Homework 7

Statistical Inference II

Noah Johnson

April 19, 2018

In this homework, the data based on a sample of n = 176 chidren within J = 10 schools in the American subsample of the PISA (Programme for International Student Assessment) and is available in PISASchools10.sav. It has intentionally been provided in an .sav format so you will have to find the function to read a .sav file.

First, we will look at the relationship between student's home education resources (HEDRES) on math achievement scores (MATHSCOR) across these 10 schools. Then, we will examine whether the schools' press for academic excellence (ACADPRES) moderates this relationship. Read each question carefully, as there are multiple parts to most questions.

HEDRES scaled from 1 to 8; lower values indicate poor home resources for education

MATHSCOR average of 50; SD of 9 points

ACADPRES scaled from 1 to 8; lower values indicate low press for academic excellence

1. Use the *sav* data file to run separate multiple regression models for each of the 10 schools (X = *hedres*; Y = *mathscor*). Review the separate regression results for each of the schools. What do you notice about the results (look at the correlations between *mathscor* and *hedres*, and the regression coefficients for each of the 10 schools)? Is it reasonable to assume that the effect of home resources on math achievement is the same in all 10 schools? Why or why not? What does your answer imply regarding how to include the *hedres* variable within our HLM?

2. MODEL 1: Using lme4 fit an unconditional random-effects ANOVA (i.e., *empty model*) with *mathscor* as the outcome. Report and interpret all the parameters as well as the ICC.

3. MODEL 2: Run a *random coefficients model* with *hedres* (**group-mean centered**) as the predictor of math achievement. Report and interpret all the parameters. **Compared to Model 1**, how much was the within-schools variability (s^2) reduced with the addition of the group-centered home resources variable?

4. Based on your results for step 3, would it make sense to eliminate the random effect for the home-resources slope (u_1)? Why or why not? Justify your decision.

5. MODEL 3: Finally, run a *contextual or conditional model*, and add the school academic press (**centered at the grand-mean**) as a level-2 predictor of both the level-1 intercepts and the home-resources slopes. Report and interpret all the parameters. **Compared to Model 2**, was the variability in the intercepts between schools reduced with the addition of the academic press variable? What was the proportion reduction in this variance? What about the variability in home-resources slopes? Compared to Model 2, what proportion of this variance was accounted for by academic press?