More Mixed-Effects Models

Andrew Zieffler

Educational Psychology

University of Minnesota

Driven to Discover℠

The Unconditional Means Model

Partitioning Total Outcome Variation Between and Within Persons

Level-1 Model $$Y_{ij} = \beta_{0j} + \epsilon_{ij} \quad \text{where} \quad \epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$$

Level-2 Model $$\beta_{0j} = \beta_{00} + \eta_{0j} \quad \text{where} \quad \eta_{0j} \sim N(0, \sigma^2_0)$$

Composite Model $$Y_{ij} = \boxed{\beta_{00}} + \boxed{\eta_{0j}} + \boxed{\epsilon_{ij}}$$

Within-school deviations

Grand mean across schools

School-specific means

```
# Unconditional means model
> lmer.u = lmer(math ~ 1 + (1|school), data = students)
```

Composite Model $Y_{ij} = \boxed{\beta_{00}} + b_{0i} + \epsilon_{ij}$

Grand mean math score across individuals and schools

σ_ϵ^2

σ_0^2

Within-school variance

Between-school variance

Let's look more closely at these variances….

```
> summary(lmer.u)

REML criterion at convergence: 47116.8

Random effects:
 Groups   Name        Variance Std.Dev.
 school   (Intercept)  8.614   2.935
 Residual             39.148   6.257
Number of obs: 7185, groups: school, 160

Fixed effects:
            Estimate Std. Error t value
(Intercept)  12.6370     0.2444   51.71
```

Using the Unconditional Means Model to Estimate the Intraclass Correlation Coefficient (ICC or ρ)

Major purpose of the unconditional means model: To **partition the variation in *Y* into two components**

```
Random effects:
 Groups   Name        Variance Std.Dev.
 school   (Intercept)  8.614   2.935
 Residual             39.148   6.257
Number of obs: 7185, groups: school, 160

Fixed effects:
            Estimate Std. Error t value
(Intercept)  12.6370     0.2444   51.71
---
```

Estimated within-school variance: Quantifies the amount of variation **within** level-2 units

Estimated between-school variance: Quantifies the amount of variation **between** level-2 units

Intraclass correlation compares the relative magnitude of these variance components (VCs) by **estimating the proportion of total variation in *Y* that lies "between" level-2 units**

$$\rho = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_\epsilon^2}$$

$$\hat{\rho} = \frac{8.614}{8.614 + 39.148} = \boxed{0.180}$$

An estimated **18.0% of the total variation in math scores is attributable to differences between schools**

Conditional Means Model

After partitioning the variation, we will include any student-level (level-1) predictors to explain within-school variation.

Level-1 Model:

$$Y_{ij} = \beta_{0j} + \beta_1(X_i) + \epsilon_{ij} \quad \text{where} \quad \epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$$

Level-2 Model:

$$\beta_{0j} = \beta_{00} + \eta_{0j} \quad \text{where} \quad \eta_{0j} \sim N(0, \sigma^2_0)$$

Composite Model:

$$\boxed{Y_{ij}} = \boxed{\beta_{00}} + \boxed{\eta_{0j}} + \boxed{\beta_1}(X_i) + \boxed{\epsilon_{ij}}$$

student-specific deviation

Mean math score for all students with $X_i = 0$

School-specific deviation in intercept

Average difference in math score for a one-unit difference in X_i across all schools

Fitting the Conditional Means Model

$$Y_{ij} = \beta_{00}(1) + \beta_1(X_i) + [\eta_{0j}(1) + \epsilon_{ij}]$$

```
# Conditional means model
> lmer.c = lmer(math ~ 1 + ses + (1|school), data = students)
```

```
REML criterion at convergence: 46645.2

Random effects:
 Groups   Name        Variance Std.Dev.
 school   (Intercept)  4.768   2.184
 Residual             37.034   6.086
Number of obs: 7185, groups: school, 160

Fixed effects:
            Estimate Std. Error t value
(Intercept)  12.6575     0.1880   67.33
ses           2.3902     0.1057   22.61
```

Between-school variance has changed because the intercept has a different meaning when we add a predictor

Within-school variance has decreased

Interpretation of the residual variance component

The estimated residual variance provides a measure of the within-school (student-to-student) variation of math scores after accounting for SES.

In our example...

There seems to be within-school variation in math scores after accounting for SES.

$$\hat{\sigma}_{\epsilon}^{2} = 37.03$$

The residual variation (level-1) decreased from 39.15 to 37.03.

$$r^2 = \frac{39.15 - 37.03}{39.15} = 0.05$$

The change in level-1 residual variation should **always** be compared to the *unconditional means* model.

This is a **Pseudo R^2**. Similar to the R^2 in OLS models, it measures the reduction in the level-1 residual variance.

Interpreting the Random-Effects

Fitted Composite Model

$$\hat{Y}_{ij} = \hat{\beta}_{00} + \hat{\eta}_{0j} + \hat{\beta}_1(X_i)$$

Interpretation of RE of intercept

The η_{0j} estimate for each school is the difference in predicted math score between the grand mean and the school mean, for a SES of 0.

Interpretation of the variance component for the REs of intercept

The variance in the η_{0j} estimates indicates the variation in average math score for a SES of 0 across all schools.

		Parameter	Model A	Model B
Fixed effects				
Intercept	Intercept	β_{00}	12.64 (0.24)	12.66 (0.19)
SES	Intercept	β_{10}		2.39 (0.11)
Variance components				
Level-1	Within-persons	σ_ϵ^2	39.148	37.03
Level-2	In intercepts	σ_0^2	8.61	4.76
	In slopes	σ_1^2		
Goodness-of-fit				
	Deviance		47116.79	46645.17
	AIC		47122.79	46653.17
	BIC		47143.43	46680.69

These are the fixed-effects from the summary output.

These are the variance components from the random-effects part of the summary output.

These are common goodness-of-fit measures reported with mixed-effects models.

Random Intercepts and Random Slopes Model

After partitioning the variation, we will include any student-level (level-1) predictors to explain within-school variation.

Level-1 Model:

$$Y_{ij} = \beta_{0i} + \beta_{1i}(X_i) + \epsilon_{ij} \quad \text{where} \quad \epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$$

Level-2 Model:

$$\begin{aligned} \beta_{0j} &= \beta_{00} + \eta_{0j} \\ \beta_{1j} &= \beta_{10} + \eta_{1j} \end{aligned} \quad \text{where} \quad \begin{bmatrix} \eta_{0j} \\ \eta_{1j} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right)$$

school-specific deviation in slope

Composite Model:

$$Y_{ij} = \beta_{00} + b_{0j} + (\beta_{10} + b_{1j})\, X_i + \epsilon_{ij}$$

student-specific deviation

Mean math score for all students with $X_i = 0$

School-specific deviation in intercept

Average difference in math score for a one-unit difference in X_i across all schools

Fitting the Random Intercepts and Random Slopes Model

$$Y_{ij} = \beta_{00}(1) + \beta_{10}(X_i) + [b_{0j}(1) + b_{1j}(X_i) + \epsilon_{ij}]$$

```
# Unconditional means model
> lmer.ri.rs = lmer(math ~ 1 + ses + (1 + ses |school),
    data = students)
```

```
REML criterion at convergence: 46640.4

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 school   (Intercept)  4.8286  2.1974
          ses          0.4129  0.6426   -0.11
 Residual             36.8302  6.0688
Number of obs: 7185, groups: school, 160

Fixed effects:
            Estimate Std. Error t value
(Intercept)  12.6650     0.1898   66.71
ses           2.3938     0.1181   20.27
```

Interpreting the Fixed-Effects

Predicted Level-1 Model

$$\hat{\text{Math}} = 12.66 + 2.39(\text{SES})$$

Interpretation of Intercept

The mean math score for students who have an average SES (`SES = 0`) is 12.66.

Interpretation of the slope

Each one-unit difference in SES is associated with a 2.39-unit difference in math score, on average.

Interpretation of the residual variance component

The estimated residual variance provides a measure of the within-school (student-to-student) variation of math scores after accounting for SES.

In our example...

There seems to be within-school variation in math scores after accounting for SES.

$$\hat{\sigma}_{\epsilon}^{2} = 36.83$$

The residual variation (level-1) decreased from 39.15 to 36.83.

$$r^2 = \frac{39.15 - 36.83}{39.15} = 0.06$$

The change in level-1 residual variation should **always** be compared to the *unconditional means* model.

This is a **Pseudo R^2**. Similar to the R^2 in OLS models, it measures the reduction in the level-1 residual variance.

Interpreting the Random-Effects

Fitted Composite Model

$$\hat{Y}_{ij} = \hat{\beta}_{00} + \hat{\eta}_{0j} + \hat{\beta}_{10}(X_i) + \hat{\eta}_{1j}(X_i)$$

Interpretation of RE of intercept

The η_{0j} estimate for each school is the difference in predicted math score between the grand mean and the school mean, for a SES of 0.

Interpretation of RE of slope

The η_{1j} estimate for each team is the difference in the effect of SES on math score between the school and the mean of all the schools' effects.

Interpretation of the variance component for the REs of intercept

The variance in the η_{0j} estimates indicates the variation in average math score for a SES of 0 across all schools.

Interpretation of the variance component for the REs of slope

The variance in the η_{1j} estimates indicates the variation in the effect of SES on math scores across all schools.

		Parameter	Model A	Model B	Model C
Fixed effects					
Intercept	Intercept	β_{00}	12.64 (0.24)	12.66 (0.19)	12.67 (0.19)
SES	Intercept	β_{10}		2.39 (0.11)	2.39 (0.12)
Variance components					
Level-1	Within-persons	σ_ϵ^2	39.148	37.03	36.83
Level-2	In intercepts	σ_0^2	8.61	4.76	4.83
	In slopes	σ_1^2			0.41
Goodness-of-fit					
	Deviance		47116.79	46645.17	46640.4
	AIC		47122.79	46653.17	46652.4
	BIC		47143.43	46680.69	46693.68

```
# Get estimates of the random-effects
> ranef(lmer.ri.rs)

$school
      (Intercept)           ses
1224 -1.604995891  0.110271946
1288  0.408590120  0.084117692
1296 -3.467719669 -0.040825210


# Estimates of the variance-covariance matrix of the random effects
> varCorr(lmer.ri.rs)$school

            (Intercept)        ses
(Intercept)   4.8286364 -0.1542761
ses          -0.1542761  0.4129288

attr(,"stddev")
(Intercept)         ses
  2.1974158   0.6425954

attr(,"correlation")
            (Intercept)        ses
(Intercept)   1.0000000 -0.1092569
ses          -0.1092569  1.0000000
```

Square roots of the variance estimates

$$\mathbf{G} = \begin{bmatrix} 4.83 & -0.15 \\ -0.15 & 0.41 \end{bmatrix}$$

The η_{0j} estimates and η_{1j} estimates are negatively related. Schools that have a lower intercept also tend to have higher effect of SES.

Assumptions

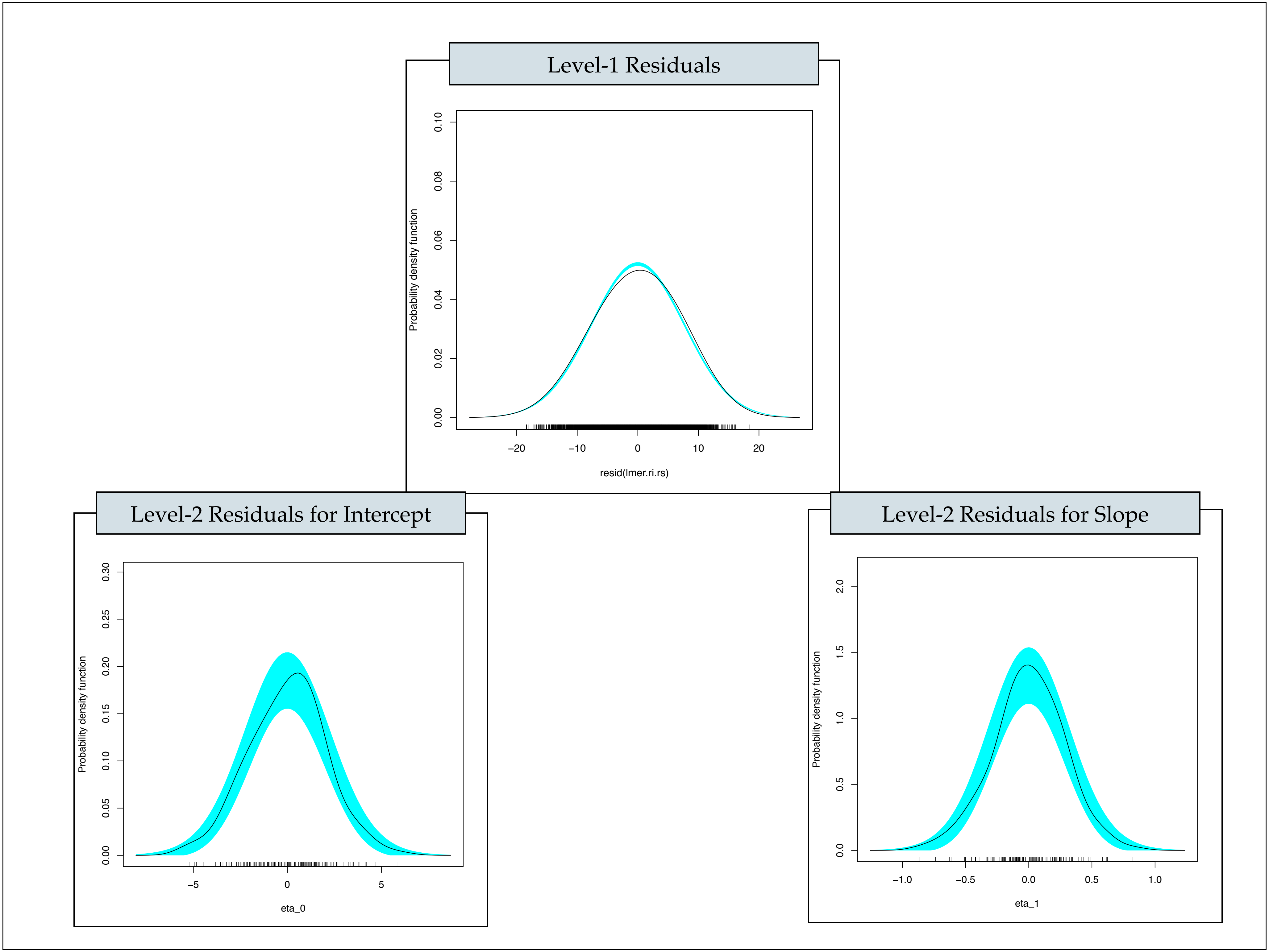
$\epsilon_{ij} \sim N(0, \sigma^2_\epsilon)$ $\eta_{0j} \sim N(0, \sigma^2_0)$ $\eta_{1j} \sim N(0, \sigma^2_1)$

Both the level-1 errors and all the random-effects need to be normally distributed.

```
# Residuals
> sm.density(resid(lmer.ri.rs), model = "normal")

# Examine the random-effects for intercept
> eta_0 = ranef(lmer.ri.rs)$school[ , 1]
> sm.density(eta_0, model = "normal")

# Examine the random-effects for slope
> eta_1 = ranef(lmer.ri.rs)$school[ , 2]
> sm.density(eta_10, model = "normal")
```


Level-1 Residuals
Probability density function
0.10
0.08
0.06
0.04
0.02
0.00
−20
−10
0
10
20
resid(lmer.ri.rs)
Level-2 Residuals for Intercept
Probability density function
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−5
0
5
eta_0
Level-2 Residuals for Slope
Probability density function
2.0
1.5
1.0
0.5
0.0
−1.0
−0.5
0.0
0.5
1.0
eta_1