정보통계학과 **백재욱** 교수

제12강(10장)

지분계획과 분할구계획

이번 시간

10.1 교차실험과 지분실험

10.2 이단지분계획

10.3 삼단지분계획

제12강(10장)

지분계획과 분할구계획

다음 시간

10.4 분할구계획

10.5 이단분할구계획

제12강 지분계획과 분할구계획

10.1 교차실험과 지분실험

10.1 교차실험과 지분실험

▪ 교차실험

어떤 요인의 각 수준마다 나타나는 또 다른 요인의 수준들이 모두 같은 실험

<합금강도데이터>

기압 \ 주조시간	10분	20분	30분
100도	1.1	1.3	1.2
	1.2	1.1	1.0
200도	1.3	1.3	1.4
	1.4	1.5	1.2
300도	1.8	2.1	2.2
	2.0	2.0	1.9

〈표 10-2〉

10.1 교차실험과 지분실험

▪ 지분실험 (계층적 실험)

어떤 요인의 각 수준마다 나타나는 또 다른 요인의 수준들이 모두 다른 실험

<표 10-1> 지분실험계획의 예 (단위 : mg)

강	1			2			3		
취수지	1	2	3	1	2	3	1	2	3
불소함량	1.1	1.3	1.2	1.3	1.3	1.4	1.8	2.1	2.2
	1.2	1.1	1.0	1.4	1.5	1.2	2.0	2.0	1.9

<표 10-3> 지분실험의 또 다른 표시법

강	1			2			3		
취수지	1	2	3	4	5	6	7	8	9
불소함량	1.1	1.3	1.2	1.3	1.3	1.4	1.8	2.1	2.2
	1.2	1.1	1.0	1.4	1.5	1.2	2.0	2.0	1.9

10.1 교차실험과 지분실험

<표 10-2> 교차 실험계획의 예

	요인 1		
요인 2	1	2	3
1	1.1 1.2	1.3 1.1	1.2 1.0
2	1.3 1.4	1.3 1.5	1.4 1.2
3	1.8 2.0	2.1 2.0	2.2 1.9

▶ 요인 1 = 주조시간
요인 2 = 기압

정보통계학과 **백재욱** 교수

제12강 지분계획과 분할구계획

10.2 이단지분계획

10.2 이단지분계획

▪ 모형

$$x_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{k(ij)}$$

$$i = 1,2, \ldots, a; \quad j = 1,2, \ldots, b; \quad k = 1,2, \ldots, n$$

- 균형지분계획(balanced nested design)임
- 상호작용효과(interaction effect)를 나타내는 항이 없음

10.2 이단지분계획

▪ 제곱합의 분할

$$SS_T = SS_A + SS_{B(A)} + SS_E \quad \cdots\cdots \quad (10.2)$$

$$SS_T = \sum_i \sum_j \sum_k (x_{ijk} - \bar{\bar{x}})^2 \quad \cdots\cdots \quad (10.3)$$

$$SS_A = \sum_i \sum_j \sum_k (\bar{x}_{i..} - \bar{\bar{x}})^2 \quad \cdots\cdots \quad (10.4)$$

$$SS_{B(A)} = \sum_i \sum_j \sum_k (\bar{x}_{ij.} - \bar{x}_{i..})^2 \quad \cdots\cdots \quad (10.5)$$

$$SS_E = \sum_i \sum_j \sum_k (x_{ijk} - \bar{x}_{ij.})^2 \quad \cdots\cdots \quad (10.6)$$

▪ 자유도의 분할

$$abn - 1 = (a-1) + a(b-1) + ab(n-1)$$

$$\cdots\cdots \quad (10.7)$$

$$\Phi_T = \Phi_A + \Phi_{B(A)} + \Phi_E$$

10.2 이단지분계획

<표 10-4> 이단지분계획에 대한 분산분석표

변인	제곱합	자유도	평균제곱
A	SS_A	$a-1$	MS_A
$B(A)$	$SS_{B(A)}$	$a(b-1)$	$MS_{B(\mathrm{A})}$
E	SS_E	$ab(n-1)$	MS_E
T	SS_T	$abn-1$	

10.2 이단지분계획

▪ 여러 가지 모형

① 고정효과 모형(A: 고정, B(A): 고정)

가정: α_i, $\beta_{j(i)}$는 상수, $\sum_i \alpha_i = \sum_j \beta_{j(i)} = 0$, $\varepsilon_{ijk} \sim N(0, \sigma^2)$ …… (10.8)

귀무가설: $H_0 : \alpha_i = 0, \quad i = 1, 2, \cdots, a$
$H_0 : \beta_{j(i)} = 0,\ i = 1, 2, \ldots, a;\ j = 1, 2, \ldots, b$ ………… (10.9)

② 랜덤효과 모형(A: 랜덤, B(A): 랜덤)

가정: $\alpha_i \sim N(0, \sigma_A^2)$, $\beta_{j(i)} \sim N(0, \sigma_{B(A)}^2)$, $\varepsilon_{ijk} \sim N(0, \sigma^2)$ …… (10.10)

귀무가설: $H_0 : \sigma_A^2 = 0$ ………………………………………… (10.11)
$H_0 : \sigma_{B(A)}^2 = 0$ ………………………………………… (10.12)

10.2 이단지분계획

▪ 여러 가지 모형(계속)

③ 혼합효과 모형(A: 고정, B(A): 랜덤)

가정: α_i는 상수, $\sum_i \alpha_i = 0$, $\beta_{j(i)} \sim N(0, \sigma^2_{B(A)})$, $\varepsilon_{ijk} \sim N(0, \sigma^2)$ ······(10.13)

귀무가설:

H_0 : $\alpha_i = 0$, $i = 1, 2, \cdots, a$ ································ (10.14)

H_0 : $\sigma^2_{B(A)} = 0$ ································ (10.15)

10.2 이단지분계획

- **여러 가지 모형(계속)**

<표 10-5> 이단지분계획에 대한 *E(MS)*

$E(MS)$	A 고정 B(A) 고정	A 랜덤 B(A) 랜덤	A 고정 B(A) 랜덤
$E(MS_A)$	$\sigma^2 + \frac{bn\sum_{i}^{a}\alpha_i^2}{a-1}$	$\sigma^2 + n\sigma^2_{B(A)} + bn\sigma^2_A$	$\sigma^2 + n\sigma^2_{B(A)} + \frac{bn\sum_{i}^{a}\alpha_i^2}{a-1}$
$E(MS_{B(A)})$	$\sigma^2 + \frac{n\sum_{i}^{a}\sum_{j}^{b}\beta_{j(i)}^2}{a(b-1)}$	$\sigma^2 + n\sigma^2_{B(A)}$	$\sigma^2 + n\sigma^2_{B(A)}$
$E(MS_E)$	σ^2	σ^2	σ^2

10.2 이단지분계획

예 10.1 약 투여 후 혈중 콜레스트롤 농축도 잼
세 가지 약은 시판되는 수십 종의 약 중에서 랜덤하게 선택
각 약을 만드는 제약회사도 여러 군데에서 두 군데씩 랜덤하게 선택

<표 10-6> 혈중 콜레스테롤 농축도

약 1		약 2		약 3	
제약회사 A	제약회사 Q	제약회사 D	제약회사 B	제약회사 L	제약회사 S
102	103	108	109	104	105
104	104	110	108	106	107

10.2 이단지분계획

풀이

$x_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}\ (i = 1,2,3;\ j = 1,2;\ k = 1,2)$..........................(10.16)

$\alpha_i \sim N(0, \sigma^2{}_A),\ \beta_{j(i)} \sim N\left(0, \sigma^2{}_{B(A)}\right),\ \varepsilon_{ijk} \sim N(0, \sigma^2)$..................................(10.17)

$H_0\colon \sigma^2{}_A = 0 \Leftrightarrow$ 콜레스테롤 농축도에 약 간 차이는 없다. ....(10.18)

$H_0\colon \sigma^2{}_{B(A)} = 0 \Leftrightarrow$ 콜레스테롤 농축도에 제약회사 간 차이는 없다. ....(10.19)

<표 10-7> 콜레스테롤 데이터에 대한 분산분석표

요인	제곱합	자유도	평균제곱	F_0
A: 약	61.167	2	30.583	61.167**
$B(A)$: 제약회사(약)	1.500	3	0.500	0.333
e: 오차	9.000	6	1.500	
T: 총	71.667	11		

10.2 이단지분계획

풀이(계속)

61.167 >> F(2, 3; 0.05) = 9.55

➡ 유의수준 5%에서 콜레스테롤 농축도에 약간 차이가 있다.

0.333 < F(3,6;0.05) = 4.76

➡ 콜레스테롤 농축도에 제약회사 간 차이가 있다고 할 수 없다.

10.2 이단지분계획

R 실습

```
nong = c(102, 104, 103, 104, 108, 110, 109, 108, 104, 106, 105, 107)
med = c(1,1,1,1,2,2,2,2,3,3,3,3)
comp = c(1,1,2,2,1,1,2,2,1,1,2,2)
ex10.1data = data.frame(med, comp, nong)
ex10.1data$med=factor(ex10.1data$med, levels=c(1, 2, 3), labels=c("med1","med2", "med3"))
ex10.1data$comp=factor(ex10.1data$comp, levels=c(1, 2), labels=c("comp1","comp2"))
attach(ex10.1data)
summary(aov(nong ~ med*comp, data=ex10.1data)) # 교차실험인 경우
summary(aov(nong ~ med+ med/comp, data=ex10.1data)) # med와 comp가 모두 고정요인인 지분실험의 경우
summary(aov(nong ~ med+ Error(med/comp), data=ex10.1data)) # med와 comp가 모두 랜덤요인인 지분실험의 경우
```

10.2 이단지분계획

R 실습

```
Error: med
      Df Sum Sq Mean Sq
med   2  61.17   30.58

Error: med:comp
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals  3    1.5     0.5

Error: Within
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals  6      9     1.5
```

10.2 이단지분계획

<표 10-5>의 랜덤모형에서 $\boldsymbol{a = 3, b = 2, n = 2}$를 대입하면

$$E(MS_A) = \sigma^2 + 2\sigma^2{}_{B(A)} + 4\sigma^2{}_A \quad \text{.........................(10.20)}$$

$$E(MS_{B(A)}) = \sigma^2 + 2\sigma^2{}_{B(A)} \quad \text{...................................(10.21)}$$

$$E(MS_E) = \sigma^2 \text{..(10.22)}$$

σ^2, $\sigma^2{}_{B(A)}$, $\sigma^2{}_A$?

10.2 이단지분계획

(10.22) → $\widehat{\sigma^2} = 1.500$ ....(10.23)

(10.21) → $\widehat{\sigma^2} + 2\,\widehat{\sigma^2{}_{B(A)}} = MS_{B(A)}$....................................(10.24)

$\widehat{\sigma^2{}_{B(A)}} = \frac{MS_{B(A)} - MS_E}{2} = \frac{0.5000 - 1.5000}{2} = -0.5000 < 0$ ......(10.25)

$\Rightarrow\ \sigma^2{}_{B(A)} = 0$ 으로 놓음

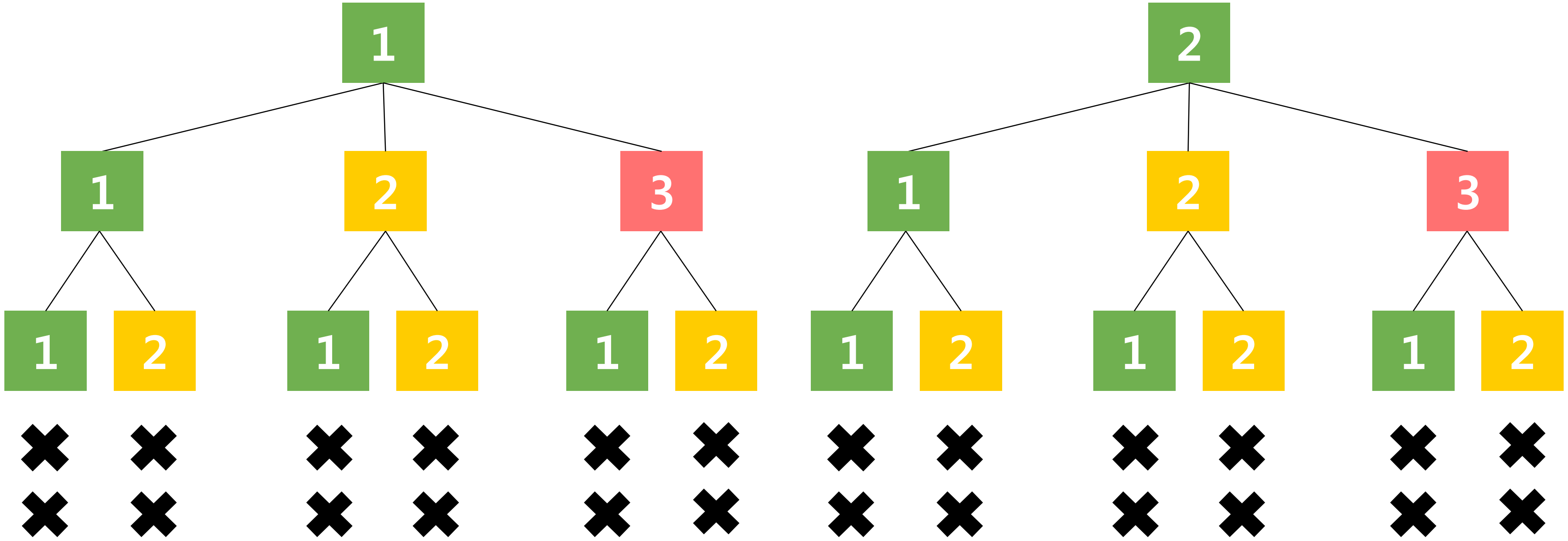
(10.20) → $\widehat{\sigma^2} + 2\,\widehat{\sigma^2{}_{B(A)}} + 4\,\widehat{\sigma^2{}_A} = MS_A$ ....(10.27)

$\widehat{\sigma^2{}_A} = \frac{MS_A - MS_E - 2\widehat{\sigma^2{}_{B(A)}}}{4} = \frac{30.583 - 1.500 - 2 \times 0.000}{4} = 7.271$ ...(10.28)

제12강 지분계획과 분할구계획

10.3 삼단지분계획

10.3 삼단지분계획



[그림 10-1] 삼단지분계획

10.3 삼단지분계획

예 많은 트럭이 들어옴! 트럭 안에 큰 박스 많음!
큰 박스 안에 작은 박스 많음!
트럭 간, 큰 박스 간, 그리고 작은 박스 간 차이가 있을까?

- **모형**

$$x_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \gamma_{k(ij)} + \varepsilon_{l(ijk)} \quad \text{............}(10.32)$$

$$(i = 1,2,\ldots,a;\ j = 1,2,\ldots,b;\ k = 1,2,\ldots,c;\ l = 1,2,\ldots,n)$$

10.3 삼단지분계획

- ## 제곱합의 분할

$$SS_T = SS_A + SS_{B(A)} + SS_{C(AB)} + SS_E \quad ....(10.33)$$

$$SS_T = \sum_i \sum_j \sum_k \sum_l (x_{ijkl} - \bar{\bar{x}})^2 \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (10.34)$$

$$SS_A = \sum_i \sum_j \sum_k \sum_l (\bar{x}_{i...} - \bar{\bar{x}})^2 \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (10.35)$$

$$SS_{B(A)} = \sum_i \sum_j \sum_k \sum_l (\bar{x}_{ij..} - \bar{x}_{i...})^2 \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (10.36)$$

$$SS_{C(AB)} = \sum_i \sum_j \sum_k \sum_l (\bar{x}_{ijk.} - \bar{x}_{ij..})^2 \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (10.37)$$

$$SS_E = \sum_i \sum_j \sum_k \sum_l (x_{ijkl} - \bar{x}_{ijk.})^2 \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (10.38)$$

- ## 자유도의 분할

$$abcn - 1 = (a-1) + a(b-1) + ab(c-1) + abc(n-1) \quad \cdots\cdots \quad (10.39)$$

$$\phi_T = \phi_A + \phi_{B(A)} + \phi_{C(AB)} + \phi_E$$

10.3 삼단지분계획

<표 10-8> 삼단지분계획에 대한 분산분석표(랜덤효과모형)

변인	제곱합	자유도	평균제곱	F_0
A	SS_A	$a-1$	MS_A	$MS_A/MS_{B(A)}$
$B(A)$	$SS_{B(A)}$	$a(b-1)$	$MS_{B(A)}$	$MS_{B(A)}/MS_{C(AB)}$
$C(AB)$	$SS_{C(AB)}$	$ab(c-1)$	$MS_{C(AB)}$	$MS_{C(AB)}/MS_E$
E	SS_E	$abc(n-1)$	MS_E	
T	SS_T	$abcn-1$		

10.3 삼단지분계획

예 **A, B, C 모두 변량인자이다. $\widehat{\boldsymbol{\sigma}^2_{C(AB)}}$ 값은 얼마인가?**

요인	SS	df
$\boldsymbol{A}$	91	1
$\boldsymbol{B(A)}$	68	6
$\boldsymbol{C(AB)}$	24	
$\boldsymbol{e}$	8	16
$\boldsymbol{T}$		31

풀이 A, B, C가 각각 수준수가 l, m, n이고 r회 반복측정이 이루어진 경우

$E\left(MS_{C(AB)}\right) = \sigma^2 + r{\sigma_{C(AB)}}^2$ 이고, $E(MS_E) = \sigma^2$이므로

$$\widehat{\boldsymbol{\sigma}^2_{C(AB)}} = \frac{MS_{C(AB)} - MS_E}{r} = \frac{3-0.5}{2} = 1.25$$ 이다.

정보통계학과 **백재욱** 교수

다음 시간 안내

제13강(10장)

지분계획과 분할구계획