**Kuang Li Report**

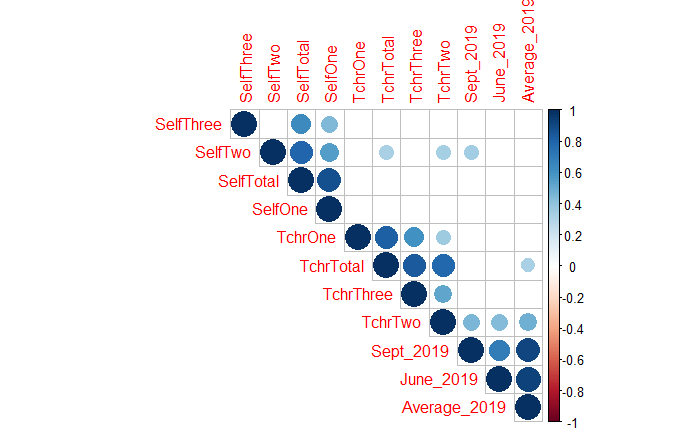
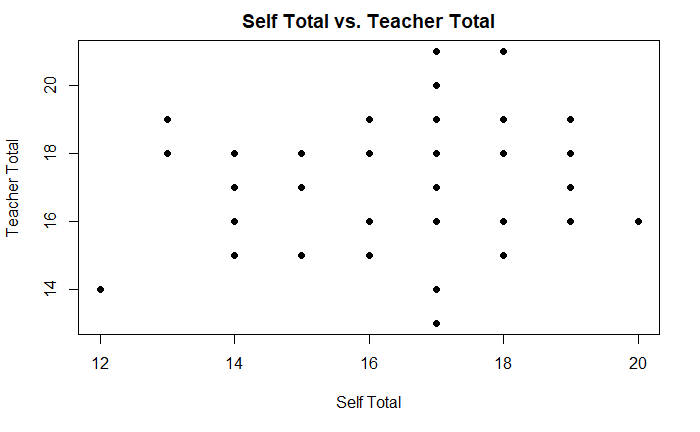
**Jason Lu, Jung Hwa Yeom, Jinzhe Zhang, Xiaozhou Lu**

1. **Introduction**

Using a quasibinomial model to compare student self-assessment and teacher assessment for a

writing task. For this study, 12th grade students from the same school, taught by the same teacher, were requested to self-assess their writing based on a rubric (10 points for Content, 10 for Language, 5 for Organization + Structure, 25 points total). Then, the teacher will grade the writing without access to student self-assessed scores. There are two datasets used in the study with the first being scores of 67 students who participated in the assessment and the second being a subset of students from the first dataset participating in the assessment again, however we will only utilize the former of the two datasets. This is because we found that students from the second session would score themselves higher than they did on the first session so the session will become a confounding variable.

1. **EDA**



An initial look of student self-evaluation score vs. teacher score as well as correlation plot of each variable against each other. The scatterplot shows that there is no clear linear relationship between student self-assessment for total score and teacher assessment for the total score.

1. **R Model**

To begin with, independent variables include:

* TchrTotal – teacher assessment for total score
* SelfTotal – student self-assessment for total score
* SelfOne – student self-assessment for Content (10 points)
* SelfTwo – student self-assessment for Language (10 points)
* SelfThree – student self-assessment for Organization (5 points)
* Average\_2019 – average points student earned from the 2019 June Exam and 2019 September

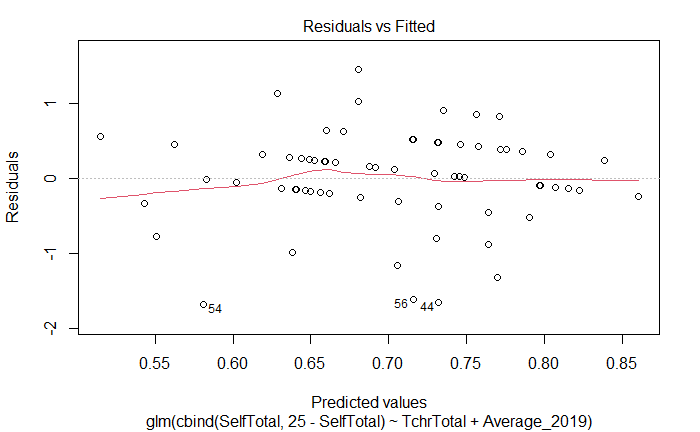
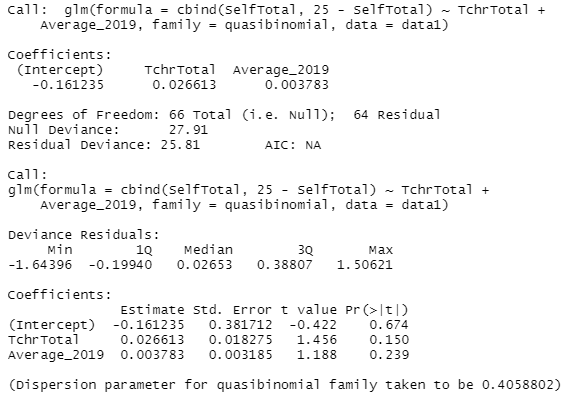
First, we plotted the scatterplot to see the distributions of the observations. From this initial look, there does not seem to be a meaningful correlation between teachers’ total assessment and students’ self-total assessment. Then we fitted a model. Because our response variable had a ceiling of 25 points, we chose to fit a binomial generalized linear model. However, there was under-dispersion after fitting the binomial model, so we used a quasi-binomial model to correct that.

The quasi-binomial model was coded as:

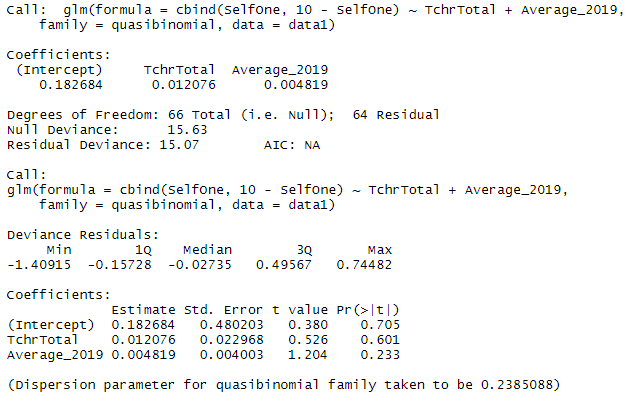
glm(cbind(SelfTotal, 25-SelfTotal)~TchrTotal + Average\_2019, data=data1, family=quasibinomial)

At this time, the use of the binomial distribution assumption did not affect the significance of the coefficient detection. The first thing was to detect whether the sample has over-dispersion or under-dispersion. Then, the alternative method was to use the quasi-binomial distribution, which fitted an extra dispersion parameter to account for that unexpected variance. We divided the residual by the residual degree of freedom. If it is lower than 1 (The actual ratio is 64/66 ≈ 0.97), it means under-dispersion that is the phenomenon that the residual variance is smaller than the conditional mean.

**Quasibinomial model: Response as SelfTotal**



**Quasibinomial model: Response as SelfOne**



For the residual plot, we can see that the red line is close to the dashed line, indicating there is a linear relationship. Therefore, it suggests the assumption is reasonable. Furthermore, points 44, 54, and 56 are the outliers since they have large residual values.

1. **Discussion**

We used a generalized linear model over correlation since it can incorporate other confounding variables such as Average2019 in the analysis. We can also check model validation assumptions for regression which is not possible to do for correlation. In our R code, we used a binomial generalized linear model since the response variable, the score for writing tasks, has a total score of 25. Due to the underdispersion in the data, we resulted in a quasi-binomial GLM to adjust the standard error calculation.

In coefficients, TchrTotal p-value was 0.150. Based on the result, we can conclude that there is no strong evidence for the association between TchrTotal and SelfTotal as the coefficients were not significant at 0.05 alpha level. To examine the validity of our model, we plotted the residual plot of the quasibinomial model. Those points on the residual plot are evenly distributed along average residual equals zero, and most of the residuals are near to zero without any specific trends (ex. fanning pattern), indicating that the model fits well.

To sum up, we could not say there is a correlation between teachers’ total assessment and students’ self-total assessment. Lastly, we can generalize only to this teacher-student class since the class is taught by the same teacher. Thus, this cannot be made in general to all classes.

1. **References**

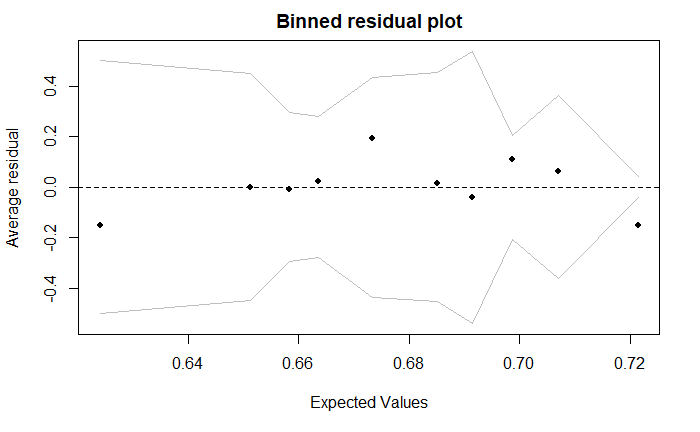
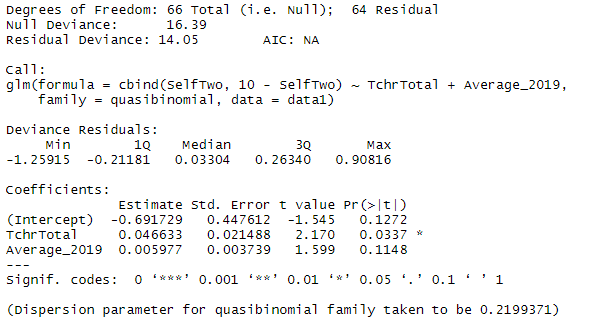
http://www.sthda.com/english/wiki/correlation-matrix-a-quick-start-guide-to-analyze-format-and-visualize-a-correlation-matrix-using-r-software

http://www.sthda.com/english/articles/39-regression-model-diagnostics/161-linear-regression-assumptions-and-diagnostics-in-r-essentials/#regression-assumptions

<http://www.john-ros.com/Rcourse/lme.html>

1. **Appendix**
   1. Model summary and binned residual plot of quasibinomial for

Self2 ~ TchrTotal + Average\_2019



* 1. Binned residual plot of quasibinomial for

Self3 ~ TchrTotal + Average\_2019

