## ■ 일원 배치법(single factorial design, one-way layout)

### ○ 용어정리

- 요인(factor) : 실험에 고려된 설명변수 (회귀분석: predictor)
  - experimental factor vs observational factor
- 수준(level) : 요인의 값
- 처리(treatment) : 요인과 수준의 조합

◉ 4 종류의 사료에 의한 병아리 무게증가 실험

  - 요인: 사료
  - 수준(처리)의 수: 4
- 하나의 요인의 여러 수준에 대한 특성치의 비교를 위한 계획법
- 실험 단위의 배치 또는 실험순서에 있어 확률화의 원리에 충실해야 한다고 해서 **완전 확률화 계획법(completely randomized design)**이라고도 불림

# □ 통계적 모형

## ○ 자료 형태

| | | 요인의 수준(처리) | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | $\cdots$ | $p$ |
| 반복수 | 1 | $Y_{11}$ | $Y_{21}$ | $\cdots$ | $Y_{p1}$ |
| | 2 | $Y_{12}$ | $Y_{22}$ | $\cdots$ | $Y_{p2}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | $n$ | $Y_{1n}$ | $Y_{2n}$ | $\cdots$ | $Y_{pn}$ |

○ 구조식

$$Y_{ij} = \mu_i + \varepsilon_{ij}, \qquad i = 1, \ ..., p, \ \ j = 1, ..., n$$
$$= \mu + \tau_i + \varepsilon_{ij}$$

- $\mu$ : 전체 모평균
- $\tau_i = \mu_i - \mu$ : $i$ 번째 처리 효과
- $\varepsilon_{ij} \sim$ iid $N(0, \sigma^2)$

○ 통계적 추론

① 처리효과에 대한 적절한 가설 검정

(예) $H_0 : \tau_1 = \ \cdots \ = \tau_p = 0$ ⇨ 평균이 같은지?

(예) $H_0 : \mu_1 = \ \cdots \ = \mu_p = 0$ ⇨ 평균이 0 인지?

② 처리효과의 추정 (단, 처리효과의 동일성 가설 기각시)

## □ 모형의 분류

### ○ 고정효과 모형(fixed effect model)

- 요인의 수준이 실험자의 의도에 의해 조정 또는 결정되는 경우
- $\tau_i$에 대한 가설검정의 결과는 분석에서 **고려된 요인의 수준에 대해서만 적용**할 수 있음

### ○ 변량효과 모형(random effect model)

- 고려할 수 있는 요인의 수준에서 random하게 선택된 경우
- $\tau_i$는 모수가 아니고 **확률변수로 취급**
- $\tau_i$의 분산에 대해 $\sigma_\tau^2 = 0$인지를 검정함
- 선택된 수준뿐만 아니라 **고려했던 모든 요인의 수준에 대해서도 결과를 확장**시킬 수 있음

# □ 고정효과모형

- Notations :

$$N = \sum_{i=1}^{p} n_i, \quad Y_{i.} = \sum_{j=1}^{n_i} Y_{ij}, \quad \overline{Y}_{i.} = Y_{i.}/n_i, \quad Y_{..} = \sum_{i=1}^{p}\sum_{j=1}^{n_i} Y_{ij}, \quad \overline{Y}_{..} = Y_{..}/N$$

- 고정효과 모수모형에서는 $\sum_{i=1}^{p} \tau_i = 0$

- 모수추정
  - $\mu \Leftarrow \overline{Y}_{..}$,
  - $\mu_i \Leftarrow \overline{Y}_{i.}$
  - $\tau_i = \mu_i - \mu \Leftarrow \overline{Y}_{i.} - \overline{Y}_{..}$
  - $\varepsilon_{ij} = Y_{ij} - \mu_i \Leftarrow e_{ij} = Y_{ij} - \overline{Y}_{i.}, \quad \sum_j e_{ij} = 0, \ \forall\, i$

○ 가설검정

◦ $H_0$ : $\tau_1 = \cdots = \tau_p = 0$ vs $H_1$ : 최소한 하나의 $i$에 대해 $\tau_i \neq 0$

◦ 귀무가설 하에서

$$Y_{ij} = \mu + \varepsilon_{ij} \;\Rightarrow\; Y_{ij} - \mu = \varepsilon_{ij} \;\Leftarrow\; \hat{\varepsilon}_{ij} = Y_{ij} - \overline{Y}_{..}$$

○ 변동분해

$$SSTO = \sum_i \sum_j \left( Y_{ij} - \overline{Y}_{..} \right)^2 = \sum_i \sum_j \left( Y_{ij} - \overline{Y}_{i.} + \overline{Y}_{i.} - \overline{Y}_{..} \right)^2$$

$$= \sum_i \sum_j ( Y_{ij} - \overline{Y}_{i.} )^2 + \sum_i \sum_j ( \overline{Y}_{i.} - \overline{Y}_{..} )^2 + 2 \sum_i \sum_j ( Y_{ij} - \overline{Y}_{i.} )( \overline{Y}_{i.} - \overline{Y}_{..} )$$

$$= \underbrace{\sum_i \sum_{j1} \left( Y_{ij} - \overline{Y}_{i.} \right)^2}_{SSE\ (N-p)} + \underbrace{\sum_i n_i \left( \overline{Y}_{i.} - \overline{Y}_{..} \right)^2}_{SSTR\ (p-1)}$$

◦ SSTO: 총제곱합, SSE: 오차제곱합, SSTR: 처리제곱합

● 평균제곱

○ MSE = SSE/(N-p)

- $E(MSE) = \sigma^2$

○ MSTR = SSTR/(p-1)

- $E(MSTR) = \sigma^2 + \sum_{i=1}^{p} n_i \tau_i^2 / (p-1)$

○ 검정통계량

$$F_0 = \frac{SSTR/(p-1)}{SSE/(N-p)} = \frac{MSTR}{MSE} \sim F_{p-1, N-p}$$

○ 대립가설 하에서는 $E(MSTR) > E(MSE)$ ⇨ 기각역 : $F_0 > F_{\alpha, p-1, N-p}$

○ 분산분석표(ANOVA table)

| 변인 | 자유도 | 제곱합(SS) | 평균제곱(MS) | F |
|---|---|---|---|---|
| 처리(모형) | $p-1$ | SSTR | MSTR | MSTR/MSE |
| 오차 | $N-p$ | SSE | MSE | |
| 전체 | $N-1$ | SSTO | | |

- 간이식
  - 보정항 $CT = Y_{..}^2/N$
  - $SSTO = \sum_i \sum_j Y_{ij}^2 - CT$
  - $SSTR = \sum_i Y_{i.}^2/n_i - CT$

## □ 처리 평균치간의 비교

- 어떻게 가설을 설정할 것인가?
  - ○ 적절한 형태의 선형식을 유도하여 가설을 설정

① $H_0 : \mu_1 + \mu_2 + \mu_3 - 3\mu_4 = 0 \quad \text{vs} \quad H_1 : \mu_1 + \mu_2 + \mu_3 - 3\mu_4 \neq 0$

② $H_0 : \mu_1 + \mu_2 - 2\mu_3 = 0 \quad \text{vs} \quad H_1 : \mu_1 + \mu_2 - 2\mu_3 \neq 0$

③ $H_0 : \mu_1 - \mu_2 = 0 \quad \text{vs} \quad H_1 : \mu_1 - \mu_2 \neq 0$

※ 각 선형식 계수의 합은 0

### ○ 대비

- p개의 평균값에 대한 임의의 선형 결합식을

$$L = c_1 \overline{Y}_{1.} + \cdots + c_p \overline{Y}_{p.} \qquad (\text{단, } \sum_i c_i = 0)$$

처리 평균값의 **비교** 또는 **대비(contrast)**

- $L \sim N(\mu_L, \sigma_L^2)$
  - $\mu_L = c_1\mu_1 + \cdots + c_p\mu_p$
  - $\sigma_L^2 = Var(c_1 \overline{Y}_{1.} + \cdots + c_p \overline{Y}_{p.}) = c_1^2 \frac{\sigma^2}{n_1} + \cdots + c_p^2 \frac{\sigma^2}{n_p} = \sigma^2 \sum_i \frac{c_i^2}{n_i}$
  - $\sigma_L^2$의 추정량: $S_L^2 = MSE \sum_i \frac{c_i^2}{n_i}$
- $T = \frac{L - \mu_L}{S_L} = \frac{L - \mu_L}{\sqrt{MSE}\sqrt{\sum_i c_i^2 / n_i}} \sim t_{N-p}$
  - 귀무가설: $T_0 = \frac{L}{\sqrt{MSE}\sqrt{\sum_i c_i^2 / n_i}} \sim t_{N-p}$ $\Rightarrow$ $T^2 = \frac{L^2}{MSE \sum_i c_i^2 / n_i} \sim F_{1, N-p}$

## □ 모든 평균치간의 다중비교

### ○ F-검정을 기초로 한 다중비교

$$H_0 \;:\; \mu_i = \mu_j \quad \text{vs} \quad H_1 \;:\; \mu_i \neq \mu_j$$

① Fisher's least significant difference(최소유의차 방법)

- ○ 모든 $i \neq j$인 모든 쌍에 대해 t-검정을 실시
- ○ 귀무가설 하에서 $T_0 = \dfrac{\overline{Y}_{i.} - \overline{Y}_{j.}}{\hat{\sigma}\sqrt{\dfrac{1}{n_i} + \dfrac{1}{n_j}}} \sim t_{N-p}$
  - $\hat{\sigma} = \sqrt{MSE}$
- ○ $\left|\overline{y}_{i.} - \overline{y}_{j.}\right| > t_{\alpha/2, N-p}\sqrt{MSE}\sqrt{\dfrac{1}{n_i} + \dfrac{1}{n_j}}$ 이면 이 쌍의 차이는 유의함

② Bonferroni 방법

○ $t_{\alpha/2,N-p}$을 $t_{\alpha/2c,N-p}$로 대신

○ $\left|\bar{y}_{i.} - \bar{y}_{j.}\right| > t_{\alpha/2c,N-p}\sqrt{MSE}\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}$ 이면 이 쌍의 차이는 유의함

③ Scheffe 방법

○ F-검정에서 $H_0$을 기각시키지 못하면 Scheffe 방법에 의한 비교에서 유의한 차이가 있는 경우가 없음

○ F-검정에서 $H_0$을 기각시킨 경우에는 최소한 하나의 비교에서 차이가 있는 것으로 나옴

○ $\left|\bar{y}_{i.} - \bar{y}_{j.}\right| > \sqrt{(p-1)F_{\alpha,p-1,N-p}}\sqrt{MSE}\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}$ 이면 이 쌍의 차이는 유의함

## ○ 평균의 범위를 기초로 한 다중비교

$$Q = \sqrt{n}\,[\max(\overline{Y}_{1.},\ ...,\ \overline{Y}_{p.}) - \min(\overline{Y}_{1.},\ ...,\ \overline{Y}_{p.})]/MSE \sim q(p, N-p)$$

- ◦ $q(p, N-p)$는 표준화 범위분포로써 표로 주어짐
- ◦ $Q > q(\alpha, p, N-p)$이면, $H_0$을 기각시킬 수 있음

① Tukey's honest significant difference(HSD)

- ◦ $HSD_{ij} = [q(\alpha, p, N-p)/\sqrt{2}]\sqrt{MSE}\sqrt{1/n_i + 1/n_j}$
- ◦ $\left|\overline{y}_{i.} - \overline{y}_{j.}\right| > HSD_{ij}$ 이면, 두 평균의 차이는 유의적임

## ■ 분산분석 검진(ANOVA Diagnostics)

- **기본가정**: $\varepsilon_{ij} \sim$ iid $N(0,\sigma^2)$ ⇦ 잔차분석(residual analysis)
  - ○ 등분산성
  - ○ 독립성 ⇨ 잔차들 간에는 항상 상관관계가 존재
  - ○ 정규성
  - ○ 이상치 유무
- 잔차
  - ○ $e_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - \overline{Y}_{i.}$
  - ○ studentized 잔차 : $r_{ij} = \dfrac{e_{ij}}{se(e_{ij})}$, $se(e_{ij}) = \sqrt{\dfrac{(n_i - 1)MSE}{n_i}}$
  - ○ studentized deleted 잔차 : $t_{ij} = e_{ij}\left[\dfrac{N - p - 1}{SSE(1 - 1/n_i) - e_{ij}^2}\right]^{1/2}$

# □ 등분산 검정

- 반복수가 같은 경우 동일한 분산을 가진다는 가정을 약간 어기는 경우 분산분석 방법은 robust함

- 반복수가 다르거나 어떤 한 분산이 다른 분산들보다 상당히 큰 경우 분산분석 방법은 robust하지 않음 ⇨ 분산들이 같은지 다른지를 검정필요

- 가설: $H_0 \ : \ \sigma_1^2 = \sigma_2^2 = \ \cdots \ = \sigma_p^2$ VS $H_1$ : 최소한 하나 이상의 분산은 다름

○ Hartley 검정

- 동일 반복수 $n$
- 검정통계량 : $H^{*} = \dfrac{\max(S_i^2)}{\min(S_i^2)} \sim H(p, n-1)$
  - ○ $S_i^2 = \sum (Y_{ij} - \overline{Y}_{i.})^2 / (n_i - 1)$
- 기각역 : $H^{*} > H(\alpha, p, n-1)$

○ Brown-Forsythe 검정

● 절대편차를 먼저 계산

$$D_{ij} = \left| Y_{ij} - \widetilde{Y}_i \right|$$

○ $\widetilde{Y}_i$: $i$ 번째 그룹의 중앙값

● 검정통계량 : $F_{BF}^* = \dfrac{MSTR^*}{MSE^*} \simeq F_{p-1, N-p}$

○ $MSTR^* = \sum n_i (\overline{D}_{i.} - \overline{D}_{..})^2 / (p-1)$, $MSE^* = \sum\sum (D_{ij} - \overline{D}_{i.})^2 / (N-p)$

□ 정규성 검정

- Shapiro-Wilk test, Kolmogorov-Smirnov test, Cramer-von Mises test, Anderson-Darling test
- Jarque-Bera test

$$JB = \frac{n}{6}\left(b_1 + \frac{1}{4}(b_2 - 3)^2\right)$$

  - ○ $\sqrt{b_1}$ : 왜도(skewness)
  - ○ $b_2$: 첨도(kurtosis)

## □ 문제 발생 시 해결방안

- 분산상수화변환(variance stablizing transformation, 분산안정화 변화)
  - ○ 잔차그림에서 잔차의 표준편차(분산)이 $\hat{Y}$의 값과 연관성을 보이는 경우
  - ○ 분산을 상수화시키기 위한 변환을 찾는 방법
    - $\sigma_i^2 = Var(Y_{ij})$와 $\mu_i = E(Y_{ij})$ 사이에 함수관계가 존재하는 경우:

$$\sigma_i^2 = f(\mu_i)$$

  - ○ $\sigma_i^2 = c\mu_i^2$ $(\sigma_i = c\mu_i)$ ⇨ 자연로그변환인 $\log(Y_{ij})$를 이용
  - ○ $\sigma_i^2 = c\mu_i$ $(\sigma_i = \sqrt{\mu_i})$ ⇨ 제곱근변환인 $\sqrt{Y_{ij}}$를 이용

- **비모수적 방법:** 자료의 값 대신 순위(rank)를 사용
  - 자료를 정렬한 후 해당 자료의 순위를 구함
  - tie가 있는 경우 순위의 중간값 사용
  - $SSTO = \sum_i \sum_j \left(R_{ij} - \overline{R}_{..}\right)^2$
  - $SSE = \sum_i \sum_{j1} \left(R_{ij} - \overline{R}_{i.}\right)^2$
  - $SSTR = \sum_i n_i \left(\overline{R}_{i.} - \overline{R}_{..}\right)^2$
  - 검정통계량

$$F_0 = \frac{SSTR/(p-1)}{SSE/(N-p)} = \frac{MSTR}{MSE} \sim F_{p-1, N-p}$$

# ■ 반복이 없는 이원배치법

## ○ 실험 설계

- ◦ 수준 수가 $p$인 요인 A, 수준 수가 $q$인 요인 B
- ◦ $p \times q$ 실험 전체를 완전 확률화

## ○ 자료구조

| 요인 B \ 요인 A | $A_1$ | $A_2$ | $\cdots$ | $A_p$ |
|---|---|---|---|---|
| $B_1$ | $Y_{11}$ | $Y_{21}$ | $\cdots$ | $Y_{p1}$ |
| $B_2$ | $Y_{12}$ | $Y_{22}$ | $\cdots$ | $Y_{p2}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| $B_q$ | $Y_{1q}$ | $Y_{2q}$ | $\cdots$ | $Y_{pq}$ |

## ○ 구조식

- ◦ 1-요인설계의 구조식

$$Y_{ij} = \mu_i + \varepsilon_{ij} = \mu + (\mu_i - \mu) + \varepsilon_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

  - $\tau_i$: 요인의 처리효과

- ◦ 2-요인설계의 구조식

$$\Rightarrow \; Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \qquad i = 1,\ \ldots, p,\ \ j = 1,\ \ldots, q$$

- ◦ $\mu$: 전체 평균, $\alpha_i$: 요인 A의 처리효과, $\beta_j$: 요인 B의 처리효과
- ◦ $\varepsilon_{ij} \sim$ iid $N(0, \sigma^2)$: 오차항
- ◦ 제약조건: $\sum_{i=1}^{p} \alpha_i = 0, \quad \sum_{j=1}^{q} \beta_j = 0$

## ○ 변동의 분해

$$Y_{ij}-\overline{Y}_{..}=(\overline{Y}_{i.}-\overline{Y}_{..})+(\overline{Y}_{.j}-\overline{Y}_{..})+(Y_{ij}-\overline{Y}_{i.}-\overline{Y}_{.j}+\overline{Y}_{..})$$

$$SSTO=SSA+SSB+SSE$$

- $SSTO=\sum_{i=1}^{p}\sum_{j=1}^{q}(Y_{ij}-\overline{Y}_{..})^2=\sum_{i=1}^{p}\sum_{j=1}^{q}Y_{ij}^2-\frac{Y_{..}^2}{N}$ : 자유도 $N-1$
- $SSA=q\sum_{i=1}^{p}(\overline{Y}_{i.}-\overline{Y}_{..})^2=\sum_{i=1}^{p}\frac{Y_{i.}^2}{q}-\frac{Y_{..}^2}{N}$ : 자유도 $p-1$
- $SSB=p\sum_{j=1}^{q}(\overline{Y}_{.j}-\overline{Y}_{..})^2=\sum_{j=1}^{q}\frac{Y_{.j}^2}{p}-\frac{Y_{..}^2}{N}$ : 자유도 $q-1$
- $SSE=\sum_{i=1}^{p}\sum_{j=1}^{q}(Y_{ij}-\overline{Y}_{i.}-\overline{Y}_{.j}+\overline{Y}_{..})^2$ : 자유도 $(p-1)(q-1)$
  - SSE 자유도: $N-(p-1)-(q-1)-1=(p-1)(q-1)$

## ○ 가설 검정

○ 요인 A의 처리 효과의 동일성 검정

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

○ 요인 B의 처리 효과의 동일성 검정

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

○ 분산분석표

| 변인 | 자유도 | 제곱합 | 평균제곱 | F |
|---|---|---|---|---|
| 모형(처리A) | $p-1$ | SSA | MSA | MSA/MSE |
| 모형(처리B) | $q-1$ | SSB | MSB | MSB/MSE |
| 오차 | $(p-1)(q-1)$ | SSE | MSE | |
| 전체 | $N-1$ | SSTO | | |

※ 분산분석에서 요인 처리에 대해 유의한 차이가 없는 것으로 나오면 일원배치법으로 재분석

## ○ $\mu(A_i)$와 $\mu(B_j)$의 구간추정

- ○ $\overline{Y}_{i.} \pm t_{\alpha/2,(p-1)(q-1)} \sqrt{MSE/q}$
- ○ $\overline{Y}_{.j} \pm t_{\alpha/2,(p-1)(q-1)} \sqrt{MSE/p}$

## ■ 반복이 있는 이원배치법

### ○ 실험 설계

- ◦ 수준 수가 $p$인 요인 A, 수준 수가 $q$인 요인 B, 반복 수가 $r$
- ◦ $p \times q \times r$ 실험 전체를 완전 확률화

### ○ 자료구조

| 요인 B \ 요인 A | $A_1$ | $A_2$ | $\cdots$ | $A_p$ |
|---|---|---|---|---|
| $B_1$ | $Y_{111}, \ldots, Y_{11r}$ | $Y_{211}, \ldots, Y_{21r}$ | $\cdots$ | $Y_{p11}, \ldots, Y_{p1r}$ |
| $B_2$ | $Y_{121}, \ldots, Y_{12r}$ | $Y_{221}, \ldots, Y_{22r}$ | $\cdots$ | $Y_{p21}, \ldots, Y_{p2r}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| $B_q$ | $Y_{1q1}, \ldots, Y_{1qr}$ | $Y_{2q1}, \ldots, Y_{2qr}$ | $\cdots$ | $Y_{pq1}, \ldots, Y_{pqr}$ |

## ○ 구조식

$$\begin{aligned} Y_{ijk} &= \mu_{ij} + \varepsilon_{ijk} \\ &= \mu + (\mu_{i.} - \mu) + (\mu_{.j} - \mu) + (\mu_{ij} - \mu_{i.} - \mu_{.j} + \mu) + \varepsilon_{ijk} \end{aligned}$$

⇨ $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$, $i = 1, \ldots, p$, $j = 1, \ldots, q$, $k = 1, \ldots, r$,

- ○ $\mu$: 전체 평균, $\alpha_i$: 요인 A의 처리효과, $\beta_j$: 요인 B의 처리효과
  - \- A와 B를 주효과(main effect)라고 함
- ○ $(\alpha\beta)_{ij}$: 요인 A와 B의 교호작용 효과 (interaction effect)
- ○ $\varepsilon_{ij} \sim$iid $N(0, \sigma^2)$: 오차항
- ○ 제약조건: $\sum_{i=1}^{p} \alpha_i = 0$, $\sum_{j=1}^{q} \beta_j = 0$
  - \- $\sum_{i=1}^{p} (\alpha\beta)_{ij} = 0$, $j = 1, \ldots, q$, $\sum_{j=1}^{q} (\alpha\beta)_{ij} = 0$, $i = 1, \ldots, p$

## ○ 교호작용

- ◦ 요인 B의 수준의 변화에 따라 요인 A의 효과가 변하는 경우 교호작용이 존재한다고 함
- ◦ 교호작용이 존재하지 않을 경우, AB의 최적조건은 A 요인의 최적조건을 구하고 B의 최적조건을 구하여 합함
- ◦ 교호작용이 존재하는 경우, 모든 수준의 조합 $A_iB_j$에서 모평균을 추정함

## ○ 변동의 분해

$$Y_{ijk} - \overline{Y}_{...} = (\overline{Y}_{i..} - \overline{Y}_{...}) + (\overline{Y}_{.j.} - \overline{Y}_{...})$$
$$+ (\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...}) + (Y_{ijk} - \overline{Y}_{ij.})$$

$$SSTO = SSA + SSB + SS(AB) + SSE$$

○ $SSTO = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}(Y_{ijk} - \overline{Y}_{...})^2 = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r} Y_{ijk}^2 - \frac{Y_{...}^2}{N}$

: 자유도=$N-1$

○ $SSA = qr\sum_{i=1}^{p}(\overline{Y}_{i..} - \overline{Y}_{...})^2 = \sum_{i=1}^{p}\frac{Y_{i..}^2}{qr} - \frac{Y_{...}^2}{N}$ : 자유도=$p-1$

○ $SSB = pr\sum_{j=1}^{q}(\overline{Y}_{.j.} - \overline{Y}_{...})^2 = \sum_{j=1}^{q}\frac{Y_{.j.}^2}{pr} - \frac{Y_{...}^2}{N}$ : 자유도=$q-1$

○ $SSTR = \sum_{i=1}^{p}\sum_{j=1}^{q}\frac{Y_{ij.}^2}{r} - \frac{Y_{...}^2}{N}$ : 자유도=$pq-1$

○ $SS(AB) = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}(\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...})^2 = SSTR - SSA - SSB$

- 자유도: $pq - 1 - (p-1) - (q-1) = (p-1)(q-1)$

○ $SSE = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}(Y_{ijk} - \overline{Y}_{ij.})^2$ : 자유도 $pq(r-1)$

- 자유도: $N - 1 - (pq-1) = pq(r-1)$

## ○ 가설 검정

- ◦ 요인 A의 처리 효과의 동일성 검정

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

- ◦ 요인 B의 처리 효과의 동일성 검정

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

- ◦ 교호작용의 효과

$$H_0 : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \ \cdots \ = (\alpha\beta)_{pq} = 0$$

- ◦ 모든 처리의 동일성 검정

$H_0$ : 모든 $\mu_{ij}$들이 같다.

○ 분산분석표

| 변인 | | 자유도 | 제곱합 | 평균제곱 | F |
|---|---|---|---|---|---|
| 처리 | | $pq-1$ | SSTR | MSTR | MSTR/MSE |
| | 처리 A | $p-1$ | SSA | MSA | MSA/MSE |
| | 처리 B | $q-1$ | SSB | MSB | MSB/MSE |
| | 교호작용 | $(p-1)(q-1)$ | SS(AB) | MS(AB) | MS(AB)/MSE |
| 오차 | | $pq(r-1)$ | SSE | MSE | |
| 전체 | | $N-1$ | SSTO | | |

- 교호작용효과가 유의하면 주효과가 유의하지 않더라도 주효과를 모형에서 생략하지 않음

| A요인 주효과 | B요인과의 교호작용효과 | A요인의 효과 |
|---|---|---|
| 있음 | 있음 | 있음 |
| 있음 | 없음 | 있음 |
| 없음 | 있음 | 있음 |
| 없음 | 없음 | 없음 |

○ B요인의 적어도 한 수준에서 A요인의 효과가 있으면 A요인은 효과 있음

○ B요인의 모든 수준에서 A요인의 효과가 없으면 A요인은 효과 없음

## ○ 분산분석후의 추정 (모수 모형)

○ $\mu(A_i)$와 $\mu(B_j)$의 구간추정

- $\overline{Y}_{i..} \pm t_{\alpha/2, pq(r-1)} \sqrt{MSE/qr}$
- $\overline{Y}_{.j.} \pm t_{\alpha/2, pq(r-1)} \sqrt{MSE/pr}$

○ $\mu(A_iB_j)$의 구간추정

- $\overline{Y}_{ij.} \pm t_{\alpha/2, pq(r-1)} \sqrt{MSE/r}$

※ 교호작용이 유의한 경우, 일반적으로 요인 A, B의 각 수준의 모평균을 추정하는 것은 의미가 없으며 수준의 조합 $A_iB_j$에서 모평균을 추정하는 것이 의미가 있음

## ○ 교호작용이 있는 경우 다중비교

- $H_0: \mu_{ij} = \mu_{kl}$ vs $H_1: \mu_{ij} \neq \mu_{kl}$ 또는 $\mu_{ij} - \mu_{kl}$의 신뢰구간
- $\overline{Y}_{ij.} - \overline{Y}_{kl.} \pm c\sqrt{MSE}\sqrt{1/r + 1/r}$
  - 최소유의차: $c = t_{\alpha/2, pq(r-1)}$
  - Bonferroni: $c = t_{\alpha/(2a), pq(r-1)}$, $a =$ 비교검정의 경우의 수
  - Scheffe: $c = \sqrt{(pq-1)F_{\alpha, pq-1, pq(r-1)}}$
  - Tukey: $\dfrac{1}{\sqrt{2}} q_{\alpha, pq, pq(r-1)}$