실험계획

2수준 요인배치법

서울시립대학교 통계학과 이용희

2021년 5월 10일

2수준 배치법

목적

- 2^k 요인배치법은 요인이 k개, 각 요인의 수준 수가 모두 2개인 요인배치법.
- 여러 가지 다양한 요인들의 효과를 탐색하기 위한 초기 실험에 적합한 계획법이다.
- 결과에 영향을 미치는 중요한 요인을 알아내기 위한 계획법

개요

- 처리의수는 $2 \times 2 \times \cdots \times 2 = 2^k$
- 2^k 요인배치법에서 주효과와 상호작용효과가 중요한 관심 사형
- 각 효과와 변동은 전체 실험자료를 1그룹과 0그룹으로 반반씩 나누어서 두 그룹간의 실험자료의 평균 차이인 대비(contrast)를 이용
- 각 효과의 크기에 대한 반정규확률그림(HALF-NORMAL PROBABILITY PLOT)에 의해 핵심요인효과 선별
- 2^k 요인배치법부터는 반응변수를 y 로 표시

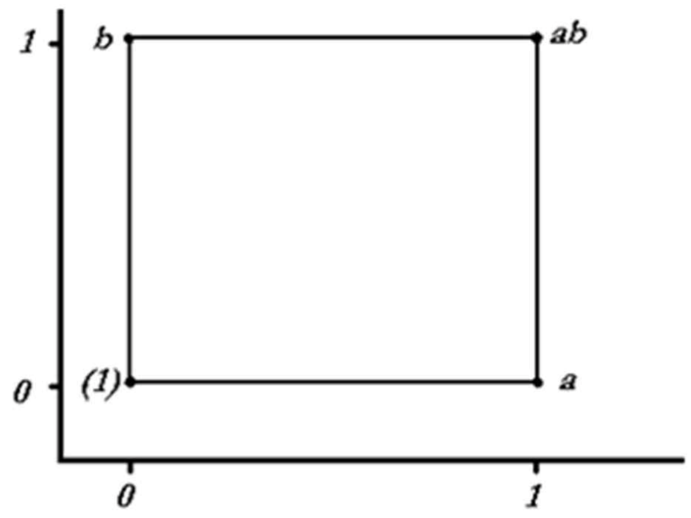
2^2 요인배치법

- 먼저 반복이 없는 경우를 고려
- 2^2 요인배치법은 요인의 수가 2개, 각 요인의 수준수가 2인 이원배치법 각각의 수준을 낮은 수준은 0 또는 -, 높은 수준은 1 또는 + 로 표시

< 표 7.2 > 2^2 요인배치법의 자료의 배열

	A_0	A_1	$T_{\cdot j}$
B_0	y_{00}	y_{10}	$T_{\cdot 0}$
B_1	y_{01}	y_{11}	$T_{\cdot 1}$
$T_{i\cdot}$	$T_{0\cdot}$	$T_{1\cdot}$	T

먼저 반복이 없는 2^2 요인배치법



< 그림 7.1 > 2^2요인배치법 실험자료의 시각적 표현

- 먼저 반복이 없는 2^2 요인배치법에서 반응변수 관측값 y_{ij} 를 다음과 같이 표시한다.

반응변수	요인 A	요인 B	요인의 조합	처리 표시
y_{00}	-	-	a^0b^0	(1)
y_{10}	+	-	a^1b^0	a
y_{01}	-	+	a^0b^1	b
y_{11}	+	+	a^1b^1	ab

- 소문자 알파벳이 나타나면 해당 요인은 높은 수준(+)으로 실험한 것
- 나타나지 않는 요임은 낮은 수준(-)에서 실험된 것

먼저 반복이 없는 2^2 요인배치법 - 효과 크기 추정

- 주효과 A 의 크기

$$\begin{aligned} A &= \bar{y}_{A+} - \bar{y}_{A-} \\ &= \frac{y_{11} + y_{10}}{2} - \frac{y_{01} + y_{00}}{2} \\ &= \frac{1}{2}[(y_{11} + y_{10}) - (y_{01} + y_{00})] \\ &= \frac{1}{2}[ab + a - b - (1)] \\ &= \frac{1}{2}(T_{1.} - T_{0.}) \end{aligned}$$

- 주효과 A 에 대한 대비

$$L = ab + a - b - (1) = y_{11} + y_{10} - y_{01} - y_{00} = T_{1.} - T_{0.}$$

- 주효과 A 에 대한 제곱합 (자유도 = 1)

$$SS_A = \frac{(T_{1.} - T_{0.})^2}{4}$$

먼저 반복이 없는 2^2 요인배치법 - 효과 크기 추정

- 주효과 B 의 크기

$$\begin{aligned} B &= \bar{y}_{B+} - \bar{y}_{B-} \\ &= \frac{y_{11} + y_{01}}{2} - \frac{y_{10} + y_{00}}{2} \\ &= \frac{1}{2}[(y_{11} + y_{01}) - (y_{10} + y_{00})] \\ &= \frac{1}{2}[ab + b - a - (1)] \\ &= \frac{1}{2}(T_{.1} - T_{.0}) \end{aligned}$$

- 주효과 B 에 대한 대비

$$L = ab + b - a - (1) = y_{11} + y_{01} - y_{10} - y_{00} = T_{.1} - T_{.0}$$

- 주효과 B 에 대한 제곱합 (자유도 = 1)

$$SS_B = \frac{(T_{.1} - T_{.0})^2}{4}$$

먼저 반복이 없는 2^2 요인배치법 - 효과 크기 추정

-상호작용 효과 $A \times B$ 의 크기

$$
\begin{aligned}
A \times B &= \frac{y_{11} - y_{10}}{2} - \frac{y_{01} - y_{00}}{2} \\
&= \frac{1}{2}[(y_{11} + y_{00}) - (y_{10} + y_{01})] \\
&= \frac{1}{2}[ab - a - b + (1)]
\end{aligned}
$$

- 상호작용 효과 $A \times B$ 에 대한 대비

$$L = ab - a - b + (1) = y_{11} - y_{01} - y_{10} + y_{00}$$

- 상호작용 효과 $A \times B$ 에 대한 제곱합 (자유도 = 1)

$$SS_{A \times B} = \frac{(y_{11} - y_{01} - y_{10} + y_{00})^2}{4}$$

먼저 반복이 없는 2^2 요인배치법 - 대비의 계산

- 각 요인 효과에 대한 대비는 다음과 같은 인수분해의 형태를 이용할 수 있다.

$$\begin{aligned} A &= (a-1)(b+1) = ab + a - b - (1) \\ B &= (a+1)(b-1) = ab - a + b - (1) \\ A \times B &= (a-1)(b-1) = ab - a - b + (1) \end{aligned}$$

<표 7.3> 2^2요인배치법에서 주효과와 상호작용효과를 구하는 표

처리조합	요인효과		
	A	B	AB
(1)	-1	-1	$+1$
a	$+1$	-1	-1
b	-1	$+1$	-1
ab	$+1$	$+1$	$+1$

반복이 없는 2^2 요인배치법 - 분산분석표

- 제곱합의 분해

$$SS_T = SS_A + SS_B + SS_{A\times B}$$

- 상호작용 $A \times B$와 오차가 교락

반복이 있는 2^2 요인배치법 - 자료의 구조

- 각 처리조합에서 r 번의 반복 측정이 있다.

반복이 있는 2^2 요인배치법 - 효과의 추정과 제곱합

- 효과의 추정

$$
\begin{aligned}
A &= \frac{1}{2r}(T_{1..} - T_{0..}) \\
B &= \frac{1}{2r}(T_{.1.} - T_{.0.}) \\
A \times B &= \frac{1}{2r}[(T_{11.} + T_{00.}) - (T_{10.} + T_{01.})]
\end{aligned}
$$

- 제곱합

$$
\begin{aligned}
SS_A &= \frac{(T_{1..} - T_{0..})^2}{4r} \\
SS_B &= \frac{(T_{.1.} - T_{.0.})^2}{4r} \\
SS_{A\times B} &= \frac{[(T_{11.} + T_{00.}) - (T_{10.} + T_{01.})]^2}{4r}
\end{aligned}
$$

반복이 있는 2^2 요인배치법 - 분산분석표

- 제곱합의 분해

$$SS_T = SS_A + SS_B + SS_{A\times B} + SS_E$$

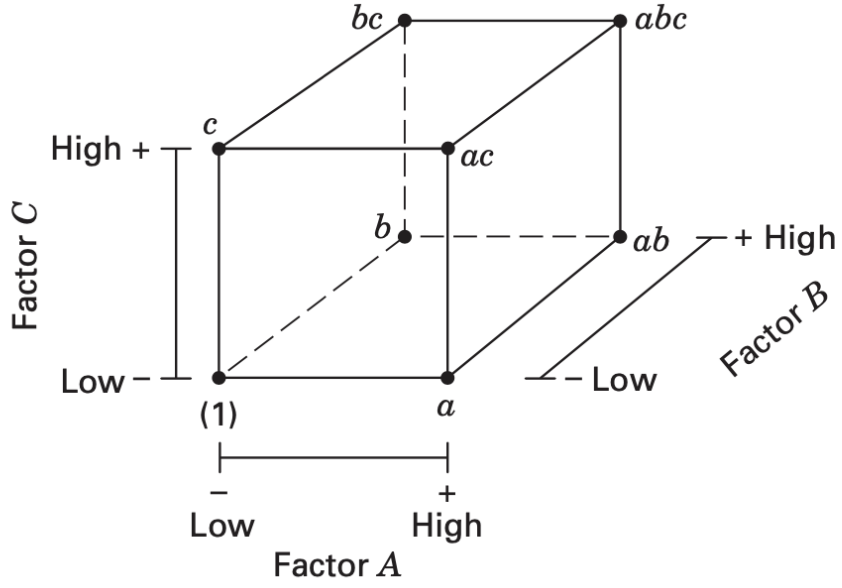
- $SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - T^2/4r$

<표 7.6> 반복이 있는 2^2요인배치법의 분산분석표

요인	제곱합	자유도	평균제곱	F_0
A	SS_A	1	MS_A	MS_A/MS_E
B	SS_B	1	MS_B	MS_B/MS_E
$A\times B$	$SS_{A\times B}$	1	$MS_{A\times B}$	$MS_{A\times B}/MS_E$
E	$SS_{A\times B}$	$4(r-1)$	MS_E	
T	SS_T	$4r-1$		

2^3 요인배치법 - 개요

- 2^3 요인배치법은 요인의 수가 3개, 각 요인의 수준수가 2인 이원배치법 각각의 수준을 낮은 수준은 0 또는 -, 높은 수준은 1 또는 + 로 표시
- 처리의 조합은 8개
- 각 처리에 r의 반복이 없는 경우를 고려



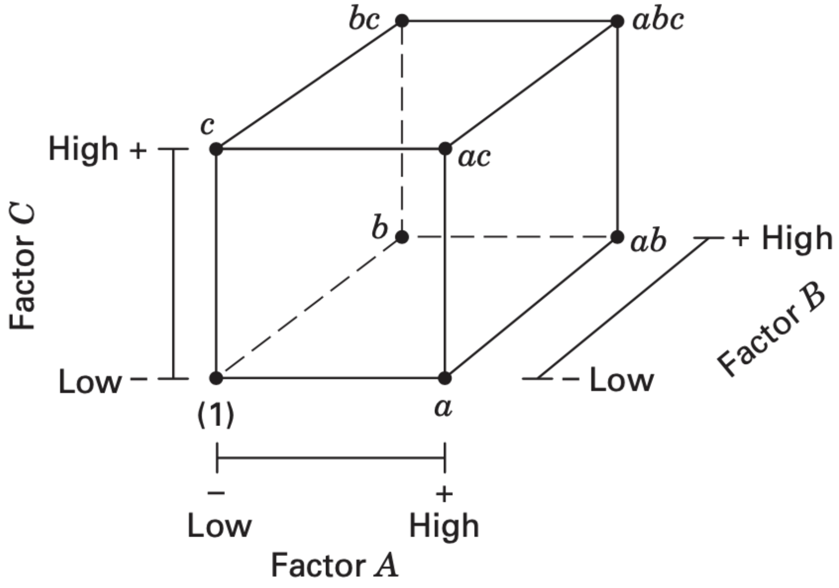
(a) Geometric view

	Factor		
Run	A	B	C
1	−	−	−
2	+	−	−
3	−	+	−
4	+	+	−
5	−	−	+
6	+	−	+
7	−	+	+
8	+	+	+

(b) Design matrix

2^3 요인배치법 - 개요

- 2^3 요인배치법은 요인의 수가 3개, 각 요인의 수준수가 2인 이원배치법 각각의 수준을 낮은 수준은 0 또는 -, 높은 수준은 1 또는 + 로 표시
- 처리의 조합은 8개
- 각 처리에 r의 반복이 없는 경우를 고려



(a) Geometric view

Run	Factor A	Factor B	Factor C
1	−	−	−
2	+	−	−
3	−	+	−
4	+	+	−
5	−	−	+
6	+	−	+
7	−	+	+
8	+	+	+

(b) Design matrix

2^3 요인배치법 - 개요

- 2^3 요인배치법은 요인의 수가 3개, 각 요인의 수준수가 2인 이원배치법 각각의 수준을 낮은 수준은 0 또는 -, 높은 수준은 1 또는 + 로 표시
- 처리의 조합은 8개
- 각 처리에 r의 반복이 없는 경우를 고려

Run	***A***	***B***	***C***	**Labels**	***A***	***B***	***C***
1	−	−	−	(1)	0	0	0
2	+	−	−	*a*	1	0	0
3	−	+	−	*b*	0	1	0
4	+	+	−	*ab*	1	1	0
5	−	−	+	*c*	0	0	1
6	+	−	+	*ac*	1	0	1
7	−	+	+	*bc*	0	1	1
8	+	+	+	*abc*	1	1	1

2^3 요인배치법 - 효과의 대비

- 효과의 대비 계산

$$\begin{aligned}
A &= (a-1)(b+1)(c+1) = (a+ab+ac+abc)-(b+c+bc+(1)) \\
B &= (a+1)(b-1)(c+1) \\
C &= (a+1)(b+1)(c-1) \\
A\times B &= (a-1)(b-1)(c+1) \\
A\times C &= (a-1)(b+1)(c-1) \\
B\times C &= (a+1)(b-1)(c-1) \\
A\times B\times C &= (a-1)(b-1)(c-1)
\end{aligned}$$

- 효과의 추정

$$\begin{aligned}
A &= \frac{1}{4r}(a-1)(b+1)(c+1) \\
&= \frac{1}{4r}[(a+ab+ac+abc)-(b+c+bc+(1))] \\
&= \frac{1}{4r}(T_{1..}-T_{0..})
\end{aligned}$$

2^3 요인배치법 - 효과의 대비

<표 7.8> 2^3요인배치법에서 주효과와 상호작용효과를 구하는 표

요인효과 / 처리조합	A	B	AB	C	AC	BC	ABC
(1)	−1	−1	+1	−1	+1	+1	−1
a	+1	−1	−1	−1	−1	+1	+1
b	−1	+1	−1	−1	+1	−1	+1
ab	+1	+1	+1	−1	−1	−1	−1
c	−1	−1	+1	+1	−1	−1	+1
ac	+1	−1	−1	+1	+1	−1	−1
bc	−1	+1	−1	+1	−1	+1	−1
abc	+1	+1	+1	+1	+1	+1	+1

2^3 요인배치법 - 효과의 대비

2^3 요인배치법 - 효과의 대비

반복 r인 2^3요인배치법에서 요인 효과와 변동을
축자적으로 계산하는 Yates 계산법

처리조합	(1)	(2)	(3)	요인효과 (3) /4r	변동 $(3)^2$ /8r
(1)	$a+(1)$	$ab+b+a+(1)$	$abc+bc+ac+c+ab+b+a+(1)$	M	CT
a	$ab+b$	$abc+bc+ac+c$	$abc-bc+ac-c+ab-b+a-(1)$	A	SS_A
b	$ac+c$	$ab-b+a-(1)$	$abc+bc-ac-c+ab+b-a-(1)$	B	SS_B
ab	$abc+bc$	$abc-bc+ac-c$	$abc-bc-ac+c+ab-b-a+(1)$	AB	$SS_{A\times B}$
c	$a-(1)$	$ab+b-a-(1)$	$abc+bc+ac+c-ab-b-a-(1)$	C	SS_C
ac	$ab-b$	$abc+bc-ac-c$	$abc-bc+ac-c-ab+b-a+(1)$	AC	$SS_{A\times C}$
bc	$ac-c$	$ab-b-a+(1)$	$abc+bc-ac-c-ab-b+a+(1)$	BC	$SS_{B\times C}$
abc	$abc-bc$	$abc-bc-ac+c$	$abc-bc-ac+c-ab+b+a-(1)$	ABC	$SS_{A\times B\times C}$

단 M은 $(3)/8r$

핵심 요인 효과의 선별

- 반복이 없는 2^3 요인배치법의 분산분석
- 귀무가설은 모든 요인효과가 존재하지 않는다. 귀무가설이 참인 경우에 다음의 각 처리에서 관측된 8개의 관측치는 $N(0, \sigma_E^2)$ 로부터 관측된 확률표본(random sample) 이다.

$$(1),\ a,\ b,\ c,\ ab,\ ac,\ bc,\ abc$$

- 7개의 각 효과의 추정치인 $\bar{y}_1 - \bar{y}_0$ 분포는 귀무가설이 참인 경우에 평균이 0이고 분산이 $\sigma_E^2/2$인 정규분포를 따르고 서로 직교한다.
- 각 효과의 추정치는 $N(0, \sigma_E^2/2)$분포를 따르는 모집단에서 뽑힌 확률표본이다.
- 그런데, 7개 요인효과의 변동을 크기 순서로 나열하는 것은 7개 요인효 과의 추정치의 절대값인 $|\bar{y}_1 - \bar{y}_0|$ 를 크기 순서로 나열하는 것과 동치.
- 선별하고 싶은 유의한 효과의 후보는 효과의 추정치의 절대값이 큰 효과.

핵심 요인 효과의 선별

- 요인효과의 추정치의 절대값인 $|\bar{y}_1 - \bar{y}_0|$ 에 대한 반정규확률 그림 그리기
  - 우선 7개 요인효과를 $|\bar{y}_1 - \bar{y}_0|$ 크기순서로 나열,
  - X 축에는 요인효과들의 $|\bar{y}_1 - \bar{y}_0|$ 값을,
  - Y축은 해당되는 요인효과의 경험적 누적확률에 대응되는 반 정규확률분포의 값인 백분위수을 기록하여 반정규확률 그림을 그리기.
- 요인효과들의 반정규확률 그림에서 원점을 지나는 선형패턴을 대략적으로 그린 후, 선형패턴을 벗어나 있는 요인효과들의 $|\bar{y}_1 - \bar{y}_0|$ 큰 즉, 오른쪽 에 있는 효과들을 핵심효과로 선별