Why?

Juan Martínez

10/23/2020

What is multilevel regression modeling?

Consider an educational study with data from students in many schools, predicting in each school the student's grades y on a standardized test given their scores on a pre-test x and other information.

- y: student's grades for each school.
- x: Student's score and other information.

A separate regression model can be fit within each school, and the parameters from these schools can themselves be modeled as depending on school characteristics such as *socioeconomic* status of the school's neighborhood, whether the school is private or public, and so on.

> "The **student-level** regression and the **school-level** regression are **two levels** of a **multilevel model.**

More generally, we consider a multilevel model to be a regression in which the parameters (regression coefficients) are given a probability model. This second-level model has parameters of its own called **hyperparameters** of the model -which are also estimated from data.

> "The two key parts if a multilevel model are **varying coefficients**, and a **model** for those varying coefficients."
>
> "The feature that distinguishes multilevel models from classical regression is in the modeling of the **varying coefficients**."

Models for regression coefficients

To keep notation simple, we assume just **one student-level** predictor x (for example, a pre-test score), and **one school level** predictor u (for example, average partent's incomes).

Varying-intercept model

This is a model in which the regressions have the **same slope** in each of the schools, and only the intercepts vary. We use the notation i for individual students and $j[i]$ for the school j containing student i:

$$\begin{aligned} y_i &= \alpha_{j[i]} + \beta_{j[i]} x_i + \epsilon_i, \quad for\ students\ i = 1, \ldots, n \\ \alpha_j &= a_0 + b_0 u_j + \nabla_{j1}, \quad for\ students\ j = 1, \ldots, J \\ \beta_j &= a_1 + b_1 u_j + \nabla_{j2}, \quad for\ students\ j = 1, \ldots, J \end{aligned} \tag{1}$$

References

Appendix