**In Class Work for Two-Way ANOVA**

**Data Set:** DrillBit.sav (an SPSS data file)

**Information:** Data were collected on the force generated with two drill speeds and four feed rates. Although the data appear quantitative, the numerical values are really categorical in nature which makes ANOVA a more appropriate method for analysis

**Variables:**

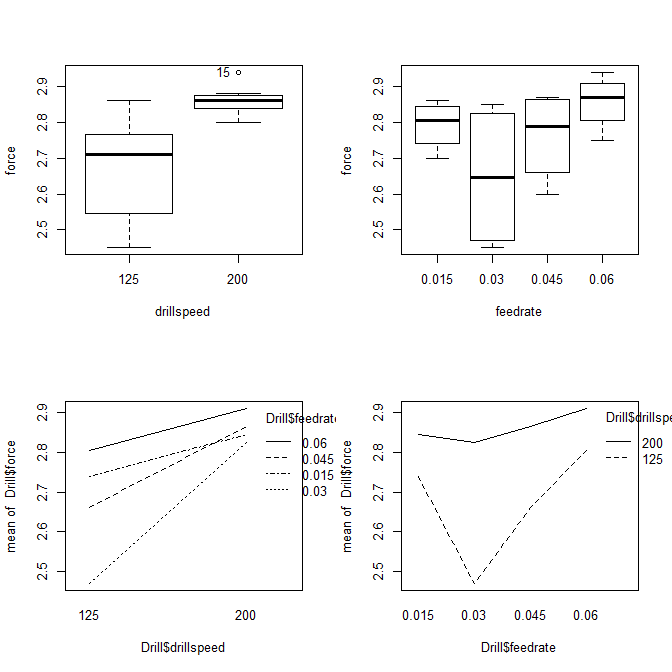
1. drillspeed
2. feedrate
3. force

**Getting the Data into R**

1. Download the data set and import it into R. Note that you will need to select SPSS data file instead of a .txt file this time.
2. View the data in R
3. Click on “Statistics”, “mean”, and then select “Two-Way ANOVA”. What happens?
4. R is viewing the categorical variables as quantitative variables because they are entered in as numbers. You will have to change them to factor variables in order to perform ANOVA. To do this
   1. Select “Data”
   2. Select “Manage Variables in Active Data Set”
   3. Select “Convert Numeric Variables into Factors …”
   4. Select “drillspeed” and “feedrate” as the two variables to convert. Click “OK”
   5. It will ask you if it is ok to overwrite “drillspeed”. Click “OK”
   6. Then it will ask you to rename the factor levels. You can rename them to the numbers that they already are. R will recognize them as categorical data instead of quantitative data even though there are still numbers listed for the variable.
   7. Repeat the process for “feedrate”. Note: when I put in the values for feedrate R gave me an error, but it seemed to complete the process just fine.
   8. The data should be ready for Two-Way ANOVA

**Preliminary Data Analysis and Testing Assumptions**

1. Divide your plot window into four panes by typing par(mfrow=c(2,2))
2. Before doing anything else, you should get a feel for your data. Examine the main effects by making side-by-side boxplots for the main effects
3. Investigate the interaction by making and interaction plot
   1. To make an interaction plot do the following:
   2. Type: interaction.plot(Drill$DrillSpeed, Drill$FeedRate, Drill$Force) into the script window. This will provide one of the plots
   3. Type: interaction.plot(Drill$FeedRate, Drill$DrillSpeed, Drill$Force) into the script window. This will provide the other plot



1. What does your preliminary data analysis tell you?

It appears that the drill speeds effect force differently and that the feed rates also likely effect force differently

The interaction plots show that there is an interaction

The distributions for both feed rate and drill speed look pretty good and there are no major outliers

1. Make QQ plots for the applicable categories. You may want to redo the par(mfrow=c(,)) setting to match the number of plots that you have.

par(mfrow=c(3,2))

qqPlot(Drill$Force[1:8], dist="norm", main="QQ-Plot for Drill Speed 125")

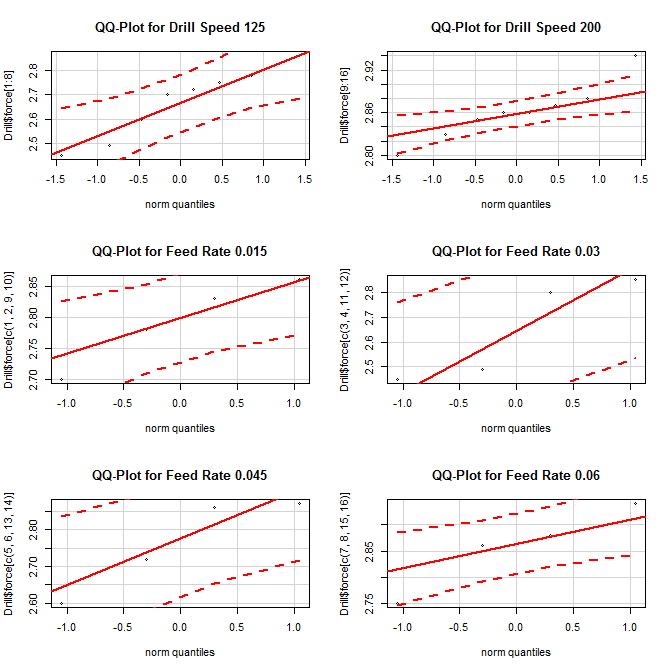
qqPlot(Drill$Force[9:16], dist="norm", main="QQ-Plot for Drill Speed 200")

qqPlot(Drill$Force[c(1,2,9,10)], dist="norm", main="QQ-Plot for Feed Rate 0.015")

qqPlot(Drill$Force[c(3,4,11,12)], dist="norm", main="QQ-Plot for Feed Rate 0.03")

qqPlot(Drill$Force[c(5,6,13,14)], dist="norm", main="QQ-Plot for Feed Rate 0.045")

qqPlot(Drill$Force[c(7,8,15,16)], dist="norm", main="QQ-Plot for Feed Rate 0.06")



1. Check to see if the variability assumption is met. This must be done independently for the main effects.

**Drill Speed**

mean sd IQR 0% 25% 50% 75% 100% data:n

125 **2.66875** 0.14327172 0.1850 2.45 2.5725 2.71 2.7575 2.86 8

200 **2.86125** 0.04051014 0.0275 2.80 2.8450 2.86 2.8725 2.94 8

**Feed Rate**

mean sd IQR 0% 25% 50% 75% 100% data:n

0.015 2.7925 0.06994045 0.0775 2.70 2.7600 2.805 2.8375 2.86 4

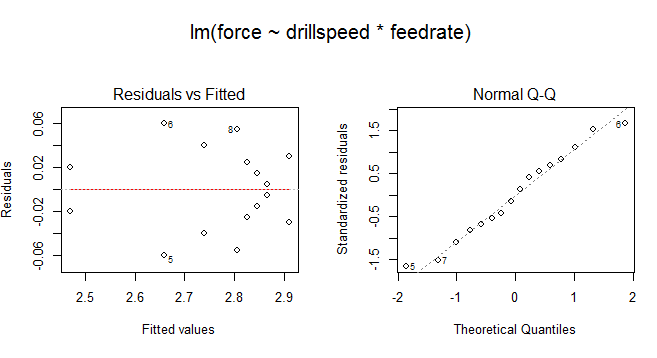
0.03 **2.6475** 0.20661962 0.3325 2.45 2.4800 2.645 2.8125 2.85 4

0.045 2.7625 0.12816006 0.1725 2.60 2.6900 2.790 2.8625 2.87 4

0.06 **2.8575** 0.07932003 0.0625 2.75 2.8325 2.870 2.8950 2.94 4

Assumption is met for both main effects (smax/smin < 2 for both factors)

1. There are other assumptions about residuals. These is the same assumption that we have in regression.
   1. The residuals are normally distributed
   2. The plot of the residuals against the predicted values should look like a cloud (no trend and no heteroscedasticity)
   3. You can verify these assumptions by clicking on “Models”, “Graphs”, “Basic Diagnostic Plots”. You only need to look at the top two plots.



The residuals appear to be normal, but the residual plot shows some heteroscedasticity.

1. Did you meet all of the assumptions?

Mostly

1. Which, if any, did you fail?

Residuals are a little dodgy, but not too worrisome

1. How bad did you fail them?

See #9

**Start the Test**

1. List the hypotheses for Two-Way ANOVA in terms of this problem

H0: µ150=µ200

Ha: µ150≠µ200

H0: µ0.015=µ0.03=µ0.045=µ0.06

Ha: The mean force is not equal for all feed rates

H0: There is no interaction between drill speed and feed rate

Ha: There is an interaction between drill speed and feed rate

1. Generate the ANOVA table
2. What are the test-statistics and p-values?
3. State your conclusions for ANOVA in terms of your hypotheses and in layman’s terms
4. Now what should you do? How do you think you should do it?

**Multiple Means Comparison Testing**

1. This has to be done through One-Way ANOVA
   1. Perform a one-way ANOVA on “drillspeed”
   2. Check the box for “Pairwise Comparison of Means”
   3. Which comparisons are significant?
   4. Repeat the process for “feedrate”
   5. Which comparisons are significant?

**Conclusions**

1. Write a paragraph or two that summarizes what you found in this analysis.