Multi-level Modeling Assignment

MATH 456 - Spring 2018

0. Initial Data Management

This homework uses the`School23` data set. Before you begin, convert the variable `school` to a factor using `labels=paste0("School", 1:23)`

1. Visualize the distribution of SES by schools. Include the school level mean in your graph, and ensure your axes labels are readable. Interpret the graph with respect to potential differences in mean SES and differences in variability of SES across schools.

2. Fit two models on the student's `math` score as the outcome (1) A model with schools and SES as fixed effects and (2) The same model but with a random intercept for each school. Report and interpret the point estimate and standard deviation for the effect of SES on math score under each model. Comment on similarities & differences.

2a. Calculate and interpret the intraclass correlation of SES within schools. Calculate this number using BOTH the output from the `summary()` function, and the output from `tbl_regression(your.ri.model.name, tidy_fun = broom.mixed::tidy)`.

2b. Use `plot_model()` to create side by side plots of regression results from both the fixed and random effects model using `plot_model(your.ri.model.name, type="re")`. Pick two schools and interpret their estimates.

2c. Compare the two models using AIC and BIC. Which one has a better fit?

3. Visualize and describe the linear relationship between `math` score and the students `ses`. Do this overall, and again within each school. Interpret the graph with respect to potential differences in the relationship between SES and math score within and across schools.

Use `#| layout-ncol: 2` in your quarto chunk to place these side by side in your rendered output. If you apply a color by school to your plot, use `theme(legend.position="none")` to remove the legend that is created.

4. Explore the other variables in the model and their relationship on math score. Add a few covariates to this model as fixed effects to improve the model fit.

5. Using your model above with a few additional fixed effects, allow the slope of `ses` to vary within school. Plot the model random effects and re-interpret the effect of school on math score for the two schools you chose earlier.

6. Does SES contribute significantly to explaining a child's math score? Does your conclusion change depending on the model you use? Justify your reasoning with specific examples and numbers (e.g. estimates, se's, ICC's etc.)