■ 반복측정 분산분석(Repeated Measures ANOVA)

- 한 실험개체에서 2회 이상 측정을 수행하는 것

- 반복측정의 장점
  - 어떤 실험에서는 충분한 수의 실험개체를 확보하기 어려울 때 적용
  - 반복측정에서는 실험개체 스스로가 대조(control)의 역할을 함
    ⇨ 실험개체가 블록이 됨

- 반복측정의 단점
  - 각 개체를 여러 번 실험처리하기 때문에 시간이 많이 소요
  - 이월효과(carry-over effect, 잔류효과)가 발생할 수 있음

□ 단일요인 반복측정 분산분석

● 반복측정 자료의 예

개체	처리 1	처리 2	처리 3	처리 4
1	30	28	16	34
2	14	18	10	22
3	24	20	18	30
4	38	34	20	44
5	26	28	14	30
평균	26.4	25.6	15.6	32.0

○ 일변량 분산분석

- 통계모형

$$Y_{ij} = \mu + \rho_i + \tau_j + \varepsilon_{ij}, \quad i = 1, \ ..., n, \ j = 1, \ ..., p$$

○ ρ_i: 개체 i의 효과(subject effect)

○ 동일한 개체의 자료들 간 상관관계가 존재할 수 있음

⇨ Y_{ij}와 Y_{ik}는 독립이라고 보기 어려움

⇒ ρ_i는 랜덤효과(변량효과): $\rho_i \sim N(0, \sigma_s^2)$

- 변동분해

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

TSS SSB SSW

  - SSB(SS due to Between subject): 개체 간 제곱합
  - SSW(SS due to Within subject): 개체 내 제곱합
  - 개체 내 제곱합 분해

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

SSW SSTR SSE

  - EMS
    - $E(MSB)=\sigma^2+p\sigma_s^2$
    - $E(MSTR)=\sigma^2+n\sum\tau_j^2/(p-1)$
    - $E(MSE)=\sigma^2$

- 가설검정
  - $H_0 : \tau_1 = \tau_2 = \cdots = \tau_p = 0$
  - $F = \dfrac{\dfrac{SSTR}{p-1}}{\dfrac{SSE}{(n-1)(p-1)}} = \dfrac{MSTR}{MSE} \sim F_{p-1,(n-1)(p-1)}$

⇨ 개체 간 변동에 영향을 받지 않고 순수하게 개체 내 변동만 평가하기 때문에 처리평균 간 비교를 정밀화할 수 있음

○ 다변량 분산분석과 자유도 수정

- 다변량적 접근방법

$$\boldsymbol{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{ip})^T, \quad i = 1, \ldots, n$$

  ○ $\boldsymbol{Y}_i \sim N_p(\boldsymbol{\mu}, \Sigma), \ \boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)^T$

- 공분산행렬

  ○ 복합대칭성(compound symmetry)

$$\sigma_{jk} = Cov(Y_{ij}, Y_{ik}) = \begin{cases} \sigma_s^2 + \sigma^2, & j = k \\ \sigma_s^2, & j \neq k \end{cases}$$

  - 일변량적 접근방법은 복합대칭성 또는 구형성(sphericity)과 같은 특수한 조건을 만족하는 경우 타당

  ○ 복합대칭성 또는 구형성 가정을 위배하는 경우 다변량 검정 또는 자유도 수정이 필요

- 다변량 분산분석(MANOVA, Multivariate ANOVA)
  - $H_0 : \mu_1 = \mu_2 = \cdots = \mu_p$
  - $\mu_j = \mu + \tau_j$ ⇨ $H_0 : \boldsymbol{\tau} = \mathbf{0}$

- Hotelling's T^2 통계량
  - $\boldsymbol{W}_i = (W_{i1}, \ldots, W_{i,p-1})^T \equiv (Y_{i1} - Y_{ip}, \ldots, Y_{i,p-1} - Y_{ip})^T$
  - $H_0 : \boldsymbol{\tau} = \mathbf{0}$ ⇨ $H_0 : \boldsymbol{\mu}_W = \mathbf{0}$
  - 검정통계량 : $T^2 = n \overline{W}^T \boldsymbol{S}_W^{-1} \overline{W}$

$$F_t = \frac{n-p+1}{(n-1)(p-1)} T^2 \sim F_{p-1, n-p+1}$$

    - $n \geq p$인 경우에만 적용가능
    - 단변량 분모의 자유도 = $(n-1)(p-1)$
    - 다변량 분모의 자유도 = $n-p+1$ < $(n-1)(p-1)$

- 자유도 수정방법
  - ○ 다변량적 방법에서는 공분산행렬에 대한 가정이 없음
    ⇨ 구형성 가정을 어느 정도 만족하는 경우 자유도 손실이 큼
  - ○ 자유도 수정계수 ϵ 이용
    - $\epsilon \in [0, 1]$
    - $\epsilon \to 1$ 구형성 만족
    - 복합대칭성(구형성) 가정 하에서 $F \sim F_{p-1,(n-1)(p-1)}$
    - 실제 F는 $F_{\epsilon(p-1),\epsilon(n-1)(p-1)}$에 근사
  - ○ ϵ의 추정
    - Greenhouse and Geisser ϵ_{GG}, Huynh and Feldt ϵ_{HF}
    - $\epsilon_{GG} \le \epsilon_{HF}$, $\epsilon_{HF} = \min(\epsilon_{HF}, 1)$
  - ○ ϵ이 충분히 크면 일변량적 방법 이용

□ 다요인 반복측정 분산분석

○ 반복요인과 분류요인이 하나씩인 경우

◉ 반복측정 2요인 실험자료: 반복요인이 B인 경우

요인 A	개체	요인 B				평균
		1	2	3	4	
1	1	0	0	5	3	2.00
	2	3	1	5	4	4.25
	3	4	3	6	2	4.75
2	1	4	2	7	8	5.20
	2	5	4	6	6	5.20
	3	7	5	8	9	5.80
평균		3.83	2.50	6.17	5.33	4.46

○ 개체(subject)는 일종의 블록요인이고 요인 A에 지분되어(nested) 있음

- 통계모형식

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + s_{i(j)} + \varepsilon_{ijk},$$
$$i = 1, ..., s,\ j = 1,\ ..., a,\ k = 1,\ ..., b$$

  - $s_{i(j)}$: 개체효과로 일반적으로 랜덤(변량)요인
  - $E(Y_{ijk}) = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk}$
  - $Var(Y_{ijk}) = \sigma^2 + \sigma_s^2$
  - $Cov(Y_{ijk}, Y_{ijk'}) = \sigma_s^2, \quad k \neq k'$
  - $Cov(Y_{ijk}, Y_{i'j'k'}) = 0, \quad i \neq i$ or/and $j \neq j'$

- 변동분해

$$TSS = \underbrace{SSA + SS_{subject(A)}}_{SS_{between}} + \underbrace{SSB + SS(AB) + SSE}_{SS_{within}}$$

○ $SSA = bs\sum_{j}(\overline{Y}_{.j.} - \overline{Y}_{...})^2 \ : \ a-1 \ \Rightarrow \ E(MSA) = \sigma^2 + b\sigma_s^2 + bs\sum \alpha_j^2/(a-1)$

○ $SSB = as\sum_{k}(\overline{Y}_{..k} - \overline{Y}_{...})^2 \ : \ b-1 \ \Rightarrow \ E(MSB) = \sigma^2 + as\sum \beta_k^2/(b-1)$

○ $SS(AB) = s\sum_{j}\sum_{k}(\overline{Y}_{.jk} - \overline{Y}_{.j.} - \overline{Y}_{..k} + \overline{Y}_{...})^2 \ : \ (a-1)(b-1)$

$\Rightarrow \ E(MS(AB)) = \sigma^2 + s\sum\sum(\alpha\beta)_{jk}^2/((a-1)(b-1))$

○ $SSS(A) = b\sum_{i}\sum_{j}(\overline{Y}_{ij.} - \overline{Y}_{.j.})^2 \ : \ a(s-1) \ \Rightarrow \ E(MSS(A)) = \sigma^2 + b\sigma_s^2$

○ $SSE = \sum_{i}\sum_{j}\sum_{k}(Y_{ijk} - \overline{Y}_{.jk} - \overline{Y}_{ij.} + \overline{Y}_{.j.})^2 \ : \ a(s-1)(b-1) \ \Rightarrow \ E(MSE) = \sigma^2$

- 유의성검정
  - 요인 A의 주효과: $H_0 : \alpha_1 = \cdots = \alpha_a = 0$

$$F_A = \frac{SSA/(a-1)}{SS_{subject(A)}/(a(n-1))} \sim F_{a-1,a(n-1)}$$

  - 요인 B의 주효과: $H_0 : \beta_1 = \cdots = \beta_b = 0$

$$F_B = \frac{SSB/(b-1)}{SSE/(a(b-1)(n-1))} \sim F_{b-1,a(b-1)(n-1)}$$

  - 상호작용 (A*B)의 효과: $H_0 : (\alpha\beta)_{ij} = 0, \quad \forall i, j$

$$F_{(AB)} = \frac{SS(AB)/(a-1)(b-1)}{SSE/(a(b-1)(n-1))} \sim F_{(a-1)(b-1),a(b-1)(n-1)}$$