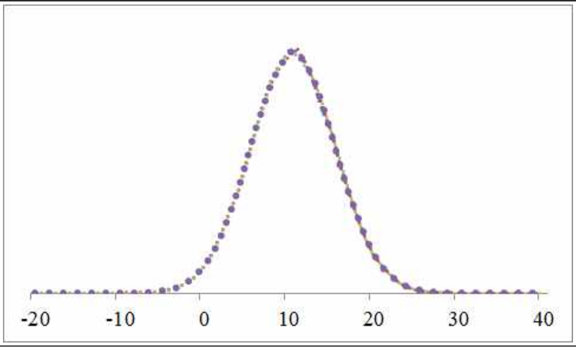
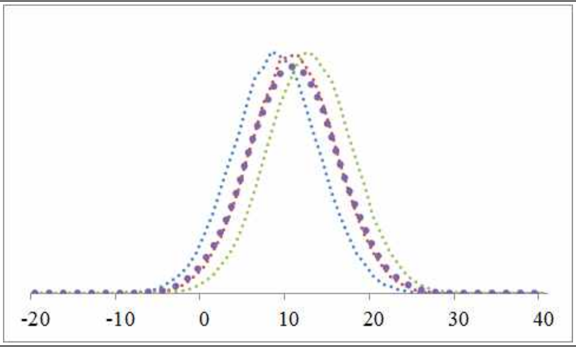
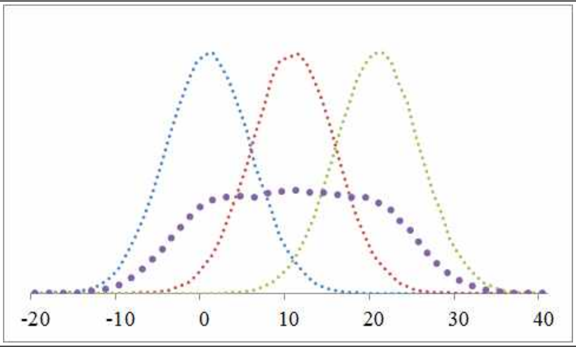
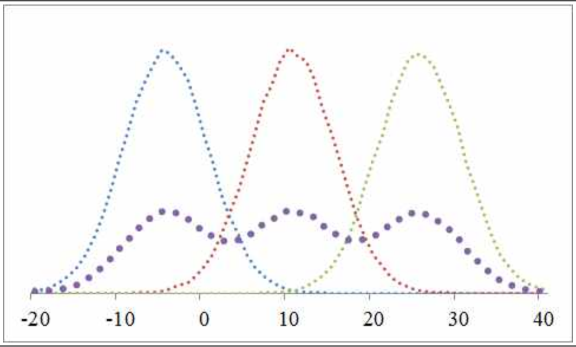
제 12 장 분산분석

분산분석이란 표본분산을 분석하여 모집단 평균들의 동질성을 검정하는 통계적 기법
→ 3개 이상의 집단 간 평균이 다른 지를 검정하는 분석방법

정규분포를 따르는 세 집단(분산=5^2)과 집단 전체의 분포도

$\mu_1 = \mu_2 = \mu_3 = 10$ $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 5^2$ $n_1 = n_2 = n_3 = 100{,}000$ **집단 전체 분산 = 25.05**	
$\mu_1 = 8,\ \mu_2 = 10,\ \mu_3 = 12$ $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 5^2$ $n_1 = n_2 = n_3 = 100{,}000$ **집단 전체 분산 = 27.71**	
$\mu_1 = 0,\ \mu_2 = 10,\ \mu_3 = 20$ $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 5^2$ $n_1 = n_2 = n_3 = 100{,}000$ **집단 전체 분산 = 91.51**	
$\mu_1 = -5,\ \mu_2 = 10,\ \mu_3 = 25$ $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 5^2$ $n_1 = n_2 = n_3 = 100{,}000$ **집단 전체 분산 = 174.92**	

1. 분산분석의 기본가정

① 각 모집단은 정규분포를 따른다.
② 각 모집단 분산은 모두 같다.
③ 표본은 각 모집단에서 무작위로 추출한다.
④ 표본은 서로 독립적이다.

2. 중요 용어

① 인자 (Factor)	변화가 가능하고 그 결과를 측정할 수 있으며 제어가 가능한 독립변수 → 집단을 구분하는 변수 - 예: 시간 , 온도, 첨가제, 작업자 - 제품특성에 영향을 미치는 원인 중 실험의 대상이 되는 원인
② 처리 (Treatment)	반응변수(종속변수)에 영향을 미치는 인자의 수준간 어떤 조합 - 예: 수율에 영향을 미치는 인자로서 온도(70, 80, 90도)와 납품업자(A, B)가 있다면, 총 가능한 처리는 $3 \times 2 = 6$. 온도 70과 납품업자 B는 처리의 한 예.
③ 반복 (Replication)	실험에서 어떤 처리조건에 대해 반복실시하는 것
④ 랜덤화 (Randomization)	같은 조건에서 미리 예견할 수 없는 실험오차를 없애기 위해 실험순서를 무작위로 결정하는 것

3. 분산분석의 기본원리

분산분석이란 인자(독립변수)를 몇 개의 수준으로 나누고 이와 같이 나누어진 그룹(Group) 간에 존재하는 평균치를 검정하는 기법

제1절 일원분산분석

1. 일원분산분석모형

분산분석모형의 가정:

① 각 처리에 대응하는 모집단은 동일한 분산을 가진다.
② 각 처리에 대응하는 모집단은 정규분포이다.
③ 각 요인 수준에 대한 관찰치들은 임의로 얻어진 것이며 서로가 독립적이다.

분산분석모형의 절차:

① 먼저 모든 처리의 평균들이 같은 가를 결정한다.

② 만약 모든 평균들이 같다면, 연구는 종료.
그렇지 않다면(평균들 중 하나라도 같지 않으면), 얼마나 다른 가를 조사하며 그리고 그 차이가 의미하는 것을 규명한다.

표 12-1

	기계 1	기계 2	기계 3
생산량	25	21	22
	20	20	20
	25	16	21
	26	15	

기호 g: 요인 수준의 수, 예) 표 12-1에서 $g = 3$

n_j: j번째 요인수준에 대한 관찰치 개수, 예) $n_1 = 4$, $n_2 = 4$, $n_3 = 3$

n: 총 관찰개수 $n = \sum_{j=1}^{g} n_j$ 예) $n = 4+4+3 = 11$

Y_{ij}: j번째 요인수준 또는 처리에 대한 i번째 관찰치;
예) Y_{23}은 기계 3에 의한 2일째 생산량

	기계 1 j=1	기계 2 j=2	기계 3 j=3
생산량	Y_{11} Y_{21} Y_{31} Y_{41}	Y_{12} Y_{22} Y_{32} Y_{42}	Y_{13} Y_{23} Y_{33}
갯수 n_j	$n_1 = 4$	$n_2 = 4$	$= 3$
합계	$\sum_{i=1}^{4} Y_{i,j=1} = \sum_{i=1}^{n_1} Y_{i,1}$	$\sum_{i=1}^{4} Y_{i,j=2} = \sum_{i=1}^{n_2} Y_{i,2}$	$\sum_{i=1}^{3} Y_{i,j=3} = \sum_{i=1}^{n_3} Y_{i,3}$
평균 $\overline{Y}_j$	$\sum_{i=1}^{n_1} Y_{i,1} \ / \ n_1$	$\sum_{i=1}^{n_2} Y_{i,2} \ / \ n_2$	$\sum_{i=1}^{n_3} Y_{i,3} \ / \ n_3$

Note: $\overline{Y}_1 = \sum_{i=1}^{n_1} Y_{i,1} \ / \ n_1$이므로, $\sum_{i=1}^{n_1} Y_{i,1} = n_1 \cdot \overline{Y}_1$

일반화하면 $\sum_{i=1}^{n_j} Y_{i,j} = n_j \cdot \overline{Y}_j$ 가 된다.

일원분산분석 모형

$Y_{ij} = \mu_j + \epsilon_{ij}$	Y_{ij} = j번째 처리에 대한 i번째 관찰치
	μ_j = j번째 처리의 모수평균
	ϵ_{ij} = 오차항; 독립적이고 $N(0,\ \sigma^2)$
	$Y_{ij} \sim N(\mu_j, \sigma^{2)}$이면, ϵ_{ij} $(= Y_{ij} - \mu_j) \sim N(0,\ \sigma^2)$

편차(standard deviation) = 관찰값 − 평균 $\rightarrow \epsilon_i = X_i - \mu$

관찰값 = 평균 + 편차 $\rightarrow X_i = \mu + \epsilon_i$

$X \sim N(\mu, \sigma^{2)}$이라 하자.

$E(X-\mu) = E(X) - \mu = 0$ Note: $E(X+b) = E(X) + b$

$V(X-\mu) = V(X) = \sigma^2$ Note: $V(X+b) = V(X)$

특정 j번째 처리의 경우,

관찰값의 합계

$$= \sum_{i=1}^{n_j} Y_{ij} = \sum_{i=1}^{n_j} (\mu_j + \epsilon_{ij}) = \sum_{i=1}^{n_j} \mu_j + \sum_{i=1}^{n_j} \epsilon_{ij} = n_j \cdot \mu_j + \sum_{i=1}^{n_j} \epsilon_{ij}$$

위의 식 $\sum_{i=1}^{n_j} Y_{ij} = n_j \cdot \mu_j + \sum_{i=1}^{n_j} \epsilon_{ij}$에서

$$\sum_{i=1}^{n_j} \epsilon_{ij} = \sum_{i=1}^{n_j} Y_{ij} - n_j \cdot \mu_j = 0 \qquad \text{since } \mu_j = \frac{\sum_{i=1}^{n_j} Y_{ij}}{n_j}$$

$$\boxed{\sum_{i=1}^{n_j} \epsilon_{ij} = 0,\ j = 1,\ 2,\ \ldots,\ g}$$

귀무가설과 대립가설

H_0: $\mu_1 = \mu_2 = \cdots = \mu_g$

H_A: 적어도 하나는 같지 않다.

$\mu_j = \mu + \alpha_j$ μ = 전체 평균,

α_j = j번째 모집단 처리 효과 Note: $\sum_{j=1}^{g} n_j \cdot \alpha_j = 0$

$Y_{ij} = \mu_j + \epsilon_{ij} \rightarrow Y_{ij} = \mu + \alpha_j + \epsilon_{ij}$

$$\sum_{j=1}^{g}\sum_{i=1}^{n_j} Y_{ij} = \sum_{j=1}^{g}\sum_{i=1}^{n_j}(\mu + \alpha_j + \epsilon_{ij})$$

$$= \sum_{j=1}^{g}\sum_{i=1}^{n_j} \mu + \sum_{j=1}^{g}\sum_{i=1}^{n_j} \alpha_j + \sum_{j=1}^{g}\sum_{i=1}^{n_j} \epsilon_{ij}$$

$$= (n_1 + n_2 + \cdots + n_g) \cdot \mu + \sum_{j=1}^{g} n_j \cdot \alpha_j + \sum_{j=1}^{g}\left(\sum_{i=1}^{n_j} \epsilon_{ij}\right)$$

$\sum_{i=1}^{n_j} \epsilon_{ij} = 0$이므로, $\sum_{j=1}^{g}\left(\sum_{i=1}^{n_j} \epsilon_{ij}\right) = 0$

위 식을 정리하면

$$\sum_{j=1}^{g} n_j \cdot \alpha_j = \sum_{j=1}^{g}\sum_{i=1}^{n_j} Y_{ij} - (n_1 + n_2 + \cdots + n_g) \cdot \mu$$

$\mu = \sum_{j=1}^{g}\sum_{i=1}^{n_j} Y_{ij} \,/\, (n_1 + n_2 + \cdots + n_g)$ 이므로, $\sum_{j=1}^{g}\sum_{i=1}^{n_j} Y_{ij} - (n_1 + n_2 + \cdots + n_g) \cdot \mu = 0$이 된다.

$$\boxed{\sum_{j=1}^{g} n_j \cdot \alpha_j = 0}$$

H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_g = 0$

H_A: 적어도 하나는 0이 아니다.

$Y_{ij} = \mu_j + \epsilon_{ij} \longrightarrow \quad Y_{ij} = \mu + \alpha_j + \epsilon_{ij}$

표본(i)	처리(j)			총계
	1	2	3	
1	$Y_{11} = 25$	$Y_{12} = 21$	$Y_{13} = 22$	
2	$Y_{21} = 20$	$Y_{22} = 20$	$Y_{23} = 20$	
3	$Y_{31} = 25$	$Y_{32} = 16$	$Y_{33} = 21$	
4	$Y_{41} = 26$	$Y_{42} = 15$		
합계	$Y_1 = 96$	$Y_2 = 72$	$Y_3 = 63$	$Y = 231$
표본수	$n_1 = 4$	$n_2 = 4$	$n_3 = 3$	$n = 11$
평균	$\overline{Y_1} = 24$	$\overline{Y_2} = 18$	$\overline{Y_3} = 21$	$\overline{Y} = 21$

처리 j의 관찰치 평균	$\overline{Y_j} = \dfrac{\sum_{i=1}^{n_j} Y_{ij}}{n_j}$
관찰치 총합	$Y = \sum_{j=1}^{g} n_j \cdot \overline{Y_j}$

Y = (처리 j=1의 관찰치의 합) + (처리 $j=2$의 관찰치의 합) + ⋯ + (처리 j=g의 관찰치의 합)

$$= \sum_{i=1}^{n_1} Y_{i,1} + \sum_{i=1}^{n_2} Y_{i,2} + \cdots + \sum_{i=1}^{n_g} Y_{i,g} = \sum_{j=1}^{g}\sum_{i=1}^{n_j} Y_{ij}$$

$$= n_1 \cdot \overline{Y_1} + n_2 \cdot \overline{Y_2} + \cdots + n_g \cdot \overline{Y_g} = \sum_{j=1}^{g} n_j \cdot \overline{Y_j}$$

관찰치 평균	$\overline{Y} = \dfrac{Y}{n}$

$$\overline{Y} = \frac{\sum_{j=1}^{g}\sum_{i=1}^{n_j} Y_{ij}}{n} = \frac{\sum_{j=1}^{g} n_j \cdot \overline{Y_j}}{n} = \frac{Y}{n}, \quad n = n_1 + \cdots + n_g$$

Y_{ij}의 기대값	$E(Y_{ij}) = \mu_j$

$E(Y_{ij}) = E(\mu_j + \epsilon_{ij}) = \mu_j$ Note: $E(\epsilon_{ij}) = 0$

처리 j의 관찰치 기댓값(평균)은 처리 j의 모집단 평균이다.

$E(Y_{ij})$의 추정치	$\overline{Y_j}$

처리 j의 관찰치 기대값은 μ_j인데, μ_j는 모수이다. μ_j를 표본에서 추정하는 값은 처리 j 관찰치들의 표본 평균인 $\overline{Y_j}$이다.

모수	추정치	모수	추정치
μ_1	$\overline{Y_1} = 24$	α_1	$\overline{Y_1} - \overline{Y} = 24 - 21 = 3$
μ_2	$\overline{Y_2} = 18$	α_2	$\overline{Y_2} - \overline{Y} = 18 - 21 = -3$
μ_3	$\overline{Y_3} = 21$	α_3	$\overline{Y_3} - \overline{Y} = 21 - 21 = 0$

잔차 $e_{ij} = Y_{ij} - \overline{Y_j}$ (잔차 = 관찰치 − 추정치)

→ $Y_{ij} = \overline{Y_j} + e_{ij}$ (관찰치 = 추정치 + 잔차)

→ $Y_{ij} = \overline{Y} + (\overline{Y_j} - \overline{Y}) + (Y_{ij} - \overline{Y_j})$ (관찰치 = 전체 평균 + 추정 처리효과 + 잔차)

관찰치 = (전체 표본평균) + (추정된 처리효과) + 잔차

관찰치	=	전체 평균	+	추정된 처리효과	+	잔차

$$\begin{vmatrix} 25 & 21 & 22 \\ 20 & 20 & 20 \\ 25 & 16 & 21 \\ 26 & 15 & \end{vmatrix} = \begin{vmatrix} 21 & 21 & 21 \\ 21 & 21 & 21 \\ 21 & 21 & 21 \\ 21 & 21 & \end{vmatrix} + \begin{vmatrix} 3 & -3 & 0 \\ 3 & -3 & 0 \\ 3 & -3 & 0 \\ 3 & -3 & \end{vmatrix} + \begin{vmatrix} 1 & 3 & 1 \\ -4 & 2 & -1 \\ 1 & -2 & 0 \\ 2 & -3 & \end{vmatrix}$$

$$Y_{ij} - \overline{Y} \quad = \quad (\overline{Y_j} - \overline{Y}) \quad + \quad (Y_{ij} - \overline{Y_j})$$

총 편차　　　처리효과　　　잔 차

총편차 = (처리효과) + 잔차

$$\begin{array}{ccccccc} Y_{ij} & - & \overline{Y} & = & (\overline{Y_j} - \overline{Y}) & + & (Y_{ij} - \overline{Y_j}) \\ \begin{vmatrix} 25 & 21 & 22 \\ 20 & 20 & 20 \\ 25 & 16 & 21 \\ 26 & 15 & \end{vmatrix} & - & \begin{vmatrix} 21 & 21 & 21 \\ 21 & 21 & 21 \\ 21 & 21 & 21 \\ 21 & 21 & \end{vmatrix} & = & \begin{vmatrix} 3 & -3 & 0 \\ 3 & -3 & 0 \\ 3 & -3 & 0 \\ 3 & -3 & \end{vmatrix} & + & \begin{vmatrix} 1 & 3 & 1 \\ -4 & 2 & -1 \\ 1 & -2 & 0 \\ 2 & -3 & \end{vmatrix} \end{array}$$

$$\begin{array}{ccccc} Y_{ij} - \overline{Y} & = & (\overline{Y_j} - \overline{Y}) & + & (Y_{ij} - \overline{Y_j}) \\ \text{총편차} & & \text{처리효과} & & \text{잔차} \\ \begin{vmatrix} 4 & 0 & 1 \\ -1 & -1 & -1 \\ 4 & -5 & 0 \\ 5 & -6 & \end{vmatrix} & = & \begin{vmatrix} 3 & -3 & 0 \\ 3 & -3 & 0 \\ 3 & -3 & 0 \\ 3 & -3 & \end{vmatrix} & + & \begin{vmatrix} 1 & 3 & 1 \\ -4 & 2 & -1 \\ 1 & -2 & 0 \\ 2 & -3 & \end{vmatrix} \end{array}$$

연습문제 1.

지역 A j=1	지역 B j=2
85	75
80	65
63	46

문제 1. $\overline{Y_A}$(지역 A의 평균), $\overline{Y_B}$(지역 B의 평균) 및 $\overline{Y}$(총평균)를 구하시오.

문제 2. 각 총편차를 구하고, 총편차의 합계를 구하시오.

문제 3. 각 처리효과를 구하고, 처리효과의 합계를 구하시오.

문제 4. 각 잔차를 구하고, 잔차의 합계를 구하시오.
문제 5. 각 관찰값을 평균, 처리효과 및 잔차로 표현하시오.

연습문제 2.

j=1	j=2
200	250
247	290
300	351

문제 1. 각 총편차, 처리효과, 잔차를 구하시오.

2. 분산분석 용어

(1) SST, SSB, SSW

총변동(sum of squares total: SST): **총편차 제곱의 합**

각 관찰치에서 전체 표본의 평균을 뺀 후에 제곱한 것을 모두 합한 것

$$\mathrm{SST} = \sum_{j=1}^{g}\sum_{i=1}^{n_j}\left(Y_{ij} - \overline{Y}\right)^2$$

그룹간 변동(variation between groups: SSB): **처리효과 제곱의 합**

'각 수준(그룹)의 평균에서 전체 평균을 뺀 후 제곱'을 해당 수준의 갯수만큼 합한 것

$$\mathrm{SSB} = \sum_{j=1}^{g}\sum_{i=1}^{n_j}\left(\overline{Y_j} - \overline{Y}\right)^2 = \sum_{j=1}^{g} n_j \cdot \left(\overline{Y_j} - \overline{Y}\right)^2$$ [1)]

그룹내 변동(variation within groups: SSW): **잔차 제곱의 합**

각 관찰치에서 해당 수준(그룹)의 평균을 뺀 후 제곱한 것을 모두 합한 것

$$\mathrm{SSW} = \sum_{j=1}^{g}\sum_{i=1}^{n_j}\left(Y_{ij} - \overline{Y_j}\right)^2$$

(2) SST = SSB + SSW

$$Y_{ij} - \overline{Y} = (Y_{ij} - \overline{Y_j}) + (\overline{Y_j} - \overline{Y}) \text{ for all } i,\ j$$

$$\sum_j \sum_i (Y_{ij} - \overline{Y})^2 = \sum_j \sum_i \left((Y_{ij} - \overline{Y_j}) + (\overline{Y_j} - \overline{Y}) \right)^2$$

오른쪽을 정리하면,

$$\sum_j \sum_i \left((Y_{ij} - \overline{Y_j}) + (\overline{Y_j} - \overline{Y}) \right)^2$$

1) $$\sum_{i=1}^{n_j}\left(\overline{Y_j} - \overline{Y}\right)^2 = \underbrace{\left(\overline{Y_j} - \overline{Y}\right)^2}_{i\,=\,1} + \underbrace{\left(\overline{Y_j} - \overline{Y}\right)^2}_{i\,=\,2} + \cdots + \underbrace{\left(\overline{Y_j} - \overline{Y}\right)^2}_{i\,=\,n_j} = n_j\left(\overline{Y_j} - \overline{Y}\right)^2$$

$$= \sum_j \sum_i \left(Y_{ij} - \overline{Y_j}\right)^2 + 2\sum_j \sum_i (Y_{ij} - \overline{Y_j}) \cdot (\overline{Y_j} - \overline{Y}) + \sum_j \sum_i \left(\overline{Y_j} - \overline{Y}\right)^2$$

여기서, $\sum_i (Y_{ij} - \overline{Y_j}) = 0 \rightarrow 2\sum_j \sum_i (Y_{ij} - \overline{Y_j}) \cdot (\overline{Y_j} - \overline{Y}) = 0$

정리하면, $\sum_j \sum_i (Y_{ij} - \overline{Y})^2 = \sum_j \sum_i \left(\overline{Y_j} - \overline{Y}\right)^2 + \sum_j \sum_i \left(Y_{ij} - \overline{Y_j}\right)^2$이 성립한다.

$\rightarrow$ SST = SSB + SSW

예제

$Y_{ij} - \overline{Y}$ 총편차			=	$(\overline{Y_j} - \overline{Y})$ 처리효과			+	$(Y_{ij} - \overline{Y_j})$ 잔차		
4	0	1		3	−3	0		1	3	1
−1	−1	−1	=	3	−3	0	+	−4	2	−1
4	−5	0		3	−3	0		1	−2	0
5	−6			3	−3			2	−3	

$\left(Y_{ij} - \overline{Y}\right)^2$ 총편차 제곱			=	$\left(\overline{Y_j} - \overline{Y}\right)^2$ 처리효과 제곱			+	$\left(Y_{ij} - \overline{Y_j}\right)^2$ 잔차 제곱		
16	0	1		9	9	0		1	9	1
1	1	1	=	9	9	0	+	16	4	1
16	25	0		9	9	0		1	4	0
25	36			9	9			4	9	
122 (SST)			=	72 (SSB)			+	50 (SSE)		

(3) 자유도와 평균제곱

변동	자유도	MS
그룹간 변동 SSB	g-1	MSB = SSB/g-1
그룹내 변동 SSW	n-g	MSW = SSW/n-g
총변동 SST	n-1	

연습문제 3. $\overline{Y_A} = 76$, $\overline{Y_B} = 62$, $\overline{Y} = 69$

지역 A j=1	지역 B j=2
85	75
80	65
63	46

문제 1. SST, SSB, SSW를 구하시오.

문제 2. 자유도와 평균제곱표를 완성하시오.

연습문제 4. $\overline{Y}_{j=1}$ = 249, $\overline{Y}_{j=2}$ = 297, $\overline{Y}$ = 273

지역 A j=1	지역 B j=2
200	250
247	290
300	351

문제 1. SST, SSB, SSW를 구하시오.

문제 2. 자유도와 평균제곱표를 완성하시오.

Note 1. $x_{j=1} \sim N(10, 10^2)$, $x_{j=2} \sim N(10, 10^2)$인 경우의 평균제곱표 ($n_{j=1}$ = $n_{j=2}$ = 10)

실험 1.

	변동(SS)	자유도	MS
그룹간(B)	0.34	1	0.34
그룹내(W)	2,127.42	18	118.19
합계(T)	2,127.76	19	

실험 2.

	변동(SS)	자유도	MS
그룹간(B)	7.38	1	7.38
그룹내(W)	927.35	18	51.52
합계(T)	934.73	19	

실험 3.

	변동(SS)	자유도	MS
그룹간(B)	63.18	1	63.18
그룹내(W)	2,081.73	18	115.65
합계(T)	2,144.91	19	

Note 2. $x_{j=1} \sim N(0, 10^2)$, $x_{j=2} \sim N(20, 10^2)$인 경우의 평균제곱표 ($n_{j=1}$ = $n_{j=2}$ = 10)

실험 1.

	변동(SS)	자유도	MS
그룹간(B)	1,949.90	1	1,949.90
그룹내(W)	2,056.49	18	114.25
합계(T)	4,006.39	19	

실험 2.

	변동(SS)	자유도	MS
그룹간(B)	953.49	1	953.49
그룹내(W)	1,812.84	18	100.71
합계(T)	2,766.33	19	

실험 3.

	변동(SS)	자유도	MS
그룹간(B)	2,594.96	1	2,594.96
그룹내(W)	1,590.74	18	88.37
합계(T)	4,185.70	19	

3. 가설검정

ANOVA(Analysis of Variance)

변동의 원천	제곱합	자유도	MS	F
그룹간	SSB $= \sum_j n_j \cdot (\overline{Y}_j - \overline{Y})^2$	g-1	MSB $= \dfrac{SSB}{g-1}$	F $= \dfrac{MSB}{MSW}$
그룹내	SSW $= \sum_j \sum_i (Y_{ij} - \overline{Y}_j)^2$	n-g	MSW $= \dfrac{SSW}{n-g}$	
합계	SST $= \sum_j \sum_i (Y_{ij} - \overline{Y})^2$	n-1		

3.1 가설검정

H_0: $\mu_1 = \mu_2 = \cdots = \mu_g$

H_A: 적어도 하나는 같지 않다.

Test Statistic: $F = \dfrac{MSB}{MSW}$, 자유도 = g-1, n-g

Rejection Region: $F > F_{\alpha,\, g-1,\, n-g}$

Note: $\mu_1 = \mu_2 = \cdots = \mu_g$이면, E(MSB) = E(MSW) = σ^2이 성립한다.

실험 1. $\mu_1 = \mu_2 = \ldots = \mu_4 = 10$, $\sigma_1^2 = \ldots = \sigma_4^2 = 10^2$; $n_1 = \ldots = n_4 = 20$;
모두 정규분포를 따른다.

1.

분산 분석 (α = 0.05)

변동요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
처리	150.330	3	50.110	**0.4466**	**0.7204**	**2.7249**
잔차	8,527.431	76	112.203			
계	8,677.760	79				

2.

분산 분석 (α = 0.05)

변동요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
처리	368.71	3	122.904	**1.1811**	**0.3226**	**2.7249**
잔차	7,908.61	76	104.061			
계	8,277.32	79				

3.

분산분석(α=0.05)

변동요인	제곱합	자유도	제곱평균	F비	P-값	F기각치
처리	93.85	3	31.284	**0.2679**	**0.8483**	**2.7249**
잔차	8,873.98	76	116.763			
계	8,967.83	79				

실험 2. $\mu_1 = \mu_2 = 30$, $\mu_3 = \mu_4 = 10$, $\sigma_1^2 = \ldots = \sigma_4^2 = 10^2$; $n_1 = \ldots = n_4 = 20$; 모두 정규분포를 따른다.

1.

분산 분석 (α = 0.05)

변동요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
처리	10,932.41	3	3,644.137	**45.7032**	**5.47E-17**	**2.7249**
잔차	6,059.85	76	79.735			
계	16,992.26	79				

2.

분산 분석 (α = 0.05)

변동요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
처리	7,514.68	3	2,504.894	**28.0548**	**2.57E-12**	**2.7249**
잔차	6,785.71	76	89.286			
계	14,300.39	79				

3.

분산분석(α=0.05)

변동요인	제곱합	자유도	제곱평균	F비	P-값	F기각치
처리	8,492.77	3	2,830.922	*35.0607*	*2.48E-14*	*2.7249*
잔차	6,136.50	76	80.743			
계	14,629.27	79				

연습문제 5.

분산분석(α=0.05)

변동요인	제곱합	자유도	제곱평균	F비	P-값	F기각치
처리	9,710.43	3	3,236.808	36.8975	8.03E-15	2.7249
잔차	6,667.04	76	87.724			
계	16,377.47	79				

문제 1. 요인의 수는 몇 개인가?

문제 2. 총 관찰갯수는 몇 개인가?

문제 3. F의 자유도는?

문제 4. $F_{0.05,3,79}$ =

문제 5. 위의 표를 기호로 표시하시오.

연습문제 6. 아래 표를 완성하시오.

문제 1.

분산분석(α=0.05)

변동요인	제곱합	자유도	제곱평균	F비	P-값	F기각치
처리	9,710.43				8.03E-15	2.7249
잔차		76				
계	16,377.47	79				

문제 2.

분산분석(α=0.01)

변동요인	제곱합	자유도	제곱평균	F비	P-값	F기각치
처리	11,478.13	4			2.91E-16	3.4795
잔차	12,464.25					
계		124				

연습문제 7.

문제 1. 실험 1의 자료 3을 사용하여 가설검정을 수행하시오. (α=0.05)

문제 2. 실험 2의 자료 3을 사용하여 가설검정을 수행하시오. (α=0.05)

3.2 μ_j 또는 (μ_j-μ_k)에 대한 (1-α)100% 신뢰구간

가정: ① 각 처리에 대응하는 모집단은 동일한 분산을 가진다.
② 각 처리에 대응하는 모집단은 정규분포이다.

②를 기호로 표현하면, $Y_{ij} \sim N(\mu_j,\ \sigma_j^2),\ j = 1,\ 2,\ ...,\ g$

①에서 동일한 분산을 가진다고 했으므로, $\sigma_1^2 = \sigma_2^2 = ... = \sigma_g^2$이다.

동일한 분산값을 σ^2이라하면, $\sigma_1^2 = \sigma_2^2 = ... = \sigma_g^2 = \sigma^2$

①과 ②를 합치면, $Y_{ij} \sim N(\mu_j,\ \sigma_j^2) \rightarrow N(\mu_j,\ \sigma^2),\ j = 1,\ 2,\ ...,\ g$

> 제10장 두 모집단의 추론에서
>
> $\sigma_1 = \sigma_2$이지만 그 값들이 알려져 있지 않을 때 분산 σ^2을 추정한 공식을 상기하자.
>
> $$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{(n_1-1)+(n_2-1)}$$

σ_1^2의 추정치 s_1^2은 $\dfrac{\sum_i^{n_1}\left(Y_{i1}-\overline{Y_1}\right)^2}{n_1-1}$, σ_2^2의 추정치 s_2^2은 $\dfrac{\sum_{i=1}^{n_2}\left(Y_{i2}-\overline{Y_2}\right)^2}{n_2-1}$,

σ_g^2의 추정치 s_g^2은 $\dfrac{\sum_i^{n_g}\left(Y_{ig}-\overline{Y_g}\right)^2}{n_g-1}$

$\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_g^2$이므로, 이들 추정치로부터 σ^2를 추정하려면 가중평균을 구하면 된다.

$$\sigma^2\text{의 추정치 } s^2 = \frac{\sigma_1^2 \text{ 추정치} \cdot (n_1-1) + \sigma_2^2 \text{ 추정치} \cdot (n_2-1) + \cdots + \sigma_g^2 \text{ 추정치} \cdot (n_g-1)}{(n_1-1) + (n_{2-1}) + \cdots + (n_g-1)}$$

$$= \frac{s_1^2 \cdot (n_1-1) + s_2^2 \cdot (n_2-1) + \cdots + s_g^2 \cdot (n_g-1)}{(n_1-1) + (n_{2-1}) + \cdots + (n_g-1)}$$

$s_j^2 \cdot (n_j\text{-}1) = \dfrac{\sum_{i=1}^{n_j}\left(Y_{ij}-\overline{Y_j}\right)^2}{n_j-1} \cdot (n_j\text{-}1) = \sum_{i=1}^{n_j}\left(Y_{ij}-\overline{Y_j}\right)^2$, $j = 1, 2, \ldots, g$이고

$(n_1\text{-}1) + (n_2\text{-}1) + \ldots + (n_g\text{-}1) = (n_1 + n_2 + \ldots + n_g - 1 - 1 \ldots. -1) = n\text{-}g$ 이므로

> σ^2의 추정치 $s^2 = \dfrac{\sum_j\sum_i\left(Y_{ij}-Y_j\right)^2}{n-g} = \dfrac{SSW}{n-g}$ = MSW이고
>
> 결국, $Y_{ij} \sim N(\mu_j, \text{MSW})$, $j = 1, 2, \ldots, g$이다.

3.2.1 μ_j에 대한 (1-α)100% 신뢰구간

> $\sigma_{\overline{X}}^2 = \dfrac{\sigma^2}{n}$이고 $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$임을 상기하자.

$\overline{Y_j}$의 분산은 $\dfrac{MSW}{n_j}$이고, 표준편차는 $\sqrt{\dfrac{MSW}{n_j}}$가 성립한다.

> μ_j에 대한 (1-α)100% 신뢰구간은 $\overline{Y_j} \pm t_{\frac{\alpha}{2}, n-g}\sqrt{\dfrac{MSW}{n_j}}$

3.2.2 [μ_j-μ_k]에 대한 [1-α]100% 신뢰구간

$\sigma^2_{\overline{X_1}-\overline{X_2}}$ = $\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$이다.

만약 σ_1^2 = σ_2^2이면 $\sigma^2_{\overline{X_1}-\overline{X_2}}$ = $\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ = $\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}$으로 표기될 수 있다.

μ_j-μ_k에 대한 추정치는 $\overline{Y_j}$-$\overline{Y_k}$이다.

($\overline{Y_j}$-$\overline{Y_k}$)의 분산은 $\frac{MSW}{n_j}+\frac{MSW}{n_k}$이고, 표준편차는 $\sqrt{\frac{MSW}{n_j}+\frac{MSW}{n_k}}$ 이다.

(μ_j-μ_k)에 대한 (1-α)100% 신뢰구간은 $\overline{Y_j}$-$\overline{Y_k}$ $\pm$ $t_{\frac{\alpha}{2},\,n-g}\sqrt{\frac{MSW}{n_j}+\frac{MSW}{n_k}}$

제2절 이원분산분석

1. 반복이 없는 이원분산분석모형

작업자	기계 1	기계 2	기계 3	합	평균
1년	25	20	21	66	22
4년	28	22	19	69	23
8년	22	18	23	63	21
합	75	60	63	198	66
평균	25	20	21	66	22

반복이 없는 경우의 실험계획모형

Y_{ij} = μ + α_i + β_j + ϵ_{ij}

여기서, Y_{ij} = 요인 1의 수준 i와 요인 2의 수준 j의 관찰치

μ = 전체 평균

α_i = 요인 1의 고정된 효과, i = 1, 2, ..., g; $\sum_{i=1}^{g}\alpha_i$ = 0

β_j = 요인 2의 고정된 효과, j = 1, 2, ..., c; $\sum_{j=1}^{c}\beta_j$ = 0

ϵ_{ij} = 요인 1의 수준 i와 요인 2의 수준 j에서의 오차항; 독립적이고 $N(0,\ \sigma^2)$

$Y_{ij} = \overline{Y} + (\overline{Y_i} - \overline{Y}) + (\overline{Y_j} - \overline{Y}) + (Y_{ij} - \overline{Y_i} - \overline{Y_j} + \overline{Y})$

여기서, $\overline{Y}$ = 전체 평균, $\overline{Y_i}$ = 요인 1의 수준 i의 평균, $\overline{Y_j}$ = 요인 2의 수준 j의 평균

위 식에서 양변에 $\overline{Y}$를 빼면

$Y_{ij} - \overline{Y} = \overline{Y} - \overline{Y} + (\overline{Y_i} - \overline{Y}) + (\overline{Y_j} - \overline{Y}) + (Y_{ij} - \overline{Y_i} - \overline{Y_j} + \overline{Y})$

모든 i, j에 대해 위의 식을 제곱한 후 더하면 (합이 0인 것은 제거하고 정리하면)

$$\sum_{i=1}^{g}\sum_{j=1}^{c}\left(Y_{ij} - \overline{Y}\right)^2 = \sum_{i=1}^{g}\sum_{j=1}^{c}\left(\overline{Y_i} - \overline{Y}\right)^2 + \sum_{i=1}^{g}\sum_{j=1}^{c}\left(\overline{Y_j} - \overline{Y}\right)^2 + \sum_{i=1}^{g}\sum_{j=1}^{c}\left(Y_{ij} - \overline{Y_i} - \overline{Y_j} + \overline{Y}\right)^2$$

$$= c \cdot \sum_{i=1}^{g}\left(\overline{Y_i} - \overline{Y}\right)^2 + g \cdot \sum_{j=1}^{c}\left(\overline{Y_j} - \overline{Y}\right)^2 + \sum_{i=1}^{g}\sum_{j=1}^{c}\left(Y_{ij} - \overline{Y_i} - \overline{Y_j} + \overline{Y}\right)^2$$

→ SST = SSA + SSB + SSW

여기서, SST = 총편차 제곱의 합
SSA = 요인 1의 처리효과 제곱의 합
SSB = 요인 2의 처리효과 제곱의 합
SSE = 잔차 제곱의 합

원천	제곱합(SS)	자유도	평균제곱(MS)	F
요인 1	SSA = $c \cdot \sum_{i=1}^{g}\left(\overline{Y_i} - \overline{Y}\right)^2$	g-1	MSA = $\frac{SSA}{g-1}$	$\frac{MSA}{MSE}$
요인 2	SSB = $g \cdot \sum_{j=1}^{c}\left(\overline{Y_j} - \overline{Y}\right)^2$	c-1	MSB = $\frac{SSB}{c-1}$	$\frac{MSB}{MSE}$
잔차	SSW = $\sum_{i=1}^{g}\sum_{j=1}^{c}\left(Y_{ij} - \overline{Y_i} - \overline{Y_j} + \overline{Y}\right)^2$	(g-1)(c-1)	MSE = $\frac{SSW}{(g-1)(c-1)}$	
합계	SST = $\sum_{i=1}^{g}\sum_{j=1}^{c}\left(Y_{ij} - \overline{Y}\right)^2$	gc-1		

요인 1 수준의 평균들이 같은 지에 관심이 있다면

H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_g = 0$

H_A: 적어도 하나는 0이 아니다.

요인 2 수준의 평균들이 같은 지에 관심이 있다면

H_0: $\beta_1 = \beta_2 = \cdots = \beta_c = 0$

H_A: 적어도 하나는 0이 아니다.

작업자(B)	(A) 기계 1 i=1	기계 2 i=2	기계 3 i=3	평균
1년 j=1	25 $Y_{i=1,j=1}$	20 $Y_{i=2,j=1}$	21 $Y_{i=3,j=1}$	22 $\overline{Y}_{j=1}$
4년 j=2	28	22	19	23 $\overline{Y}_{j=2}$
8년 j=3	22	18	23	21 $\overline{Y}_{j=3}$
평균	25 $\overline{Y}_{i=1}$	20 $\overline{Y}_{i=2}$	21 $\overline{Y}_{i=3}$	22 $\overline{Y}$

SST $\sum_{i=1}^{g}\sum_{j=1}^{c}\left(Y_{ij}-\overline{Y}\right)^2$ 계산: 합계 76 → 총편차 제곱의 합

	i=1	i=2	i=3
j=1	9	4	1
j=2	36	0	9
j=3	0	16	1

i=1	i=2	i=3
$(25-22)^2$	$(20-22)^2$	$(21-22)^2$
$(28-22)^2$	$(22-22)^2$	$(19-22)^2$
$(22-22)^2$	$(18-22)^2$	$(23-22)^2$

SSA $\sum_{i=1}^{g}\sum_{j=1}^{c}\left(\overline{Y}_i-\overline{Y}\right)^2 = c \cdot \sum_{i=1}^{g}\left(\overline{Y}_i-\overline{Y}\right)^2$ 계산: 합계 42 → 요인 1의 처리효과 제곱의 합

	i=1	i=2	i=3
j=1	9	4	1
j=2	9	4	1
j=3	9	4	1

i=1	i=2	i=3
$(25-22)^2$	$(20-22)^2$	$(21-22)^2$
$(25-22)^2$	$(20-22)^2$	$(21-22)^2$
$(25-22)^2$	$(20-22)^2$	$(21-22)^2$

SSB = $\sum_{i=1}^{g}\sum_{j=1}^{c}\left(\overline{Y}_j-\overline{Y}\right)^2 = g \cdot \sum_{j=1}^{c}\left(\overline{Y}_j-\overline{Y}\right)^2$ 계산: 합계 6 → 요인 2의 처리효과 제곱의 합

	i=1	i=2	i=3
j=1	0	0	0
j=2	1	1	1
j=3	1	1	1

i=1	i=2	i=3
$(22-22)^2$	$(22-22)^2$	$(22-22)^2$
$(23-22)^2$	$(23-22)^2$	$(23-22)^2$
$(21-22)^2$	$(21-22)^2$	$(21-22)^2$

SSW = $\sum_{i=1}^{g}\sum_{j=1}^{c}\left(Y_{ij}-\overline{Y}_i-\overline{Y}_j+\overline{Y}\right)^2$ 계산: 합계 28 → 잔차 제곱의 합

	i=1	i=2	i=3
j=1	0	0	0
j=2	4	1	9
j=3	4	1	9

	i=1	i=2	i=3
j=1	$(25-25-22-22)^2$	$(20-20-22-22)^2$	$(21-21-22-22)^2$
j=2	$(28-25-23-22)^2$	$(22-20-23-22)^2$	$(19-21-23-22)^2$
j=3	$(22-25-21-22)^2$	$(18-20-21-22)^2$	$(23-21-21-22)^2$

원천	제곱합(SS)	자유도	평균제곱(MS)	F
요인 1	SSA $= c \cdot \sum_{i=1}^{g}\left(\overline{Y_i}-\overline{Y}\right)^2$	g-1	MSA $= \dfrac{SSA}{g-1}$	$\dfrac{MSA}{MSE}$
요인 2	SSB $= g \cdot \sum_{j=1}^{c}\left(\overline{Y_j}-\overline{Y}\right)^2$	c-1	MSB $= \dfrac{SSB}{c-1}$	$\dfrac{MSB}{MSE}$
잔차	SSW $= \sum_{i=1}^{g}\sum_{j=1}^{c}\left(Y_{ij}-\overline{Y_i}-\overline{Y_j}+\overline{Y}\right)^2$	(g-1)(c-1)	MSE $= \dfrac{SSW}{(g-1)(c-1)}$	
합계	SST $= \sum_{i=1}^{g}\sum_{j=1}^{c}\left(Y_{ij}-\overline{Y}\right)^2$	gc-1		

원천	제곱합(SS)	자유도	평균제곱(MS)	F	F(0.05)	F(p-value)
요인 1	42	2	21	3.0	6.94	0.1600
요인 2	6	2	3	0.43	6.94	0.6782
잔차	28	4	7			
합계	76	8				

실험: 모든 자료가 $N(10,10^2)$을 따르는 경우 **(자료는 5행, 3열)**

실험 1.

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	114.69	4	28.67	0.2021	0.9302	3.8379
인자 B(열)	342.50	2	171.25	1.2071	0.3482	4.4590
잔차	1,134.99	8	141.87			
계	1,592.18	14				

실험 2.

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	740.24	4	185.06	1.3818	0.3223	3.8379
인자 B(열)	437.55	2	218.77	1.6335	0.2542	4.4590
잔차	1,071.45	8	133.93			
계	2,249.24	14				

실험: 모든 자료가 정규분포, 분산은 10^2; 모평균이 아래와 같은 경우(자료는 5행, 3열)

10	**50**	10
110	**150**	**110**
10	**50**	10
10	**50**	10
10	**50**	10

1.

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	27,168.20	4	6,792.05	41.2658	2.20E-05	3.8379
인자 B(열)	6,001.58	2	3,000.79	18.2316	1.05E-03	4.4590
잔차	1,316.74	8	164.59			
계	34,486.52	14				

2.

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	24,446.49	4	6,111.62	179.7000	7.27E-08	3.8379
인자 B(열)	5,067.17	2	2,533.58	74.4949	6.74E-06	4.4590
잔차	272.08	8	34.01			
계	29,785.73	14				

연습문제 8. 반복이 없는 이원분산분석

	A(i)	
B(j)	95	88
	129	87

문제 1. $\overline{Y}$, $\overline{Y}_{i=1}$, $\overline{Y}_{i=2}$, $\overline{Y}_{j=1}$, $\overline{Y}_{j=2}$을 구하시오.

문제 2. 95를 $\overline{Y} = 99.75$, $\overline{Y}_{i=1} = 112.0$, $\overline{Y}_{i=2} = 87.5$, $\overline{Y}_{j=1} = 91.5$, $\overline{Y}_{j=2} = 108$로 설명하시오.

문제 3. 87을 $\overline{Y} = 99.75$, $\overline{Y}_{i=1} = 112.0$, $\overline{Y}_{i=2} = 87.5$, $\overline{Y}_{j=1} = 91.5$, $\overline{Y}_{j=2} = 108$로 설명하시오.

문제 4. SSA, SSB, SSW, SST를 구하시오.

문제 5. 분산분석을 수행하시오. 단, $\alpha = 0.05$, $F_{0.05,1,1} = 161.4476$이다.

연습문제 9. 반복이 없는 이원분산분석

	A(i)	
B(j)	5	4
	7	10

문제 1. $\overline{Y}$, $\overline{Y}_{i=1}$, $\overline{Y}_{i=2}$, $\overline{Y}_{j=1}$, $\overline{Y}_{j=2}$을 구하시오.

문제 2. SSA, SSB, SSW, SST를 구하시오.

문제 3. 분산분석을 수행하시오. 단, α = 0.05, $F_{0.05,1,1}$ = 161.4476이다.

문제 4. 가설검정을 수행하시오. 단, α = 0.05, $F_{0.05,1,1}$ = 161.4476이다.

연습문제 10. 아래 표를 완성하시오.

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비
인자 A(행)	17,606.32	4		
인자 B(열)		2	2,566.74	
잔차				
계	23,767.16	14		

2. 반복이 있는 이원분산분석 모형

온도	압력 200	압력 250	압력 300
저온	98	108	104
	89	99	111
	86	114	100
고온	99	115	106
	102	109	99
	102	121	92

반복이 있는 경우의 실험계획모형

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}$$

여기서, Y_{ijk} = 요인 1의 수준 i와 요인 2의 수준 j의 k번째 관찰치

μ = 전체 평균

α_i = 요인 1의 고정된 효과, $i = 1, 2, ..., g$; $\sum_{i=1}^{g}\alpha_i = 0$

β_j = 요인 2의 고정된 효과, $j = 1, 2, ..., c$; $\sum_{j=1}^{c}\beta_j = 0$

$\alpha\beta_{ij}$ = 요인 1과 요인 2의 상호작용효과, $\sum_{i=1}^{g}\sum_{j=1}^{c}\alpha\beta_{ij} = 0$

ϵ_{ijk} = 오차항; $i = 1, 2, ..., g$; $j = 1, 2, ..., c$; $k = 1, 2, ..., h$

서로 독립적이고 $N(0, \sigma^2)$

$E(Y_{ijk})$	$= \mu$	$+ \alpha_i$	$+ \beta_j$	$+ \alpha\beta_{ij}$
(평균반응)	(전체반응)	(요인 1의 주효과)	(요인 2의 주효과)	(요인 1과 2의 상호효과)

$$Y_{ijk} = \overline{Y} + (\overline{Y_i} - \overline{Y}) + (\overline{Y_j} - \overline{Y}) + (\overline{Y_{ij}} - \overline{Y_i} - \overline{Y_j} + \overline{Y}) + (Y_{ijk} - \overline{Y_{ij}})$$

여기서, $\overline{Y}$ = 전체 평균

$\overline{Y_i}$ = 요인 1의 수준 i의 평균

$\overline{Y_j}$ = 요인 2의 수준 j의 평균

$\overline{Y_{ij}}$ = 요인 1의 수준 i와 요인 2의 수준 j의 평균

$$Y_{ijk} - \overline{Y} = \overline{Y} - \overline{Y} + (\overline{Y}_i - \overline{Y}) + (\overline{Y}_j - \overline{Y}) + (\overline{Y}_{ij} - \overline{Y}_i - \overline{Y}_j + \overline{Y}) + (Y_{ijk} - \overline{Y}_{ij})$$

$$\rightarrow\ Y_{ijk} - \overline{Y} = (\overline{Y}_i - \overline{Y}) + (\overline{Y}_j - \overline{Y}) + (\overline{Y}_{ij} - \overline{Y}_i - \overline{Y}_j + \overline{Y}) + (Y_{ijk} - \overline{Y}_{ij})$$

모든 i,j,k에 대해 위의 식을 제곱한 후 더하면

$$\sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(Y_{ijk} - \overline{Y}\right)^2$$

$$= \sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(\overline{Y}_i - \overline{Y}\right)^2 + \sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(\overline{Y}_j - \overline{Y}\right)^2$$

$$+ \sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(\overline{Y}_{ij} - \overline{Y}_i - \overline{Y}_j + \overline{Y}\right)^2 + \sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(Y_{ijk} - \overline{Y}_{ij}\right)^2$$

$$= c \cdot h \cdot \sum_{i=1}^{g}\left(\overline{Y}_i - \overline{Y}\right)^2 + g \cdot h \cdot \sum_{j=1}^{c}\left(\overline{Y}_j - \overline{Y}\right)^2$$

$$+ h \cdot \sum_{i=1}^{g}\sum_{j=1}^{c}\left(\overline{Y}_{ij} - \overline{Y}_i - \overline{Y}_j + \overline{Y}\right)^2 + \sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(Y_{ijk} - \overline{Y}_{ij}\right)^2$$

→ SST = SSA + SSB + SSAB + SSW

원천	제곱합(SS)	자유도	평균제곱(MS)	F
요인 1	SSA $= c \cdot h \cdot \sum_{i=1}^{g}\left(\overline{Y}_i - \overline{Y}\right)^2$	g-1	MSA $= \dfrac{SSA}{g-1}$	$\dfrac{MSA}{MSE}$
요인 2	SSB $= g \cdot h \cdot \sum_{j=1}^{c}\left(\overline{Y}_j - \overline{Y}\right)^2$	c-1	MSB $= \dfrac{SSB}{c-1}$	$\dfrac{MSB}{MSE}$
상호작용	SSAB $= h \cdot \sum_{i=1}^{g}\sum_{j=1}^{c}\left(\overline{Y}_{ij} - \overline{Y}_i - \overline{Y}_j + \overline{Y}\right)^2$	(g-1)(c-1)	MSAB $= \dfrac{SSAB}{(g-1)(c-1)}$	$\dfrac{MSAB}{MSE}$
잔차	SSW $= \sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(Y_{ijk} - \overline{Y}_{ij}\right)^2$	$gc(h-1)$	MSE $= \dfrac{SSW}{gc(h-1)}$	
합계	SST $= \sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(Y_{ijk} - \overline{Y}\right)^2$	gch-1		

두 요인 사이에 상호작용이 있는지에 관심이 있다면

H_0: 모든 $\alpha\beta_{ij} = 0 \quad (i=1,2,...,\ g;\ j=1,2,...,\ c)$

H_A: 적어도 하나는 0이 아니다.

요인 1 수준의 평균들이 같은 지에 관심이 있다면

H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_g = 0$

H_A: 적어도 하나는 0이 아니다.

요인 2 수준의 평균들이 같은 지에 관심이 있다면

H_0: $\beta_1 = \beta_2 = \cdots = \beta_c = 0$

H_A: 적어도 하나는 0이 아니다.

평균 계산 1. $\overline{Y_i}$, $\overline{Y_j}$, $\overline{Y}$

온도		압력 200 j=1	압력 250 j=2	압력 300 j=3	평균
저온 i=1	k=1	98	108	104	101
	k=2	89	99	111	
	k=3	86	114	100	$\overline{Y}_{i=1}$
고온 i=2	k=1	99	115	106	105
	k=2	102	109	99	
	k=3	102	121	92	$\overline{Y}_{i=2}$
	평균	96 $\overline{Y}_{j=1}$	111 $\overline{Y}_{j=2}$	102 $\overline{Y}_{j=3}$	103 $\overline{Y}$

평균 계산 2. $\overline{Y}_{ij}$

온도		압력 200 j=1	압력 250 j=2	압력 300 j=3
저온 i=1	k=1	98	108	104
	k=2	89	99	111
	k=3	86	114	100
	평균	91 $\overline{Y}_{11}$	107 $\overline{Y}_{12}$	105 $\overline{Y}_{13}$
고온 i=2	k=1	99	115	106
	k=2	102	109	99
	k=3	102	121	92
	평균	101 $\overline{Y}_{21}$	115 $\overline{Y}_{22}$	99 $\overline{Y}_{23}$

SST $\sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(Y_{ijk}-\overline{Y}\right)^2$ 계산: 1,414 (Note: $\overline{Y}$=103)

		j=1	j=2	j=3
i=1	k=1	25	25	1
	k=2	196	16	64
	k=3	289	121	9
i=2	k=1	16	144	9
	k=2	1	36	16
	k=3	1	324	121

j=1	j=2	j=3
$(98-103)^2$	$(108-103)^2$	$(104-103)^2$
$(89-103)^2$	$(99-103)^2$	$(111-103)^2$
$(86-103)^2$	$(114-103)^2$	$(100-103)^2$
$(99-103)^2$	$(115-103)^2$	$(106-103)^2$
$(102-103)^2$	$(109-103)^2$	$(99-103)^2$
$(102-103)^2$	$(121-103)^2$	$(92-103)^2$

SSA(온도) $\sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(\overline{Y}_i-\overline{Y}\right)^2= c\cdot h\cdot\sum_{i=1}^{g}\left(\overline{Y}_i-\overline{Y}\right)^2$ 계산: 72

Note: $\overline{Y}_{i=1}$ = 101, $\overline{Y}_{i=2}$ = 105

		j=1	j=2	j=3
i=1	k=1	4	4	4
	k=2	4	4	4
	k=3	4	4	4
i=2	k=1	4	4	4
	k=2	4	4	4
	k=3	4	4	4

j=1	j=2	j=3
$(101-103)^2$	$(101-103)^2$	$(101-103)^2$
$(101-103)^2$	$(101-103)^2$	$(101-103)^2$
$(101-103)^2$	$(101-103)^2$	$(101-103)^2$
$(105-103)^2$	$(105-103)^2$	$(105-103)^2$
$(105-103)^2$	$(105-103)^2$	$(105-103)^2$
$(105-103)^2$	$(105-103)^2$	$(105-103)^2$

SSB(압력) = $\sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(\overline{Y_j}-\overline{Y}\right)^2 = g \cdot h \cdot \sum_{j=1}^{c}\left(\overline{Y_j}-\overline{Y}\right)^2$ 계산: 684

Note: $\overline{Y}_{j=1}$ = 96, $\overline{Y}_{j=2}$ = 111, $\overline{Y}_{j=3}$ = 102

		j=1	j=2	j=3
i=1	k=1	49	64	1
	k=2	49	64	1
	k=3	49	64	1
i=2	k=1	49	64	1
	k=2	49	64	1
	k=3	49	64	1

j=1	j=2	j=3
$(96-103)^2$	$(111-103)^2$	$(102-103)^2$
$(96-103)^2$	$(111-103)^2$	$(102-103)^2$
$(96-103)^2$	$(111-103)^2$	$(102-103)^2$
$(96-103)^2$	$(111-103)^2$	$(102-103)^2$
$(96-103)^2$	$(111-103)^2$	$(102-103)^2$
$(96-103)^2$	$(111-103)^2$	$(102-103)^2$

SSAB(상호작용) = $\sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(\overline{Y_{ij}}-\overline{Y_i}-\overline{Y_j}+\overline{Y}\right)^2 = h \cdot \sum_{i=1}^{g}\sum_{j=1}^{c}\left(\overline{Y_{ij}}-\overline{Y_i}-\overline{Y_j}+\overline{Y}\right)^2$ 계산: 228

Note: $\overline{Y}_{11}$ = 91, $\overline{Y}_{12}$ = 107, $\overline{Y}_{13}$ = 105

$\overline{Y}_{21}$ = 101, $\overline{Y}_{22}$ = 115, $\overline{Y}_{23}$ = 99

		j=1	j=2	j=3
i=1	k=1	9	4	25
	k=2	9	4	25
	k=3	9	4	25
i=2	k=1	9	4	25
	k=2	9	4	25
	k=3	9	4	25

	j=1	j=2	j=3
i=1	$(91-101-96+103)^2$	$(107-101-111+103)^2$	$(105-101-102+103)^2$
	$(91-101-96+103)^2$	$(107-101-111+103)^2$	$(105-101-102+103)^2$
	$(91-101-96+103)^2$	$(107-101-111+103)^2$	$(105-101-102+103)^2$
i=2	$(101-105-96+103)^2$	$(115-105-111+103)^2$	$(99-105-102+103)^2$
	$(101-105-96+103)^2$	$(115-105-111+103)^2$	$(99-105-102+103)^2$
	$(101-105-96+103)^2$	$(115-105-111+103)^2$	$(99-105-102+103)^2$

SSW(잔차) $\sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h}\left(Y_{ijk}-\overline{Y}_{ij}\right)^2$ 계산: 430

Note: $\overline{Y}_{11}$ = 91, $\overline{Y}_{12}$ = 107, $\overline{Y}_{13}$ = 105

$\overline{Y}_{21}$ = 101, $\overline{Y}_{22}$ = 115, $\overline{Y}_{23}$ = 99

		j=1	j=2	j=3
i=1	k=1	49	1	1
	k=2	4	64	36
	k=3	25	49	25
i=2	k=1	4	0	49
	k=2	1	36	0
	k=3	1	36	49

j=1	j=2	j=3
$(98-91)^2$	$(108-107)^2$	$(101-105)^2$
$(89-91)^2$	$(99-107)^2$	$(101-105)^2$
$(86-91)^2$	$(114-107)^2$	$(101-105)^2$
$(99-101)^2$	$(115-115)^2$	$(106-99)^2$
$(102-101)^2$	$(109-115)^2$	$(99-99)^2$
$(102-101)^2$	$(121-115)^2$	$(92-99)^2$

원천	제곱합(SS)	자유도	평균제곱(MS)	F
요인 1	SSA $= c \cdot h \cdot \sum_{i=1}^{g} \left(\overline{Y_i} - \overline{Y}\right)^2$	g-1	MSA $= \dfrac{SSA}{g-1}$	$\dfrac{MSA}{MSE}$
요인 2	SSB $= g \cdot h \cdot \sum_{j=1}^{c} \left(\overline{Y_j} - \overline{Y}\right)^2$	c-1	MSB $= \dfrac{SSB}{c-1}$	$\dfrac{MSB}{MSE}$
상호 작용	SSAB $= h \cdot \sum_{i=1}^{g}\sum_{j=1}^{c} \left(\overline{Y_{ij}} - \overline{Y_i} - \overline{Y_j} + \overline{Y}\right)^2$	(g-1)(c-1)	MSAB $= \dfrac{SSAB}{(g-1)(c-1)}$	$\dfrac{MSAB}{MSE}$
잔차	SSW $= \sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h} \left(Y_{ijk} - \overline{Y_{ij}}\right)^2$	$gc(h-1)$	MSE $= \dfrac{SSW}{gc(h-1)}$	
합계	SST $= \sum_{i=1}^{g}\sum_{j=1}^{c}\sum_{k=1}^{h} \left(Y_{ijk} - \overline{Y}\right)^2$	gch-1		

원천	제곱합(SS)	자유도	평균제곱(MS)	F	F(0.05)	F(p-value)
압력	648	2	342	9.55	3.89	0.0033
온도	72	1	72	2.01	4.75	0.1818
상호작용	228	2	114	3.18	3.89	0.0779
잔차	430	12	35.83			
합계	1,414	17				

모의실험 1. 수준처리 효과는 존재하지만 상호작용은 없는 경우

요인수준 조합별로 20개씩의 자료

각 요인 조합별 모집단 평균은 아래 표와 같음(분산은 모두 3^2)

α_1 = -10, α_2 = 0, α_3 = 10, β_1 = -20, β_2 = 0, β_3 = 20, $\alpha\beta_{ij}$ = 0

$\mu = 50$

$\mu_{11} = \mu + \alpha_1 + \beta_1 + \alpha_{11} = 50 - 10 - 20 + 0 = 20$

$\mu_{12} = \mu + \alpha_1 + \beta_2 + \alpha_{12} = 50 - 10 + 0 + 0 = 40$

$\mu_{13} = \mu + \alpha_1 + \beta_3 + \alpha_{13} = 50 - 10 + 20 + 0 = 60$

$\mu_{21} = \mu + \alpha_2 + \beta_1 + \alpha_{21} = 50 + 0 - 20 + 0 = 30$

$\mu_{22} = \mu + \alpha_2 + \beta_2 + \alpha_{22} = 50 + 0 + 0 + 0 = 50$

$\mu_{23} = \mu + \alpha_2 + \beta_3 + \alpha_{23} = 50 + 0 + 20 + 0 = 70$

$\mu_{31} = \mu + \alpha_3 + \beta_1 + \alpha_{31} = 50 + 10 - 20 + 0 = 40$

$\mu_{32} = \mu + \alpha_3 + \beta_1 + \alpha_{32} = 50 + 10 - 0 + 0 = 60$

$\mu_{12} = \mu + \alpha_3 + \beta_2 + \alpha_{33} = 50 + 10 + 20 + 0 = 80$

수준 조합별 모집단 평균 μ_{ij}

	$j = 1$	$j = 2$	$j = 3$
$i = 1$	20	40	60
$i = 2$	30	50	70
$i = 3$	40	60	80

분산 분석 (α=0.05)

실험 1.

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	11,818.40	2	5,909.20	569.51	2.48E-76	3.0488
인자 B(열)	48,921.32	2	24,460.66	2,357.44	3.29E-125	3.0488
상호작용	25.57	4	6.39	0.62	0.6517	2.4245
잔차	1,774.29	171	10.38			
계	62,539.57	179				

실험 2.

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	13,066.64	2	6,533.32	701.30	3.86E-83	3.0488
인자 B(열)	48,518.42	2	24,259.21	2,604.03	8.83E-129	3.0488
상호작용	18.71	4	4.68	0.50	0.7343	2.4245
잔차	1,593.04	171	9.32			
계	63,196.81	179				

모의실험 2. 수준처리 효과와 상호작용이 모두 있는 경우

요인수준 조합별로 20개씩의 자료

각 요인 조합별 모집단 평균은 아래 표와 같음(분산은 모두 3^2)

α_1 = -10, α_2 = 0, α_3 = 10, β_1 = -20, β_2 = 0, β_3 = 20,

$\alpha\beta_{11} = \alpha\beta_{22} = \alpha\beta_{33} = 20$, 기타 $\alpha\beta_{ij}$는 -10

$\mu = 50$

$\mu_{11} = \mu + \alpha_1 + \beta_1 + \alpha_{11} = 50 - 10 - 20 + 20 = 40$

$\mu_{12} = \mu + \alpha_1 + \beta_2 + \alpha_{12} = 50 - 10 + 0 - 10 = 30$

$\mu_{13} = \mu + \alpha_1 + \beta_3 + \alpha_{13} = 50 - 10 + 20 - 10 = 50$

$\mu_{21} = \mu + \alpha_2 + \beta_1 + \alpha_{21} = 50 + 0 - 20 - 10 = 20$

$\mu_{22} = \mu + \alpha_2 + \beta_2 + \alpha_{22} = 50 + 0 + 0 + 20 = 70$

$\mu_{23} = \mu + \alpha_2 + \beta_3 + \alpha_{23} = 50 + 0 + 20 - 10 = 60$

$\mu_{31} = \mu + \alpha_3 + \beta_1 + \alpha_{31} = 50 + 10 - 20 - 10 = 30$

$\mu_{32} = \mu + \alpha_3 + \beta_1 + \alpha_{32} = 50 + 10 - 0 + 20 = 50$

$\mu_{12} = \mu + \alpha_3 + \beta_2 + \alpha_{33} = 50 + 10 + 20 - 10 = 100$

수준 조합별 모집단 평균 μ_{ij}

	$j = 1$	$j = 2$	$j = 3$
$i = 1$	40	30	50
$i = 2$	20	70	60
$i = 3$	30	50	100

실험 1.

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	11,897.66	2	5,948.83	552.30	2.41E-75	3.0488
인자 B(열)	47,495.24	2	23,747.62	2,204.76	8.18E-123	3.0488
상호작용	34,738.65	4	8,684.66	806.30	8.64E-110	2.4245
잔차	1,841.85	171	10.77			
계	95,973.40	179				

실험 2.

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	12,191.13	2	6,095.57	718.07	6.36E-84	3.0488
인자 B(열)	45,532.96	2	22,766.48	2,681.94	7.68E-130	3.0488
상호작용	33,399.08	4	8,349.77	983.62	7.89E-117	2.4245
잔차	1,451.59	171	8.49			
계	92,574.75	179				

연습문제 11.

	$j=1$	$j=2$
$i=1$	98	104
	86	100
$i=2$	99	108
	101	96

문제 1. $\overline{Y}$, $\overline{Y}_{i=1}$, $\overline{Y}_{i=2}$, $\overline{Y}_{j=1}$, $\overline{Y}_{j=2}$, $\overline{Y}_{11}$, $\overline{Y}_{12}$, $\overline{Y}_{21}$, $\overline{Y}_{22}$를 구하시오.

문제 2. 빈 칸을 채우시오.

	SST		SS행		SS열		SS행열		잔차	
	1	25	4	4	9	9				
		1	4	4	9	9				
	0	81			9	9	4	4		
	4	9			9	9	4	4		
합계										

문제 3. 분산분석을 실시하시오.

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비	F 기각치
인자 A(행)					7.709
인자 B(열)					7.709
상호작용					7.709
잔차					
계					

연습문제 정답

1. (1)

	지역 A	지역 B	합계
	85	75	
	80	65	
	63	46	
합계	228	186	414
평균	76	62	69

$\overline{Y_A} = 76, \ \overline{Y_B} = 62, \ \overline{Y} = 69$

2.

(2)

	지역 A	지역 B	합계
	16	**6**	
	11	**-4**	
	-6	**-23**	
합계	21	-21	**0**
평균	7	-7	0

지역 A	지역 B
85-69	75-69
80-69	65-69
63-69	46-69

Note: 총편차의 합 $\sum_{j=1}^{g}\sum_{i=1}^{n_j}(Y_{ij}-\overline{Y})$은 항상 0이다.

(3)

	지역 A	지역 B	합계
	7	**-7**	
	7	**-7**	
	7	**-7**	
합계	21	-21	**0**
평균	7	-7	0

지역 A	지역 B
76-69	62-69
76-69	62-69
76-69	62-69

Note: 처리효과의 합 $\sum_{j=1}^{g}\sum_{i=1}^{n_j}(\overline{Y_j}-\overline{Y}) = n_j \cdot \sum_{j=1}^{g}(\overline{Y_j}-\overline{Y})$은 항상 0이다.

(4)

	지역 A	지역 B	합계
	9	13	
	4	3	
	-13	-16	
합계	0	0	0
평균	0	0	0

지역 A	지역 B
85-76	75-62
80-76	65-62
63-76	46-62

Note: 잔차의 합 $\sum_{j=1}^{g}\sum_{i=1}^{n_j}(Y_{ij}-\overline{Y_j})$은 항상 0이다.

(5)

	관찰값
지역 A	**85**
	80
	63
지역 B	**75**
	65
	46
합계	**228**
평균	**76**

평균	처리 효과	잔차	**합계 (평+처+잔)**
69	7	9	**85**
69	7	4	**80**
69	7	-13	**63**
69	-7	13	**75**
69	-7	3	**65**
69	-7	-16	**46**
228	0	0	**228**
76	0	0	**76**

2. (1) $\overline{Y}_{j=1}$, $\overline{Y}_{j=2}$, $\overline{Y}$ 구하기

	j=1	j=2	
	200	250	
	247	290	
	300	351	
합계	747	891	1,638.0
평균	249.0	297.0	273.0

	총편차		처리효과		잔차	
	j=1	j=2	j=1	j=2	j=1	j=2
	-73	-23	-24	24	-49	-47
	-26	17	-24	24	-2	-7
	27	78	-24	24	51	54
합계	-72	72	-72	72	0	0
평균	-24	24	-24	24	0	0
합계	0		0		0	

3. (1)

지역 A j=1	지역 B j=2	SST		SSB		SSW	
85	75	256	36	49	49	81	169
80	65	121	16	49	49	16	9
63	46	36	529	49	49	169	256
합계		994		294		700	

Note: SST = SSB + SSW

참고: 연습문제 1의 총편차, 처리효과 및 잔차

지역 A j=1	지역 B j=2	총편차		처리효과		잔차	
85	75	16	6	7	-7	9	13
80	65	11	-4	7	-7	4	3
63	46	-6	-23	7	-7	-13	-16

(2)

변동	자유도	MS
SSB = 294	2-1=1	294/1 = 294
SSW = 700	6-2=4	700/4 = 175
SST = 994	6-1=5	

4. (1)

j=1	j=2	SST		SSB		SSW	
200	250	5,329	529	576	576	2,401	2,209
247	290	676	289	576	576	4	49
300	351	729	6,084	576	576	2,601	2,916
합계		13,636		3,456		10,180	

Note: SST = SSB + SSW

참고: 연습문제 1의 총편차, 처리효과 및 잔차

지역 A j=1	지역 B j=2	총편차		처리효과		잔차	
200	250	-73	-23	-24	24	-49	-47
247	290	-26	17	-24	24	-2	-7
300	351	27	78	-24	24	51	54

(2)

변동	자유도	MS
SSB = 3,456	2-1=1	3,456/1 = 3,456
SSW = 10,180	6-2=4	10,180/4 = 2,545
SST = 13,636	6-1=5	

5. (1) 3+1 = 4 (2) 79+1 = 80 (3) 3, 76 (4) 2.7249

(5)

분산분석(α=0.05)

변동요인	제곱합	자유도	제곱평균	F비	F기각치
처리	SSB	g-1	MSB = $\frac{SSB}{g-1}$	F = $\frac{MSB}{MSW}$	$F_{\alpha, g-1, n-g}$
잔차	SSW	n-g	MSW = $\frac{SSW}{n-g}$		
계	SST	n-1			

6. (1)

분산분석(α=0.05)

변동요인	제곱합	자유도	제곱평균	F비	P-값	F기각치
처리	9,710.43	3	3,236.808	36.8975	8.03E-15	2.7249
잔차	6,667.04	76	87.724			
계	16,377.47	79				

(2)

분산분석(α=0.01)

변동요인	제곱합	자유도	제곱평균	F비	P-값	F기각치
처리	11,478.13	4	2,869.532	27.6265	2.91E-16	3.4795
잔차	12,464.25	120	103.869			
계	23,942.38	124				

7. (1) ① H_0: $\mu_1 = \mu_2 = \cdots = \mu_4$

H_A: 적어도 하나는 같지 않다.

② Test Statistic: $F = \dfrac{MSB}{MSW}$, 자유도 = g-1=3, n-g=76

③ Rejection Region: $F > F_{\alpha,\, g-1,\, n-g} = F_{0.05,\, 3,\, 76} = 2.7249$

④ Value of the Test Statistic: $F = \dfrac{MSB}{MSW} = \dfrac{31.284}{116.763} = 0.2679$

⑤ Conclusion: Do not reject H_0

(2) ① H_0: $\mu_1 = \mu_2 = \cdots = \mu_4$

H_A: 적어도 하나는 같지 않다.

② Test Statistic: $F = \dfrac{MSB}{MSW}$, 자유도 = g-1=3, n-g=76

③ Rejection Region: $F > F_{\alpha,\, g-1,\, n-g} = F_{0.05,\, 3,\, 76} = 2.7249$

④ Value of the Test Statistic: $F = \dfrac{MSB}{MSW} = \dfrac{2{,}830.922}{80.743} = 35.0607$

⑤ Conclusion: Reject H_0

8. (1)

	A(i)		평균
B(j)	95	88	**91.50**
	129	87	**108.00**
평균	**112.00**	**87.50**	**99.75**

$\overline{Y} = 99.75,\ \overline{Y}_{i=1} = 112.0,\ \overline{Y}_{i=2} = 87.5,\ \overline{Y}_{j=1} = 91.5,\ \overline{Y}_{j=2} = 108$

(2) $Y_{ij} = \overline{Y} + (\overline{Y}_i - \overline{Y}) + (\overline{Y}_j - \overline{Y}) + (Y_{ij} - \overline{Y}_i - \overline{Y}_j + \overline{Y})$이므로

$Y_{11} = \overline{Y} + (\overline{Y}_{i=1} - \overline{Y}) + (\overline{Y}_{j=1} - \overline{Y}) + (Y_{11} - \overline{Y}_{i=1} - \overline{Y}_{j=1} + \overline{Y})$

= 99.75 + (112 - 99.75) + (91.5 - 99.75) + (95 − 112 − 91.5 + 99.75)
= 99.75(전체평균) + 12.25(i=1 처리효과) -8.25(j=1 처리효과) -8.75(잔차)
= 95

(3) $Y_{22} = \overline{Y} + (\overline{Y}_{i=2} - \overline{Y}) + (\overline{Y}_{j=2} - \overline{Y}) + (Y_{22} - \overline{Y}_{i=2} - \overline{Y}_{j=2} + \overline{Y})$
= 99.75 + (87.5 - 99.75) + (108 - 99.75) + (87 − 87.5 − 108 + 99.75)
= 99.75(전체평균) - 12.25(i=2 처리효과) + 8.25(j=2 처리효과) -8.75(잔차)
= 87

(4)

	SSA		SSB		SSW		SST	
	150.06	150.06	68.06	68.06	76.56	76.56	22.56	138.06
	150.06	150.06	68.06	68.06	76.56	76.56	855.56	162.56
합계		600.25		272.25		306.25		1,178.75

(5)

분산 분석

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(열)	600.25	1	600.25	1.9600	0.3949	161.4476
인자 B(행)	272.25	1	272.25	0.8890	0.5187	161.4476
잔차	306.25	1	306.25			
계	1178.75	3				

9. (1)

	A(i)		평균
B(j)	5	4	**4.50**
	7	10	**8.50**
평균	**6.00**	**7.00**	**6.50**

$\overline{Y} = 6.5$, $\overline{Y}_{i=1} = 6$, $\overline{Y}_{i=2} = 7$, $\overline{Y}_{j=1} = 4.5$, $\overline{Y}_{j=2} = 8.5$

(2)

	SSA		SSB		SSW		SST	
	0.25	0.25	4.00	4.00	1.00	1.00	2.25	6.25
	0.25	0.25	4.00	4.00	1.00	1.00	0.25	12.25
합계		1.00		16.00		4.00		21.00

(3)

분산 분석

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(열)	1.00	1	1.00	0.2500	0.7048	161.4476
인자 B(행)	16.00	1	16.00	4.0000	0.2952	161.4476
잔차	4.00	1	4.00			
계	21.00	3				

(4) 요인 1 수준의 평균들이 같은 지에 관심이 있다면

① H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_g = 0$

H_A: 적어도 하나는 0이 아니다.

② Test Statistic: $F_{g-1,(g-1)(c-1)} = \dfrac{MSA}{MSE}$

③ Reject Region $F > F_{\alpha,g-1,(g-1)(c-1)} = F_{0.05,1,1} = 161.4476$

④ Value of the Test Statistic $F = \dfrac{MSA}{MSE} = 0.2500$

⑤ Conclusion Do not reject H_0

요인 2 수준의 평균들이 같은 지에 관심이 있다면

① H_0: $\beta_1 = \beta_2 = \cdots = \beta_c = 0$

H_A: 적어도 하나는 0이 아니다.

② Test Statistic: $F_{c-1,(g-1)(c-1)} = \dfrac{MSA}{MSE}$

③ Reject Region $F > F_{\alpha,g-1,(g-1)(c-1)} = F_{0.05,1,1} = 161.4476$

④ Value of the Test Statistic $F = \dfrac{MSB}{MSE} = 4.0000$

⑤ Conclusion Do not reject H_0

10. (1)

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비
인자 A(행)	17,606.32	4	4,401.58	34.2743
인자 B(열)	5,133.47	2	2,566.74	19.9867
잔차	1,027.38	8	128.42	
계	23,767.16	14		

11. (1) $\overline{Y} = 99$, $\overline{Y}_{i=1} = 97$, $\overline{Y}_{i=2} = 101$, $\overline{Y}_{j=1} = 96$, $\overline{Y}_{j=2} = 102$,

$\overline{Y}_{11} = 92,\ \overline{Y}_{12} = 102,\ \overline{Y}_{21} = 100,\ \overline{Y}_{22} = 102$

(2)

	SST		SS행		SS열		SS행열		잔차	
	1	25	4	4	9	9	4	4	36	4
	169	1	4	4	9	9	4	4	36	4
	0	81	4	4	9	9	4	4	1	36
	4	9	4	4	9	9	4	4	1	36
합계	290		32		72		32		154	

(3)

분산 분석 (α=0.05)

변동의 요인	제곱합	자유도	제곱 평균	F 비	F 기각치
인자 A(행)	32	1	32.00	0.8312	7.709
인자 B(열)	72	1	72.00	1.8701	7.709
상호작용	32	1	32.00	0.8312	7.709
잔차	154	4	38.50		
계	290	7			