STAT347: Generalized Linear Models
Lecture 14

Today's topics:

- Examples for linear mixed effect models

1 Two examples for LMM

1.1 Multilevel model for smoking prevention and cessation study (Chapter 9.2.3)

1600 students are collected from 135 classrooms in 28 schools. We want to understand the effect of SC (exposure to a school-based curriculum or not), TV (exposure to a television-based prevention program or not) and previous THK scale on the current THK scale. We have 1600 samples, but some share the same school and some share the same classroom.

The multilevel model:

$$y_{ics} = \beta_0 + \beta_1 \mathrm{PTHK}_{ics} + \beta_2 \mathrm{SC}_{ics} + \beta_3 \mathrm{TV}_{ics} + u_s + v_{cs} + \epsilon_{ics}$$

Please see the R Data example 8, example 1

1.2 Multi-subject, multi-group example

We try to understand the relationship between a student's GPA on his/her test scores.

- Each student has a GPA x_i
- For $j = 1, 2, \cdots, p$th type of exam, student i has a test score y_{ij}

Here are a few related modeling ideas from different perspectives

- Assume that y_{ij} are i.i.d. across students for each exam j

$$y_{ij} = \beta_{0j} + \beta_{1j} x_i + \epsilon_{ij}$$

- To consider the fact that each student can have different ability/background, that affects scores across all of his/her exams, there are two perspectives
  - Each student has a student-specific baseline score:

$$y_{ij} = (\beta_{0j} + u_i) + \beta_{1j} x_i + \epsilon_{ij}$$

    which shows that the model has a student-specific intercept (baseline). Here there is both a student indicator and an exam type indicator.

- scores are correlated within each student by sharing a latent variable u_i
$$y_{ij} = \beta_{0j} + \beta_{1j} x_i + u_i + \epsilon_{ij}$$
where $u_i \overset{i.i.d.}{\sim} N(0, \sigma_u^2)$
- The above two ideas are very similar. In the first idea, we can add a prior of u_i to borrow across students, and then we have the same LMM as from the second idea. From the perspective of the first idea, u_i can also be fixed p different parameters.
- Treating u_i fixed we assume less model assumptions while by treating u_i random we obtain more efficient estimate of both μ_i and β.

2 More examples

See more examples in R Data Example 8.