**CB[1] – Enrichment Key**

**Type in your score here 🡪 \_\_\_\_ out of 49 points possible**

1. (5 points) Ponder/Reflect Exercise – Reflect on what you have learned from this portion of the class. Examples of what you can do are: a brief outline of material covered, insights you gained from class or personal study, or items you feel that you need to follow up or work on. (3-5 sentences)

**Any thoughtful answer is sufficient.**

2. #B1 on pages 256-7

a) (4 points)



b) (2 points)

**Roughage – 695**

**Partial Grain – 811**

**Full Grain – 957**

c) (3 points)

**It appears on average that milk yield went up in the middle period (weeks 7-12), but dropped back down in the last period (weeks 13-18). It doesn’t appear to be conclusive however that milk yield over time drops.**

d) (2 points)

**It appears that the cows factor has a bigger difference, particularly with Cow I that has an average yield of 723 vs. Cow II and III with average milk yields of 894 and 846 respectively.**

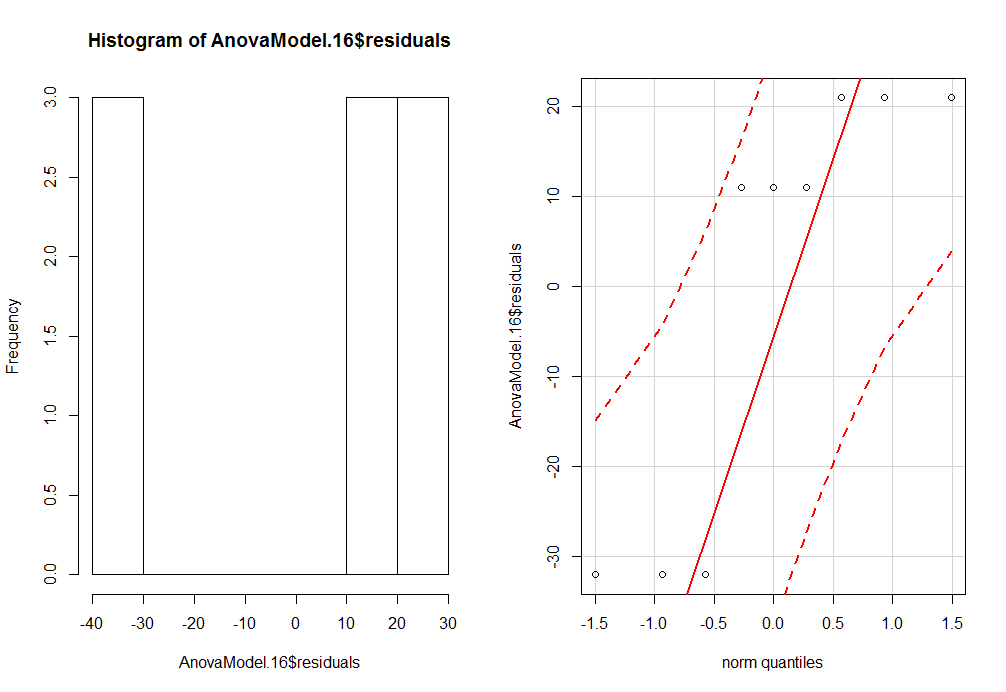
e) (2 points)

**The factor of interest is experimental. We are randomly assigning when a cow gets a treatment by randomly reordering the cows and then the time periods when the cows get certain treatments.**

**It is unusual for the factor of interest to be observational since we typically randomize the material that we currently have (like cows) for an LS design rather than randomly selecting material each from several populations to be a part of the analysis.**

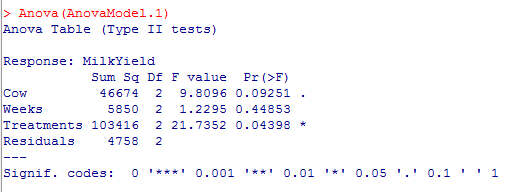
3. Now use the data from the previous problem (found in Figure 7.7 on page 254) to conduct the formal analysis using software (SAS or R). For R, you can either (1) type the data into an excel spreadsheet, save the file as a .csv format, and read in the data; or (2) type the data directly into R. For SAS, the SAS filename is *milk*. Please do the following:

1. (5 points) Check the assumption of residuals being normally distributed using **software**.



**The data does not appear to be normally distributed,**

1. (4 points) Get an ANOVA table using **software**



1. (5 points) For the effect of interest: i) state the null and alternative hypotheses, ii) give the test statistic, iii) give the degrees of freedom, iv) state the p-value, v) determine whether you should reject or not reject the null hypothesis, and vi) write a sentence which gives an appropriate conclusion.

**Ho: µMethod1= µMethod2= µMethod3 Ha: at least one of the population means is different**

**Test Statistic F=21.735**

**Degrees of Freedom Num df=2 Den df=2**

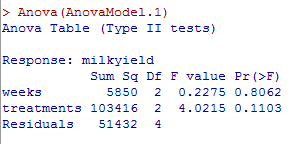
**P-value = 0.044**

**Since the p-value is less than 0.05, we would reject the null hypothesis**

**Therefore, we have sufficient evidence that at least one of the population means is different**

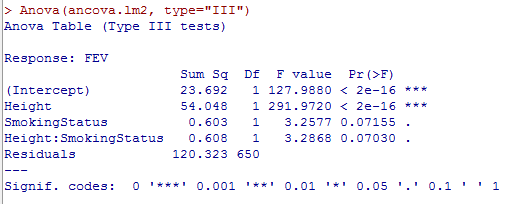
1. (2 points) Did the nuisance variables (cow and time period) have substantial impact on the yield?

**It appears that cow had more of a substantial impact on yield. If we would have pulled out the cow factor and run the analysis, then the treatment factor would no longer be significant. (see below)**



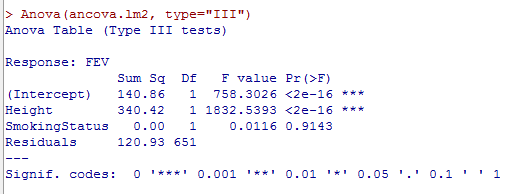
4. A cross-sectional study was conducted in East Boston to determine if smoking affects the pulmonary (lung) function of youths. The subjects in this observational study reported their age, gender, and if they currently smoke. The researchers also measured the height and forced expiratory volume (FEV) of each of the subjects. FEV is a measure of pulmonary function, and higher values are more favorable. FEV is measured as the volume of air (in liters) that is expelled in the first second after the lungs are filled to capacity. The primary objective of this study was to determine if smoking (SmokingStatus) affects the subjects’ FEV. Use height as a covariate in the analysis. Use α=0.05.

1. Check if there is an interaction between the covariate and Smoking Status (Show the table)(3 pts).

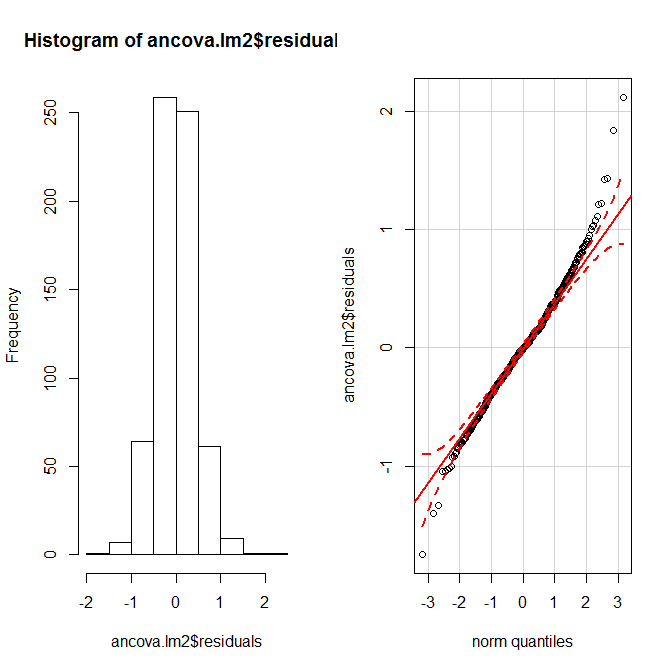


**Since the interaction is not significant we can do an ANCOVA.**

1. Run the ANCOVA model (show the table) and summarize the results (3 pts).

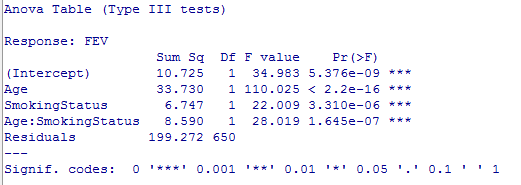


1. **Ho: µs= µns Ha: at least one of the population means is different (µs ≠ µns )**
2. **Test Statistic F=0.012**
3. **Degrees of Freedom Num df=1 Den df=651**
4. **P-value = 0.914**
5. **Since the p-value is greater than 0.05, we would NOT reject the null hypothesis.**
6. **Therefore, we have insufficient evidence that at least one of the population means is different**
7. Do a qqplot and histogram of the residuals to see if we can assume normality (3 pts).



**There are a few outliers but for the most part, the residuals look normally distributed.**

1. Now use Age as a covariate. Check to see if there is an interaction between the covariate and Smoking Status (Show the table). What do you conclude based on this (3 pts)?



**The interaction between smoking status and age is very significant so, ANCOVA would not be appropriate, but this would be a great model that you might likely see in Math 425 (Regression) ;-)**