Project 2

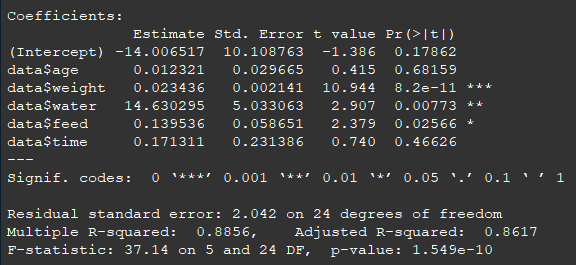
MA3740

Ruben Pena

1a. All 5 predictors:

y= .0123(*Age*) + 0.0234(*Weight)* + 14.6303(*Water*) + 0.1395(*Feed*) + 0.1713(*Time*) – 14.0065

R2 =.8856, RSE=2.042

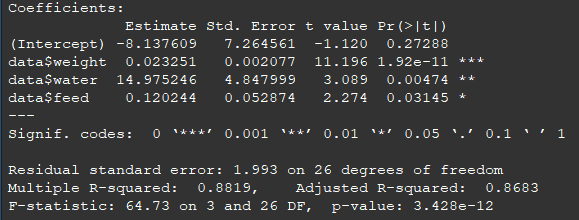


I ran 4 different models dropping 1) *Age*, 2) *Time*, 3) *Age* and *Time*, for all the R2 values were ~88%, but the model with the optimal RSE was a model without *Age* and *Time*.

These variables can be dropped because their p-values indicate that they are insignificant in the prediction of the size of eggs compared to the other variables. This can be seen above in the R output above.

1b. Without *Age* and *Time*:

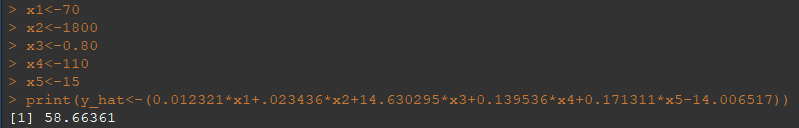
y= 0.0232(*Weight)* + 14.9752(*Water*) + 0.1202(*Feed*) – 8.1376

R2=88.19, RSE=1.993

The R2 value decreased very slightly, but the RSE decreased meaning that the regression line has fewer residuals. With both models having similar R2, optimizing the RSE is beneficial and dropping the variables makes the linear model slightly easier to interpret.

1c. Prediction

The average egg weight for a chicken with the values given is: 58.66361 grams



2a. What non-parametric test is appropriate for this situation?

Non-parametric, independent samples: **Mann-Whitney**.

2b and 2c. See Attached.

3a. What non-parametric test is appropriate for this situation?

Non-parametric, paired samples: **Wilcoxon.**

3b and 3c. See Attached.