ASSIGNMENT #5

EPsy 8252

This assignment is intended to give you experience in using the `lmer()` function for fitting multi-level regression models. You will use the two datasets *popularLevel1.csv*, and *popularLevel2.csv*, which come from Hox (2002). The codebook for these data is also available. You will use the data to examine predictors of students' popularity, which is the average popularity of a student as rated by his/her classmates. Since the sociometric procedure used to assign a popularity measure asked all pupils in a class to rate all the other pupils, and then assigned the average popularity rating to each pupil, there are likely classroom-level effects. To deal with this, the models you will fit in this assignment will need to account for the within-class variation by including a class-level random effect.

Please submit your responses to each of the questions below in a printed document. Please adhere to the following guidelines for formatting your assignment:

- All graphics should be resized so that they do not take up more room than necessary and all should have an appropriate caption.
- Any typed mathematics (equations, matrices, vectors, etc.) should be appropriately typeset within the document using Equation Editor, Markdown, or LaTeX.
- All syntax included should be typeset in a monospaced font, appropriately commented and follow the Data Camp Style Guide (`https://teach.datacamp.com/style-guide`).

PREPARATION

To begin the assignment, you will need to merge the `popularLevel2` data into the `popularLevel1` data. This should result in a data frame with 2000 rows and 7 variables.

LMER I

To begin the analysis, you will fit the intercept-only model with a random-effect of intercept. This will provide a benchmark for any subsequent models that you fit. You can label this section as "LMER I" in your Markdown document.

1. Using symbolic notation, including variable names where appropriate, write the equations for the multi-level model.
2. Using symbolic notation, variable names where appropriate, write the composite (mixed-effects) equation for the model.
3. Fit the model in R. Add the results into a summary regression table that will present the results of this model as well as the subsequent models you will fit in the assignment.
4. Compute and interpret the value of the intraclass correlation.

LMER II

Fit a model that includes the random effects of gender and level of extraversion (and intercept) to explain variation in popularity. Answer the following questions. You can label this section as "LMER II" in your Markdown document.

5. Using symbolic notation, including variable names where appropriate, write the equations for the multi-level model.
6. Using symbolic notation, including variable names where appropriate, write the composite (mixed-effects) equation for the model.
7. Fit the model in R. Add the results into your summary regression table.
8. Consider the `Random effects` table in the output. Explain each of the numbers in the `Variance` column (i.e., interpret the values).
9. Consider the `Random effects` table in the output. Explain what the value in the `Corr` column represents and what it means in this situation (i.e., provide an interpretation for this example).
10. Based on the estimated values, explain why there is evidence that the random effect for `female` does not belong in the model.
11. Provide an appropriate interpretation of the fixed effect estimate in the `extra` row.
12. Provide an appropriate interpretation of the fixed effect estimate in the `female` row.
13. Based on the estimated values, compute a 95% confidence interval for the fixed effect of `extra`.

LMER III

Add the Level-2 predictor of teacher expereience to the model fitted in the LMER I section. Answer the following questions. You can label this section as "LMER III" in your Markdown document.

14. Using symbolic notation, including variable names where appropriate, write the equations for the multi-level model.
15. Using symbolic notation, including variable names where appropriate, write the composite (mixed-effects) equation for the model.
16. Fit the model in R. Add the results into your summary regression table.
17. Which of the five fixed effect estimates should be interpreted? Explain.

Model Comparison

18. Based on the estimated value of the deviance, which of the three models fitted thus far seems to fit the best? Explain.
19. Based on the estimated value of the AIC, which of the three models fitted thus far seems to fit the best? Explain.
20. Based on the estimated value of the BIC, which of the three models fitted thus far seems to fit the best? Explain.