**Assignment 2**

**Multilevel analysis – MSBBSS – Utrecht University**

Piet Jonker (4281306)

Jovan Kalyango (7014430)

Daphne Weemering (3239480)

March 10, 2021

1. **Convert the data into long format. Check the data and recode if necessary.**

The data is converted from wide to long format and the relevant variables are recoded such that they include an interpretable zero value. The dependent variable measuring antisocial behavior (*anti*) already included a meaningful zero value. For the variable time, 1 is subtracted from every value in order to have a meaningful zero value. Furthermore, the variables measuring reading skills (*read*), cognitive stimulation (*homecog*) and mother’s age (*momage*) are all centered around their grand mean.

1. **Check the linearity assumption, report and include plots.**

In order to check whether there is a linear relationship between the dependent variable (*anti*) and the predictors (*time*, *read*, *homecog*, *momage*), we look at the scatterplots between the dependent variable and the respective predictor. The plots are shown in appendix A.

Both *time* and *homecog* show a clear linear relationship with *anti*. The linear and quadratic line overlap, which implies that a linear relationship suffices. However, there is less overlap between the linear and quadratic line for the relationship between *read* and *anti* and between *momage* and *anti*. This implies that there might be a quadratic relationship between these respective predictors and the dependent variable. However, if we look at the curvature of the line, we see that it only slightly differs from the linear line, and hence, the quadratic relationship is not large enough to be included into the model. Also, we do not have a substantive argument to assume that there is a non-linear relationship between these respective predictors and the dependent variable, and hence, we will take them into account as linear relationships.

Another interesting note on the relationship between *read* and *anti* is that there is actually no trend; the linear line is horizontal. Even though there is not really any trend visible between reading skills and antisocial behaviour, we do not yet know how this relationship changes when we control for other variables. Therefore, it is still interesting and meaningful to take this variable into the model.

1. **Check for outliers (don’t perform analyses, just look in the scatterplots), report.**

Looking at the separate scatterplots, we can see that there are just a few students who scored particularly high on the antisocial behaviors score. However, it is unknown whether these high scores are the actual true values of these individuals or whether these high scores are due to measurement error. If we look at the first plot (X = time, Y = antisocial behaviors), we see that it is essentially one individual who scores higher at every measurement occasion. We deem it safe to assume that these values are within a reasonable value and therefore conclude that there are no outliers.

1. **Answer the question: should you perform a multilevel analysis?**
2. **What is the mixed model equation?**

The mixed model equation for the random intercept model:

Where subscript *t* (*t* = 0, …, *T*) denotes the measurement occasion (level 1), and subscript *i* (*i* = 1, …, *n*) denotes the individuals (level 2).

1. **Provide and interpret the relevant results (don’t just copy the output, report the relevant results in APA style).**

In order to determine whether we should perform a multilevel analysis, we compare the intercept only model with the intercept only model that allows for random intercepts. If the model including the random intercepts has a better fit, we should indeed perform a multilevel analysis to these data.

Model 0, the “regular” intercept only model, shows that the *deviance* (2) = 3569.5 and *AIC* (2) = 3573.5. Moving from model 0 to model 1, the goodness-of-fit statistics decrease, with *deviance* (3) = 3337.5 and *AIC* (3) = 3343.5. Moving from a intercept-only to a random intercept model also shows to be a significant improvement of model fit ((1) = 231.97, *p* < .001). The intercepts thus vary significantly across individuals.

1. **What is the intraclass correlation?**

The intraclass correlation is , which implies that 48 percent of the variance in antisocial behavior is variance between the children, and the remaining variance is the variance within children across time.

1. **What is your conclusion regarding the overall question regarding the necessity of performing a multilevel analysis?**

Based on the knowledge we obtained by comparing the intercept only model with the random intercept model, we can conclude that it is necessary for these data to perform a multilevel analysis. The model including the random intercepts shows to be a significantly better model fit compared to the “regular” intercept only model ((1) = 231.97, *p* < .001). This implies that the intercepts vary significantly across the different children (i.e. the children have different starting scores on antisocial behavior), and multilevel analysis is useful for analyzing the data. Also, an intraclass correlation of 48 percent shows that there is dependence of observations, which can cause underestimation of standard errors when we use more conventional tools for analyzing the data (e.g., ANOVA, multiple regression). This underestimation of standard errors can cause an inflation of type I errors. This problem with dependent observations can be solved by analyzing the data using multilevel analysis.

1. **Add the time-varying predictor(s). Provide and interpret relevant results, and provide your overall conclusion.**

In model 2, time is added as a linear predictor with the same coefficient for all the school children. At the first measurement occasion, the model predicts a value for antisocial behavior of 1.55, which increases with 0.18 at every succeeding measurement occasion. Time is a significant predictor (*t* = 4.51, *p* < .001) in this model. Looking at the goodness-of-fit statistics, the *deviance* (4) = 3317.5 and *AIC* (4) = 3325.5, we see that both the deviance and the AIC decreased when we moved from the random intercept model (model 1) to the model including time as predictor (model 2). This improvement in goodness-of-fit of the model also shows to be a significant improvement ((1) = 20.06, *p* < .001). Based on the results we can conclude that adding time as a predictor significantly improves the model, and therefore, time will be included in the succeeding models.

Then, in the next model (model 3), we add the time-varying predictor reading skills of the children. The effect of reading skills is non-significant (*t* = -0.54, *p* = 0.59). The goodness-of-fit statistics of model 3 also do not show a very convincing decrease, and the AIC even increased compared to model 2 (*deviance* (5) = 3317.2 and *AIC* (5) = 3327.2). Not surprisingly, then, the difference of deviance test shows to be non-significant ((1) = 0.29, *p* = 0.59). Since this predictor has no significant effect on antisocial behavior, we will remove it from the succeeding models. Notice that this model is included in the table because it is discussed here in the text. However, due to the insignificance of the model, it will not be used to compare other models with (for the difference in deviance).

1. **On which level or levels can you expect explained variance? Calculate and interpret the relevant results, and provide your overall conclusion.**

Model 2 is added as a baseline model for calculating the explained variance. Multilevel modelling assumes a hierarchical sampling model. For a time series design, the variability between subjects in the measurement occasion variable is in practice much higher than the hierarchical sampling model assumes. As a consequence, the random intercept model overestimates the variance at the occasion level (level 1) and underestimates the variance at the subject level (level 2) (Hox et al., 2017). The model including only the time variable as a predictor (model 2) uses this predictor to model the occasion level variance in the dependent variable, antisocial behavior. Therefore, the model including time as the only predictor, which is model 2, is used as a baseline model for calculating the explained variance, so that the variances estimated at the measurement occasions and the subject level are more realistic.

Then, for answering the question, we can expect explained variance on both the measurement occasions level and the subject level for the time-varying predictor reading score. We can explain variance at both levels, because reading score is measured at both the measurement occasion level (it is measured at several points in time), and on the subject level (reading score per individual child). The reading score explains (1.689 -1.693) / 1.689 = -0.002 or -0.2 percent of the variance, indicating that the effect of reading score on antisocial behavior does not vary across the different measurement occasions. The time-varying predictor reading score explains (1.592 – 1.576) / 1.592 = 0.01 or 1 percent of the variance between the school children, indicating that the children differ on their reading scores. We already showed that adding the predictor reading score does not significantly improve the model. Here, we also show that this predictor does not explain a lot of variance, it even has a negative explained variance on the occasion level. Therefore, we can rightfully remove the predictor reading score from the succeeding models.

1. **Add the time invariant predictor(s) in the model. Provide and interpret the relevant results, and provide your overall conclusion.**

The time invariant predictors are cognitive stimulation and mother’s age. These predictors are time invariant because they are only measured at the first measurement occasion; the value for these predictors remains the same over all the measurement occasions.

The effect of mother’s age on antisocial behavior is non-significant (*t* = -0.02, *p* = 0.98). This predictor will be removed from the model. However, the effect of cognitive stimulation does show to have a significant effect on antisocial behavior (*t* = -3.35, *p* < .001). The negative relationship (*b* = -0.13) implies that children who get more cognitive stimulation at home will have a lower score on antisocial behavior. Then, we can look at the goodness-of-fit statistics: *deviance* (5) = 3305.8 and *AIC* (5) = 3315.8, which are both lower compared to the previous model (model 2, since we do not compare with model 3). Moving from model 2 to model 4 is a significant improvement in the goodness-of-fit of the model ((1) = 11.64, *p* < .001). Cognitive behavior does have as significant effect, and hence, this variable will remain in the succeeding models.

1. **On which level or levels can you expect explained variance? Calculate and interpret the explained variances.**

After we added the time-invariant predictors, we can only expect additional explained variance on the subject level (level 2). As mentioned before, cognitive stimulation and mother’s age are both time-invariant predictors, meaning that they are only measured at the subject level and not at the occasion level. This is due to the fact that they are only measured once, and thus, they do not vary over time. No additional variance is therefore explained on the measurement occasion level. The model including cognitive stimulation (mother’s age is removed from the model) explains (1.592 – 1.488) / 1.592 = 0.065 or 6.5 percent of the variance between the school children. This implies that the addition of the predictor cognitive stimulation explains an additional (0.068 – 0.010) 5.8 percent of the variance (compared to model 2).

1. **For the time-varying predictor(s), check if the slope is fixed or random.**
2. **What are the null- and alternative hypotheses?**

Until this point, we have assumed that the slopes for the time-varying predictors are the same for all the children. Now, we are going to determine whether the slopes of the reading score and the slopes of the time variable vary across the children. The null hypothesis is that the slopes for the reading scores and the time variable do not vary across children. The alternative hypothesis is that the slopes for the reading scores and the time variable do vary across children.

1. **Provide and interpret relevant results.**

First, we checked whether the time predictor has varying slopes across the children (i.e. do the rates of change differ across the individuals?). In order to determine whether adding random slopes for the time variable is a significant improvement, we need to compare this current model with the previous model, namely model 4 where cognitive behavior was included as (time-invariant) predictor. Comparing the current model, including the random slopes for time, with the previous model, we can conclude that moving to the current model is a significant improvement ((2) = 26.56, *p* < .001). The addition of random slopes for time improves the model fit, and therefore we can assume that time has varying slopes. Or, in other words, the rates of change in antisocial behavior differ across individuals.

Then, we can add random slopes for the predictor reading scores. First of all, if we add random slopes for the variable reading score, the model does not converge anymore. Lack of convergence is an indication that the data does not fit the model well. This might be due to too many poorly fitting observations. This is already an indication that the addition of random slopes for the reading scores is not great. Then, if we look at the difference in deviance between this model and the previous model, which includes random slopes for time, we see that this difference is not significant ((3) = 0.24, *p* = 0.97). The addition of random slopes for the reading score does not improve the model fit, and therefore, no random slopes for the reading scores are included in the succeeding models. Notice that this model is also added to the table, because it is discussed here in the text. However, this model will not be used to compare other models with (in difference of deviance).

1. **Provide an overall conclusion.**

Based on the performed analyses, we conclude that the random slopes for time should be included, because this significantly improves the model fit. Adding random slopes for time is thus better than a model without random slopes for time. Substantively, this means that the scores on the antisocial behavior scale over time are allowed to differ across the children; some children might have a steeper growth or decrease in antisocial behavior over time than others. Adding random slopes for the reading score would not improve the model fit, and therefore we have decided to not include random slopes for the reading scores in the succeeding models.

1. **If there is a random slope, set up a model that predicts the slope variation. Provide and interpret the relevant results, and provide your overall conclusion.**

The effect of the random slope of time could be predicted by individual-level predictors cognitive stimulation and/or mother’s age. We add these predictors as cross-level interactions to see if this is actually the case.

First, the interaction between time and cognitive stimulation (*time\*homecog*) is added to the model. Looking at the relevant goodness-of-fit statistics (*deviance* (8) = 3272.4 and *AIC* (8) = 3288.4), we see that these decreased compared to the previous model, which included random slopes for time. Comparing these two models, we can conclude that this current model is a significant improvement compared to the previous model ((1) = 6.88, *p* < .001).

Then, the interaction between time and mother’s age (*time\*momage*) was added to the model. Again, we look at the relevant goodness-of-fit statistics (*deviance* (10) = 3272.1, *AIC* (10) = 3292.1), and we see that the deviance slightly decreased, but the AIC increased compared to the previous model. This already indicates that adding the interaction between time and mother’s age does not improve the model fit. Then, looking at the difference in deviance test, we can conclude that this model does not significantly improve the model ((2) = 0.30, *p* = 0.86).

Substantively, the addition of the interaction between time and cognitive stimulation means that the amount of change over time for each child also depends on the amount of cognitive stimulation at home. The explained variance for the cross-level interaction for the slopes of time is (0.096 – 0.084) / 0.096 = 0.125, which means that we can explain 12.5 percent of the variance in the random slopes.

1. **Decide on the final model.**
2. **Provide the separate level 1 and level 2 equations, as well as the mixed model equation.**

Level 1:

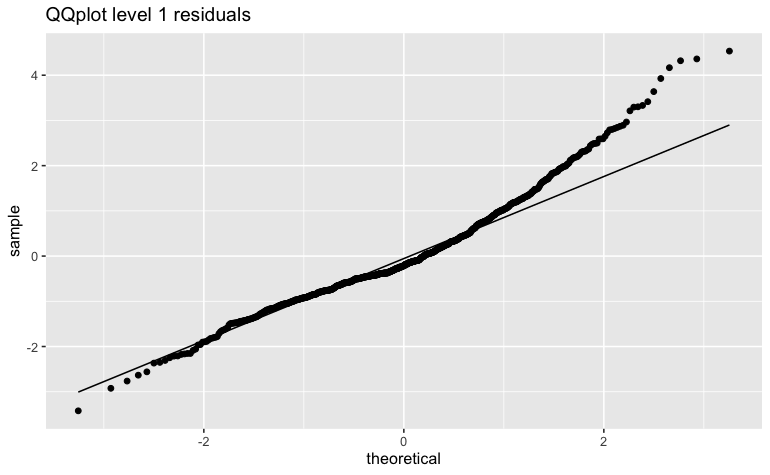
Level 2:

Mixed:

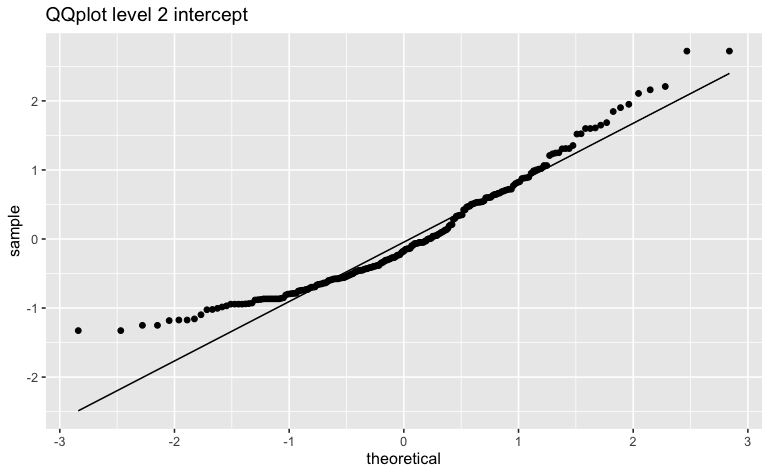
Where subscript *t* (*t* = 0, …, *T*) denotes the measurement occasion (level 1), and subscript *i* (*i* = 1, …, *n*) denotes the individuals (level 2).

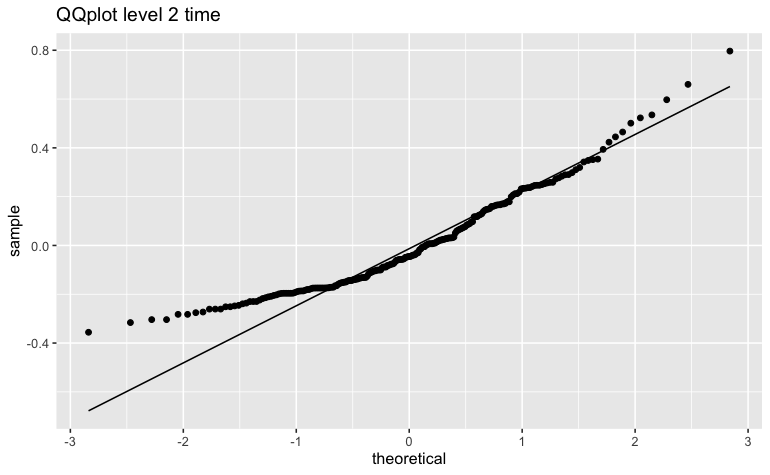
1. **Check the normality assumption for both the level 1 and level 2 errors, report.**

Level 1 errors:



Level 2 errors:





The level 1 residual errors seem to violate the assumption of being normally distributed. Also, the level 2 residuals of intercept and time seem to violate the assumption of normally distributed errors. We have tried natural log, square root and cube root transformations. None of these transformations resulted in normally distributed residual errors. We have to be aware of this flaw in our model and interpret it carefully. For all the plots, we see that the beginnings and the ends of the plots move away from the diagonal line. For the level 1 errors, we can see that the distribution of the residuals is skewed to the right, since the upper part of the plot is particularly moving away from the diagonal straight line. Then, if we look at the first plot for the second level residuals, we see that the residuals are essentially peaked at the middle of the distribution. The distribution will have a somewhat ‘fatter’ tail. The second plot for the second level residuals shows that these residuals are essentially skewed to the left, since the lower part of the plot deviates from the straight diagonal line.

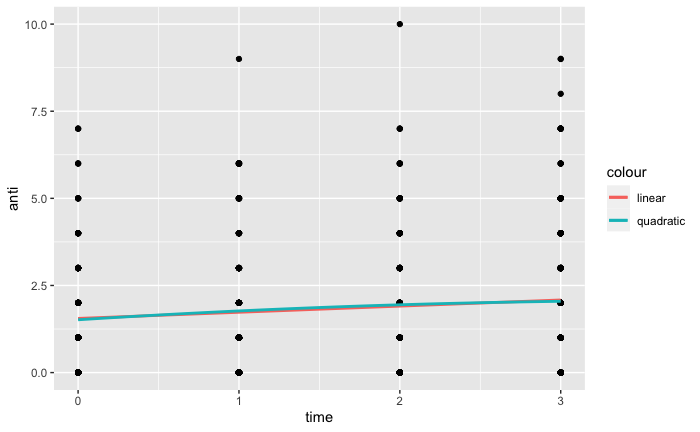
The fact that the residuals do not lie perfectly on the straight diagonal line of the Q-Q tells us something about the skewness and kurtosis of the distribution. However, none of these residuals show directly to be perfectly normally distributed.

**References**

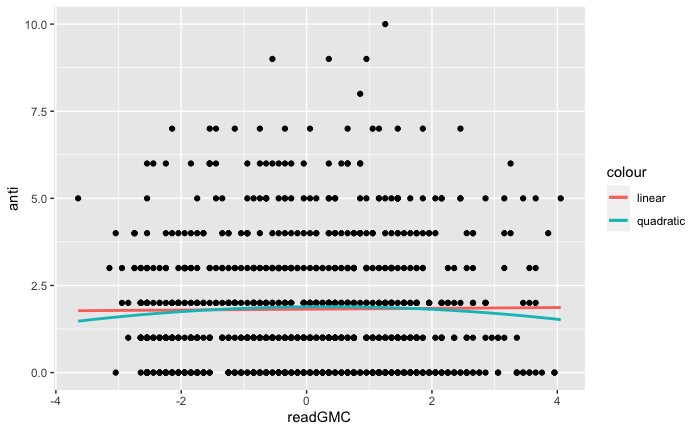
Hox, J. J., Moerbeek, M., & Schoot, V. R. (2017). *Multilevel Analysis: Techniques and Applications, Third Edition (Quantitative Methodology Series)* (3rd ed.). Routledge.

**Appendix A: Scatterplots between predictors and dependent variable**

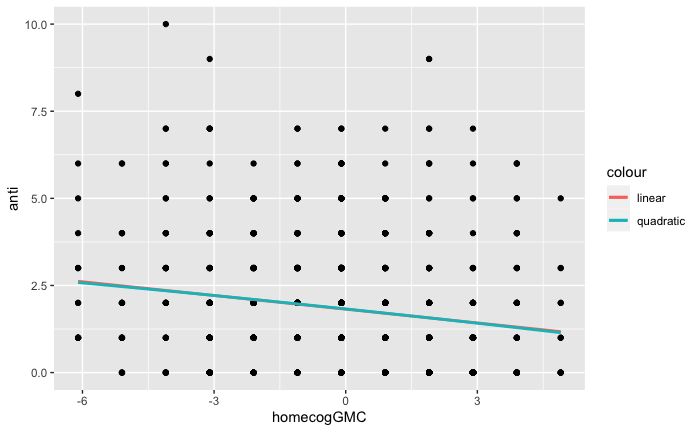
1. Time (X) and antisocial behavior (Y).



1. Reading score (X) and antisocial behavior (Y).



1. Cognitive stimulation (X) and antisocial behavior (Y).



1. Mother’s age (X) and antisocial behavior (Y).

