정보통계학과 **백재욱** 교수

제7강 (7장)

요인배치법 1

이번시간

정보통계학과 **백재욱** 교수

제7강 (7장)

요인배치법 1

다음시간

7.4 2^n 요인배치법

7.5 3^2 요인배치법

7.6 회귀모형

제7강 요인배치법1

7.1 대비와 직교분해

7.1 대비와 직교분해

◆ k^n 요인배치법contrast)

- 요인의 수가 n이고 각 요인의 수준수가 k인 실험
- 요인의 수 및 수준수가 늘어나면 실험의 횟수가 비약적으로 증가함

◆ 대비($contrast$)

$T_1,\ T_2,\ \cdots,\ T_a$가 각 처리수준에서 측정값들의 합인 경우
선형식 $L = c_1 T_1 + c_2 T_2 + \cdots + c_a T_a$ 에서

$$c_1 + c_2 + \cdots + c_a = 0 \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (7.3)$$

인 조건이 만족될 때 이 선형식을 **대비**라 한다.

대비의 변동 $SS_L = \frac{L^2}{D \times r}$ 으로 자유도 1 ($D = c_1^2 + c_2^2 + \cdots + c_a^2$, r은 반복수)이다.

7.1 대비와 직교분해

◆ 직교 (*orthogonal*)

2개의 대비가

$$L_1 = c_1 T_1 + c_2 T_2 + \cdots + c_a T_a$$
$$L_2 = d_1 T_1 + d_2 T_2 + \cdots + d_a T_a$$

로 주어질 때

$$c_1 d_1 + c_2 d_2 + \cdots + c_a d_a = 0$$

이 성립하면, 2개의 대비 L_1과 L_2는 서로 **직교(orthogonal)**한다고 말한다.

7.1 대비와 직교분해

◆ 직교분해 (*orthogonal decomposition*)

수준수가 a인 처리 A 의 변동 SS_A는 각각의 자유도가 1인 서로 직교하는 $(a-1)$개의 대비에 의한 변동으로 분해할 수 있다.
이때 서로 직교하는 대비를

$$L_1,\ L_2,\ \cdots,\ L_{a-1}$$

이라 하면 다음이 성립한다.

$$SS_A = SS_{L_1} + SS_{L_2} + \cdots + SS_{L_{a-1}}$$

7.1 대비와 직교분해

예제 7.1

1) 원료의 제조회사 ***A*** (A_0 : 자회사, A_1 : 국내 타회사, A_2 : 외국회사)와 성형온도 B (B_0 : 100℃, B_1 : 110℃ B_2 : 120℃)는 플라스틱 강도에 영향을 미치는가?

2) 제조회사(A)가 영향을 미친다면 이는 '국산과 외제 간의 차이' 때문인가 아니면 '자사와 국내 타회사 간의 차이' 때문인가?

3) 계량요인(B)이 영향을 미친다면 온도의 효과는 1차적인가 2차적인가?

<표 7-1> 플라스틱 강도

A \ B	B_0	B_1	B_2	$T_{i.}$
A_0	11	18	25	54
A_1	1	6	14	21
A_2	6	15	18	39
$T_{.j}$	18	39	57	T = 114

7.1 대비와 직교분해

풀이 1) 원료의 제조회사 A와 성형온도 B는 플라스틱강도에 유의한 영향을 미치는가?

< 분산분석표 >

요 인	제곱합	자유도	평균제곱	F_0
A	182	2	91	45.5**
B	254	2	127	63.5**
E	8	4	2	
T	444	8		

7.1 대비와 직교분해

풀이 2) '국산과 외제 간의 차이' 때문인가?
또는 '자회사와 국내 타회사 간의 차이' 때문인가?

$$L_1 = \text{국산과 외제 간의 차이} = \frac{1}{6}(T_{0.} + T_{1.}) - \frac{1}{3}T_{2.} = -0.5$$

$$L_2 = \text{자회사와 국내 타회사 간의 차이} = \frac{1}{3}T_{0.} - \frac{1}{3}T_{1.} = 11$$

2개의 선형식 L_1과 L_2 의 계수는 $c_1 = \frac{1}{6}$, $c_2 = \frac{1}{6}$, $c_3 = -\frac{1}{3}$

$d_1 = \frac{1}{3}$, $d_2 = -\frac{1}{3}$, $d_3 = 0$ 이므로

7.1 대비와 직교분해

풀이 (계속)

선형식 L_1 과 L_2 는 대비이며 **서로 직교**한다.

$$SS_A = SS_{L_1} + SS_{L_2}$$

$$SS_{L_1} = \frac{(L_1)^2}{(\sum c_i^2) \times r} = \frac{(-0.5)^2}{\left\{(\frac{1}{6})^2+(\frac{1}{6})^2+(-\frac{1}{3})^2\right\} \times 3} = 0.5$$

$$SS_{L_2} = \frac{(L_2)^2}{(\sum d_i^2) \times r} = \frac{(11)^2}{\left\{(\frac{1}{3})^2+(-\frac{1}{3})^2\right\} \times 3} = 181.5$$

$SS_{L_1} <<< SS_{L_2}$ ➡ 제조회사 A에 의한 변동은 '국산과 외제 간의 차이' 때문이기 보다는 '자회사와 국내 타 회사 간의 차이' 때문이다.

7.1 대비와 직교분해

풀이 3) 성형온도의 경우 1차적(linear) 효과가 큰가?
또는 2차적(quadratic)인 효과가 큰가?

1차 효과 : $L_l = (T_{.1} - T_{.0}) + (T_{.2} - T_{.1}) = T_{.2} - T_{.0} = 39$

로 주어지고 $c_1 = -1, c_2 = 0, c_3 = +1$ 이므로 대비이다.

2차 효과 : $L_q = (T_{.2} - T_{.1}) - (T_{.1} - T_{.0}) = T_{.2} - 2T_{.1} + T_{.0} = -3$

으로 주어지고 $d_1 = 1,\ d_2 = -2,\ d_3 = 1$이므로 대비이고

L_l 과 L_q 는 서로 직교한다(직접 확인!).

7.1 대비와 직교분해

풀이 (계속)

각 대비의 변동: $SS_l = \frac{(39)^2}{2 \times 3} = 253.5$

$$SS_q = \frac{(-3)^2}{6 \times 3} = 0.5$$

$$SS_B = SS_l + SS_q = 254$$

$$SS_l >>> SS_q$$

➡ 변동 중에서 1차 효과가 대부분 차지함

7.1 대비와 직교분해

풀이 (계속)

요 인	제곱합	자유도	평균제곱	F_0
A	182	2	91	45.5**
L_1	0.5	1	0.5	0.25
L_2	181.5	1	181.5	90.75**
B	254	2	127	63.5**
L_l	253.5	1	253.5	117.75**
L_q	0.5	1	0.5	0.25
E	8	4	2	
T	444	8		

7.1 대비와 직교분해

R 실습

```
gang <- c(11, 18, 25, 1, 6, 14, 6, 15, 18)
wol <- c(rep(0, 3), rep(1, 3), rep(2, 3))
temp <- c(rep(c(0, 1, 2), 3))
plastic.data <- data.frame(gang, wol, temp)
plastic.data$wol <- factor(plastic.data$wol, levels=c(0, 1, 2), labels=c("a0", "a1", "a2"))
plastic.data$temp <- factor(plastic.data$temp, levels=c(0, 1, 2), labels=c("b0", "b1", "b2"))
anova <- aov(gang ~ wol +  temp, data=plastic.data)
summary(anova)
```

```
            Df   Sum Sq   Mean Sq   F value   Pr(>F)
wol          2      182        91      45.5   0.001773 **
temp         2      254       127      63.5   0.000932 ***
Residuals    4        8         2
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

7.1 대비와 직교분해

R 실습

```
c1 = c(1/6,1/6,-1/3)
c2 = c(1/3, -1/3, 0)
mat.wol = cbind(c1, c2)
contrasts(plastic.data$wol) = mat.wol
c3 = c(-1, 0, 1)
c4 = c(1, -2, 1)
mat.temp = cbind(c3, c4)
contrasts(plastic.data$temp) = mat.temp
mod.contrast = aov(gang ~ wol + temp, data = plastic.data)
summary(mod.contrast, split=list(wol=list("국산과 외제"=1, "자사와 국내 타회사"=2),
temp=list("linear"=1, "quadratic"=2)))
```

```
                          Df  Sum Sq  Mean Sq  F value  Pr(>F)
wol                        2   182.0     91.0    45.50  0.001773 **
   wol: 국산과 외제          1     0.5      0.5     0.25  0.643330
   wol: 자사와 국내 타회사    1   181.5    181.5    90.75  0.000678 ***
temp                       2   254.0    127.0    63.50  0.000932 ***
   temp: linear            1   253.5    253.5   126.75  0.000355 ***
   temp: quadratic         1     0.5      0.5     0.25  0.643330
Residuals                  4     8.0      2.0
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
```

정보통계학과 **백재욱** 교수

제7강 요인배치법1

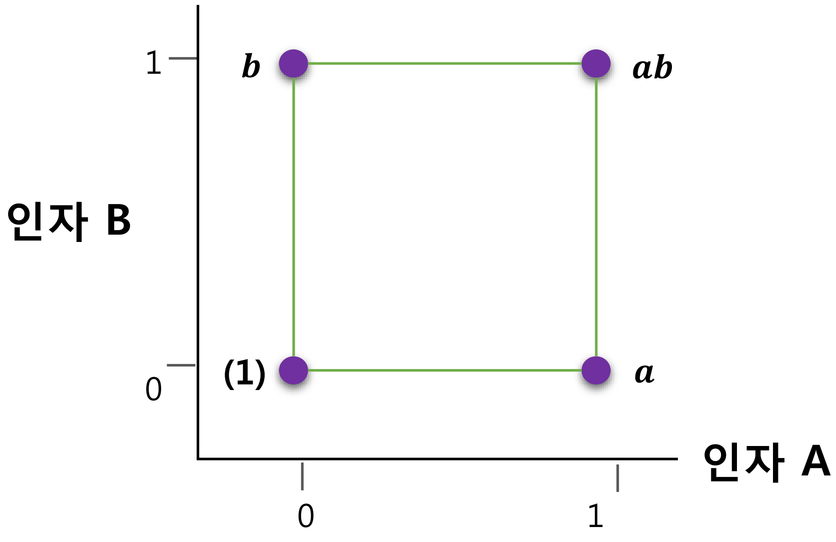
7.2 2^2 요인배치법

7.2 2^2 요인배치법

◆ 반복이 없는 경우

〈표 7-2〉 2^2 요인배치법의 자료배열

	A_0	A_1	$T_{.j}$
B_0	x_{00}	x_{10}	$T_{.0}$
B_1	x_{01}	x_{11}	$T_{.1}$
$T_{i.}$	$T_{0.}$	$T_{1.}$	T



〈그림 7-1〉 2^2요인배치법

- 주효과(main effect) 및 상호작용효과(interaction effect)

$$A = \frac{1}{2}(x_{11} + x_{10} - x_{01} - x_{00}) = \frac{1}{2}\left(ab + a - b - (1)\right) = \frac{1}{2}(T_{1.} - T_{0.})$$

$$B = \frac{1}{2}(x_{11} + x_{01} + x_{10} - x_{00}) = \frac{1}{2}\left(ab + b - a - (1)\right) = \frac{1}{2}(T_{.1} - T_{.0})$$

$$AB = \frac{1}{2}[(x_{11} - x_{10}) - (x_{01} - x_{00})] = \frac{1}{2}(ab + (1) - a - b)$$

7.2 2^2 요인배치법

- 주효과(main effect) 및 상호작용효과(interaction effect)의 파악

〈표 7-4〉 Yates 계산법 (반복이 없으면 r=1임) (2^2 요인배치 이므로 n=2임)

처리 조합	(1)	(2)	요인효과 $(2)/(2^{n-1}r)$	변동 $(2)^2/(2^n r)$
(1)	(1) + a	$(1)+a+\mathrm{b}+ab$	$(2)/(2^n r)$ = M	CT
a	$\mathrm{b}+ab$	$a-(1)+ab-\mathrm{b}$	$(2)/(2^{n-1}r)$ = A	SS_A
b	$a-(1)$	$b+ab-(1)-a$	$(2)/(2^{n-1}r)$ = B	SS_B
ab	$ab-\mathrm{b}$	$ab-b-a+(1)$	$(2)/(2^{n-1}r)$ = AB	$SS_{A\times B}$

단, M=$(2)/(2^n r)$

7.2 2^2 요인배치법

〈표 7-5〉 2^2 요인배치법의 분산분석표

요인	제곱합	자유도	평균제곱
A	SS_A	1	MS_A
B	SS_B	1	MS_B
$A \times B$ (또는 E)	$SS_{A\times B}$	1	$MS_{A\times B}$
T	SS_T	3	

7.2 2^2 요인배치법

◆ 반복이 있는 경우

<표 7-6> 2^2 요인배치법의 자료의 배열

	A_0	A_1	$T_{.j.}$
B_0	$\left.\begin{matrix} x_{001} \\ x_{002} \\ \vdots \\ x_{00r} \end{matrix}\right\} T_{00.}$	$\left.\begin{matrix} x_{101} \\ x_{102} \\ \vdots \\ x_{10r} \end{matrix}\right\} T_{10.}$	$T_{.0.}$
B_1	$\left.\begin{matrix} x_{011} \\ x_{012} \\ \vdots \\ x_{01r} \end{matrix}\right\} T_{01.}$	$\left.\begin{matrix} x_{111} \\ x_{112} \\ \vdots \\ x_{11r} \end{matrix}\right\} T_{11.}$	$T_{.1.}$
$T_{i..}$	$T_{0..}$	$T_{1..}$	T

<표 7-7> 반복이 있는 2^2 요인배치법의 분산분석표

(F_0를 보면 A, B가 고정인자의 경우임)

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_E	$4(r-1)$	MS_E	
T	SS_T	$4r-1$		

7.2 2^2 요인배치법

예제 7.2 온도(A), 습도(B) → 강도에 미치는 영향?

<표 7-8> 반복이 2회인 요인배치법의 자료

	A_0	A_1	$T_{.j.}$
B_0	$\left.\begin{matrix}4\\6\end{matrix}\right\}10$	$\left.\begin{matrix}-2\\2\end{matrix}\right\}0$	10
B_1	$\left.\begin{matrix}3\\7\end{matrix}\right\}10$	$\left.\begin{matrix}-4\\-6\end{matrix}\right\}-10$	0
$T_{i..}$	20	-10	10=T

7.2 2^2 요인배치법

풀이

< 예이츠(Yates)계산 >

처리조합	강도(합계)	(1)	(2)	요인효과	요인변동
				(2) /(2×2)	$(2)^2/(4\times 2)$
(1)	10	10	10	1.25=M	$12.5 = CT$
a	0	0	-30	-7.5=A	$112.5 = SS_A$
b	10	-10	-10	-2.5=B	$12.5 = SS_B$
ab	-10	-20	-10	-2.5=AB	$12.5 = SS_{A\times B}$

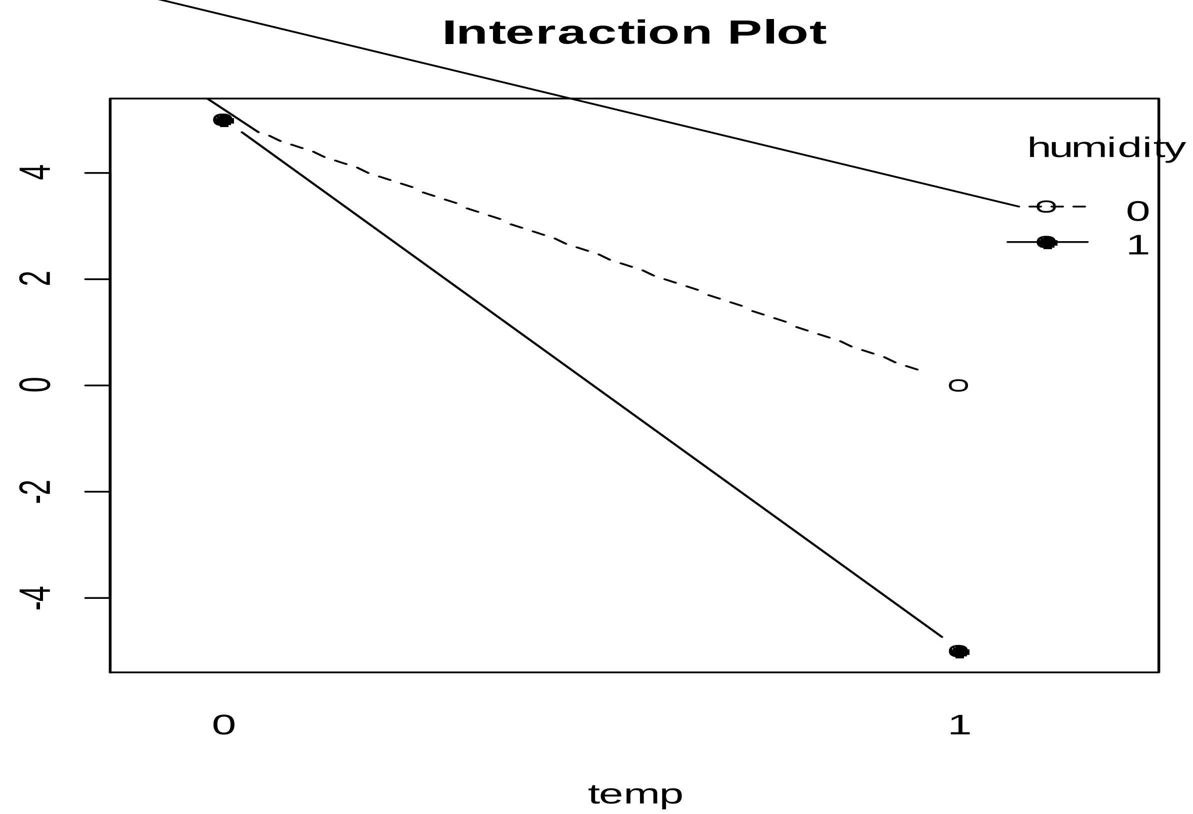
< 분산분석표 >

요 인	제곱합	자유도	평균제곱	F_0
A	112.5	1	112.5	22.5^{**}
B	12.5	1	12.5	2.5
$A\times B$	12.5	1	12.5	2.5
E	20.0	4	5	
T	157.5	7		

7.2 2^2 요인배치법

R 실습

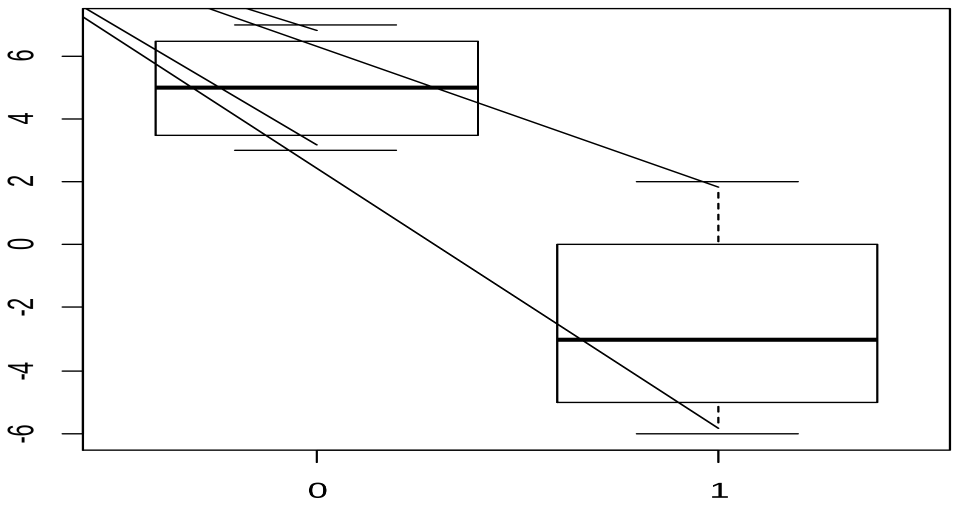
```
gang = c(4, 6, 3, 7, -2, 2, -4, -6)
temp = c(0, 0, 0, 0, 1, 1, 1, 1)
humidity = c(0, 0, 1, 1, 0, 0, 1, 1)
ex7.2data = data.frame(temp, humidity, gang)
ex7.2data$a = factor(ex7.2data$temp, levels=c(0, 1),
labels=c("A0","A1"))
ex7.2data$b = factor(ex7.2data$humidity, levels=c(0, 1),
labels=c("B0","B1"))
attach(ex7.2data)
with(ex7.2data, interaction.plot(x.factor=temp, trace.factor=humidity,
response=gang, fun=mean, type="b", legend=T, ylab="강도",
main="Interaction Plot", pch=c(1,19)))
```



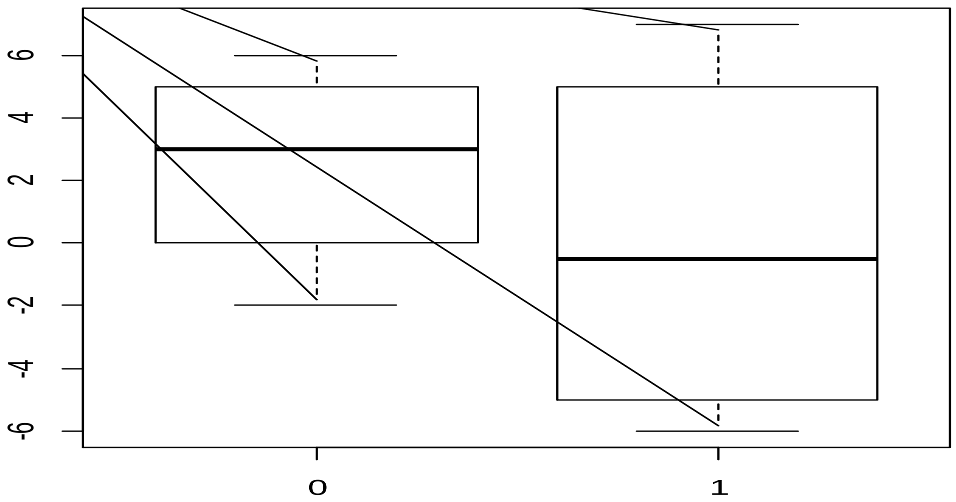
7.2 2^2 요인배치법

R 실습

boxplot(gang ~ temp)



boxplot(gang ~ humidity)



aov.out = aov(gang ~ temp * humidity, data=ex7.2data)

summary(aov.out)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
temp	1	112.5	112.5	22.5	0.00901**
humidity	1	12.5	12.5	2.5	0.18900
temp:humidity	1	12.5	12.5	2.5	0.18900
Residuals	4	20.0	5.0		

Signif. codes:

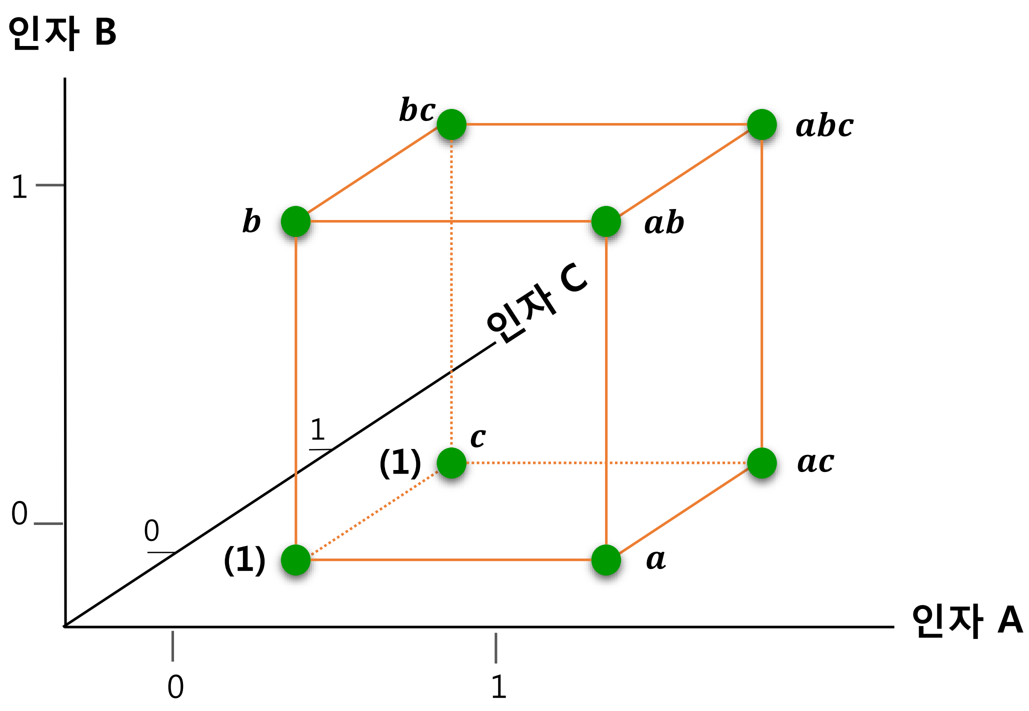
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0

정보통계학과 **백재욱** 교수

제7강 요인배치법1

7.3 2^3 요인배치법

7.3 2^3 요인배치법



[그림 7-2] 2^3요인배치법

▪ 주효과와 상호작용효과

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

$$B = \frac{1}{4r}(b + bc + ab + abc - (1) - a - c - ac) = \frac{1}{4r}(T_{.1.} - T_{.0.})$$

C = …

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

$$= \frac{1}{4r}(T_{11.} + T_{00.} - T_{01.} - T_{10.})$$

AC = …

BC = …

ABC = …

7.3 2^3 요인배치법

<표 7-10> 예이츠계산법

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

7.3 2^3 요인배치법

<표 7-11> 2^3 요인배치법의 분산분석표(r=1인 경우)

요　　인	제곱합	자유도	평균제곱
A	SS_A	1	MS_A
B	SS_B	1	MS_B
C	SS_C	1	MC_C
$A \times B$	$SS_{A \times B}$	1	$MS_{A \times B}$
$A \times C$	$SS_{A \times C}$	1	$MS_{A \times C}$
$B \times C$	$SS_{B \times C}$	1	$MS_{B \times C}$
$E(=A \times B \times C)$	SS_E	1	MS_E
T	SS_T	7	

7.3 2^3 요인배치법

예제 7.3 온도(A), 습도(B), 압력(C) 세 요인에 대한 실험을 한 결과 제품의 강도가 다음 표와 같았다. 각 요인의 효과와 변동을 구하고 분산분석표를 작성하라 (반복이 없는 r=1인 경우임).

		A_0	A_1
B_0	C_0	(1)=2	a=-5
	C_1	c=-12	ac=-17
B_1	C_0	b=15	ab=13
	C_1	bc=-2	$a$bc=-7

7.3 2^3 요인배치법

풀이 < 예이츠계산법 >

처리조합	자료	(1)	(2)	(3)	요인효과	요인변동
(1)	2	-3	25	-13	$-1.625 = M$	$21.125 = CT$
a	-5	28	-38	-19	$-4.75 = A$	$45.125 = SS_A$
b	15	-29	-9	51	$12.75 = B$	$325.125 = SS_B$
ab	13	-9	-10	5	$1.25 = AB$	$3.125 = SS_{A\times B}$
c	-12	-7	31	-63	$-15.75 = C$	$496.125 = SS_C$
ac	-17	-2	20	-1	$-0.25 = AC$	$0.125 = SS_{A\times C}$
bc	-2	-5	5	-11	$-2.75 = BC$	$15.125 = SS_{B\times C}$
abc	-7	-5	0	-5	$-1.25 = ABC$	$3.125 = SS_{A\times B\times C}$

7.3 2^3 요인배치법

풀이 < 분산분석표 >

요 인	제곱합	자유도	평균제곱	F_0
A	45.125	1	45.125	21.235^{*}
B	325.125	1	325.125	153^{**}
C	496.125	1	496.125	233.47^{**}
$B\times C$	15.125	1	15.125	7.118
E'	6.375	3	2.125	
	887.875	7		

7.3 2^3 요인배치법

R 실습

```
gang <- c(2, -5, -12, -17, 15, 13, -2, -7)
temp <- rep(c(0, 1), 4)
humidity <- c(rep(0, 4), rep(1, 4))
pressure <- rep(c(0, 0, 1, 1), 2)
data <- data.frame(gang, temp, humidity, pressure)
data$temp <- factor(data$temp, levels=c(0, 1), labels=c("a0", "a1"))
data$humidity <- factor(data$humidity, levels=c(0, 1), labels=c("b0", "b1"))
data$pressure <- factor(data$pressure, levels=c(0, 1), labels=c("c0", "c1"))
ano <- aov(gang ~ temp*humidity*pressure, data=data)
summary(ano)
```

	Df	Sum Sq	Mean Sq
temp	1	45.1	45.1
Humidity	1	325.1	325.1
Pressure	1	496.1	496.1
temp:humidity	1	3.1	3.1
temp:pressure	1	0.1	0.1
humidity:pressure	1	15.1	15.1
temp:humidity:pressure	1	3.1	3.1

7.3 2^3 요인배치법

R 실습

```
anova <- aov(gang ~ temp+ humidity+ pressure+ humidity:pressure,
data=data)
summary(anova)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
temp	1	45.1	45.1	21.235	0.019220 *
humidity	1	325.1	325.1	153.000	0.001138 **
Pressure	1	496.1	496.1	233.471	0.000609 ***
humidity:pressure	1	15.1	15.1	7.118	0.075826 .
Residuals	3	6.4	2.1		

7.3 2^3 요인배치법

<표 7-12> **반복이 r회**인 **2^3**요인배치법의 분산분석표

요 인	제곱합	자유도	평균제곱
A	SS_A	1	MS_A
B	SS_B	1	MS_B
C	SS_C	1	MS_C
$A \times B$	$SS_{A \times B}$	1	$MS_{A \times B}$
$A \times C$	$SS_{A \times C}$	1	$MS_{A \times C}$
$B \times C$	$SS_{B \times C}$	1	$MS_{B \times C}$
$A \times B \times C$	$SS_{A \times B \times C}$	1	$MS_{A \times B \times C}$
E	SS_E	$8(r-1)$	MS_E
T	SS_T	$8r-1$	

정보통계학과 **백재욱** 교수

다음 시간 안내

제8강 (7장)

요인배치법 2