9.3절

13주2강 분산분석3

9.3절

9.3 확률화 블럭계획법의 분산분석

- 확률화 블럭계획법 : 실험단위들이 동질적인 것 끼리 블럭화 되어 있으며, 각 블록 내에서 실험단위들은 각각의 처리에 랜덤하게 배치되는 실험

- 블럭효과
- 처리효과
- 오차효과 : 블럭효과와 처리효과로 설명되지 않는 효과

확률화 블럭계획법의 자료

- 블럭의 수가 b이고 처리의 수가 t 이며 각 블럭과 처리의 조합에는 모두 하나의 관측값이 배치되어 있다고 할 때, 전체관측값의 수는 bt 개 이다. 이와 같은 자료는 다음과 같이 표현될 수 있다.

표 11-8 확률화 블럭계획법의 자료

블럭 \ 처리	1	2	⋯	t	합	평균
1	X_{11}	X_{12}	⋯	X_{1t}	$X_{1\cdot}$	$\bar{X}_{1\cdot}$
2	X_{21}	X_{22}	⋯	X_{2t}	$X_{2\cdot}$	$\bar{X}_{2\cdot}$
⋮	⋮	⋮	⋱	⋮	⋮	⋮
b	X_{b1}	X_{b2}	⋯	X_{bt}	$X_{b\cdot}$	$\bar{X}_{b\cdot}$
합	$X_{\cdot 1}$	$X_{\cdot 2}$	⋯	$X_{\cdot t}$	$X_{\cdot\cdot}$	
평균	$\bar{X}_{\cdot 1}$	$\bar{X}_{\cdot 2}$	⋯	$\bar{X}_{\cdot t}$		$\bar{X}_{\cdot\cdot}$

① 제곱합 계산 ($i = 1, 2, \dots, b, \quad j = 1, 2, \dots, t$)

X_{ij} : i번째 블럭의 j번째 처리에 의한 반응

$X_{i\cdot}$: i번째 블럭의 합 $= \sum_{j=1}^{t} X_{ij}$

$\bar{X}_{i\cdot}$: i번째 블럭의 평균 $= \frac{1}{t}\sum_{j=1}^{t} X_{ij}$

$X_{\cdot j}$: j번째 처리의 합 $= \sum_{i=1}^{b} X_{ij}$

$\bar{X}_{\cdot j}$: j번째 처리의 평균 $= \frac{1}{b}\sum_{i=1}^{b} X_{ij}$

$X_{\cdot\cdot}$: 전체 관측값의 합 $= \sum_{i=1}^{b}\sum_{j=1}^{t} X_{ij}$

$\bar{X}_{\cdot\cdot}$: 전체 관측값의 평균 $= \frac{1}{bt}\sum_{i=1}^{b}\sum_{j=1}^{t} X_{ij}$

- 확률화 블럭계획법에서의 총제곱합은 다음과 같이 분할된다.

$$\sum_{i=1}^{b}\sum_{j=1}^{t}(X_{ij}-\overline{X}_{..})^2 = t\sum_{i=1}^{b}(\overline{X}_{i.}-\overline{X}_{..})^2 + b\sum_{j=1}^{t}(\overline{X}_{.j}-\overline{X}_{..})^2 + \sum_{i=1}^{b}\sum_{j=1}^{t}(X_{ij}-\overline{X}_{i.}-\overline{X}_{.j}+\overline{X}_{..})^2$$

총제곱합(TSS)=블럭제곱합(SS_B)+처리제곱합(SS_T)+오차제곱합(SS_E)

* 각 제곱합의 자유도는 다음과 같다.

TSS 의 자유도 : $bt-1$

SS_B 의 자유도 : $b-1$

SS_T 의 자유도 : $t-1$

SS_E 의 자유도 : $(b-1)(t-1)$

* 평균제곱합은 각각의 제곱합에 자유도로 나눈 것으로 다음과 같다.

$$MS_B = SSB/(b-1)$$

$$MS_T = SST/(t-1)$$

$$MS_E = SSE/(b-1)(t-1)$$

② 분산분석표의 작성

확률화 블럭계획법에서의 분산분석표는 다음과 같다.

표 11-9 확률화 블럭계획법의 분산분석표

변인	$d.f.$	SS	MS	F
블럭	$b-1$	SS_B	$SS_B/b-1$	MS_B/MS_E
처리	$t-1$	SS_T	$SS_T/t-1$	MS_T/MS_E
오차	$(b-1)(t-1)$	SS_E	$SS_E/(b-1)(t-1)$	
전체	$bt-1$	TSS		

$$CM = \frac{1}{bt}\left(\sum_{i=1}^{b}\sum_{j=1}^{t} X_{ij}\right)^2$$

$$TSS = \sum_{i=1}^{b}\sum_{j=1}^{t} {X_{ij}}^2 - CM$$

$$SS_B = \frac{1}{t}\sum_{i=1}^{b} {X_{i.}}^2 - CM$$

$$SS_T = \frac{1}{b}\sum_{j=1}^{t} {X_{.j}}^2 - CM$$

블럭효과에 대한 검정

- $\beta_i \ (i = 1, 2, \dots, b)$: i번째 블럭집단의 모평균
- H_0 : $\beta_1 = \beta_2 = \cdots = \beta_b$
- H_1 : H_0가 사실이 아니다.

$$F = \frac{MS_B}{MS_E} \sim F(b-1, (b-1)(t-1))$$

- 유의수준 α를 이용할 때 $F > F(\alpha : b-1, (b-1)(t-1))$이면 귀무가설을 기각.

처리효과에 대한 검정

- $\tau_j \ (j = 1, 2, \dots, t)$: j번째 처리집단의 모평균
- $H_0 : \tau_1 = \tau_2 = \cdots = \tau_t$
- $H_1 : H_0$가 사실이 아니다.

$$F = \frac{MS_B}{MS_E} \sim F(t-1, (b-1)(t-1))$$

- 유의수준 α를 이용할 때 $F > F(\alpha : t-1, (b-1)(t-1))$이면 귀무가설 기각.

예9-4.

다음 자료는 4명의 사람에게 각각 3가지의 서로 다른 자극(A, B, C)을 랜덤하게 순서를 정하여 준 후에 반응하는 데까지 걸린 시간(단위 : 초)를 측정한 값이다. 3가지 자극에 있어서의 평균반응시간에 차이가 있는가와 각 사람에 따라서 평균반응시간에 차이가 있는가를 검정하라.

사람 1	사람 2	사람 3	사람 4
A 1.7	C 2.1	A 0.1	B 2.2
C 2.3	A 1.5	B 2.3	A 0.6
B 3.4	B 2.6	C 0.8	C 1.6

(sol) 분석을 하기 위하여 주어진 표를 블럭(사람)과 처리(자극)에 대한 이원분류표를 만들면 다음과 같다.

표 11-10 자극실험의 이원분류표

블럭 \ 처리	A	B	C	합	평균
1	1.7	3.4	2.3	7.4	2.467
2	1.5	2.6	2.1	6.2	2.067
3	0.1	2.3	0.8	3.2	1.067
4	0.6	2.2	1.6	4.4	1.467
합	3.9	10.5	6.8	21.2	
평균	0.975	2.625	1.7		1.767

- 또한 위의 자료와 9.3.1절에 주어진 공식을 이용하여 각각의 제곱합을 구하면 다음과 같다.

$$CM = \frac{(총합)^2}{b \cdot t} = \frac{(21.2)^2}{12} = 37.453$$

$$\begin{aligned} TSS &= \sum_{i=1}^{4}\sum_{j=1}^{3} X_{ij}^2 - CM \\ &= (1.7)^2 + (2.3)^2 + \cdots + (1.6)^2 - CM \\ &= 9.4067 \end{aligned}$$

$$\begin{aligned} SS_B &= \frac{1}{3}\sum_{i=1}^{4} X_{i\cdot}^2 - CM \\ &= \frac{1}{3}\left\{(7.4)^2 + (6.2)^2 + (3.2)^2 + (4.4)^2\right\} - CM \\ &= 3.4767 \end{aligned}$$

$$\begin{aligned} SS_T &= \frac{1}{4}\sum_{j=1}^{3} X_{\cdot j}^2 - CM \\ &= \frac{1}{4}\left\{(3.9)^2 + (10.5)^2 + (6.8)^2\right\} - CM \\ &= 5.4717 \end{aligned}$$

$$\begin{aligned} SS_E &= TSS - SS_B - SS_T \\ &= 9.4069 - 3.48 - 5.4717 \\ &= 0.4552 \end{aligned}$$

❑ 계산한 결과값을 바탕으로 분산분석표를 작성하면 다음과 같다.

표 11-11 자극실험의 분산분석표

변인	$d.f.$	SS	MS	F
블럭	3	3.4767	1.16	15.26
처리	2	5.4767	2.74	36.05
오차	6	0.4533	0.076	
전체	11	9.4067		

- 각각의 처리집단에 있어서 반응시간의 모평균이 동일한가에 대한 가설 검정은 다음과 같다. 검정통계량의 값은 분산분석표에서 F값이다.

$$H_0 : \tau_A = \tau_B = \tau_C$$

H_1 : H_0이 사실이 아니다.

$$F = \frac{MS_T}{MS_E} = \frac{2.74}{0.076} = 36.05$$

- 이 검정통계량 F는 자유도가 $(2, 6)$인 F-분포를 따르며 유의수준을 5%로 할 때 부록 V의 [표 6]에 의하여 $F_{0.05,2,6} = 5.14$ 이다. 따라서 $F = 36.05 > F_{0.05,2,6}$ 이므로 귀무가설은 기각된다. 즉, 위의 자료에 의할 때 각 자극에 대한 평균반응시간이 모두 같다고 주장할 수 없다.

- 각각의 사람에 따른 반응시간의 모평균이 동일한가에 대한 검정은 각 사람에 대한 반응시간의 모평균을 $\beta_1, \beta_2, \beta_3, \beta_4$ 라고 할 때 가설과 검정통계량은 다음과 같다.

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4$$

$H_2 : H_0$이 사실이 아니다.

$$F = \frac{MS_B}{MS_E} = \frac{1.16}{0.076} = 15.26$$

- 이 검정통계량 F는 자유도가 $(3, 6)$인 F-분포를 따르며 유의수준을 5%로 할 때 부록 V의 [표 6]에 의하여 $F_{0.05,3,6} = 4.76$ 이다. 따라서 $F = 15.26 > F_{0.05,3,6}$ 이므로 귀무가설은 기각된다. 즉 위의 자료에 근거할 때 각 사람에 따라서 나타나는 반응의 평균시간이 모두 같다고 볼 수 없다.

확률화 블럭계획법에서 모수의 추정

- τ_j : j번째 처리집단의 모평균 $(j = 1, 2, \dots, t)$
- β_i : i번째 블럭집단의 모평균 $(i = 1, 2, \dots, b)$

- $100(1-\alpha)\%$ 신뢰구간

① 처리의 단일 모평균 τ_j : $\bar{X}_{\cdot j} \pm t_{\alpha/2} \frac{S}{\sqrt{b}}$

② 처리의 두 모평균의 차 $\tau_j - \tau_k$: $\left(\bar{X}_{\cdot j} - \bar{X}_{\cdot k}\right) \pm t_{\alpha/2} \frac{S}{\sqrt{2/b}}$

③ 블록의 단일 모평균 β_i : $\bar{X}_{i\cdot} \pm t_{\alpha/2} \frac{S}{\sqrt{t}}$

④ 블록의 두 모평균의 차 $\beta_i - \beta_l$: $\left(\bar{X}_{i\cdot} - \bar{X}_{l\cdot}\right) \pm t_{\alpha/2} \frac{S}{\sqrt{2/t}}$

(여기서, $S = \sqrt{MS_E}$, $t_{\alpha/2}$는 자유도 MS_E의 자유도인 $(b-1)(t-1)$을 갖는 t-분포로부터 구한 값으로 $P_r\left(t > t_{\alpha/2}\right) = \frac{\alpha}{2}$)

예9-5.

- [예 9.4]에 주어진 자료를 이용하여 자극 A과 자극 B의 모평균의 차($\tau_1 - \tau_2$)에 대한 95% 신뢰구간을 구하고, 첫 번째 사람의 반응의 평균시간(β_1)에 대한 95% 신뢰구간을 구하라.

- (sol) $\overline{X}_{\cdot 1} = 0.975$, $\overline{X}_{\cdot 2} = 2.625$, $S = \sqrt{MS_E} = \sqrt{0.076} = 0.276$이고, S의 자유도가 6이므로 자유도가 6인 t-분포에서 $t_{0.025,6} = 2.447$임을 알 수 있다.

- 따라서 $\tau_1 - \tau_2$에 대한 95% 신뢰구간은 다음과 같다.

$$(\overline{X}_{\cdot 1} - \overline{X}_{\cdot 2}) \pm t_{\alpha/2} \cdot S/\sqrt{\frac{2}{b}} = (0.975 - 2.625) \pm (2.447)(0.276)/\sqrt{\frac{2}{4}} = (-1.65 \pm 0.48) = (-2.13, -1.17)$$

- 또한 β_1에 대한 95% 신뢰구간은 다음과 같이 구할 수 있다.

$$(\overline{X}_{1\cdot} \pm t_{\alpha/2} \cdot S/\sqrt{t}) = (2.467 \pm (2.447)(0.276)/\sqrt{3}) = (2.467 \pm 0.39) = (2.077, 2.857)$$

끝~~ ♥♥