

13강. LMM 3-수준 군집자료분석

■ 주요용어

| 용어 | 해설 |
|--|--|
| 상향식 모형 구축 | 모형을 구축할 때 변량요인과 상수항만 있는 모형에서 출발하여 1-수준 공변량의 고정효과 모형을 선택하고 다음으로 2-수준 공변량 또는 1-수준 변량계수 모형을 선택하는 순서로 모형을 구축해 가는 방법이다. |
| 분산성분모형 (variance component models) | 상수항과 수준 별 변량요인만 있는 모형. 3-수준 군집자료에서 ‘mean-only models’ 또는 ‘unconditional models’이라고도 한다. |
| 변량효과의 유의성 검정 | <p>변량효과는 확률변수이므로 변량효과의 유의성 검정은 변량효과 분산(σ_u^2)이 0인지를 검정한다. 즉,</p> $H_0 : \sigma_u^2 = 0 \text{ vs. } H_1 : \sigma_u^2 > 0.$ |
| 급 내 상관계수 (ICC:intraclass correlation coefficient) | <p>군집 내 반응변수들 간의 유사성 또는 동질성을 나타내는 척도로 사용된다. Y_{ijk}는 k번째 학교의 j번째 학급에서의 i번째 학생의 종속변수이다.</p> <p>① 학교 내 상관계수</p> $ICC_{sch} = Corr(Y_{ijk}, Y_{i'j'k}) = \frac{Cov(Y_{ijk}, Y_{i'j'k})}{Var(Y_{ijk})}$ <p>여기서 $i \neq i', j \neq j'$</p> <p>② 학급 내 상관계수</p> $ICC_{cl} = Corr(Y_{ijk}, Y_{i'jk}) = \frac{Cov(Y_{ijk}, Y_{i'jk})}{Var(Y_{ijk})}$ <p>여기서 $i \neq i'$</p> |

정리하기

■ 요약하기

1. 자료의 특성 파악하기: 기술통계량 및 그래프 상용
2. 모형의 구축
 - 2.1 상향식(Step-up) 모형 구축
 - ▷ Step1: ‘mean-only’ 모형 적합
 - 고정효과는 “상수항”만 포함
 - 2-수준 및 3-수준 변량요인(각각 학급, 학교) 포함
 - 분산성분모형(variance components models)
 - 3-수준 변량절편모형
 - 변량요인에 대한 유의성 검정
 - ▷ Step2
 - Step1에서 선택된 모형에 Level 1 공변량 추가
 - Step2에서 모형 검정의 내용: Level 1 공변량 모수에 대한 검정(가능도비검정, F-검정, t-검정)의 유의성
 - ▷ Step3
 - Step2에서 선택된 모형에 Level 2 공변량 추가
 - Step3에서 모형 검정의 내용: Level 2 공변량 모수에 대한 검정(가능도비검정, F-검정, t-검정)의 유의성 등
 - ▷ Step4
 - Step3에서 선택된 모형에 Level 3 공변량 추가
 - Step4에서 모형 검정의 내용: Level 3 공변량 모수에 대한 검정(가능도비검정, F-검정, t-검정)의 유의성 등

2.2 3-수준 변량절편모형

상향식 모형구축을 통하여 아래 모형 선택.

$$Y_{ijk} = \beta_0 + \beta_1^* x_{1ijk} + \beta_2^* x_{2ijk} + \beta_3^* x_{3ijk} + \beta_4^* x_{4ijk} + u_{j|k} + u_k + \epsilon_{ijk}$$

여기서, $u_{j|k} \sim^{iid} N(0, \sigma_{cl}^2)$, $u_k \sim^{iid} N(0, \sigma_{sch}^2)$, $\epsilon_{ijk} \sim^{iid} N(0, \sigma^2)$ 이고 $u_{j|k}$, u_k , ϵ_{ijk} 는 서로 독립.

3. 모수 추정치의 해석

3.1 고정효과 모수의 해석

$$\begin{aligned} E(\widehat{Y}_{ijk}) &= \hat{\beta}_0 + \hat{\beta}_1 * x_{1ijk} + \hat{\beta}_2 * x_{2ijk} + \hat{\beta}_3 * x_{3ijk} + \hat{\beta}_4 * x_{4ijk} \\ &= 282.79 - 1.25 * x_{1ijk} - 8.26 * x_{2ijk} - 0.47 * x_{3ijk} + 5.35 * x_{4ijk} \end{aligned}$$

▷ 고정효과 모수추정치(REML추정방법 사용)와 해석

① “사회경제적상태(ses: x_{4ijk})” 추정치 해석

- “나머지 공변량들(sex, minority, mathkind)이 보정되었을 때 ses가 1단위 높을 수록 수학적성취점수는 평균적으로 5.35 (SE=1.24) 높다 (p<0.0001).”

② 나머지 공변량의 계수 추정치에 대해서도 유사한 방법으로 해석 가능

3.2 공분산 모수의 해석

▷ 급내상관계수(ICC: intraclass correlation coefficient)

- 분산성분모형

$$Y_{ijk} = \beta_0 + u_{jlk} + u_k + \epsilon_{ijk}$$

여기서, $u_{jlk} \sim^{iid} N(0, \sigma_{cl}^2)$, $u_k \sim^{iid} N(0, \sigma_{sch}^2)$, $\epsilon_{ijk} \sim^{iid} N(0, \sigma^2)$ 이고 u_{jlk} , u_k , ϵ_{ijk} 는 서로 독립.

- 학교 내 상관계수

$$\widehat{ICC}_{sch} = \widehat{Corr}(\widehat{Y}_{ijk}, \widehat{Y}_{i'j'k}) = \frac{\widehat{\sigma}_{sch}^2}{\widehat{\sigma}^2 + \widehat{\sigma}_{cl}^2 + \widehat{\sigma}_{sch}^2} = \frac{77.44}{1028.49 + 99.20 + 77.44} = 0.06$$

- 학급 내 상관계수

$$\widehat{ICC}_{sch} = \widehat{Corr}(\widehat{Y}_{ijk}, \widehat{Y}_{i'j'k}) = \frac{\widehat{\sigma}_{cl}^2 + \widehat{\sigma}_{sch}^2}{\widehat{\sigma}^2 + \widehat{\sigma}_{cl}^2 + \widehat{\sigma}_{sch}^2} = \frac{99.20 + 77.44}{1028.49 + 99.20 + 77.44} = 0.15$$

▷ 분산성분모형에 Level 1 공변량을추가하면

- 오차분산은 $1028.49 \Rightarrow 734.41$ (29%↓)

- 학교 내 학급분산 $99.20 \Rightarrow 83.36$ (16%↓)

- 학교 분산 $77.44 \Rightarrow 75.17$ (3%↓)

감소. 즉 Level 1 공변량, 자료 변동을 효과적으로 설명

4. 모형진단: 잔차분석

▷ 조건부 잔차를 이용한 오차분포 가정에 대한 검토

- 양쪽(특히 왼쪽)긴 꼬리를 가지는 분포

- 정규분포에서의 이탈 시사

- 이들에 대한 추가적인 탐구 시사

√ 변량요인(예: 변량계수 등)추가 검토

- √ 오차분산구조(예: 이분산등)탐구
- ▷ Loess 곡선 큰 예측값에서 증가하는 경향
 - 큰(작은)예측값을 설명할 수 있는 공변량 또는 변량요인의 결여 시사
- √ 각 공변량별로 산점도탐구
 - ▶ 비선형항 검토
 - ▶ 변량계수모형 검토

참고하기

교재

1. Linear Mixed Models A practical Guide Using Statistics Software, 3rd ed. West, BT., Welch, KB., Galecki, AT., 2022, Chapman & Hall/CRC

교재 사이트

2. <http://www-personal.umich.edu/~bwest/almmussp.html>

| | |
|--|-------------|
| | 과제하기 |
|--|-------------|

| 구분 | 내용 |
|---------|--|
| 과제 주제 | <p>1. 상향식 모형(step-up)구축 전략을 각 단계별로 설명하시오.</p> <p>2. ‘mathknow’ 공변량에 결측치가 없다고 가정하자. 모형4.3(Level 1, Level 2 공변량 포함 모형)에서 Level 2 공변량의 유의성을 검정하는 가설을 설정하고 가능도비 검정 방법을 제시하시오.</p> <p>3. 최종분석모형(모형4.2) 적합 결과로부터 ‘성별’, ‘mathkind’, ‘minority’ 가 수학성취도에 미치는 영향을 해석하시오,</p> <p>4. Level 1 공변량을 보정했을 때 ‘학교 내 상관계수’와 ‘학급 내 상관계수’를 각각 구하시오. 이들을 분산성분모형에서의 그것들과 비교하시오.</p> |
| 목적 | 13주차 강의 내용을 복습하고, 3-수준 변량절편 모형의 구축과 모형의 선택에 대한 이해를 심화 하기 위함. 또한 모수 추정치의 의미와 설명법에 대한 이해도를 높이기 위함. |
| 제출 기간 | 13주차 강의 후 1주 후 토요일 밤 10시까지 |
| 참고 자료 | 교재와 강의자료를 참고하기 바람 |
| 기타 유의사항 | |