[10주차-2차시]

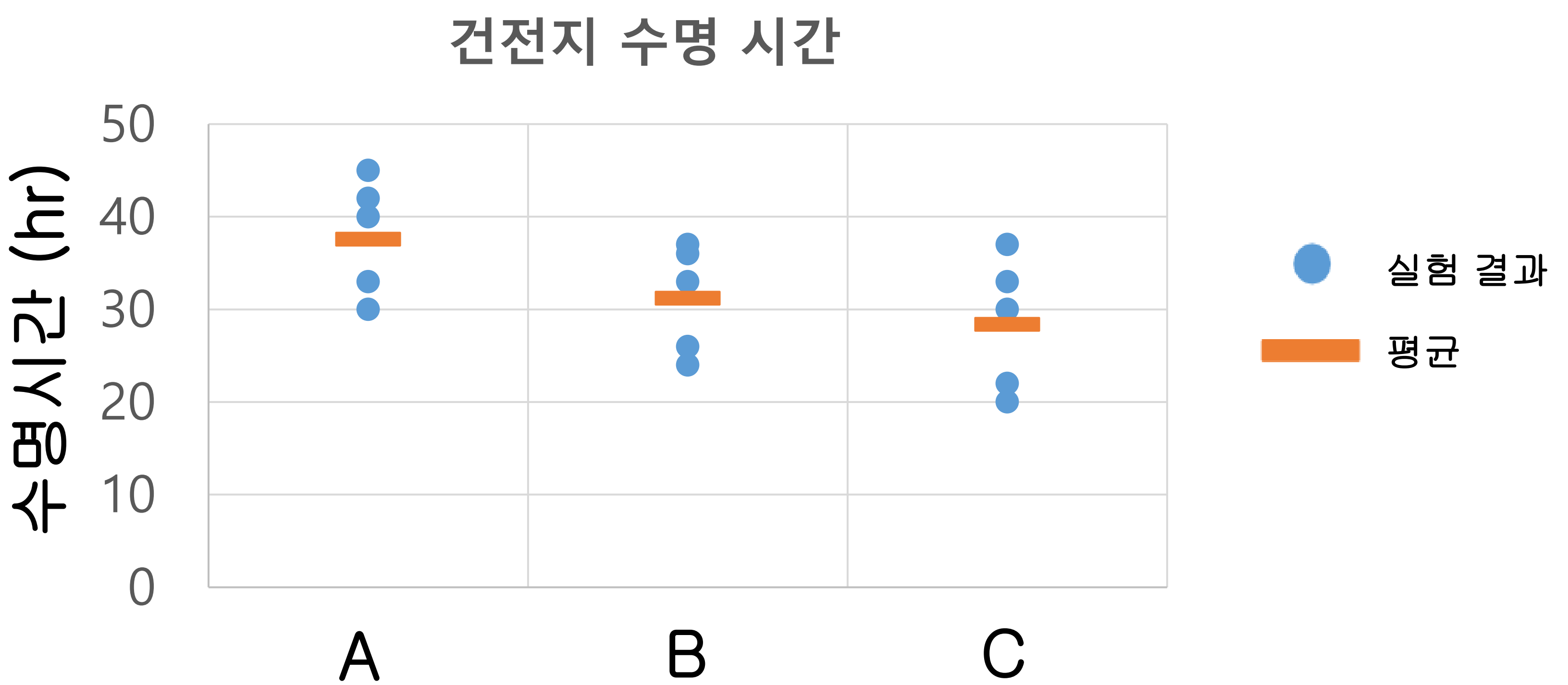
분산분석(1)

일원배치 분산분석

중앙대학교 기계공학부 김석민 교수
smkim@cau.ac.kr

분산분석

- 분산분석 (Analysis of variance, ANOVA)
  - 둘 이상의 모집단간의 평균을 비교하는 방법
  - Ex1) A, B, C 사에서 제작한 건전지중 어느건전지의 수명이 오래 가는가?



- Ex1) A, B, C 사에서 제작한 건전지중
  어느건전지의 수명이 오래 가는가?

- 각 회사별 건전지 5개를 이용한 실험결과 평균기준 A,B,C 순으로 수명이 길다.
- 5개 표본의 평균을 모평균의 평균으로 판단할 수 있나?

- 모평균의 구간 추정 결과를 통해 A, B, C 모평균의 추정 범위를 비교하여 A, B, C 모집단의 모수를 비교할 수 있음.

- 검정의 방법을 통해 A, B, C 간의 유의미한 차이가 있는지 확인할 수 있음.

분산분석 용어 정리

- **반응 (Response)** : 실험의 결과 (통계에서 관측, 실험되는 수치)
- **요인 (Factor)** : 실험에서 반응에 영향을 미칠것으로 판단되어 독립적으로 그 값을 변화시키는 항목
- **수준 (Level)** : 각 요인이 갖는 값
- **처리 (Treatment)** : 요인의 각 값을 시행하는 행위
- Ex. 1 A, B, C 사에서 제작한 건전지중 어느건전지의 수명이 오래 가는가?
  - 반응 : 건전지의 수명, 요인 : 건전지 제작사, Level : A, B, C
- Ex. 2 화분에서 당근을 키우는데 물의 양을 매일 1컵, 2컵, 3컵 화분에 주는 거름의 양을 0.5kg, 1kg, 1.5kg, 2kg 씩 다르게 주고 수확기에 당근의 무게를 측정하였다. 가장 무거운 당근을 생산하는 물과 거름의 양은?
  - 반응 : 당근의 무게, 요인: 물/거름의 양 ,
    물의 Level : 1컵, 2컵, 3컵, 거름의 Level : 0.5 ~ 5kg

일원배치 분산분석법 (요인 1개에 대한 실험)

- **완전랜덤화 설계 (Randomize design)**
  실험에서 모집단 (어떤 특정한 처리 하에서 나타날 수 있는 반응값 (실험값) 들의 전체 모임) 에서 독립적으로 표본을 추출하는 방법
- Ex. 온도를 제어할 수 있고 내부 압력을 측정할 수 있는 챔버를 이용하여 기체의 압력변화를 분석하고자 한다.
  실험은 상온상태에서 챔버 내부에 일정량의 공기를 주입하고 온도를 100도, 200도, 300도, 400도로 변경하며 각각의 압력을 측정하는 형태로 진행되며 (4 level : I=4)
  1회 실험에서 한가지 온도만 setting 이 가능하다.

  각 온도별로 5회의 반복 (replication, n=5) 실험을 진행하고자 한다. 실험의 순서를 계획하여라.

일반적인 경우 : 1~5회차 100도, 6~10회차 200도 …

완전랜덤화 설계 : 400, 300, 400, 100, 100, 200, 400 … 등

k개의 처리, n 회 반복 실험의 결과에 대한 수식화

y_{ij} : i 번째 treatment 에서 j 번째 반복 실험 결과

	관측자료	평균	오차의 제곱합
처리 1	$y_{11}, y_{12}, \cdots, y_{1n}$	$\bar{y}_{1\cdot} = \mu_1$	$\sum_{j=1}^{n} (y_{1j} - \bar{y}_{1\cdot})^2$
처리 2	$y_{21}, y_{22}, \cdots, y_{2n}$	$\bar{y}_{2\cdot} = \mu_2$	$\sum_{j=1}^{n} (y_{2j} - \bar{y}_{2\cdot})^2$
…			
처리 k	$y_{k1}, y_{k2}, \cdots, y_{kn}$	$\bar{y}_{3\cdot} = \mu_3$	$\sum_{j=1}^{n} (y_{kj} - \bar{y}_{k\cdot})^2$

$$y_{i\cdot} = \sum_{j=1}^{n} y_{ij}, \bar{y}_{i\cdot} = \frac{y_{i\cdot}}{n}$$

$$\bar{y} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n} y_{ij}}{n \times k}$$

$$\varepsilon_{ij} = y_{ij} - \bar{y}_{i\cdot}$$

$$y_{ij} = \overline{y}_{i\bullet} + \varepsilon_{ij} = \mu_i + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim N(0, \sigma^2)$$

$$\alpha_i = \overline{y}_{i\bullet} - \overline{y} = \mu - \mu_i$$: i 번째 처리의 효과

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

$$y_{ij} - \mu = (\mu_i - \mu) + (y_{ij} - \mu_i)$$

$$(y_{ij} - \mu)^2 = \{(\mu_i - \mu) + (y_{ij} - \mu_i)\}^2$$

$$= (\mu_i - \mu)^2 + 2(\mu_i - \mu)(y_{ij} - \mu_i) + (y_{ij} - \mu_i)^2$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \mu)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}(\mu_i - \mu)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \mu_i)^2$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(\mu_i-\mu)(y_{ij}-\mu_i)=\sum_{i=1}^{k}\sum_{j=1}^{n}\mu_i y_{ij}-\mu y_{ij}-\mu_i{}^2+\mu\mu_i$$

$$=\sum_{i=1}^{k}\sum_{j=1}^{n}\mu_i y_{ij}-\sum_{i=1}^{k}\sum_{j=1}^{n}\mu y_{ij}-\sum_{i=1}^{k}\sum_{j=1}^{n}\mu_i{}^2+\sum_{i=1}^{k}\sum_{j=1}^{n}\mu\mu_i$$

$$=\sum_{i=1}^{k}\mu_i\sum_{j=1}^{n}y_{ij}-\mu\sum_{i=1}^{k}\sum_{j=1}^{n}y_{ij}-\sum_{i=1}^{k}\mu_i{}^2\sum_{j=1}^{n}1-\mu\sum_{i=1}^{k}\mu_i\sum_{j=1}^{n}1$$

$$=\sum_{i=1}^{k}\mu_i y_{i\bullet}-\mu y_{\bullet\bullet}-\sum_{i=1}^{k}\mu_i{}^2 n-\mu\sum_{i=1}^{k}n\mu_i$$

$$=\sum_{i=1}^{k}\mu_i(n\mu_i)-\mu(kn\mu)-n\sum_{i=1}^{k}\mu_i{}^2-n\mu\sum_{i=1}^{k}\mu_i$$

$$=n\sum_{i=1}^{k}\mu_i{}^2-kn\mu^2-n\sum_{i=1}^{k}\mu_i{}^2-n\mu(k\mu)=0$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij}-\bar{y})^2=\sum_{i=1}^{k}\sum_{j=1}^{n}(\bar{y}_{i\bullet}-\bar{y})^2+\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i\bullet})^2$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij}-\bar{y})^2=n\sum_{i=1}^{k}(\bar{y}_{i\bullet}-\bar{y})^2+\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i\bullet})^2$$

총 제곱합 (sum of squares total, SST)

처리 제곱합 (sum of squares treatment, SStr)

오차 제곱합 (sum of squares errors, SSE)

SST = SStr + SSE

일원배치 실험에 대한 분산분석

- **귀무가설 (H_0):** $\mu_1 = \mu_2 = \cdots = \mu_k = \mu$

$\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$ **처리의 효과가 없다.**

- **대립가설 (H_1): 적어도 하나의 i 에 대해** $\alpha_i \neq 0$
- **검정 통계량**

귀무가설이 참이라면 각 treatment 반복 결과를 각각의 표본으로 생각하고 $\bar{y}_{i.}$ 을 각 표본의 평균으로 판단할 수 있음.

$$y_{11}, y_{12}, \cdots, y_{1n} \quad \bar{y}_{1.}$$

$$y_{21}, y_{22}, \cdots, y_{2n} \quad \bar{y}_{2.}$$

$$\vdots$$

귀무가설 하에서 모 표준편차의 예측

$$y_{11}, y_{12}, \cdots, y_{kn} \sim N(\bar{y}, \sigma^2) \quad \rightarrow \quad \hat{\sigma}^2 = \frac{\sum_i \sum_j (y_{ij} - \bar{y})^2}{N-1}$$

SST

전체실험의 자유도

$$\bar{y}_{1\cdot}, \bar{y}_{2\cdot}, \cdots, \bar{y}_{k\cdot} \sim N(\bar{y}, \frac{\sigma^2}{n}) \quad \rightarrow \quad \frac{\hat{\sigma}^2}{n} = \frac{\sum_i (\bar{y}_{i\cdot} - \bar{y})^2}{k-1}$$

$$\hat{\sigma}^2 = \frac{n\sum_i (\bar{y}_{i\cdot} - \bar{y})^2}{k-1}$$

SStr

처리의 자유도

$$y_{11}, y_{12}, \cdots, y_{1n} \sim N(\bar{y}_{1.}, \sigma_1^2) \;\rightarrow\; \sigma_1^{\,2} = \frac{\sum_j (y_{1j} - \bar{y}_{1.})^2}{n-1}$$

$$y_{21}, y_{22}, \cdots, y_{2n} \sim N(\bar{y}_{2.}, \sigma_2^2) \;\rightarrow\; \sigma_2^{\,2} = \frac{\sum_j (y_{2j} - \bar{y}_{2.})^2}{n-1}$$

$$y_{k1}, y_{k2}, \cdots, y_{kn} \sim N(\bar{y}_{k.}, \sigma_k^2) \;\rightarrow\; \sigma_k^{\,2} = \frac{\sum_j (y_{kj} - \bar{y}_{k.})^2}{n-1}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{k} \sigma_i^{\,2}}{k} = \frac{\sum_i \frac{\sum_j (y_{ij} - \bar{y}_{i.})^2}{n-1}}{k} = \frac{\sum_i \sum_j (y_{ij} - \bar{y}_{i.})^2}{k(n-1)}$$

SSE

오차의 자유도

일원배치 실험에 대한 분산분석

검정 통계량 : 모표준편차의 비에 대한 검정에 사용된 F분포 이용

F 분포 (서로다른 χ^2 분포의 비에 대한 분포)

두개의 모집단이 정규분포를 따르고 $\sigma_1 = \sigma_2$ 일때

$$F = \frac{S_1^{\,2} / \sigma_1^{\,2}}{S_2^{\,2} / \sigma_2^{\,2}} = \frac{S_1^{\,2}}{S_2^{\,2}} \sim F(n_1 - 1, n_2 - 1)$$

귀무가설 하에서
SStr 을 이용한 모표준편차 예측 결과와,
SSE 를 이용한 모표준편차 예측 결과가 동일한 모수임.

$$F = \frac{SStr / (k-1)}{SSE / (k(n-1))} \sim F(k-1, k(n-1))$$

F 값이 F 함수의 일반적인 범위를 벗어나면 귀무가설 기각

[11주차-1차시]

분산분석(2)

일원배치 분산분석 표

중앙대학교 기계공학부 김석민 교수
smkim@cau.ac.kr

일원배치 분산분석

- **가설 :** $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$

  $H_1 :$ 적어도 하나의 i에 대해 $\alpha_i \neq 0$

- **검정통계량**

$$F = \frac{SStr/(k-1)}{SSE/(k(n-1))} \sim F(k-1, k(n-1))$$

- **기각역**

$$P : F \geq F_\alpha(k-1, k(n-1))$$

귀무가설이 기각되면 SStr 은 error 보다 큰 값이 되어야 하므로 단측 검정을 실시함.

분산분석표

- **분산분석표 (ANOVA table)**

요인	제곱합	자유도	평균제곱합
처리	$SStr = \mathrm{n}\sum_i (\bar{y}_{i\cdot} - \bar{y})^2$	$k-1$	$MStr = \dfrac{SStr}{k-1}$
오차	$SSE = \sum_i \sum_j (y_{ij} - \bar{y}_{i\cdot})^2$	$N-k$	$MSE = \dfrac{SSE}{N-k}$
합계	$SST = \sum_i \sum_j (y_{ij} - \bar{y})^2$	$N-1$	

$$N = kn, \quad F = \frac{MStr}{MSE} \sim F(k-1, N-k)$$

예제 7.1

- **CD의 음질에 영향을 미치는 네종류의 코팅방법, A, B, C, D 가 음질의 재생에 얼마나 차이를 유발하는지를 비교하려고 한다. 다음의 결과는 네종류의 방법으로 코팅 처리된 CD 로 부터 소리의 일그러짐을 5회씩 측정하여 얻은 자료이다. 이 자료로부터 네 종류의 코팅방법에 따라 음질에 차이가 있다고 할 수 있는지 유의수준 5% 로 검정하라.**

코팅물질	관측자료	평균
A	10, 14, 10, 12, 14	12.0
B	18, 17, 18, 19, 18	18.0
C	13, 13, 14, 15, 15	14.0
D	12, 15, 17, 15, 16	15.0
		$\bar{y} = 14.75$

$$SStr = 5\sum_{i=1}^{4}(\bar{y}_{i\cdot} - \bar{y})^2 = 5\{(12-14.75)^2 + \bullet\bullet\bullet + (15-14.75)^2\} = 93.75$$

$$SSE = \sum_{i=1}^{4}\sum_{j=1}^{n_i}(y_{ij} - \overline{y_i})^2 = 36, \quad SST = SStr + SSE$$

- 분산분석표 (ANOVA table)

요인	제곱합	자유도	평균제곱합	F
처리	93.75	3	31.25	13.8889
오차	36	16	2.25	
합계	129.75	19		

```
>>finv(0.95, 3,16)
ans = 3.2389
fcdf(13.8889, 3, 16)
ans =0.9999
```

P value < 0.0001

$F > F_{\alpha}(3,16)$ 이므로 귀무가설 기각

코팅방법에 따라 음질에 영향 있다

MATLAB 을 이용한 분석

x = [10 18 13 12;
14 17 13 15;
10 18 14 17;
12 19 15 15;
14 18 15 16]

각 treatment 별로 세로(열)로 data 표기

>> anova1(x)

ans =

1.0175e-04

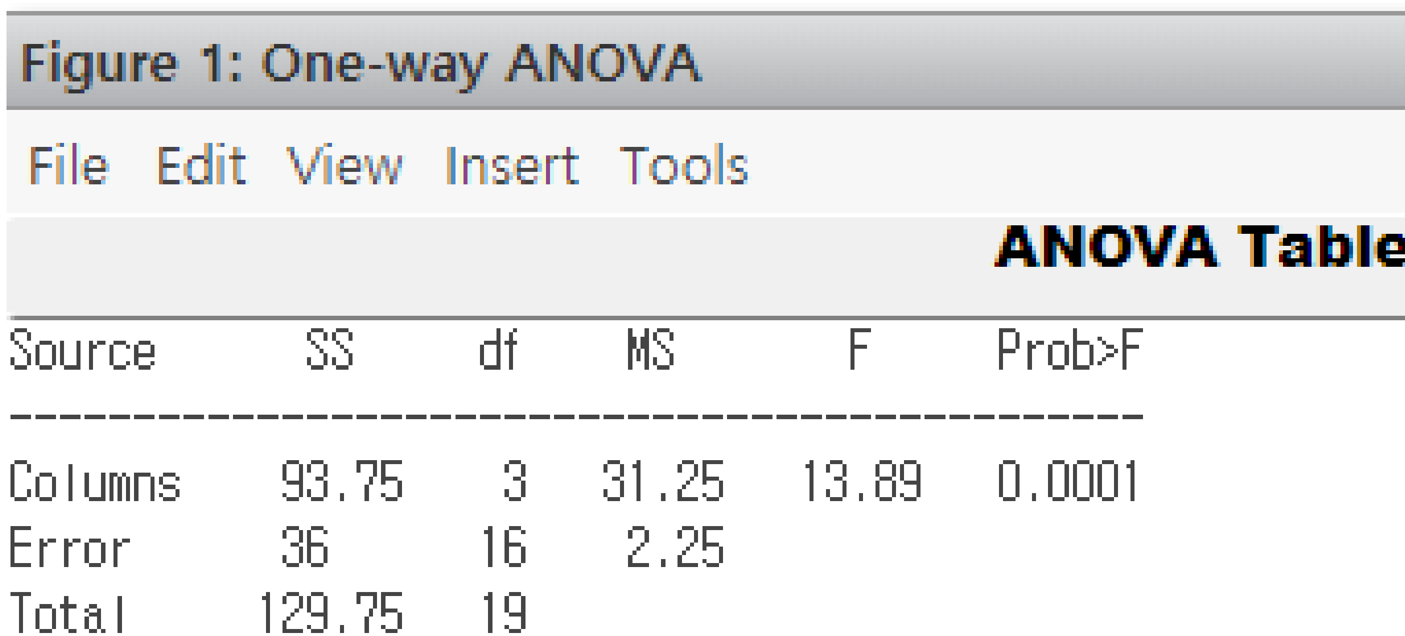


Figure 1: One-way ANOVA

File Edit View Insert Tools

ANOVA Table

Source	SS	df	MS	F	Prob>F
Columns	93.75	3	31.25	13.89	0.0001
Error	36	16	2.25		
Total	129.75	19			

[11주차-2차시]

분산분석(3)

이원배치 분산분석

중앙대학교 기계공학부 김석민 교수
smkim@cau.ac.kr

본 콘텐츠는 교육부와 한국연구재단의 재원으로 지원을 받아 제작되었으며,

반복이 없는 이원배치법

y_{ij} : Factor A은 i 번째 level, Factor B 는 j 번째 level 에서 관측한 결과값

모형 : $y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \qquad i = 1, \cdots, a \quad j = 1, \cdots, b$

$H_0(A) : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$

v.s. $H_1(A)$:적어도 하나 $\alpha_i \neq 0$

$H_0(B) : \beta_1 = \beta_2 = \cdots = \beta_b = 0$

v.s. $H_1(B)$:적어도 하나 $\beta_j \neq 0$

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

$$y_{ij} - \bar{y} = (\bar{y}_{i\bullet} - \bar{y}) + (\bar{y}_{\bullet j} - \bar{y}) + \varepsilon_{ij}$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{i\bullet} - \bar{y})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{\bullet j} - \bar{y})^2 + SSE$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y})^2 = l\sum_{i=1}^{a}(\bar{y}_{i\bullet} - \bar{y})^2 + k\sum_{j=1}^{b}(\bar{y}_{\bullet j} - \bar{y})^2 + SSE$$

SST SS(A) SS(B)

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i\bullet} - \bar{y}_{\bullet j} + \bar{y})^2$$

반복이 없는 이원배치법 분산분석표

분산분석표 (ANOVA table)

요인	제곱합	자유도	평균제곱합
A	$SS(A) = l\sum_i (\bar{y}_{i\cdot} - \bar{y})^2$	a−1	SS(A)/(a−1)
B	$SS(B) = k\sum_j (\bar{y}_{\cdot j} - \bar{y})^2$	$\boldsymbol{b}$−1	SS(B)/($\boldsymbol{b}$−1)
오차	SSE	(a−1)($\boldsymbol{b}$−1)	SSE/(a−1)($\boldsymbol{b}$−1)
합계	$SST = \sum_i \sum_j (y_{ij} - \bar{y})^2$	$\boldsymbol{ab}$−1	

$$F_A = \frac{MS(A)}{MSE} \sim F(a-1, (a-1)(b-1))$$

$$F_B = \frac{MS(B)}{MSE} \sim F(b-1, (a-1)(b-1))$$

예제 7.2

- 코팅 물질 (요인 A)와 코팅방법 (요인B)이 음질에 미치는 영향을 알아보기 위하여 반복이 없는 이원배치모형의 실험을 수행하였다. 실험은 12회를 완전랜덤화 하여 실시한 후 다음과 같은 자료를 얻었다. (유의수준 5% 에서 A, B 의 영향 검정)

	B1	B2	B3	B4	평균
A1	10	14	17	12	13.25
A2	15	18	17	17	16.75
A3	14	19	19	22	18.50
평균	13	17	17.67	17	16.17

$$SS(A) = l\sum_i (\bar{y}_{i\bullet} - \bar{y})^2 = 4\{(13.25 - 16.17)^2 + \ldots\} = 57.2$$

$$SS(B) = k\sum_i (\bar{y}_{\bullet j} - \bar{y})^2 = 3\{(13 - 16.17)^2 + \ldots\} = 41.0$$

$$SST = \sum_i \sum_j (\bar{y}_{ij} - \bar{y})^2 = (10 - 16.17)^2 + \ldots = 121.7$$

- **분산분석표 (ANOVA table)**

요인	제곱합	자유도	평균제곱합	F	P
A	57.2	2	28.6	7.30	0.0247
B	41.0	3	13.7	3.49	0.0901
오차	23.5	6	3.92		
합계	121.7	11			

```
>>1-fcdf(7.30,2,6)
ans = 0.0247
>>1-fcdf(3.49,3,6)
ans = 0.0901
```

```
>>finv(0.95,2,6)
ans =5.1433
>> finv(0.95,3,6)
ans = 4.7571
```

A는 영향을 주지만
B는 영향을 주지 않는다

```
>> x = [10 15 14
        14 18 19
        17 17 19
        12 17 22];
>> anova2(x)
ans = 0.0247 0.0901
```

Figure 1: Two-way ANOVA

File Edit View Insert Tools

ANOVA Table

Source	SS	df	MS	F	Prob>F
Columns	57.167	2	28.5833	7.3	0.0247
Rows	41	3	13.6667	3.49	0.0901
Error	23.5	6	3.9167		
Total	121.667	11			

반복이 있는 이원배치법

y_{ijk} : Factor A은 i 번째 level, Factor B 는 j 번째 level 에서 k 번째 반복 실험에서 관측한 결과값

Factor A level : a, Factor B level : b, 반복 : n

Factor B

Factor A	1	2	. . .	b
1	$y_{111}, y_{112}, \ldots, y_{11n}$	$y_{121}, y_{122}, \ldots, y_{12n}$		$y_{1b1}, y_{1b2}, \ldots, y_{1bn}$
2	$y_{211}, y_{212}, \ldots, y_{21n}$	$y_{221}, y_{222}, \ldots, y_{22n}$		$y_{2b1}, y_{2b2}, \ldots, y_{2bn}$
⋮				
a	$y_{a11}, y_{a12}, \ldots, y_{a1n}$	$y_{a21}, y_{a22}, \ldots, y_{a2n}$		$y_{ab1}, y_{ab2}, \ldots, y_{abn}$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y})^2=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{i..}-\bar{y})^2+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{.j.}-\bar{y})^2$$

$$+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y})^2+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij.})^2$$

SST SS(A) SS(B)

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y})^2=bn\sum_{i=1}^{a}(\bar{y}_{i..}-\bar{y})^2+an\sum_{j=1}^{b}(\bar{y}_{.j.}-\bar{y})^2$$

SSE

$$+n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y})^2+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij.})^2$$

SS(AB)

$$SS_T=SS_A+SS_B+SS_{AB}+SS_E$$

[자유도]

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

반복이 있는 이원배치법 가설

$$H_0(A): \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$$

v.s. $H_1(A)$:적어도 하나 $\alpha_i \neq 0$

$$H_0(B): \beta_1 = \beta_2 = \cdots = \beta_b = 0$$

v.s. $H_1(B)$:적어도 하나 $\beta_j \neq 0$

$$H_0(AB): \alpha\beta_{11} = \alpha\beta_{12} = \cdots = \alpha\beta_{ab} = 0$$

v.s. $H_1(AB)$:적어도 하나 $\alpha\beta_{ij} \neq 0$

반복이 있는 이원배치법 분산분석표

요인	제곱합	자유도	평균제곱합	F
A	$SS(A) = bn\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y})^2$	a-1	SS(A) /a-1	MS(A) /MSE
B	$SS(B) = an\sum_{j=1}^{b}(\bar{y}_{.j.} - \bar{y})^2$	b-1	SS(B) /b-1	MS(B) /MSE
AB	SS(AB)	(a-1)(b-1)	SS(AB) /(a-1)(b-1)	MS(AB) /MSE
오차	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2$	ab(n-1)	SSE /ab(n-1)	
합계	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y})^2$	abn-1		

7.3

반도체공정에서 감광제(photoresist) spin coating 공정

Factor A : 회전속도 (7500 or 6500 rpm)

Factor B : 감광제 용량 (3 or 5 cc)

Response : 코팅 두께

	B1	B2	평균
A1	432,430,449 ,422,454	453,445,426 ,452,430	439.3
A2	447,466,448 ,447,450	449,454,463 ,463,445	453.2
평균	444.5	448.0	446.25

$$SS(A) = bn\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y})^2 = 2\times 5\{(439.3 - 446.25)^2 + \ldots\} = 966.05$$

$$SS(B) = an\sum_{j=1}^{b}(\bar{y}_{.j.} - \bar{y})^2 = 2\times 5\{(444.5 - 446.25)^2 + \ldots\} = 61.25$$

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2 = \{(432 - 437.4)^2 + \ldots\} = 1888.0$$

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y})^2 = \{(432 - 446.25)^2 + \ldots\} = 2915.75$$

$$SS(AB) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{i..} - \bar{y}_{,j.} + \bar{y})^2 = SST - SS(A) - SS(B) - SSE = 0.45$$

분산분석표

요인	제곱합	자유도	평균제곱합	F
A	966.05	1	966.05	8.1869
B	61.25	1	61.25	0.5191
AB	0.45	1	0.45	0.0038
오차	1888.0	16	118.0	
합계	2915.75	19		

```
>>finv(0.95,1,16)
ans =4.4940
```

```
>>1-fcdf(8.1869,1,16)
ans = 0.0113
>>1-fcdf(0.5191,1,16)
ans = 0.4816
>>1-fcdf(0.0038,1,16)
ans = 0.9516
```

A는 두께에 영향을 주나
B, 및 AB 는 영향을 주지 않음.

```
>> x = [432 447
        430 466
        449 448
        422 447
        454 450
        453 449
        445 454
        426 463
        452 463
        430 445 ];
>> anova2(x, 5)
```

B1 조건에서
A1 및 A2 조건5회 반복 결과

B2 조건에서
A1 및 A2 조건5회 반복 결과

ans = 0.0113 0.4816 0.9515

Figure 2: Two-way ANOVA

File Edit View Insert Tools

ANOVA Table

Source	SS	df	MS	F	Prob>F
Columns	966.05	1	966.05	8.19	0.0113
Rows	61.25	1	61.25	0.52	0.4816
Interaction	0.45	1	0.45	0	0.9515
Error	1888	16	118		
Total	2915.75	19			