Multilevel (Mixed Effects) Models

Multi-Level Model

Level-1: $$\text{life satisfaction}_{ij} = \beta_0^* + \beta_1^*(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

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

$$\text{life satisfaction}_{ij} = \beta_0^* + \beta_1^*(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

- Level-1 model specifies the regression equation across players
  - ✓ β_0 is the intercept across players
  - ✓ β_1 is the slope across players
  - ✓ ε_{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, ε_{ij}

Level-2 Model

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

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

- 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: $$\text{life satisfaction}_{ij} = \beta_0^* + \beta_1^*(\text{Shots on five}_{ij}) + \epsilon_{ij}$$

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

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

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.

$$\text{life satisfaction}_{ij} = \beta_0^* + \beta_1^*(\text{Shots on five}_{ij}) + \boxed{\epsilon_{ij}}$$

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

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

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

```
> library(lme4)
> lmer.1 = lmer(Life_Satisfaction ~ 1 + Shots_on_five +
    (1 + Shots_on_five | Team_ID), data = nba)
```

Level-2 ID

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

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

```
> summary(lmer.1)

Fixed effects:
              Estimate Std. Error t value
(Intercept)     6.4296     0.3169   20.29
Shots_on_five   3.2887     0.1340   24.55
```

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 free-throw 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.

```
Random effects:
 Groups   Name          Variance Std.Dev. Corr
 Team_ID  (Intercept)   0.09279  0.3046
          Shots_on_five 0.09913  0.3148   1.00
 Residual               5.10616  2.2597
Number of obs: 300, groups:  Team_ID, 30
```

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.

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

$$\beta_{0j} \sim \mathcal{N}\left(\beta^*_0, \hat{\sigma}^2_{\beta_{0j}}\right) \qquad \hat{\sigma}^2_{\beta_{0j}} = 0.093$$

$$\beta_{1j} \sim \mathcal{N}\left(\beta^*_1, \hat{\sigma}^2_{\beta_{1j}}\right) \qquad \hat{\sigma}^2_{\beta_{1j}} = 0.099$$

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