ASSIGNMENT #8

EPsy 8252

Spring 2015

This assignment is intended to give you more experience in working with 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.

Assignment Guidelines: Please submit your responses to each of the questions below in word-processed (or Markdown) document. Please label the sections as indicated below within your printed document. All graphics should be resized so that they don't take up more room than necessary and have an appropriate caption. Any equations should be appropriately typeset within the document. There are 15 points possible for the assignment (each question is worth one point).

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. Next, fit the following three models (henceforth 'Model A', 'Model B', and 'Model C').

```
# Model A
popularity ~ 1 + (1 | class)

# Model B
popularity ~ 1 + extra + female + (1 + extra | class)

# Model C
popularity ~ 1 + extra + female + teacherExp + teacherExp:extra + (1 + extra | class)
```

Add these models to an appropriate summary regression table. Include the log-likelihood and AICc values to this table for each model as well.

PSEUDO R^2 MEASURES

1. Compute the proportional reduction in the Level-1 (residual) variance component for Model B and Model C. Add these to your summary regression table.

2. Interpret the proportional reduction in the Level-1 variance component for Model C.
3. Compute the proportional reduction in the Level-2 (intercept, extra) variance components for Model C. Add these to your summary regression table.
4. Interpret the proportional reduction in the Level-2 variance components for Model C.
5. Compute the squared correlation between the observed and fitted popularity values for all three models. Add these to your summary regression table.
6. Interpret the squared correlation between the observed and fitted popularity values for Model C.
7. Explain why it is not unexpected that the proportional reduction in the Level-1 (residual) variance for Model B is greater than zero.
8. Explain why it is not unexpected that the proportional reduction in the Level-1 (residual) variance for Model B and Model C are similar.

EXAMINATION OF RESIDUALS

9. Create boxplots of the level-1 residuals conditioned on classroom for the three models. Be sure that the plots are directly comparable by using the same limits or facetting.
10. In 1–2 sentences comment on what you see in the plot you just created.
11. Create boxplots of the level-2 residuals conditioned on classroom for Models B and C. Be sure that the plots are directly comparable by using the same limits or facetting.
12. In 1–2 sentences comment on what you see in the plot you just created.
13. Create density plots of the level-1 and level-2 residuals from Model C to help you evaluate the assumption of normality for the mixed-effects model.
14. In 1–2 sentences comment on whether you believe the normailty assumption has been satisfied.

PLOTTING FITTED VALUES

15. Create a single display that you believe is the best visual representation of the results of this analysis. There are many viable alternatives for this visualization. In choosing among them, be sure to consider the substantive points you want to make and create a graph that best allows you to highlight these conclusions.