실험계획 - 3주차 강의

4장: 이원배치법

시립대학교 통계학과

2021년 3월 16일

이원배치법의 개요

실험의 랜덤화

이원배치실험의 통계적 모형

상호작용

변동의 분해

가설 검정

반복이 없는 이원배치법

이원배치법의 개요

이원배치법

- 이원배치법은 실험에서 고려되는 **요인의 수가 2개**인 경우이다.
- 이원배치법에서 가장 중요한 관심 사항 중 하나는 두 요인의 상호작용(interaction)이다.
- 물론 일원배치에서 처럼 각 요인에 대한 주효과(main effects)도 중요하다
- 실험단위가 동질적. 처리를 제외한 나머지 조건들이 차이가 없음. 실험순서의 완전 랜덤화.
- [예제] 화합물 공정에서 반응변수인 효율이 반응온도 뿐만 아니라 촉매의 양에 따라서 영향을 받으리라 기대되는 경우

실험의 랜덤화

실험의 랜덤화

- 랜덤화를 통해서 분석의 타당성을 제공
- 반응치의 모평균들 간의 차이가 난 원인이 실험계획에서 고려된 요인 수준들의 차이에만 기인하도록 랜덤화 사용
- 각 처리가 두 요인 A와 B의 수준 조합인 A_iB_j에 의해서 결정된다는 점 만을 제외하고, 일원배치법과 동일

이원배치의 실험 배정 순서

- 반복 2회의 2×3 이원배치법. $N = 2 \times 3 \times 2 = 12$

요인 A \ 요인 B	B_1	B_2	B_3
A_1	1 2 ⑤	3 ① 4	5 ④ 6
A_2	7 8 ②	9 10	11 ③ 12

단계 1 : 각각의 실험조건에 1에서 12사이의 일련번호를 할당. (std order)

단계 2 : 1에서 12까지의 12개 숫자의 랜덤한 배열 구하기 (run order)
3, 8, 11, 5, 2, ⋯

단계 3 : 랜덤한 배열에서 나온 숫자에 해당되는 실험조건의 순서로 실험 실시.
실험실시의 순서가 원형숫자.
$A_1B_2, A_2B_1, A_2B_3, A_1B_3, A_1B_1, \cdots$ 순서로 실험 실시

자료의 구조 - 반복이 있는 균형자료

- 요인의 수가 2개: 요인 A 와 B
- 각 요인 A 와 B에 대한 수준 수가 각각 a개와 b개

$$A_1, A_2, A_3, \ldots, A_a$$

$$B_1, B_2, B_3, \ldots, B_b$$

- 이때 처리의 수는 ab 개
- 각 처리마다 반복 수가 r 인 실험
- 전체 실험의 횟수 $N = abr$

자료의 구조 - 반복이 있는 균형자료

< 표 4.2 > 이원배치법의 자료 구조

요인 A / 요인 B	B_1	B_2	$\cdots$	B_b
A_1	x_{111} x_{112} $\vdots$ x_{11r}	x_{121} x_{122} $\vdots$ x_{12r}	$\cdots$ $\cdots$ $\cdots$	x_{1b1} x_{1b2} $\vdots$ x_{1br}
A_2	x_{211} x_{212} $\vdots$ x_{21r}	x_{221} x_{222} $\vdots$ x_{22r}	$\cdots$ $\cdots$ $\cdots$	x_{2b1} x_{2b2} $\vdots$ x_{2br}
$\cdots$		$\cdots$		
A_a	x_{a11} x_{a12} $\vdots$ x_{a1r}	x_{a21} x_{a22} $\vdots$ x_{a2r}	$\cdots$ $\cdots$ $\cdots$	x_{ab1} x_{ab2} $\vdots$ x_{abr}

x_{ijk} : A의 i번째 수준과 B의 j번째 수준에서 k번째 반복으로 관측된 반응값

자료의 구조 - 일원배치와 사실상 동일

< 표 4.2 > 일원배치법의 자료 구조로 표현

실험반복 \ 처리수준	A_1B_1	A_1B_2	$\cdots$	A_aB_b
1	x_{111}	x_{121}	$\cdots$	x_{ab1}
2	x_{112}	x_{122}	$\cdots$	x_{ab2}
$\vdots$	$\vdots$	$\vdots$		
r	x_{11r}	x_{12r}	$\cdots$	x_{abr}

이원배치실험의 통계적 모형

반복이 있는 균형자료에 대한 통계적 모형

- 평균모형

$$x_{ijk} = \mu + \tau_{ij} + e_{ijk} \text{ or } x_{ijk} = \mu_{ij} + e_{ijk}$$

- 여기서 오차항 e_{ijk}는 정규분포 $N(0, \sigma_E^2)$를 따르며 모두 독립이다.
- **효과모형**

$$x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$$

1. 모수 α_i는 요인 A의 i 번째 수준의 효과
2. 모수 β_i는 요인 B의 j 번째 수준의 효과
3. 모수 $(\alpha\beta)_{ij}$는 A_i와 B_j간의 상호작용 효과

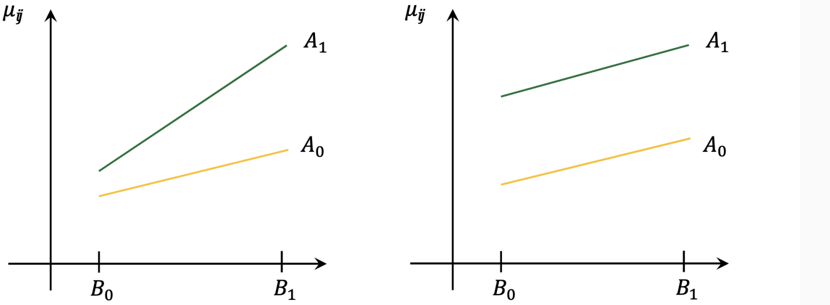
통계적 추론의 목적

$$x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$$

- 주효과와 상호작용효과의 해석
  1. 요인 A 의 수준에 따라서 반응치의 모평균에 차이가 나는가?
  $$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$$
  2. 요인 B 의 수준에 따라서 반응치의 모평균에 차이가 나는가?
  $$H_0 : \beta_1 = \beta_2 = \cdots = \beta_b = 0$$
  3. 상호작용효과: 주효과 A의 크기가 B 의 수준에 따라서 차이가 나는가?
  $$H_0 : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{ab} = 0$$
- 상호작용효과의 유의성 검정을 우선적으로 실시

상호작용

상호작용



상호작용 효과 $A \times B$가 존재

상호작용 효과 $A \times B$가 존재하지 않음

상호작용

다음과 같은 2개의 요인 A, B을 고려하고 각각 2개의 수준이 있다고 하자.

Table 1: 상호작용이 없는 경우

요인 A/B	A_0	A_1
B_0	0	2
B_1	2	4

Table 2: 상호작용이 있는 경우

요인 A/B	A_0	A_1
B_0	0	2
B_1	2	**6**

변동의 분해

총편차의 분해와 제곱합

$$\underbrace{(x_{ijk} - \bar{\bar{x}})}_{\text{total deviation}} = \underbrace{(\bar{x}_{ij.} - \bar{\bar{x}})}_{\text{between-treatment deviation}} + \underbrace{(x_{ijk} - \bar{x}_{ij.})}_{\text{within-treatment deviation}}$$

$$\begin{aligned} SS_T &= \sum_{i=1}^{a}\sum_{i=1}^{b}\sum_{k=1}^{r}(x_{ijk} - \bar{\bar{x}})^2 \\ &= \sum_{i=1}^{a}\sum_{i=1}^{b}\sum_{k=1}^{r}(\bar{x}_{ij.} - \bar{\bar{x}})^2 + \sum_{i=1}^{a}\sum_{i=1}^{b}\sum_{k=1}^{r}(x_{ijk} - \bar{x}_{ij.})^2 \\ &= SS_{AB} + SS_E \end{aligned}$$

처리그룹간의 편차의 분해와 제곱합

$$(\bar{x}_{ij.} - \bar{\bar{x}}) = \underbrace{(\bar{x}_{i..} - \bar{\bar{x}})}_{\text{A effect}} + \underbrace{(\bar{x}_{.j.} - \bar{\bar{x}})}_{\text{B effect}} + \underbrace{(\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{\bar{x}})}_{\text{interaction}}$$

$$\begin{aligned} SS_{AB} &= \sum_{i=1}^{a}\sum_{i=1}^{b}\sum_{k=1}^{r}(\bar{x}_{ij.} - \bar{\bar{x}})^2 \\ &= \sum_{i=1}^{a}\sum_{i=1}^{b}\sum_{k=1}^{r}(\bar{x}_{i..} - \bar{\bar{x}})^2 + \sum_{i=1}^{a}\sum_{i=1}^{b}\sum_{k=1}^{r}(\bar{x}_{.j.} - \bar{\bar{x}})^2 \\ &\quad + \sum_{i=1}^{a}\sum_{i=1}^{b}\sum_{k=1}^{r}(\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{\bar{x}})^2 \\ &= SS_A + SS_B + SS_{A\times B} \end{aligned}$$

자유도의 분할

제곱합	자유도
총변동 SS_T	$\phi_T = abr - 1$
AB간 변동 SS_{AB}	$\phi_{AB} = ab - 1$
잔차제곱합 SS_E	$\phi_E = \phi_T - \phi_{AB} = ab(r-1)$
요인 A 제곱합 SS_A	$\phi_A = a - 1$
요인 B 제곱합 SS_B	$\phi_B = b - 1$
상호작용 제곱합 $SS_{A\times B}$	$\phi_{A\times B} = (a-1)(b-1)$

여기서

$$\phi_{A\times B} = \phi_{AB} - \phi_A - \phi_B = \phi_A \times \phi_B = (a-1)(b-1)$$

가설 검정

가설검정의 순서

- 우선 상호작용효과의 유의성을 검토.

$$H_0 : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{ab} = 0$$

- $A \times B$ 상호작용효과가 유의하면, 주효과 A와 B의 유의성 검정은 기술적 의미가 없음
- 이미 주효과 A의크기가 B의수준에 따라서 다름

반복이 있는 이원배치법의 분산분석표

요인	제곱합	자유도	평균제곱합	F_0
요인 A	SS_A	$a-1$	MS_A	MS_A/MS_E
요인 B	SS_B	$b-1$	MS_B	MS_B/MS_E
상호작용 $A \times B$	$SS_{A\times B}$	$(a-1)(b-1)$	$MS_{A\times B}$	$MS_{A\times B}/MS$
잔차 E	SS_E	$ab(r-1)$	MS_E	
총합	SS_T	$abr-1$		

반복이 없는 이원배치법

반복이 없는 이원배치법에 대한 통계적 모형

- $r = 1$, 즉 아래의 효과모형에서 모든 $k = 1$

$$x_{ij1} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ij1}$$

- 반복이 없으면 상호작용 효과와 오차가 구분이 불가능
- 오차항의 자유도가 0 ($ab(r-1) = 0$)
- 상호작용 효과가 없는 모형을 사용

$$x_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

처리그룹간의 편차의 분해와 제곱합

$$\underbrace{(x_{ij} - \bar{\bar{x}})}_{\text{total deviation}} = \underbrace{(\bar{x}_{i.} - \bar{\bar{x}})}_{\text{A effect}} + \underbrace{(\bar{x}_{.j} - \bar{\bar{x}})}_{\text{B effect}} + \underbrace{(x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{\bar{x}})}_{\text{residual}}$$

$$\begin{aligned} SS_T &= \sum_{i=1}^{a}\sum_{i=1}^{b}(x_{ij} - \bar{\bar{x}})^2 \\ &= \sum_{i=1}^{a}\sum_{i=1}^{b}(\bar{x}_{i.} - \bar{\bar{x}})^2 + \sum_{i=1}^{a}\sum_{i=1}^{b}(\bar{x}_{.j} - \bar{\bar{x}})^2 \\ &\quad + \sum_{i=1}^{a}\sum_{i=1}^{b}(x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{\bar{x}})^2 \\ &= SS_A + SS_B + SS_E \end{aligned}$$

반복이 없는 이원배치법의 분산분석표

요인	제곱합	자유도	평균제곱합	F_0
요인 A	SS_A	$a-1$	MS_A	MS_A/MS_E
요인 B	SS_B	$b-1$	MS_B	MS_B/MS_E
잔차 E	SS_E	$(a-1)(b-1)$	MS_E	
총합	SS_T	$ab-1$		