실험계획

2수준 요인배치법- 교락법

서울시립대학교 통계학과 이용희

2021년 5월 18일

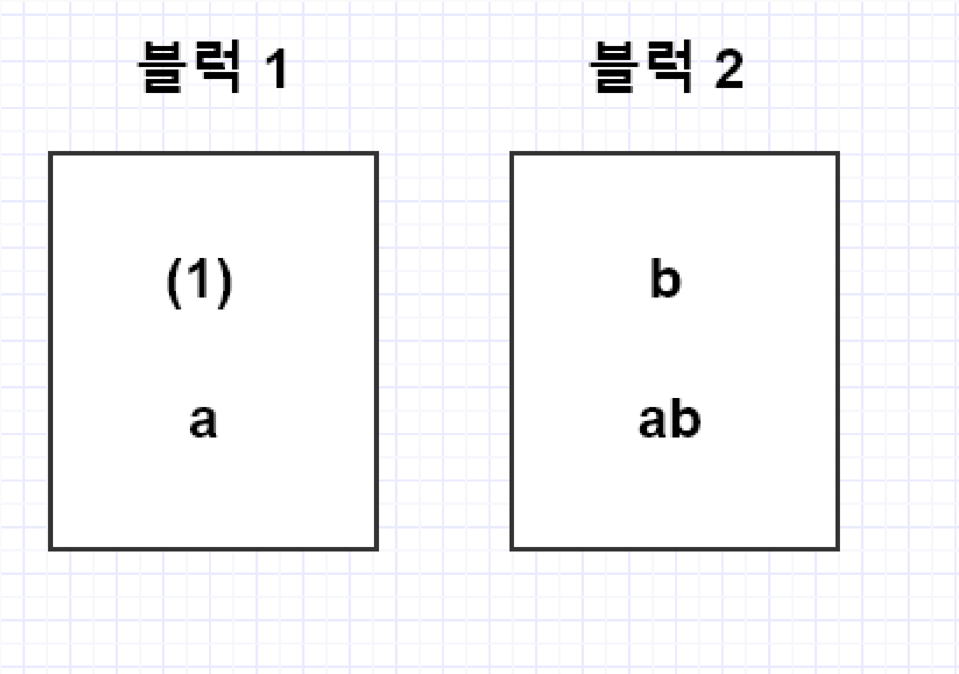
교락법

목적

- 2^k 요인배치법은 요인의 수 k가 증가하면 처리의 수 2^k 이 급격히 증가
- 처리의 개수가 증가하면 시간이 길어지고 여러 장소에서 실험을 수행해야 하므로 교락(confounding)의 요인이 증가
- 해결책은 2^k 개의 실험을 2^p 개의 블럭에서 실시
- 블럭들은 블럭 안에서는 동질적, 블럭 간에는 이질적으로 선정
- 교락법(confounding)은 완전 랜덤화 2수준 요인배치법의 처리들을 목적에 맞게 블럭에 배정하는 기술
- 교락법의 목적은 예상되는 주요 핵심요인의 효과와 블럭효과가 구별될 수 있게 실험을 배치하는 것

2^2 요인배치법과 블럭효과

- 2^2 요인배치법은 요인의 수가 2개, 각 요인의 수준수가 2개 이므로 4개의 처리에 대한 실험이 필요
- 실제로 실험을 시행할 때 피할 수 없는 블럭 효과가 생기기도 한다.
- 예를 들어 원료의 배치 1개를 이용해서 실험을 최대 2회 할 수 있다.
- 4개의 처리를 2개의 배치에 완전 랜덤화하여 배치할 수 있다.
- 만약 처리가 다음과 같이 배정되었을 때.... 블럭 효과가 존재한다면?
- 주요인 A, B 의 효과는 알아낼 수 있나?

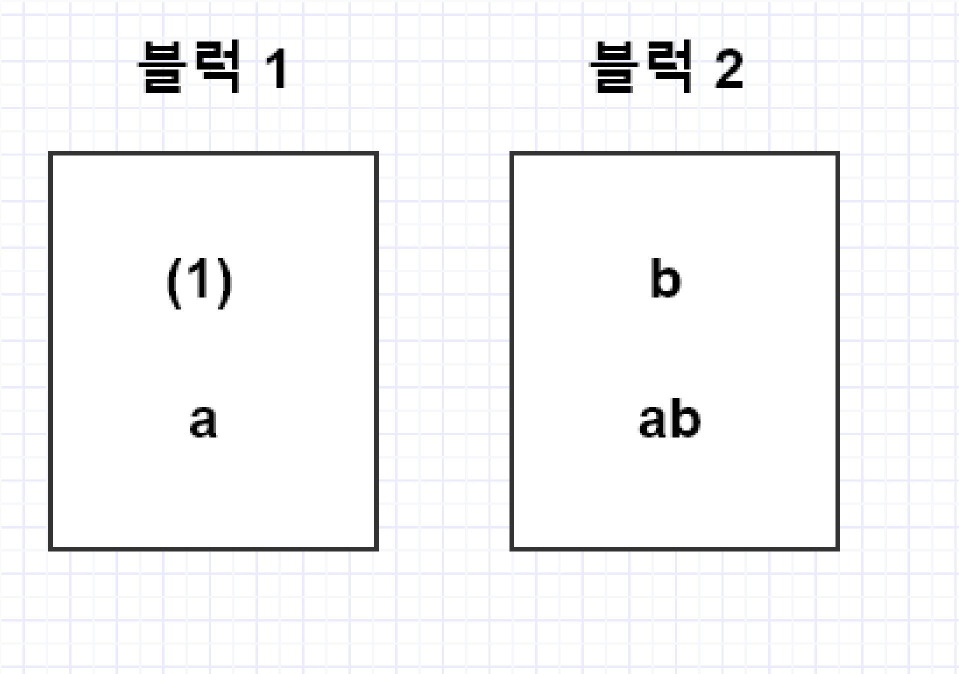


2^2 요인배치법과 블럭효과

- 먼저 블럭 효과는 어떻게 구할 수 있나?
- 만약 처리들의 효과가 모두 동일하다고 가정하면 블럭효과는 아래와 같이 추정할 수 있다.

$$\begin{aligned} Block\text{ effect} &= Block_2 - Block1 \\ &= \frac{1}{2}[b+ab] - \frac{1}{2}[a+(1)] \\ &= \frac{1}{2}[ab-a+b-1] \\ &= \frac{1}{2}(a+1)(b-1) \end{aligned}$$

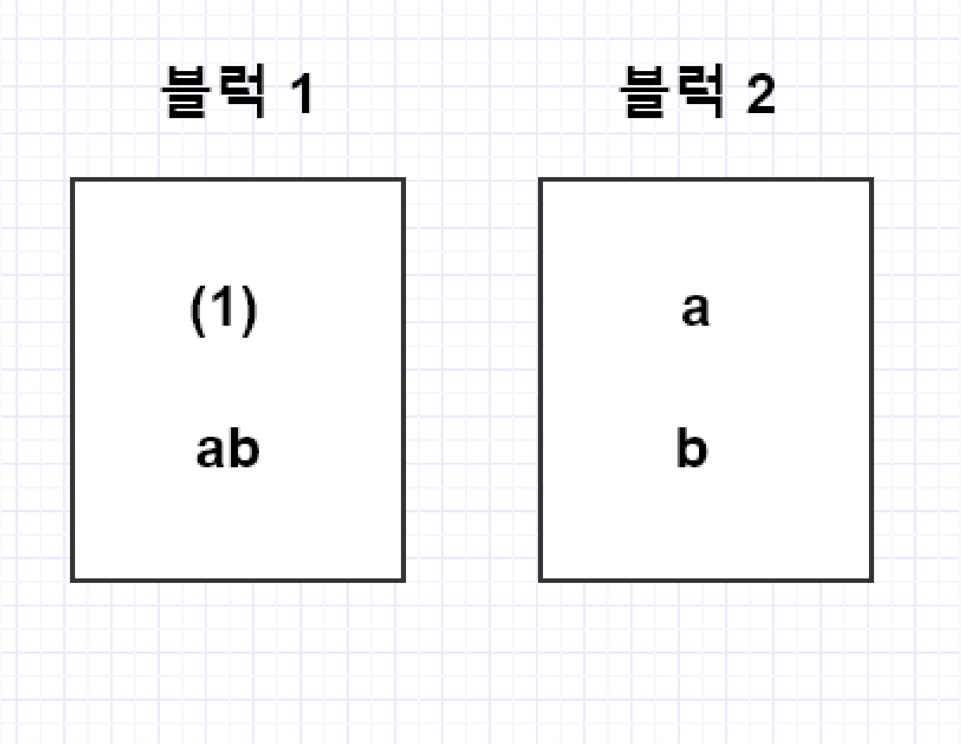
- 그런데 $(a+1)(b-1)$ 는 주효과 B 의 표기법과 동일하다.
- 블럭 효과와 주효과 B 는 교락되어서 구별할 수 없다.
- 만약 주효과 B 가 예상 핵심 요인이었다면..... ㅜㅜ



2^2 요인배치법과 블럭효과

- 실험계획의 핵심은 블럭을 인지하고, 블럭을 제어하고, 블럭을 이용하는 것!!!
- 이제 완전 랜덤회를 하지 않고 오른쪽과 같이 처리를 블럭에 배치하여 블럭안에서 랜덤하게 배치

$$\begin{aligned}
Block\text{ effect} &= Block_2 - Block1 \\
&= \frac{1}{2}[a+b] - \frac{1}{2}[ab+(1)] \\
&= -\frac{1}{2}[ab-a-b+1] \\
&= -\frac{1}{2}(a-1)(b-1)
\end{aligned}$$



- $(a-1)(b-1)$ 의 상호작용 $A \times B$ 는 블럭 효과와 교락
- 그러나 주효과 A 와 B 는 추정가능

2^2 요인배치법과 블럭효과

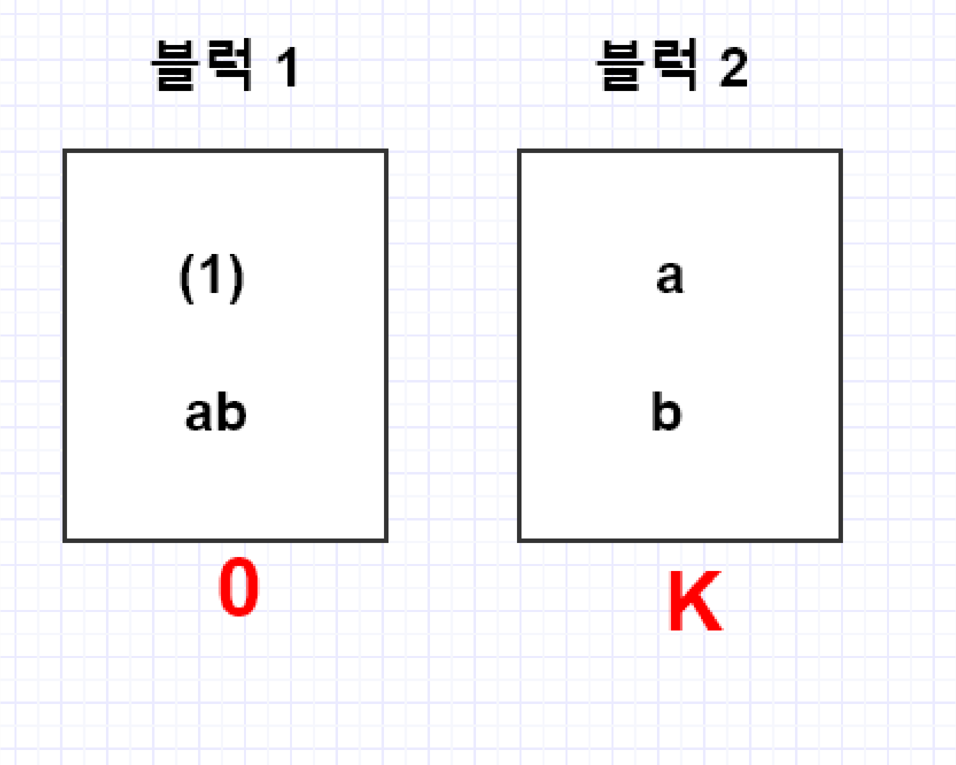
- 주효과 A 의 추정: 블럭효과가 없는 경우

$$A = \frac{1}{2}(a+1)(b-1)$$
$$= \frac{1}{2}[ab + b - a - (1)]$$

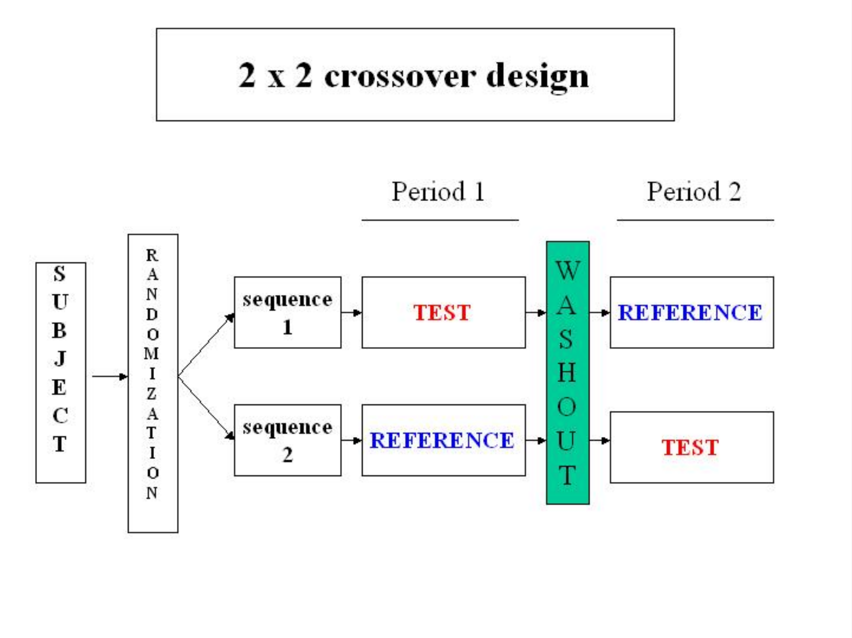
- 주효과 A 의 추정: 블럭효과가 있는 경우

$$A = \frac{1}{2}[ab + (b+k) - (a+k) - (1)]$$
$$= \frac{1}{2}[ab + b - a - (1)]$$

- 오른쪽의 실험배치는 블럭효과의 유무에 관계없이 주요인 A 와 B 효과 추정 가능



교차실험에서의 블럭효과



- 교차실험에서는 블럭효과가 잔여효과(Residual effect, Period effect)
- 잔여효과는 먼저 받은 치료(Period 1)가 두 번째 받은 치료 (Period 2)의 효과에 영향을 주는 효과
- 잔여효과가 치료에 관계없이 동일하다면 두 개의 처리를 비교할 수 있다.
- 잔여효과가 치료에 따라 다르다면 두 개의 처리 효과의 차이는 잔여효과와 교락
- 잔여효과를 최소화하기 위하여 washout 기간을 충분하게 가진다.

2^3 요인배치법과 블럭효과

- 2^3실험을 2개의 블록으로 나누어 실시. 각 블록에 2^{3-1}개의 처리를 효율적으로 배치? 그런데, SS_{Bblock} 의 자유도는 1. 블록과 교락될 효과?

배치 1: 주효과 $\boldsymbol{A}$가 블록과 교락된 배치법.

블록 1	블록 2
a	(1)
ab	b
ac	c
abc	bc

$SS_{Bblock} = SS_A$

$\Leftrightarrow$ 주효과 A와 블록효과가 서로 교락 & 구분 불가능.

주효과 A 추정 불가!!

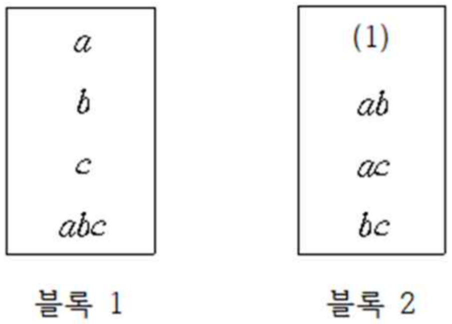
<그림 8.1> 주효과 A가 블록과 교락된 2^3요인배치법

- A 효과는 블럭효과와 교락

$$A = \frac{1}{4} = (a-1)(b+1)(c+1) = (a + ab + ac + abc - (1) - b - c - bc)$$

2^3 요인배치법과 블럭효과

배치 2: 상호작용효과 $\boldsymbol{A \times B \times C}$가 블록과 교락된 배치법.



<그림 8.2> 상호작용효과 $A \times B \times C$가 블록과 교락된 2^3요인배치법

$$SS_{Bbck} = SS_{A\times B\times C}$$

$\Leftrightarrow$ 상호작용효과 $A \times B \times C$와 블록효과가 서로 교락 & 구분 불가능.

- $A \times B \times C$ 효과는 블럭효과와 교락

$$A \times B \times C = \frac{1}{4} = (a-1)(b-1)(c-1) = (a+b+c+abc-(1)-ab-ac-bc)$$

- 블럭효과가 존재해도 A 효과는 추정 가능! 예를 들어 블럭 효과가 k 라고 하면(블럭1=0, 블럭2=k)

$$(a+[ab+k]+[ac+k]+abc-[(1)+k]-b-c-[bc+k]) = (a+ab+ac+abc-(1)-b-c-bc)$$

2^k 요인배치법과 교락법

- 교락법에서 효율적 실험 설계?
- **관심이 없는 요인효과인 고차의 상호작용효과를 블록효과와 교락 되도록 배치**
- 먼저 실험을 설계할 때 블럭 효과가 의심되는 요인이 있는지 판단(예: 다른 실험일, 실험자, 재료의 배치, 다른 공정라인 ...)
- 블럭이 있다고 판단되면 관심이 적은 상호작용효과를 블록효과와 교락 되도록 처리를 배치
- 블럭에 실험을 배치하는 방법
    - 인수분해식을 사용하는 방법
    - 선형표현식을 사용하는 방법

교락법: 인수분해식을 사용하는 방법

예제 8.1 $\mathbf{2^4}$요인배치법에서 최고차의 상호작용 $\boldsymbol{A \times B \times C \times D}$를 블록효과와 교락시켜 2개의 블록으로 나누어 실험 배치하기.

< 풀이>

(1) 인수분해식을 이용하는 방법

$$
\begin{aligned}
A \times B \times C \times D &= \frac{1}{8}(a-1)(b-1)(c-1)(d-1) \\
&= \frac{1}{8}((1) + ab + ac + ad + bc + bd + cd + abcd \\
&\qquad -a - b - c - d - abc - abd - acd - bcd\ )
\end{aligned}
$$

\+ 부호의 처리조합인 $(1), ab, ac, ad, bc, bd, cd, abcd$ 를 블록 1에,
\- 부호의 처리조합인 $a, b, c, d, abc, abd, acd, bcd$ 를 블록 2에 배치.

교락법: 인수분해식을 사용하는 방법

- $\boldsymbol{2^4}$실험을 $\boldsymbol{2^2}$개의 블록으로 나누어 실시. 각 블록에 $\boldsymbol{2^{4-2}}$개의 처리를 효율적으로 배치? $\boldsymbol{SS_{Block}}$ 의 자유도는 3. 블록과 교락될 효과?

  교락요인으로 상호작용효과 $A \times B \times C$와 상호작용 $B \times C \times D$를 선택.

  $$\begin{aligned} ABC &= \frac{1}{8}(a-1)(b-1)(c-1)(d+1) \\ &= \frac{1}{8}(a+b+c+ad+bd+cd+abc+abcd \\ &\qquad -(1)-d-ab-ac-bc-abd-acd-bcd) \end{aligned}$$

  \+ 부호를 갖는 처리조합 8개와 −부호를 갖는 처리조합 8개

  $$\begin{aligned} BCD &= \frac{1}{8}(a+1)(b-1)(c-1)(d-1) \\ &= \frac{1}{8}(b+c+d+ab+ac+ad+bcd+abcd \\ &\qquad -(1)-a-bc-bd-cd-abc-acd-abd) \end{aligned}$$

  ABC와 BCD 각각 + 부호를 갖는 처리조합 8개와 −부호를 갖는 처리조합 8개로 나누기. ABC, BCD에서 각각 (+,+), (+,-), (-,+), (-,-)로 되어 있는 4개의 처리조합을 각각 하나의 묶음으로 하여 4개의 블록으로 나누기.

교락법: 인수분해식을 사용하는 방법

- 블럭의 개수가 4개
- 블럭의 자유도는 3
- 효과의 자유도는 1

따라서

- 3개의 요인효과가 블럭 효과와 교락

예제에서 AD 는 ABC 와 BCD 의 결합요인

블록 1	블록 2	블록 3	블록 4
c	a	d	(1)
b	bd	ab	bc
ad	cd	ac	abd
$abcd$	abc	bcd	acd

< 그림 8.3 > ABC, BCD를 교락 시킨 2^4요인배치법의 교락법

블록과 교락될 나머지 효과?
ABC와 BCD의 상호작용효과인 $ABC \times BCD = AB^2C^2D = AD$

※블록효과와 2개 이상의 상호작용효과를 교락 시킬 때는 주효과나 구하고자 하는 상호작용효과가 결합상호작용효과가 되지 않도록 주의.

예> $ABCD$와 ABC가 교락 요인효과로 선택.
$\Rightarrow ABCD \times ABC = A^2B^2C^2D = D.$
주효과 D가 블록과 교락되어 주효과 D의 추정이 불가능

교락법: 선형표현식을 사용하는 방법

- 나머지 함수(MOD 함수, $\texttt{MOD}(x, y)$) : 정수 x 를 정수 y 로 나눈 나머지를 구하는 함수
  - 예: $\texttt{MOD}(1, 2) = 0$, $\texttt{MOD}(3, 2) = 1$, $\texttt{MOD}(7, 3) = 1$, $\texttt{MOD}(0, 2) = 0$
  - 모든 x 에 대하여 $\texttt{MOD}(x + x, 2) = 0$
  - 모든 x 에 대하여 $\texttt{MOD}(\texttt{MOD}(x, 2) + \texttt{MOD}(x, 2), 2) = 0$
- 요인에 대한 변수 x의 정의
  - 요인의 수준(처리의 유무)를 0 과 1 을 가지는 변수 x_i 로 표현
  - 예를 들어 3개의 요인 A, B, C 에 대하여 각각 수준을 x_1, x_2, x_3 로 표현
  - $A = 1$, $B = 0$, $C = 1$ 을 $x_1 = 1$, $x_2 = 0$, $x_3 = 1$
- **정의대비**(defining contrast) I
  - 어떤 효과가 블럭과 교락되는 가를 나타내는 대비이며 I로 표현
  - 예를 들어 상호작용 $A \times B \times C$를 블럭과 교락시키고 싶다면 정의대비는 다음과 같다.

$$I = ABC$$

교락법: 선형표현식을 사용하는 방법

- **선형표현식** L: 정의대비를 해당하는 변수 x 들의 합을 2로 나눈 나머지
- 예를 들어 정의대비가 $I = ABC$ 이면

$$L = x_1 + x_2 + x_3 (\mathrm{mod} 2) = \mathtt{MOD}(x_1 + x_2 + x_3, 2)$$

- 예를 들어 정의대비가 $I = AC$ 이면

$$L = x_1 + x_3 (\mathrm{mod} 2) = \mathtt{MOD}(x_1 + x_3, 2)$$

- 블럭에 처리를 배치하는 방법
    - 2^k 요인배치법에서 정의대비 I가 정한다.
    - 정의대비 I에 해당하는 선형표현식 L을 만든다.
    - 각 처리에 대한 선형표현식 L 값을 계산
    - 선형표현식 L 의 0 과 1 을 이용하여 처리를 블럭에 배치

교락법: 선형표현식을 사용하는 방법

<표 8.1> $I = ABC$일 때의 L의 값

처리조합	A B C	L의 값
(1)	0 0 0	L=0+0+0=0
a	1 0 0	L=1+0+0=1
b	0 1 0	L=0+1+0=1
ab	1 1 0	L=1+1+0=2=0
c	0 0 1	L=0+0+1=1
ac	1 0 1	L=1+0+1=2=0
bc	0 1 1	L=0+1+1=2=0
abc	1 1 1	L=1+1+1=3=1

따라서 L의 값이 0인 처리조합 $(1), ab, ac, bc$를 블록 1에 배치하고
L의 값이 1인 처리조합 a, b, c, abc 를 블록 2에 배치.

이 배치에서, 상호작용효과 $A \times B \times C$는 블록효과와 교락.

교락법: 선형표현식을 사용하는 방법

- 2^4 실험을 2^2 개의 블럭에 나누어 배치하려고 한다.
- 블럭에 교락할 효과는 ABC, BCD 로 선택
- 따라서 정의대비는

$$I = ABC, \quad I = BCD$$

- 또한 정의대비에 대한 선형표현식은

$$L_1 = x_1 + x_2 + x_3 (\mathrm{mod} 2), \quad L_2 = x_2 + x_3 + x_4 (\mathrm{mod} 2)$$

- 그런데 결합요인은 AD

$$(ABC)(BCD) = AB^2C^2D = AD$$

- 결합요인은 대한 정의대비는

$$I = AD$$

- 또한 결합요인에 대한 선형표현식은

$$L_3 = L_1 + L_2 (\mathrm{mod} 2)$$

교락법: 선형표현식을 사용하는 방법

<표 8.2> $I = ABC = BCD = AD$일 때, L_1, L_2, L_3의 값

처리조합	A	B	C	D	L_1	L_2	L_3
(1)	0	0	0	0	0	0	0
a	1	0	0	0	1	0	1
b	0	1	0	0	1	1	0
ab	1	1	0	0	0	1	1
c	0	0	1	0	1	1	0
ac	1	0	1	0	0	1	1
bc	0	1	1	0	0	0	0
abc	1	1	1	0	1	0	1
d	0	0	0	1	0	1	1
ad	1	0	0	1	1	1	0
bd	0	1	0	1	1	0	1
abd	1	1	0	1	0	0	0
cd	0	0	1	1	1	0	1
acd	1	0	1	1	0	0	0
bcd	0	1	1	1	0	1	1
abcd	1	1	1	1	1	1	0

교락법: 선형표현식을 사용하는 방법

$$
\begin{aligned}
L_3 &= L_1 + L_2 (\text{mod} 2) \\
&= \texttt{MOD}(\texttt{MOD}(x_1 + x_2 + x_3, 2) + \texttt{MOD}(x_2 + x_3 + x_4, 2), 2) \\
&= \texttt{MOD}([x_1 + x_2 + x_3] + [x_2 + x_3 + x_4], 2) \quad \text{(Distributive law)} \\
&= \texttt{MOD}(x_1 + 2x_2 + 2x_3 + x_4, 2) \\
&= \texttt{MOD}(x_1 + x_4, 2) \\
&= x_1 + x_4 (\text{mod} 2)
\end{aligned}
$$

교락법: 선형표현식을 사용하는 방법

블록1	블록2	블록3	블록4
(1)	d	a	c
bc	ab	bd	b
abd	ac	cd	ad
acd	bcd	abc	abcd
$L_1 = L_2 = 0$	$L_1 = 0, L_2 = 1$	$L_1 = 1, L_2 = 0$	$L_1 = L_2 = 1$

< 그림 8.4 > ABC, BCD, AD를 교락 시킨 2^4요인배치법

$$SS_{block} = SS_{A\times B\times C} + SS_{B\times C\times D} + SS_{A\times D}$$

- 다른 모든 요인효과들은 블록효과와 관계없이 독립적으로 추정
- 모든 요인들의 변동은 2^4 요인배치법에서와 동일.

교락법에 대한 R 패키지 `conf.design`

```r
> #install.packages("conf.design")
> library(conf.design)
```

- 2^3 요인배치법에서 $I = ABC$ 인 경우 배치법

```r
> Def.contrast1 <- c(1,1,1)
> conf.design(G = Def.contrast1 , p=2, block.name = "블럭", treatment.names=c("A", "B", "C"))
```

```
  블럭 A B C
1    0 0 0 0
2    0 1 1 0
3    0 1 0 1
4    0 0 1 1
5    1 1 0 0
6    1 0 1 0
7    1 0 0 1
8    1 1 1 1
```

교락법에 대한 R 패키지 `conf.design`

- 함수 `conf.design`의 인자
  - `G` : 정의대비 행렬, 각 행에 정의대비에 속하는 요인을 1로, 아니면 0
  - `p` : 수준의 개수
  - `block.name` : 블럭 이름
  - `treatment.names` : 요인의 이름
- 2^4 요인배치법에서 2^2 블럭에 $I = ABC, I = BCD$ 인 경우 배치법

```
> Def.contrast2 <- matrix(c(1,1,1,0, 0,1,1,1), 2,4, byrow=TRUE)
> Def.contrast2
```

```
     [,1] [,2] [,3] [,4]
[1,]    1    1    1    0
[2,]    0    1    1    1
```

교락법에 대한 R 패키지 `conf.design`

```
> conf.design(G=Def.contrast2 , p=2, block.name = "블럭", treatment.names=c("A", "B", "C", "D"))
```

```
   블럭 A B C D
1    00 0 0 0 0
2    00 0 1 1 0
3    00 1 1 0 1
4    00 1 0 1 1
5    01 1 1 0 0
6    01 1 0 1 0
7    01 0 0 0 1
8    01 0 1 1 1
9    10 1 0 0 0
10   10 1 1 1 0
11   10 0 1 0 1
12   10 0 0 1 1
13   11 0 1 0 0
14   11 0 0 1 0
15   11 1 0 0 1
16   11 1 1 1 1
```