators can cause systematic difference in battery lifetime. Hence operators should be treated as blocks. The blocking scheme is every operator conduct a single replicate of the full factorial design.

Recall: Blocking in Factorial Design-Example II

For each treatment (treatment combination), the observations were in the order of the operators 1, 2, 3, and 4. This is a **blocked factorial design**.

Statistical Model for Blocked Factorial Experiment I

Model:

Variations in different operators

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \delta_k + \epsilon_{ijk}$$

for $i = 1, 2, \cdots, a$, $j = 1, 2, \cdots, b$ and $k = 1, 2, \cdots, n$, δ_k is the effect of the kth block.

- randomization restriction is imposed. (complete block factorial design).
- interactions between blocks and treatment effects are assumed to be negligible.

Statistical Model for Blocked Factorial Experiment II

- The previous ANOVA table for the experiment should be modified as follows:
  ADD: Block Sum of Squares

$$SS_{Blocks} = \frac{1}{ab}\sum_k y_{\cdot\cdot k}^2 - \frac{y_{\cdots}^2}{abn}, \quad df = n - 1.$$

  MODIFY: Error Sum of Squares

$$(new)SSE = (old)SSE - SS_{Blocks}, \quad df = (ab-1)(n-1).$$

- other inferences should be modified accordingly.

Statistical Model for Blocked Factorial Experiment III

■ **TABLE 5.20**

Analysis of Variance for a Two-Factor Factorial in a Randomized Complete Block

Source of Variation	Sum of Squares	Degrees of Freedom	Expected Mean Square	F_0
Blocks	$\frac{1}{ab}\sum_k y_{..k}^2 - \frac{y_{...}^2}{abn}$	$n-1$	$\sigma^2 + ab\sigma_\delta^2$	
A	$\frac{1}{bn}\sum_i y_{i..}^2 - \frac{y_{...}^2}{abn}$	$a-1$	$\sigma^2 + \frac{bn\sum \tau_i^2}{a-1}$	$\frac{MS_A}{MS_E}$
B	$\frac{1}{an}\sum_j y_{.j.}^2 - \frac{y_{...}^2}{abn}$	$b-1$	$\sigma^2 + \frac{an\sum \beta_j^2}{b-1}$	$\frac{MS_B}{MS_E}$
AB	$\frac{1}{n}\sum_i\sum_j y_{ij.}^2 - \frac{y_{...}^2}{abn} - SS_A - SS_B$	$(a-1)(b-1)$	$\sigma^2 + \frac{n\sum\sum(\tau\beta)_{ij}^2}{(a-1)(b-1)}$	$\frac{MS_{AB}}{MS_E}$
Error	Subtraction	$(ab-1)(n-1)$	σ^2	
Total	$\sum_i\sum_j\sum_k y_{ijk}^2 - \frac{y_{...}^2}{abn}$	$abn-1$		

Statistical Model for Blocked Factorial Experiment IV

- This is a **blocking replicated design**.
- If there are n replicates of the design, then each replicate is a block.
- Each **replicate** is run in one of the **blocks** (time periods, batches of raw material, etc.).
- Runs within the block are randomized.
- How about **unreplicated** case?

$n = 1$

Confounding in blocks I

- Now consider the **unreplicated** case.
- Clearly the previous discussion does not apply, since there is only one replicate.
- To illustrate, consider the situation of Example 6.2, the pilot plant filtration rate experiment.
- This is a 2^4, $n = 1$ replicate.

Confounding in blocks II

Filtration Rate Experiment (revisited)

factor				
A	B	C	D	original response
$-$	$-$	$-$	$-$	45
$+$	$-$	$-$	$-$	71
$-$	$+$	$-$	$-$	48
$+$	$+$	$-$	$-$	65
$-$	$-$	$+$	$-$	68
$+$	$-$	$+$	$-$	60
$-$	$+$	$+$	$-$	80
$+$	$+$	$+$	$-$	65
$-$	$-$	$-$	$+$	43
$+$	$-$	$-$	$+$	100
$-$	$+$	$-$	$+$	45
$+$	$+$	$-$	$+$	104
$-$	$-$	$+$	$+$	75
$+$	$-$	$+$	$+$	86
$-$	$+$	$+$	$+$	70
$+$	$+$	$+$	$+$	96

Confounding in blocks III

- Suppose there are two batches of raw material. Each batch can be used for only 8 runs. It is known these two batches are very different. Blocking should be employed to eliminate this variability.
- How to select 8 treatments (level combinations, or runs) for each block?

Confounding the 2^k Factorial Design with two blocks I

Suppose there are two factors (A, B) each with 2 levels, and two blocks (b_1, b_2) each containing two runs (treatments). Since b_1 and b_2 are interchangeable, there are three possible blocking scheme:

			Blocking scheme		
A	B	Response	1	2	3
-	-	y_{--}	b_1	b_1	b_2
+	-	y_{+-}	b_1	b_2	b_1
-	+	y_{-+}	b_2	b_1	b_1
+	+	y_{++}	b_2	b_2	b_2

Interaction Term

Confounding the 2^k Factorial Design with two blocks II

Comparing blocking schemes:

Scheme 1:

- block effect: $b = \bar{y}_{b_2} - \bar{y}_{b_1} = \frac{1}{2}(-y_{--} - y_{+-} + y_{-+} + y_{++})$ $\sim B$
- main effect: $B = \frac{1}{2}(-y_{--} - y_{+-} + y_{-+} + y_{++})$
- B and b are not distinguishable or confounded.

Scheme 2:

- block effect: $b = \bar{y}_{b_2} - \bar{y}_{b_1} = \frac{1}{2}(-y_{--} + y_{+-} - y_{-+} + y_{++})$ $\sim A$
- main effect: $A = \frac{1}{2}(-y_{--} + y_{+-} - y_{-+} + y_{++})$
- A and b are not distinguishable or confounded.

Scheme 3:

- block effect: $b = \bar{y}_{b_2} - \bar{y}_{b_1} = \frac{1}{2}(y_{--} - y_{+-} - y_{-+} + y_{++})$ $\sim AB$
- main effect: $AB = \frac{1}{2}(y_{--} - y_{+-} - y_{-+} + y_{++})$
- AB and b are not distinguishable or confounded.

Confounding the 2^k Factorial Design with two blocks III

- The reason for confounding: the block arrangement matches the contrast of some factorial effect.
- Confounding makes the effect **Inestimable**.
- **Question**: **which scheme is the best (or causes the least damage)**?

Confounding the 2^k Factorial Design with two blocks IV

2^k **Design with Two Blocks via Confounding**

- Confound blocks with the effect (contrast) of the highest order
- Block 1 consists of all treatments with the contrast coefficient equal to -1
- Block 2 consists of all treatments with the contrast coefficient equal to 1

Confounding the 2^k Factorial Design with two blocks V

Example 1. Block 2^3 Design

$X =$

	factorial effects (contrasts)							
I	A	B	C	AB	AC	BC	ABC	
1	-1	-1	-1	1	1	1	-1	$E(y_{---})$
1	1	-1	-1	-1	-1	1	1	$E(y_{+--})$
1	-1	1	-1	-1	1	-1	1	$E(y_{-++})$
1	1	1	-1	1	-1	-1	-1	
1	-1	-1	1	1	-1	-1	1	
1	1	-1	1	-1	1	-1	-1	
1	-1	1	1	-1	-1	1	-1	
1	1	1	1	1	1	1	1	$E(y_{+++})$

Defining relation: $b = ABC$

Block 1: $(- - -),(+ + -),(+ - +),(- + +)$

Block 2: $(+ - -),(- + -),(- - +),(+ + +)$

Confounding the 2^k Factorial Design with two blocks VI

Example 2: For 2^4 design with factors, A, B, C, D, the defining contrast (optimal) for blocking factor (b) is

$$b = ABCD$$

In general, the optimal blocking scheme for 2^k design with two blocks is given by $b = AB \cdots K$, where $A, B, \cdots, K$ are the factors.

Analyze Unreplicated Block 2^k Experiment

Filtration Experiment (four factors: A, B, C, D):

- Use defining relation: $b = ABCD$, i.e., if a treatment satisfies $ABCD = -1$, it is allocated to block 1 (b_1); if $ABCD = 1$, it is allocated to block 2 (b_2).
- Assume that, all the observations in block 2 will be reduced by 20 because of the poor quality of the second batch of material, i.e., the true block effect$= -20$.

Confounding the 2^k Factorial Design with two blocks VIII

factor				blocks	
A	B	C	D	$b = ABCD$	response
$-$	$-$	$-$	$-$	1=b_2	45-20=25
$+$	$-$	$-$	$-$	-1=b_1	71
$-$	$+$	$-$	$-$	-1=b_1	48
$+$	$+$	$-$	$-$	1=b_2	65-20=45
$-$	$-$	$+$	$-$	-1=b_1	68
$+$	$-$	$+$	$-$	1=b_2	60-20=40
$-$	$+$	$+$	$-$	1=b_2	80-20=60
$+$	$+$	$+$	$-$	-1=b_1	65
$-$	$-$	$-$	$+$	-1=b_1	43
$+$	$-$	$-$	$+$	1=b_2	100-20=80
$-$	$+$	$-$	$+$	1=b_2	45-20=25
$+$	$+$	$-$	$+$	-1=b_1	104
$-$	$-$	$+$	$+$	1=b_2	75-20=55
$+$	$-$	$+$	$+$	-1=b_1	86
$-$	$+$	$+$	$+$	-1=b_1	70
$+$	$+$	$+$	$+$	1=b_2	96-20=76

See Unreplicated-Block.SAS.

2^k Factorial Design with Four Blocks I

Need two 2-level blocking factors to generate 4 different blocks.
Confound each blocking factors with a high order factorial effect.
The interaction between these two blocking factors matters. The interaction will be confounded with another factorial effect.
In another word, for four blocks, select two effects to confound, automatically confounding a third effect.
Optimal blocking scheme has least confounding severity.

2^k Factorial Design with Four Blocks II

2^4 design with four blocks: factors are A, B, C, D and the blocking factors are $b1$ and $b2$

A	B	C	D	AB	AC	....	CD	ABC	ABD	ACD	BCD	ABCD
-1	-1	-1	-1	1	1		1	-1	-1	-1	-1	1
1	-1	-1	-1	-1	-1		1	1	1	1	-1	-1
-1	1	-1	-1	-1	1		1	1	1	-1	1	-1
1	1	-1	-1	1	-1		1	-1	-1	1	1	1
...	...	...	...	...	...	...	...	...	...	...	...	...
-1	-1	1	1	1	-1		1	1	1	-1	-1	1
1	1	1	1	-1	1		1	-1	-1	1	-1	-1
-1	-1	1	1	-1	-1		1	-1	-1	-1	1	-1
1	1	1	1	1	1		1	1	1	1	1	1

b1	b2	blocks
-1	-1	1
1	-1	2
-1	1	3
1	1	4

2^k Factorial Design with Four Blocks III

Possible blocking schemes:
Scheme 1: Defining relationships: $b_1 = ABC$, $b_2 = ACD$ induce confounding

$$b_1 b_2 = ABC * ACD = A^2BC^2D = BD$$

Scheme 2: Defining relationships: $b_1 = ABCD$, $b_2 = ABC$ induce confounding

$$b_1 b_2 = ABCD * ABC = D$$

which is better?

Confounding 2^k factorial design in 2^p blocks I

- k factors: $A, B, \cdots, K$, and p is usually much less than k.
- p blocking factors: $b_1, b_2, \cdots, b_p$ with levels -1 and 1.
- confound blocking factors with k chosen high-order factorial effects, i.e., $b_1 =$ effect 1, $b_2 =$ effect 2, etc. (p defining relations).
- These p defining relations induce another $2^p - p - 1$ confounding.
- treatment combinations with the same values of $b_1, \cdots, b_p$ are allocated to the dame block. Within each block.
- each block consists of 2^{k-p} treatment combinations (runs).
- Given k and p, optimal schemes are tabulated, e.g. Table 7.9

Confounding 2^k factorial design in 2^p blocks II

TABLE 7.9
Suggested Blocking Arrangements for the 2^k Factorial Design

Number of Factors, k	Number of Blocks, 2^p	Block Size, 2^{k-p}	Effects Chosen to Generate the Blocks	Interactions Confounded with Blocks
3	2	4	*ABC*	*ABC*
	4	2	*AB, AC*	*AB, AC, BC*
4	2	8	*ABCD*	*ABCD*
	4	4	*ABC, ACD*	*ABC, ACD, BD*
	8	2	*AB, BC, CD*	*AB, BC, CD, AC, BD, AD, ABCD*
5	2	16	*ABCDE*	*ABCDE*
	4	8	*ABC, CDE*	*ABC, CDE, ABDE*
	8	4	*ABE, BCE, CDE*	*ABE, BCE, CDE, AC, ABCD, BD, ADE*
	16	2	*AB, AC, CD, DE*	All two- and four-factor interactions (15 effects)
6	2	32	*ABCDEF*	*ABCDEF*
	4	16	*ABCF, CDEF*	*ABCF, CDEF, ABDE*
	8	8	*ABEF, ABCD, ACE*	*ABEF, ABCD, ACE, BCF, BDE, CDEF, ADF*
	16	4	*ABF, ACF, BDF, DEF*	*ABF, ACF, BDF, DEF, BC, ABCD, ABDE, AD, ACDE, CE, CDF, BCDEF, ABCEF, AEF, BE*
	32	2	*AB, BC, CD, DE, EF*	All two-, four-, and six-factor interactions (31 effects)
7	2	64	*ABCDEFG*	*ABCDEFG*
	4	32	*ABCFG, CDEFG*	*ABCFG, CDEFG, ABDE*
	8	16	*ABC, DEF, AFG*	*ABC, DEF, AFG, ABCDEF, BCFG, ADEG, BCDEG*
	16	8	*ABCD, EFG, CDE, ADG*	*ABCD, EFG, CDE, ADG, ABCDEFG, ABE, BCG, CDFG, ADEF, ACEG, ABFG, BCEF, BDEG, ACF, BDF*
	32	4	*ABG, BCG, CDG, DEG, EFG*	*ABG, BCG, CDG, DEG, EFG, AC, BD, CE, DF, AE, BF, ABCD, ABDE, ABEF, BCDE, BCEF, CDEF, ABCDEFG, ADG, ACDEG, ACEFG, ABDFG, ABCEG, BEG, BDEFG, CFG, ADEF, ACDF, ABCF, AFG, BCDFG*
	64	2	*AB, BC, CD, DE, EF, FG*	All two-, four-, and six-factor interactions (63 effects)

General advice about blocking I

- When in doubt, block.
- Block out the nuisance variables you know about, randomize as much as possible and rely on randomization to help balance out unknown nuisance effects.
- Measure the nuisance factors you know about but cannot control (ANCOVA).
- It may be a good idea to conduct the experiment in blocks even if there is not an obvious nuisance factor, just to protect against the loss of data or situations where the complete experiment can't be finished