# Multilevel (Mixed Effects) Models

# Multi-Level Model

Level-1: life satisfaction<sub>ij</sub> =  $\beta_0^* + \beta_1^*$  (Shots on five<sub>ij</sub>) +  $\epsilon_{ij}$ 

Level-2: 
$$\beta_0^* = \beta_{00} + \eta_{0j}$$
  
 $\beta_1^* = \beta_{01} + \eta_{1j}$ 

The **level-1 model** specifies how the player-level predictors relate to the outcome (life satisfaction scores)

The **level-2 model** specifies how each coefficient in the level-1 model is predicted by team-level predictors

This is called a multi-level model.

## Level-1 Model

life satisfaction<sub>ij</sub> = 
$$\beta_0^* + \beta_1^*$$
 (Shots on five<sub>ij</sub>) +  $\epsilon_{ij}$ 

- Level-1 model specifies the regression equation across players
  - ✓  $\beta_0$  is the intercept across players
  - ✓  $\beta_1$  is the slope across players
  - ✓  $\epsilon_{ij}$  are the random errors for the *i*th player to the regression line

- Only source of variation in Level-1 model is player-to-player (within-team) variation
- Any unaccounted for player-to-player variation is absorbed into the residual,  $\varepsilon_{ij}$

## Level-2 Model

$$\beta_0^* = \beta_{00} + \eta_{0j}$$
$$\beta_1^* = \beta_{01} + \eta_{1j}$$

$$\beta_1^* = \beta_{01} + \eta_{1j}$$

- Level-2 models are team-level regression models
  - Outcome variables are the Level-1 regression parameters (e.g., intercepts and slopes)
  - Number of Level-2 equations determined by number of regression parameters in the Level-1 model
  - Covariates that account for team-to-team (between-team) variation exclusively appear in Level-2 model

- Error terms in Level-2 model are called random-effects
  - Can only be as many random-effects as there are Level-2 equations
  - Not every Level-2 equation needs to have random-effect

# Mixed-Effects Model

Level-1: life satisfaction<sub>ij</sub> =  $\beta_0^* + \beta_1^* (Shots on five_{ij}) + \epsilon_{ij}$ 

Level-2: 
$$\beta_0^* = \beta_{00} + \eta_{0j}$$
  
 $\beta_1^* = \beta_{01} + \eta_{1j}$ 

Substitute the Level-2 equations into the Level-1 equation to get the mixed-effects model

life satisfaction<sub>ij</sub> =  $[\beta_{00} + \eta_{0j}] + [\beta_{01} + \eta_{1j}]$  (Shots on five<sub>ij</sub>) +  $\epsilon_{ij}$ 

The model essentially helps us to partition the variation in the residual...that which is due to between-school (school-to-school) variation (the random effects) and that which is within-school (student-to-student) error.

life satisfaction<sub>ij</sub> = 
$$\beta_0^* + \beta_1^*$$
 (Shots on five<sub>ij</sub>) +  $\epsilon_{ij}$ 

life satisfaction<sub>ij</sub> = 
$$\beta_{00} + \beta_{01}(\text{Shots on five}_{ij}) + \eta_{0j} + \eta_{1j}(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

The mixed-effects model is used to specify the R syntax for fitting the model

$$\boxed{\text{life satisfaction}_{ij} = \beta_{00}(1) + \beta_{01}(\text{Shots on five}_{ij}) + \eta_{0j}(1) + \eta_{1j}(\text{Shots on five}_{ij}) + \epsilon_{ij}}$$

- > library(lme4)

Level-2 ID

life satisfaction<sub>ij</sub> = 
$$\beta_{00}(1) + \beta_{01}(\text{Shots on five}_{ij}) + \eta_{0j}(1) + \eta_{1j}(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

The lmer() function uses syntax similar to the lm() function.

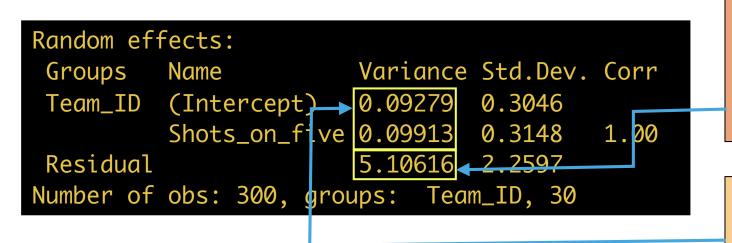
- The fixed-effects use the exact same syntax
- The random-effects appear in parentheses and reference the level-2 ID variable

The fixed-effects are interpreted exactly the same as we interpret the coefficients in an LM.

- The predicted life satisfaction score for a player who shoots zero free-throws is 6.42.
- The average difference in life satisfaction scores between students who are one freethrow different is predicted to be 3.29.

There are no *p*-values given for results from a mixed-effects analysis in R. That is because it is theoretically unclear what value for the *df* should be used for the *t*-test.

The output associated with the random-effects provide variance estimates for the associated random-effects in the model and also for the residuals.



### Level-1 variance:

Quantifies the amount of life satisfaction variation **within** teams

#### Level-2 variance:

Quantifies the amount of variation in life satisfaction **between** teams.

$$\epsilon_{ij} \sim \mathcal{N}\left(0, \hat{\sigma}_{\epsilon}^2\right) \qquad \qquad \hat{\sigma}_{\epsilon}^2 = 5.106$$

$$\beta_{0j} \sim \mathcal{N}\left(\beta_0^*, \hat{\sigma}_{\beta_{0j}}^2\right) \qquad \hat{\sigma}_{\beta_{0j}}^2 = 0.093$$

$$\beta_{1j} \sim \mathcal{N}\left(\beta_1^*, \hat{\sigma}_{\beta_{1j}}^2\right) \qquad \hat{\sigma}_{\beta_{1j}}^2 = 0.099$$

The level-2 (betweenteam) variance estimates tell us how much the school-level intercept and slopes vary.