Multilevel Linear Models

Lindsay Xu

2022-11-04

Please click on **"See Complete Visualization"** to see the complete visualization. I also upload the complete visualization to Sakai in case you cannot open the link.

1. Introduction

Multilevel model is an useful modeling tool for estimating variance components using the observational data having a hierarchical or clustered structure and make prediction based on these estimations.

Multilevel model compromises between the two extremes of excluding a categorical predictor from a model (**complete pooling**), or estimating separate models within each level of the categorical predictor (**no pooling**) (Gelman and Hill, 2007). Multilevel models allow a technique called "**shrinkage**", also know as "**partial pooling**" to partially pools the group-level parameters α_j toward their mean level μ_α.

In specific, there is more pooling when the group-level standard deviation σ_α is small, and more smoothing for groups with fewer observations. In general, the multilevel-modeling estimate of the group-level parameters α_j can be expressed as a weighted average of the no-pooling for its group $(\bar{y_j} - \beta\bar{x_j})$ and the mean μ_α (this expression is adopted from Gelman and Hill, 2007):

$$\alpha_j \approx \frac{n_j/\sigma_y^2}{n_j/\sigma_y^2 + 1/\sigma_\alpha^2}(\bar{y_j} - \beta\bar{x_j}) + \frac{1/\sigma_\alpha^2}{n_j/\sigma_y^2 + 1/\sigma_\alpha^2}\mu_\alpha$$

2. Goals

The aim of this tool is to illustrate random intercepts as well as fixed and random slopes in multilevel modelling. I first discussed the problems with the no-pooling and complete-pooling analyses. Then, I gradually build on the model and introduce the complexity of multilevel linear regression step by step with a figures of predicted model values. Specifically, I am going to explain the properties of the multilevel models by visualizing how the predicted values change under different model specifications. In addition, I discuss and compare model fit.

Overall, the goals of this tool are to:
- Give a general idea of limitations of no-pooling and complete pooling estimations for clustered

data
- Give a general understanding of random intercepts across clusters in hierarchical data
- Give a general understanding of random slope model, which allow the explanatory variable to have a different effect for each group
- Be able to interpret parameter estimates of fixed and random effects in multilevel models
- Be able to assess the model fit of multilevel models

3. Data

I used a simulated data for 2000 pupils in 100 schools. To better demonstrate the visualization, I randomly selected 15 schools out of these 100 schools with 307 pupils in total. The main outcome variable is the pupil popularity, a popularity rating on a scale of 1-10 derived by a sociometric procedure. Typically, a sociometric procedure asks all pupils in a class to rate all the other pupils, and then assigns the average received popularity rating to each pupil. Because of the sociometric procedure, group effects as apparent from higher-level variance components are rather strong. The explanatory variables on individual level are pupil gender (boy = 0, girl = 1) and pupil extraversion (10-point scale); the explanatory variables on group level (class level) is teacher's experience (the data was adopted from: https://multilevel-analysis.sites.uu.nl/datasets/).

3.1 Level 1 observational unites (i)

Therefore, in this case, the level one refers to the individual pupils clustered in the classes. The level-one predictors are students' gender and extraversion.

3.2 Level 2 observational units (j)

The level two refers to the classes and the level-two predictors are teachers' experience.

4. Research questions

In this research, I am especially interested in studying how pupils' extraversion (level-one predictor) would affect their popularity by controlling their gender (level-one predictor). I am also exploring whether teacher's experience (level-two predictor) will affect students' popularity.

To address this question, I mainly fitted three models. **Model 1** is a random intercept model, where popularity is the outcome, extraversion and gender are fixed effect coefficient. **Model 2** is a random slope model, where popularity is the outcome, extraversion is a random effect coefficient and gender is a fixed effect coefficient. **Model 3** is a random slope model, where popularity is the outcome, extraversion is a random effect coefficient and gender is a fixed effect coefficient.I also include the level-two predictor–teacher's experience as a fixed effect coefficient.

5. Analysis

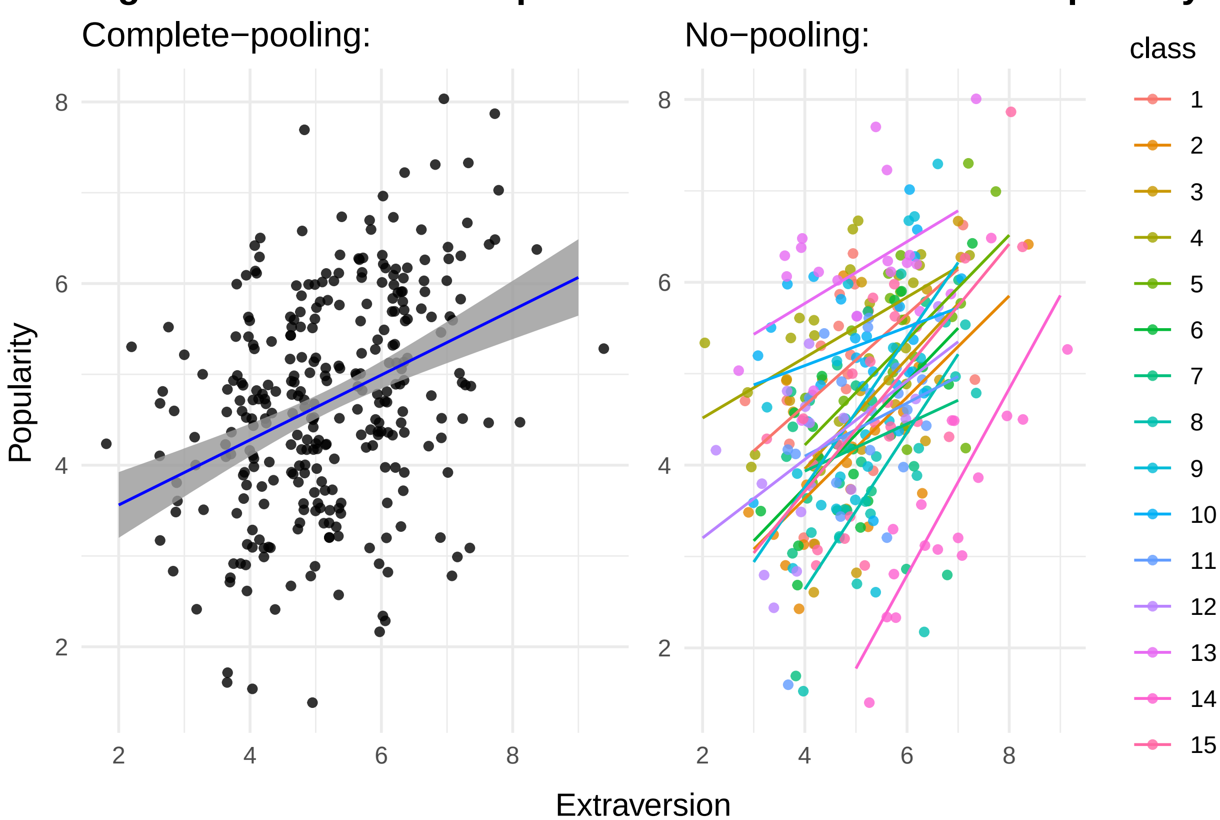
5.1 problems with the no-pooling and complete-pooling analyses

To illustrate the problems with no-pooling and complete-pooling analyses, I started with investigating the association between pupils' extraversion and their popularity among peers.

Figure 1 shows the estimations using complete-pooling and no-pooling method.

However, both analyses shown in Figure 1 have problems. The complete-pooling method ignores any variation in pupils' popularity between classes. The no-pooling method overestimates the variability of pupils' popularity among the classes, especially when we have a relatively small sample size of observations in each class. Again, **Multilevel model** compromises between the two extremes of excluding a categorical predictor from a model (**complete pooling**), or estimating separate models within each level of the categorical predictor (**no pooling**) (Gelman and Hill, 2007). It borrows information both from the variability in the groups and information across all the groups.

Figure 1. The relationship between Extraversion and Popularity



5.2 Random intercept models

Random Intercept model shows us how much variation is at each level. In this case, random intercept model allows me to study how much an effect class has on students' popularity after controlling for students' gender and extraversion. The model can be expressed as below:

$$y_i = \alpha_{j[i]} + \beta 1 x_{1i} + \beta 2 x_{2i} + \epsilon_i$$
$$\alpha_j = \mu_\alpha + \eta_j$$
$$\eta_j \sim N(0, \sigma_\alpha^2)$$

Here, I fitted a random-intercept model. y_{ij} is the dependent variable–students' popularity, x_{1i} is the key predictor–extraversion. $\beta 1$ is the regression slope coefficient of extraversion (fixed effect). x_{2i} is another individual-level predictor–gender. $\beta 2$ is the regression slope coefficient of gender (fixed effect). ϵ_i is the unobservable error term. α_j represents the intercepts that varies across classes. $\alpha_j = \mu_\alpha + \eta_j$ represent the second level (class) of the model.η_j represents group-level errors.

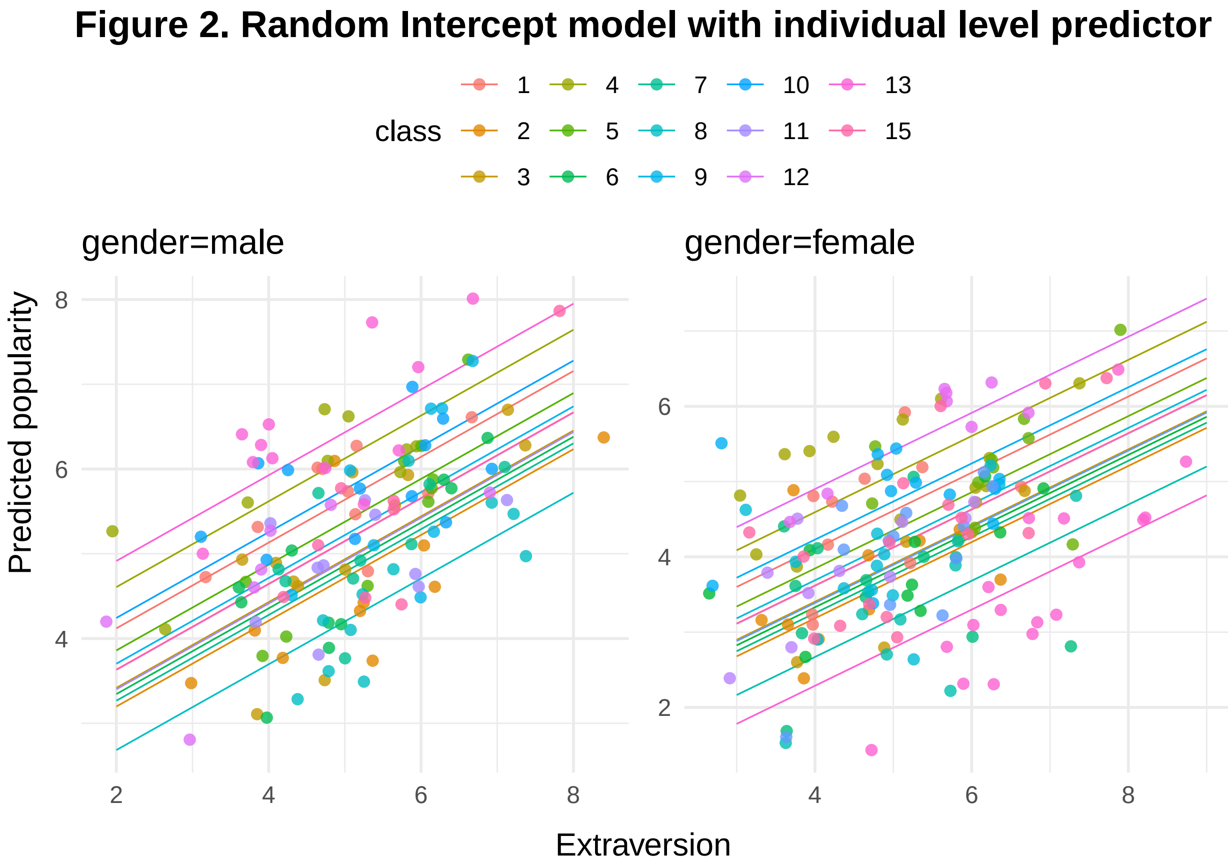


Figure 2 shows the random intercept model controlling for gender. Overall, extraversion is positively related to the popularity. However, male pupils have a relatively high popularity compared to the female pupils.

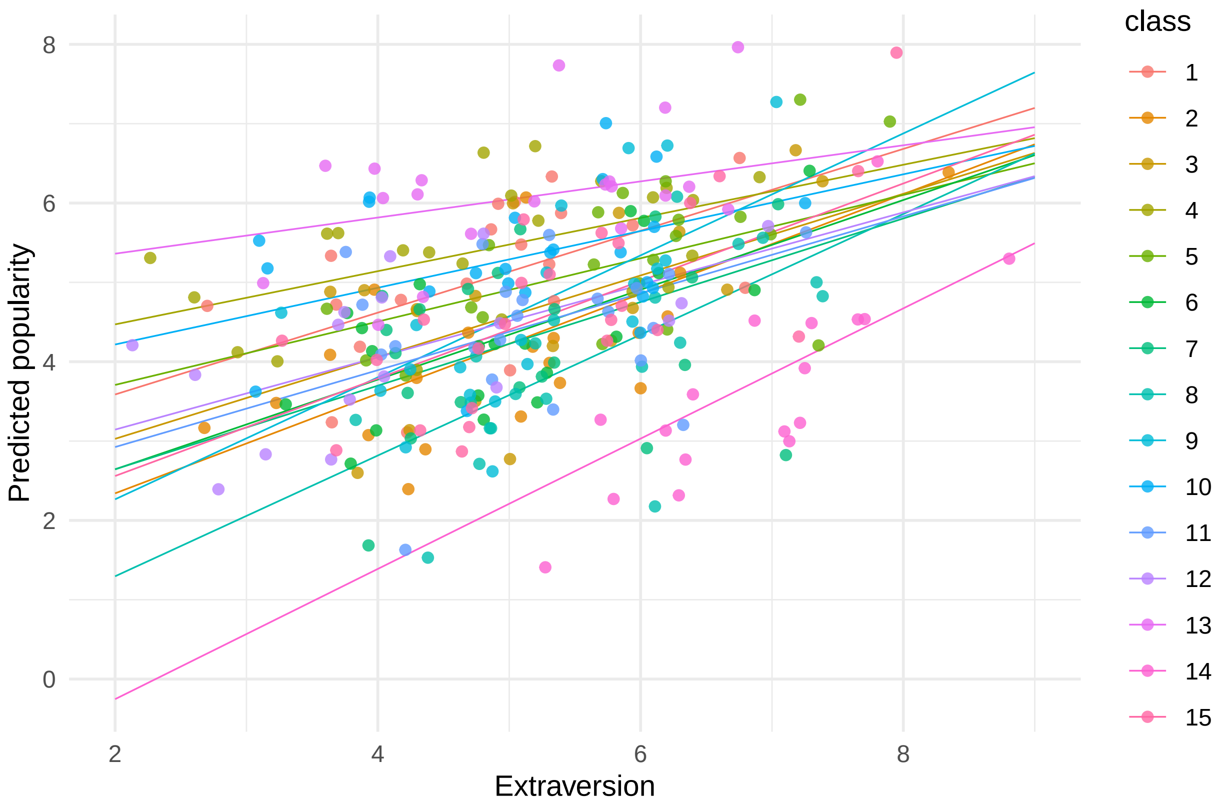
5.3 Random slope models

Random slope models allows each group line to have a different slope and that means that the random slope model allows the explanatory variable to have a different effect for each group. In this case, I include a random slope coefficient for extraversion, which means that the effects of extraversion on popularity differ across different groups. The model can be expressed as below:

$$y_i = \alpha_{j[i]} + \beta_{j[i]} x_{1i} + \beta_2 x_{2i} + \epsilon_i, i = 1, ..., n$$
$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \sim N(\begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \alpha_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix}), j = 1, ..., J$$

Here, I fitted a random-slope model. y_{ij} is the dependent variable–students' popularity, x_{1i} is the key predictor–extraversion. β_j allows the regression slope coefficients of extraversion to vary across the classes. x_{2i} is another individual-level predictor–gender. $\beta 2$ is the regression slope coefficient of gender (fixed effect). ϵ_i is the unobservable error term. α_j represents the intercepts that varies across claases. α_j and β_j follow a multivariate normal distribution, where ρ represents the between-group correlation parameters.

Figure 3. Random coefficients model with individual level predictor



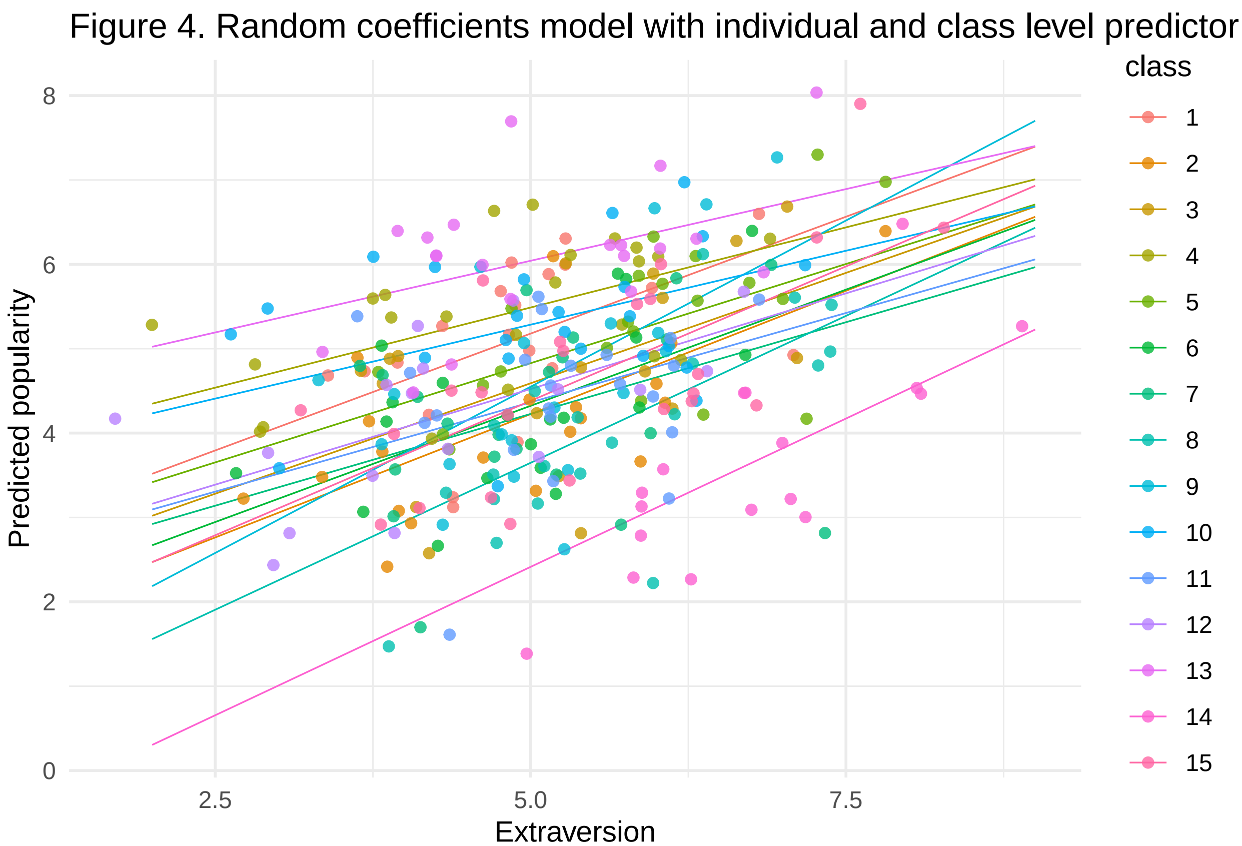


Figure 3 and Figure 4 illustrate the random-slope models. Overall, extraversion is a positive predictors for pupils' popularity. However, the effects of extraversion on popularity differ across different class. A lot of variations are explained by the classes level.

6. Comparing models and interpretation

According to the table 1, model 3–random slope model with individual- and class-level predictors fits the best and the random effects are significant. Because model 3 has a lowest AIC and the P-value is significant.

Table 1: Comparing three models

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m1	5	711.8461	730.4803	-350.9230	701.8461	NA	NA	NA
m2	7	707.8031	733.8910	-346.9015	693.8031	8.043012	2	0.0179260
m3	8	705.6621	735.4768	-344.8310	689.6621	4.141001	1	0.0418564

According to the results from three models (table 2), extraversion, gender and teacher's experience are all positive predictors for pupil's popularity. In particular, the increasing in extraversion predicts the increasing in popularity. In addition, male pupils have a relatively high popuylarity compared to female pupils. Meanwhile, pupils from a class with a more experienced teacher have a relatively high popularity compared those from a class with less experienced teacher.

	Dependent variable:		
	popular		
	M1	M2	M3
extrav	0.506*** (0.038)	0.524*** (0.055)	0.518*** (0.055)
sex	1.026*** (0.084)	1.026*** (0.083)	1.036*** (0.083)
texp			0.080*** (0.026)
Constant	1.584*** (0.272)	1.440*** (0.433)	0.414 (0.485)
Observations	307	307	307
Log Likelihood	−355.630	−351.334	−352.193
Akaike Inf. Crit.	721.260	716.667	720.387
Bayesian Inf. Crit.	739.894	742.755	750.201
Note:			*p<0.1; **p<0.05; ***p<0.01

Reference

Gelman, A., & Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge university press.

https://m-clark.github.io/mixed-models-with-R/random_intercepts.html

http://www.bristol.ac.uk/cmm/learning/videos/random-slopes.html