Dichotomous Data Exercise Instructions

Beth Grandfield

3/2/2022

Exercise 3: General Instructions

This exercise is to be completed in lab and uploaded to blackboard. Complete the exercise in groups of 3.

While completing the exercise, fill in the excel file for each model. Note that you need a different excel file for this than before, which can be found on BlackBoard. Upload the completed excel file at the end of the lab via Blackboard.

Exercise 3: Dichotomous Data

In this exercise, you will practice multilevel analysis on dichotomous data.

The data file UTHAI2.csv contains the Thai educational data example. The dependent variable is rep1, a dichotomous variable indicating whether the pupil has ever repeated a class. Independent variables are at the pupil level having had pre-primary school education (pped; yes = 1) and sex (male = 1). At the school level we have independent variable msesc; the mean SES score of each school.

Now we proceed with the actual multilevel analysis. Read in the data and check a summary of the data. To analyze dichotomous multilevel data in R, we need to use the function glmer() instead of lmer(), where 'g' stands for generalized of the generalized linear model. Check the helpfile of glmer() to see which argument you need to provide to glmer() to analyze dichotomous data. Also, you have to adjust the number of points per axis for evaluating the adaptive Gauss-Hermite approximation to the log-likelihood using nAGQ = 10 in order to ensure convergence of the model.

1. Specify the model with the random intercept and no predictors.
    - Inspect the output and compare the results with the model from the lecture. When they are the same, calculate and interpret the ICC.
2. Set up the model with the level 1 predictors male and pped.
    - Inspect and compare the results with the model from the lecture.
3. Obtain the variance of the linear predictor using the obtained regression coefficients for the level 1 predictors (and intercept).
    - Calculate and interpret the explained variance, and the unexplained variance at level 1 and level 2. Inspect and compare the results with the model from the lecture, When they are the same, continue with the next step.
4. Set up the model with the level 2 predictor msesc (grand mean centered).
    - Inspect the output and compare the results with table 6.3, column labelled 'Laplace ML' in the book of Joop Hox. When they are the same (note that the estimated intercept variance differs, this is correct), go on to the next step.

5. Again, obtain the variance of the linear predictor using the obtained regression coefficients for the level 1 and 2 predictors (and intercept).
    - Calculate and interpret the explained variance, and the unexplained variance at level 1 and level 2.
    - Note that if this would be an actual analysis, we would not recalculate the explained variance, as the predictor mses is not significant. But, as it is nice exercise on how to calculate the variance of the linear predictor, the explained variance and the unexplained variance variance at both levels, we will do this just to practice.
    - The correct answer to this question will be seen in the excel answer sheet uploaded **after** class.
6. For the level 1 predictors, check if the slopes are fixed or random.
    - You only have to check for the random slope of the variable male, as convergence of a model that includes a random slope for pped takes very very long.

Note: As we have no random slopes, we also do not calculate cross level interactions: there is no slope variance to explain.