실험계획과 분석

심송용(한림대학교 데이터과학스쿨)

http://jupiter.hallym.ac.kr

임의효과와 고정효과

실험에 사용되는 요인들 중에는

- 몇가지 고정된 수준만 가지고 있거나, 다양한 수준이 있더라도 실험의 목적이 특정한 수준들에서의 차이에 대한 추론인 경우도 있고,
- 수준의 수가 아주 많거나 경우에 따라 무한대인 경우 많은 수준들 중에서 일부 수준을 임의로 선택하여 실험을 한 후 그 결과를 모든 수준에서의 차이로 확장해서 해석

하는 두 가지 경우가 있다.

이 두 가지의 수준 구성에서

- 전자는 실험의 수준이 고정되어 있어서 해당 요인을 **고정효과(fixed effect)**라고 하고
- 후자는 많은 가능한 수준들 중 (임의로) 선택된 일부 수준에서 실험을 하므로 **임의효과 (random effect)**

라고 한다.

임의효과와 고정효과

요인이 임의효과인지 고정효과인지는 임의효과인 경우와 고정효과인 경우 **평균제곱의 기댓값이 달라져서** 가설검정에서 사용할 검정통계량 F_0의 계산이 달라지기 때문임.

보기:
성별(남, 여)에 따른 차이: (   )
자동차 구동시스템에 따른 차이(전륜구동, 후륜구동, 4륜구동): (   )
온도(30C, 60C, 90C)에 따른 차이: (   )
시중에 판매되는 비료에 따른 수확량 차이 (   )

지금까지 살펴본 일원배치, 이원배치 모형은 모두 요인이 고정효과인 경우임.

임의효과와 고정효과

임의효과 일원배치법

$y_{ij} = \mu + \tau_i + \epsilon_{ij}$,

$\underline{\tau_i \sim N(0, \sigma_\tau^2)\ \text{이고 독립}\ (i = 1, 2, \ldots, a)}$

$\epsilon_{ij} \sim N(0, \sigma^2)$ 이고 독립 $(j = 1, 2, \ldots, n_i)$

τ_i와 ϵ_{ij}들도 독립

요인에 의한 효과가 차이여부는

$$\boxed{H_0 : \sigma_\tau^2 = 0\ (\text{차이 없음})\ \text{대}\ H_1 : \sigma_\tau^2 \neq 0}$$

의 검정임. (분산이 0이면 모든 값이 평균(=0)과 같음. 즉 τ_i 들이 모두 같음)

제곱합의 분해, 자유도 ⇨ 고정모형일 때와 동일

분산분석표 ⇨ 고정모형일 때와 동일.

임의효과와 고정효과

임의효과 일원배치 분산분석표

요인	제곱합	자유도	평균제곱	F	유의확률
처리	$\mathrm{SSTrt} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(\bar{y}_{i.} - \bar{y}_{..})^2$	$a-1$	$\mathrm{MSTrt} = \frac{\mathrm{SSTrt}}{a-1}$	$F_0 = \frac{\mathrm{MSTrt}}{\mathrm{MSE}}$	$P = \Pr[F_{a-1,N-a} > F_0]$
오차	$\mathrm{SSE} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2$	$N-a$	$\mathrm{MSE} = \frac{\mathrm{SSE}}{N-a}$		
전체	$\mathrm{SST} = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2$	$N-1$			

평균제곱의 기댓값($n_1 = n_2 = \cdots = n_a = n$ 일 때)

$E(MSTrt) = \sigma^2 + n\sigma_\tau^2$

$E(MSE) = \sigma^2$

임의효과와 고정효과

귀무가설 $H_0 : \sigma_\tau^2 = 0$ 이 참이면 $F_0 = \dfrac{MSTrt}{MSE}$의 분자의 기댓값, 분모의 기댓값이 모두 σ^2

귀무가설 $H_0 : \sigma_\tau^2 = 0$ 이 거짓이면 $F_0 = \dfrac{MSTrt}{MSE}$의 분자의 기댓값이 분모의 기댓값보다 커서 검정통계량이 클 가능성.

이원배치법의 경우 요인이 두 개이므로

- 둘다 고정효과 ⇨ 고정효과모형(fixed effect model; 앞에서 다룸)
- 둘다 임의효과 ⇨ 임의효과모형(random effect model)
- 둘중 하나만 임의효과 ⇨ 혼합(효과)모형(mixed effect model)

임의효과와 고정효과

임의효과모형

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk},$$

$i = 1,2,\ldots,a;\ j = 1,2,\ldots,b;\ k = 1,2,\ldots,k$

$\alpha_i \sim N(0,\sigma_\alpha^2)$ 이고 독립

$\beta_j \sim N(0,\sigma_\beta^2)$ 이고 독립

$(\alpha\beta)_{ij} \sim N(0,\sigma_{\alpha\beta}^2)$ 이고 독립

$\epsilon_{ijk} \sim N(0,\sigma^2)$ 이고 모두 독립

$\alpha_i,\ \beta_j,\ (\alpha\beta)_{ij},\ \epsilon_{ijk}$ 도 모두 독립

효과의 유의성에 대한 가설

- 주효과 A의 유의성 검정: 귀무가설 $H_0 : \sigma_\alpha^2 = 0$ 대 대립가설 $H_1 : \sigma_\alpha^2 \neq 0$
- 주효과 B의 유의성 검정: 귀무가설 $H_0 : \sigma_\beta^2 = 0$ 대 대립가설 $H_1 : \sigma_\beta^2 \neq 0$
- 상호작용의 유의성 검정: 귀무가설 $H_0 : \sigma_{\alpha\beta}^2 = 0$ 대 대립가설 $H_1 : \sigma_{\alpha\beta}^2 \neq 0$

임의효과와 고정효과

임의효과모형의 제곱합의 분해, 자유도

제곱합의 분해, 각 제곱합의 자유도는 고정효과모형과 동일

임의효과모형인 경우 평균제곱의 기댓값

$$E(MSA) = \sigma^2 + n\sigma^2_{\alpha\beta} + bn\sigma^2_{\alpha}$$

$$E(MSB) = \sigma^2 + n\sigma^2_{\alpha\beta} + an\sigma^2_{\beta}$$

$$E(MSAB) = \sigma^2 + n\sigma^2_{\alpha\beta}$$

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

각 가설검정의 검정

귀무가설이 참일 때 두 평균제곱의 비율인 F 값의 기댓값이 분자 분모가 같은 값이 갖도록 함.

임의효과와 고정효과

혼합모형의 분산분석표

요인	제곱합	자유도	평균제곱(MS)	F	유의확률
A	SSA	$a-1$	$\mathrm{MSA}=\dfrac{\mathrm{SSA}}{a-1}$	$F_0=\dfrac{\mathrm{MSA}}{\mathrm{MSAB}}$	$\Pr[F_{a-1,(a-1)(b-1)} > F_0]$
B	SSB	$b-1$	$\mathrm{MSB}=\dfrac{\mathrm{SSB}}{b-1}$	$F_0=\dfrac{\mathrm{MSB}}{\mathrm{MSAB}}$	$\Pr[F_{b-1,(a-1)(b-1)} > F_0]$
A×B	SSAB	$(a-1)(b-1)$	$\mathrm{MSAB}=\dfrac{\mathrm{SSAB}}{(a-1)(b-1)}$	$F_0=\dfrac{\mathrm{MSAB}}{\mathrm{MSE}}$	$\Pr[F_{(a-1)(b-1),ab(n-1)} > F_0]$
오차	SSE	$ab(n-1)$	$\mathrm{MSE}=\dfrac{\mathrm{SSE}}{ab(n-1)}$		
전체	SST	$abn-1$			

- A의 주효과 검정: 귀무가설 $H_0 : \sigma_\alpha^2 = 0$ 대 대립가설 $H_1 : \sigma_\alpha^2 \neq 0$

  $F_0 = \dfrac{\mathrm{MSA}}{\mathrm{MSAB}} > F_{a-1,(a-1)(b-1);\alpha}$ 이면 H_0 기각

임의효과와 고정효과

- B의 주효과 검정: 귀무가설 $H_0 : \sigma_\beta^2 = 0$ 대 대립가설 $H_1 : \sigma_\beta^2 \neq 0$

$$F_0 = \frac{\mathrm{MSB}}{\mathrm{MSAB}} > F_{b-1,(a-1)(b-1);\alpha}$$ 이면 H_0 기각

- 상호작용의 유의성 검정: 귀무가설 $H_0 : \sigma_{\alpha\beta}^2 = 0$ 대 대립가설 $H_1 : \sigma_{\alpha\beta}^2 \neq 0$

$$F_0 = \frac{\mathrm{MSAB}}{\mathrm{MSE}} > F_{(a-1)(b-1),ab(n-1);\alpha}$$ 이면 H_0 기각

임의효과와 고정효과

혼합효과모형(요인 A는 고정, 요인 B는 임의 효과)

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk},$$

$$i = 1,2,...,a;\ j = 1,2,...,b;\ k = 1,2,...,k$$

$$\sum_{i=1}^{a} \alpha_i = 0$$

$\beta_j \sim N(0, \sigma_\beta^2)$ 이고 독립

$(\alpha\beta)_{ij} \sim N(0, \sigma_{\alpha\beta}^2)$ 이고 독립

$\epsilon_{ijk} \sim N(0, \sigma^2)$ 이고 모두 독립

$\beta_j,\ (\alpha\beta)_{ij},\ \epsilon_{ijk}$ 도 모두 독립

효과의 유의성에 대한 가설

- 주효과 A의 유의성 검정: 귀무가설 $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$ 대 대립가설 $H_1 :$ 적어도 하나의 $\alpha_i \neq 0$

임의효과와 고정효과

- 주효과 B의 유의성 검정: 귀무가설 $H_0 : \sigma_\beta^2 = 0$ 대 대립가설 $H_1 : \sigma_\beta^2 \neq 0$
- 상호작용의 유의성 검정: 귀무가설 $H_0 : \sigma_{\alpha\beta}^2 = 0$ 대 대립가설 $H_1 : \sigma_{\alpha\beta}^2 \neq 0$

혼합모형의 제곱합의 분해, 자유도

제곱합의 분해, 각 제곱합의 자유도는 고정효과모형과 동일

혼합모형인 경우 평균제곱의 기댓값

$$E(MSA) = \sigma^2 + n\sigma_{\alpha\beta}^2 + \frac{bn}{a-1}\sum_{i=1}^{a}\alpha_i^2$$

$$E(MSB) = \sigma^2 + n\sigma_{\alpha\beta}^2 + an\sigma_\beta^2$$

$$E(MSAB) = \sigma^2 + n\sigma_{\alpha\beta}^2$$

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

임의효과와 고정효과

각 가설검정의 검정

귀무가설이 참일 때 두 평균제곱의 비율인 F 값의 기댓값이 분자 분모가 같은 값이 갖도록 함.

혼합모형의 분산분석표

요인	제곱합	자유도	평균제곱(MS)	F	유의확률
A	SSA	$a-1$	$\mathrm{MSA}=\dfrac{\mathrm{SSA}}{a-1}$	$F_0=\dfrac{\mathrm{MSA}}{\mathrm{MSAB}}$	$\Pr[F_{a-1,(a-1)(b-1)}>F_0]$
B	SSB	$b-1$	$\mathrm{MSB}=\dfrac{\mathrm{SSB}}{b-1}$	$F_0=\dfrac{\mathrm{MSB}}{\mathrm{MSAB}}$	$\Pr[F_{b-1,(a-1)(b-1)}>F_0]$
A×B	SSAB	$(a-1)(b-1)$	$\mathrm{MSAB}=\dfrac{\mathrm{SSAB}}{(a-1)(b-1)}$	$F_0=\dfrac{\mathrm{MSAB}}{\mathrm{MSE}}$	$\Pr[F_{(a-1)(b-1),ab(n-1)}>F_0]$
오차	SSE	$ab(n-1)$	$\mathrm{MSE}=\dfrac{\mathrm{SSE}}{ab(n-1)}$		
전체	SST	$abn-1$			

임의효과와 고정효과

- A의 주효과 검정: 귀무가설 $H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$ 대 대립가설 H_1 : 적어도 하나의 $\alpha_i \neq 0$

  $F_0 = \dfrac{\text{MSA}}{\text{MSAB}} > F_{a-1,(a-1)(b-1);\alpha}$ 이면 H_0 기각

- B의 주효과 검정: 귀무가설 $H_0 : \sigma_\beta^2 = 0$ 대 대립가설 $H_1 : \sigma_\beta^2 \neq 0$

  $F_0 = \dfrac{\text{MSB}}{\text{MSAB}} > F_{b-1,(a-1)(b-1);\alpha}$ 이면 H_0 기각

- 상호작용의 유의성 검정: 귀무가설 $H_0 : \sigma_{\alpha\beta}^2 = 0$ 대 대립가설 $H_1 : \sigma_{\alpha\beta}^2 \neq 0$

  $F_0 = \dfrac{\text{MSAB}}{\text{MSE}} > F_{(a-1)(b-1),ab(n-1);\alpha}$ 이면 H_0 기각

혼합모형일 때와 임의효과 모형일 때 이론적으로 MSA의 기댓값은 다르지만 분산분석표상으로 얻어지는 결과는 동일