



실험계획과 분석

심송용(한림대학교 데이터과학스쿨)

<http://jupiter.hallym.ac.kr>

이원배치법

이원배치분산분석(twoway ANOVA)

- ➡ 종속변수에 영향을 미치는 설명변수가 2개인 경우.
- ➡ 옥수수 수확량에 영향을 미치는 요인으로 비료의 종류만 고려하면 일원배치, 비료의 종류와 농약의 종류 두 개의 설명변수이면 이원배치
- ➡ 설명변수가 두 개 이상이면 설명변수들 사이의 상호작용이 존재할 수 있으므로 이에 대한 고려가 있어야 함.

비료의 종류 1, 2, 농약의 종류 1, 2 두 가지가 있을 때

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \epsilon_{ij} \sim N(0, \sigma^2), i = 1, 2, 3, 4.$$

τ_1 : 비료1, 농약1일 때의 처리효과

τ_2 : 비료1, 농약2일 때의 처리효과

τ_3 : 비료2, 농약1일 때의 처리효과

τ_4 : 비료2, 농약2일 때의 처리효과

이원배치법

인 모형을 생각할 수 있음.

비료 i , 농약 j 일 때의 k 번째 관측치를 y_{ijk} 라고 하면 모형을 ($i = 1, 2; j = 1, 2; k = 1, 2, \dots, n$)

$$y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}, \epsilon_{ijk} \sim N(0, \sigma^2)$$

로 사용할 수 있으며 $\alpha_i + \beta_j$ 가 앞 모형식의 τ_i 의 역할을 함.

이 두 모형은 상호작용이 존재할 때 적용하기 어려움.

상호작용(교호작용; interaction)

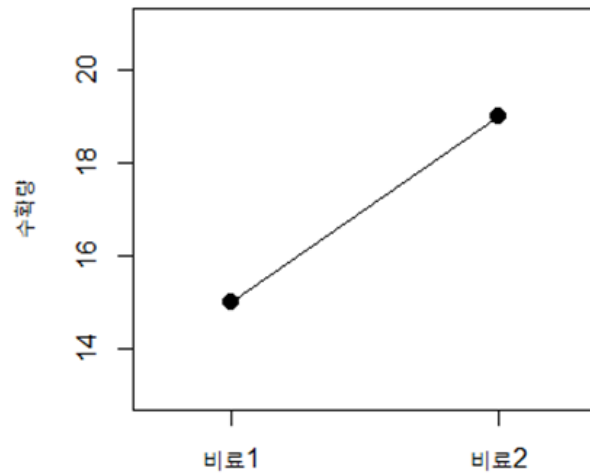
둘 이상의 요인이 있을 때 각 요인의 수준이 서로 교차작용하여 종속변수의 값에 주는 영향.
각 요인이 종속변수에 주는 영향은 **주효과**라고 함.

이원배치법

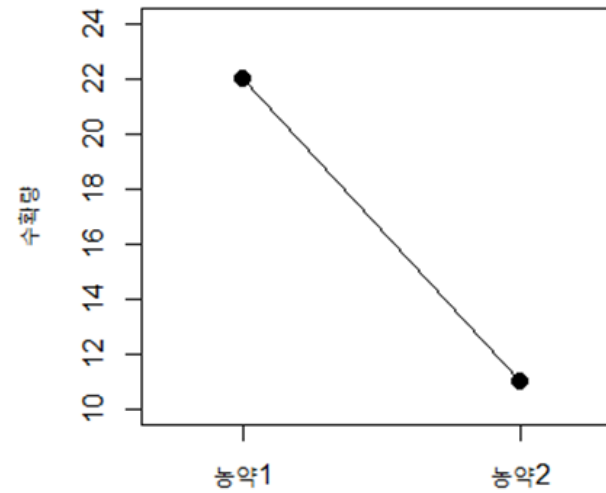
보기:

평균 수확량	농약1	농약2	비료별평균
비료1	20	10	15
비료2	24	14	19
농약별 평균	22	12	17

비료에 따른 평균



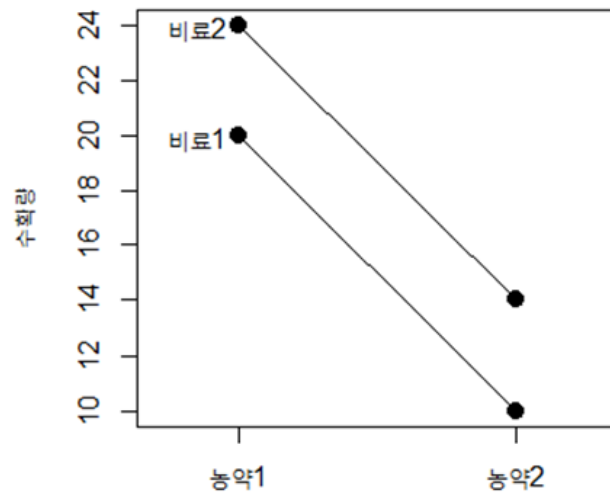
농약에 따른 평균



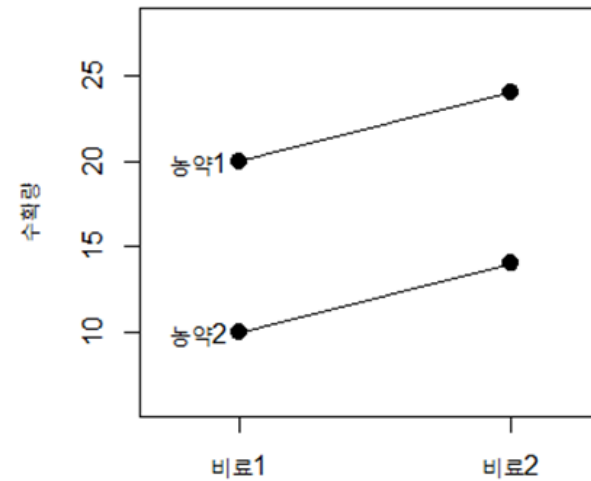
이원배치법

평균 수확량	농약1	농약2	비료별평균
비료1	20	10	15
비료2	24	14	19
농약별 평균	22	12	17

비료와 농약에 따른 수확량



또는

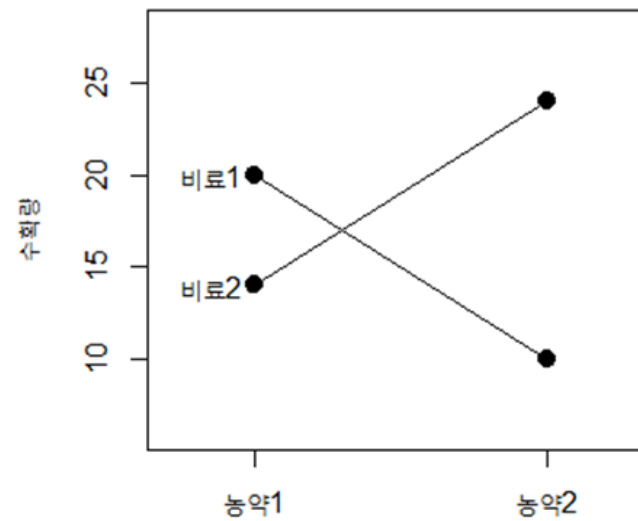
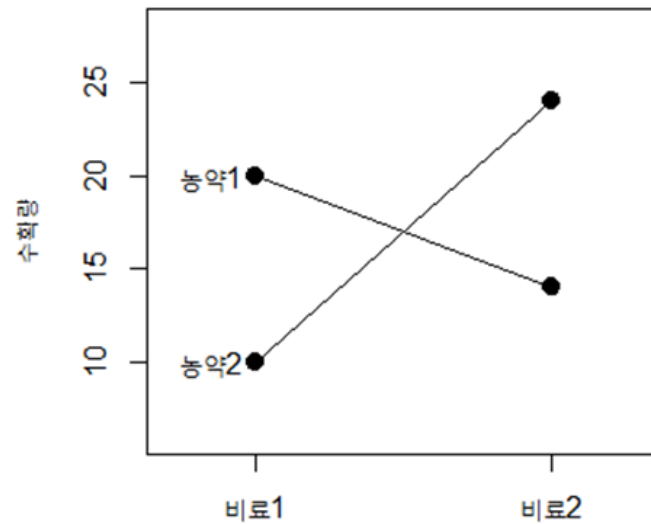


상호작용이 없는 경우임.

이원배치법

상호작용이 있는 경우

평균 수확량	농약1	농약2	비료별평균
비료1	20	10	15
비료2	14	24	19
농약별 평균	17	17	17



이원배치법

두 요인 A, B에 의한 이원배치법 자료

B A	1	2	...	b	평균
1	$y_{111}, y_{112}, \dots, y_{11n}$	$y_{121}, y_{122}, \dots, y_{12n}$...	$y_{1b1}, y_{1b2}, \dots, y_{1bn}$	$\bar{y}_{1..}$
2	$y_{211}, y_{212}, \dots, y_{21n}$	$y_{221}, y_{222}, \dots, y_{22n}$...	$y_{2b1}, y_{2b2}, \dots, y_{2bn}$	$\bar{y}_{2..}$
\vdots	\vdots	\vdots	...	\vdots	
a	$y_{a11}, y_{a12}, \dots, y_{a1n}$	$y_{a21}, y_{a22}, \dots, y_{a2n}$...	$y_{ab1}, y_{ab2}, \dots, y_{abn}$	$\bar{y}_{a..}$
평균	$\bar{y}_{.1.}$	$\bar{y}_{.2.}$...	$\bar{y}_{.b.}$	$\bar{y}_{...}$

$\bar{y}_{ij.}$: i 번째 행, j 번째 열에 해당하는 칸의 평균

- y_{ijk} : 요인 A의 i 번째 수준, 요인 B의 j 번째 수준에서 k 번째 관측값
- $y_{ij.}$: 요인 A의 i 번째 수준, 요인 B의 j 번째 수준에서의 합($y_{ij.} = \sum_{k=1}^n y_{ijk}$)

이원배치법

- $y_{i..}$: 요인 A의 i 번째 수준에서의 합($y_{i..} = \sum_{j=1}^b \sum_{k=1}^n y_{ijk}$)
- $y_{.j.}$: 요인 B의 j 번째 수준에서의 합($y_{.j.} = \sum_{i=1}^a \sum_{k=1}^n y_{ijk}$)
- $y_{...}$: 전체 자료합($y_{...} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}$)
- $\bar{y}_{ij.}$: 요인 A의 i 번째 수준, 요인 B의 j 번째 수준에서의 평균($\bar{y}_{ij.} = y_{ij.}/n = \sum_{k=1}^n y_{ijk}/n$)
- $\bar{y}_{i..}$: 요인 A의 i 번째 수준에서의 평균($\bar{y}_{i..} = y_{i..}/bn = \sum_{j=1}^b \sum_{k=1}^n y_{ijk}/bn$)
- $\bar{y}_{.j.}$: 요인 B의 j 번째 수준에서의 평균($\bar{y}_{.j.} = y_{.j.}/an = \sum_{i=1}^a \sum_{k=1}^n y_{ijk}/an$)
- $\bar{y}_{...}$: 전체 평균($\bar{y}_{...} = y_{...}/abn = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}/abn$)