Name: Md Allama Ikbal Sijan

**Hierarchical Linear Modeling**

Psyc 741, Spring 2025

**Due Date**: March 3, 2025 (11:00 AM)

***Homework #2***

For this homework assignment, you will be carrying out null, random intercept, and random slope models and interpreting the relevant output.

You will be completing the following questions, some of which involve conducting analyses in R. You will hand in the answers to the questions; enter them into this word document. Some of the answers involve you copying and pasting your R code. However, you also need to submit your *complete* R code too (saved as a .R script file). Submit both this completed document and your R script file to the submission portal for the Homework #2 Submission portal on Canvas.

1. Import the **schools.csv** data file on Canvas into RStudio. This data file contains information about samples of students nested within various schools.

**Null Model** [30 points]

1. Execute a null multilevel model with the ID values for the schools as the random intercept variable and student math scores as the dependent variable. Make sure that maximum likelihood (ML) is used as the estimator. Paste your R syntax below.

null\_model <- lmer(math\_score ~ (1|school), REML = F, data = df)

summary(null\_model)

1. Report the following sample sizes from the data:
   1. Number of individual participants: 7185
   2. Number of groups: 160
2. Report the following fit statistics from the null model:
   1. AIC: 47121.8
   2. BIC: 47142.4
   3. -2LL: 47115.8
3. Report the following variance components from the model output:
   1. Between-group variance: 8.553
   2. Within-group variance: 39.148
4. Calculate and report the intraclass correlation coefficient (ICC) below and interpret its value.

ICC = 0.1793044, which indicates that 17.9% of the variance in the DV is explained by the grouping variable “school”.

**Random Intercept Model** [50 points]

1. There is a variable in the dataset that contains the individual socioeconomic status of the students on a standardized scale. The models below also include the school average (mean) socioeconomic status scores. Create an aggregate version of the socioeconomic status variable such that the mean socioeconomic status of the students for each school is included in the data frame. Copy/paste your R syntax below.

df <- df %>%

group\_by(school) %>%

mutate(gr\_ses = mean(ses, na.rm = T))

1. Execute a random intercept model with the same random intercept from the null model as above and the following fixed effects predictors: the socioeconomic status of the student, the school average socioeconomic status variable you created in Q7, the size of the school, and whether the school is private (0 = private, 1 = public). Make sure that maximum likelihood (ML) is used as the estimator. Paste your R syntax below.

ri\_model <- lmer(math\_score ~ ses + gr\_ses + school\_size + private + (1|school), REML = F, data = df)

summary(ri\_model)

1. Report the following fit statistics from the random intercept model:
   1. AIC: 46559.2
   2. BIC: 46607.4
   3. -2LL: 46545.2
2. Report the following variance components from the random intercept model output:
   1. Between-group variance: 2.256
   2. Within-group variance: 37.015
3. Report and interpret all the fixed effects slope results below in APA format. Make sure to include the relevant statistical information in your reporting.

The fixed intercept was statistically significant, B = 11.53, SE = 0.39, *t* (156.28) = 29.44, *p* < .001, indicating that the predicted value of the dependent variable is 11.53 when individual and school-level socioeconomic status are at their average values, the school is private, and school size is zero.

Socioeconomic status (SES) was a significant predictor of the dependent variable, B = 2.19, SE = 0.11, *t* (7021.96) = 20.17, *p* < .001, indicating that for each one-unit (or one standard deviation as SES is standardized) increase in SES, math scores increased by approximately 2.19 units. Group-level SES (gr\_ses) also had a significant positive effect, B = 3.12, SE = 0.38, *t* (183.50) = 8.24, *p* < .001, suggesting that one unit increase in average SES within a school was associated with 3.11 unit increase in the student’s math scores.

Additionally, private school status was a significant predictor, B = 1.48, SE = 0.33, *t* (150.76) = 4.44, *p* < .001, indicating that students in public school scored on average 1.48 units higher in math in comparison to private school students. However, school size was not a significant predictor of math scores, B = 0.00045, SE = 0.00025, *t* (158.70) = 1.77, *p* = .079.

1. To help demonstrate the effect size of the random intercept model, calculate and report below the between-group proportional reduction in variance (BG-PRV) and within-group proportional reduction in variance (WG-PRV) between the null model and the random intercept model. What is your final determination of the effect size of this random intercept model?

The between-group proportional reduction in variance (BG-PRV) was 0.736, indicating that the inclusion of predictors in the random intercept model reduced approximately 73.6% of the between-group variance compared to the null model. The within-group proportional reduction in variance (WG-PRV) was 0.054, suggesting that 5.4% of the within-group (residual) variance is reduced by added predictors in the random intercept model. Overall, these results indicate that the random intercept model substantially reduces between-group variance, demonstrating a strong effect size at the group level, while the within-group effect size is relatively small.

**Random Slope Model** [20 points]

1. Execute a random slope model that is the same as the random intercept model above but also has a random effect for the individual socioeconomic predictor included. Paste your R syntax below.

rs\_model <- lmer(math\_score ~ ses + gr\_ses + school\_size + private + (1|school) + (1|school:ses), REML = F, data = df)

summary(rs\_model)

1. Report the random slope estimate below.

The random slope estimate for SES within schools was not statistically significant, *χ²* (1) = 1.39, *p* = .24, indicating minimal variation in the SES effect across schools.

1. What is your determination of this random slope effect? Use statistical information to support your determination.

Given the non-significant random slope effect (*p* = .24), there is insufficient evidence to conclude that the relationship between SES and math scores varies significantly across schools. This suggests that a random slope for SES may not be necessary in the model.