

통계학 특강: 과제 2번 (10월 11일까지 제출)

서울시립대 통계학과

October 11, 2019

동대문구에서 4개의 초등학교($i = 1, 2, 3, 4$)를 임의로 선택하여 2개의 학교에서는 1학년 학생들에게 A 교수법을 이용하여 글쓰기를 가르치고 나머지 2개의 학교에서는 B교수법을 사용하여 글쓰기를 가르쳤다. 교수법의 효과를 비교하기 위하여 각 학교에서 10명의 학생들($j = 1, 2, \dots, 10$)을 임의로 추출하여 시험을 보았다. 시험성적의 결과는 화일 `hw-02-data.csv`에 주어져 있다.

글쓰기 교육에서 얻은 시험점수 y_{ij} 에 대한 모형으로 다음과 같은 일원배치 임의효과모형(random effects model)을 고려하자. 여기서 $\tau_{(i)}$ 는 교수법을 나타내는 고정효과로서 A교수법을 받으면 0, B 교수법을 받으면 τ 로 나타난다. A_i 는 학교의 효과를 나타내는 임의효과이며 서로 독립이고 $N(0, \sigma_a^2)$ 를 따른다. 또한 오차항 e_{ij} 는 서로 독립이며 $N(0, \sigma_e^2)$ 를 따른다.

$$y_{ij} = \mu + \tau_{(i)} + A_i + e_{ij}, \quad i = 1, 2, \dots, 4 \text{ and } j = 1, 2, \dots, 10 \quad (1)$$

1. 시험점수 y_{ij} 의 기대값을 구하시오.
2. 시험점수 y_{ij} 의 분산을 구하시오.
3. 다른 학교에 속한 학생들의 시험점수 y_{ij} 에 대한 상관계수를 구하시오.
4. 같은 학교에 속한 학생들의 시험점수 y_{ij} 에 대한 상관계수를 구하시오.
5. 모형 (1)을 다음과 같은 선형혼합모형으로 나타낼 때 벡터 β , b 와 행렬 X , Z 의 형태를 구하시오.

$$y = X\beta + Zb + e \quad (2)$$

6. 반응값 벡터 y 의 공분산행렬을 구하시오.

7. 자료화일에 주어진 시험점수 자료와 다음 R 프로그램을 이용하여 모형 (1)을 적합시키고 모든 모수의 추정값을 구하시오.

```
library(lme4)
data1 <- read.csv("hw-02-data.csv", header=T)
fm1 <- lmer(score~teach+(1|school), data=data1)
summary(fm1)
```