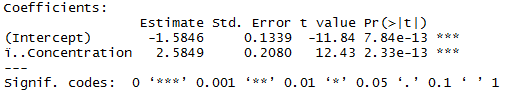
**James Young Statistics Take Home 2**

All included snips are from R

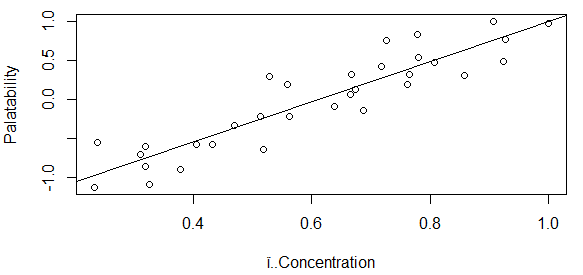
**a. State the linear regression model and define all terms.**

,

Where and represent the values of independent and dependent variables, respectively. represents the intercept, represents the slope of the regression line, and , is variables coming from a normal distribution with mean zero and variance 2.

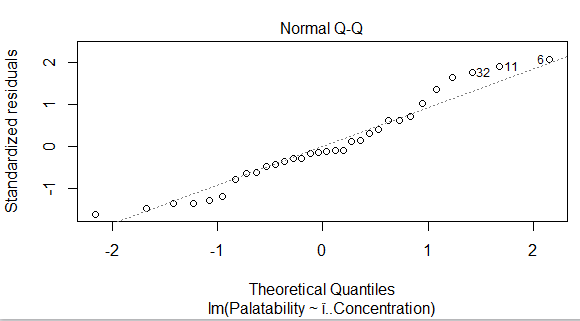
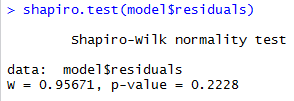
**b. Calculate the estimated regression line to predict Palatability based on Concentration and give the estimated regression function.**

= -1.5846 + 2.5849x

**c. Plot the estimated line and the actual data points. Comment about the strength of the linear relationship.**

The linear relationship appears consistent

and balanced throughout the data.

**d. State the assumptions necessary for linear regression as given in Section 7.8. Perform and interpret regression diagnostics and comment on the validity of the assumptions.**

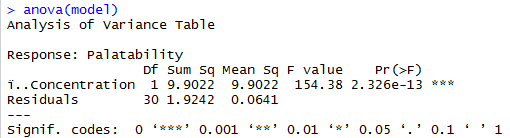
1. The linear regression is appropriate

2. The error terms are independent

3. The error terms are normally distributed

4. The error terms have common variance, .

Distribution and residuals seem normal with a small distortion on the tails of the qq plot.

**e. Test if 𝛽1 is equal to zero. State the hypotheses, calculate the test statistic, provide the decision rule, and state your conclusion in the context of this scenario.**

H0: = 0

H1: 0

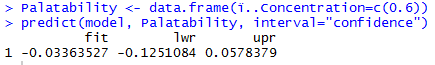
F = = 9.9022/0.0641 = 154.38

The calculated F value is 154.38. The critical F value with 1 df in the numerator and 13 df in the denominator is 4.67. 154.38 > 4.67, therefore we reject H0, and conclude that does not equal 0.

**f. Calculate and interpret a 95% confidence interval for the slope of the true regression line.**

We can say with 95% confidence that the slope

is [2.160,3.010].

**g. Estimate true average Palatability when Concentration is .6 and do so in a way that conveys information about reliability and precision.**

We predict the true average palatability score is -0.0336

when concentration is 0.6. We can say with 95% confidence that the palatability is

[-0.1251,0.0578]. This is reliable for an average but less reliable for a single sample with a precision better than the prediction interval and relative to all possible response values within the model.

**h. Predict Palatability for a single beverage sample when Concentration is .6 and do so in a way that conveys information about reliability and precision.**

We can predict that a single beverage sample at

concentration 0.6 will have a palatability

score [-0.5589,0.4916]. This is reliable but with a precision lower than the 95% CI.

**i. Does it appear that true average Palatability for a Concentration of .7 is something other than 0? State and test the appropriate hypotheses.**

H0:

H1: 0

0 does not fall in the 95% confidence

interval so reject H0 and conclude true average palatability is something other than 0 when concentration is 0.7.

**R Code Appendix**

