**Math 326 – Experimental Design**

**Exam 2 (100 Points)**

**Instructions:**

* Do not look at the exam until you are ready to start.
* **Record your start and stop times on the exam (even if it is multiple times on this page)**. You can take the exam in several sittings.
* You cannot use your class notes, homework solutions, your book, or anything (or anyone) else (so don't take them with you when you take the exam). You can use two formula sheet that you prepare (8.5x11 – both sides). Once you start reading the test, you cannot update your notes. **Attach the notes to your exam when you submit it. If you did not use notes, please state on the exam “DID NOT USE NOTES.”**
* Please submit all item in I-Learn by July 6th at 11:55 pm. **This includes the test, notes, this page signed (image if needed), graphs, and R output.**
* Once you have taken the exam, you will not talk to anyone about the contents of the exam (except Brother Cromar) until after July 6th at 11:55 pm.
* **For problems #7-10 you will be using R.**
* You may use the following: a calculator, your two page notes, scratch paper, R data file is provided for you in I-Learn in Test 2 for these problems.
* Please read and sign the following statement: You have read the instructions above and have obeyed all the rules with exactness (If you need friendly reminders on the importance of obedience and integrity, please refer to 1 Nep. 3:7 and 2 Nep. 9:34 ☺).

I agree to the terms of this test, and I kept them as I took this exam.

Print name: \_\_Cortland Watson\_\_\_ Start Date/Time: \_\_\_\_4:05pm\_\_\_\_\_

Signature: \_\_\_\_on notes/bottom of test\_\_\_ End Date/Time: \_\_\_\_\_6:23pm\_\_\_\_\_

1. (3 pts.) You are designing a Latin Square study with one factor of interest and using one square. The factor has 4 levels (A, B, C, and D). Give one possible Latin Square design by filling in the rows and columns of the table below with the factor level labels (A, B, C, and D).

|  |  |  |  |
| --- | --- | --- | --- |
| A | B | C | D |
| B | C | D | A |
| C | D | A | B |
| D | A | B | C |

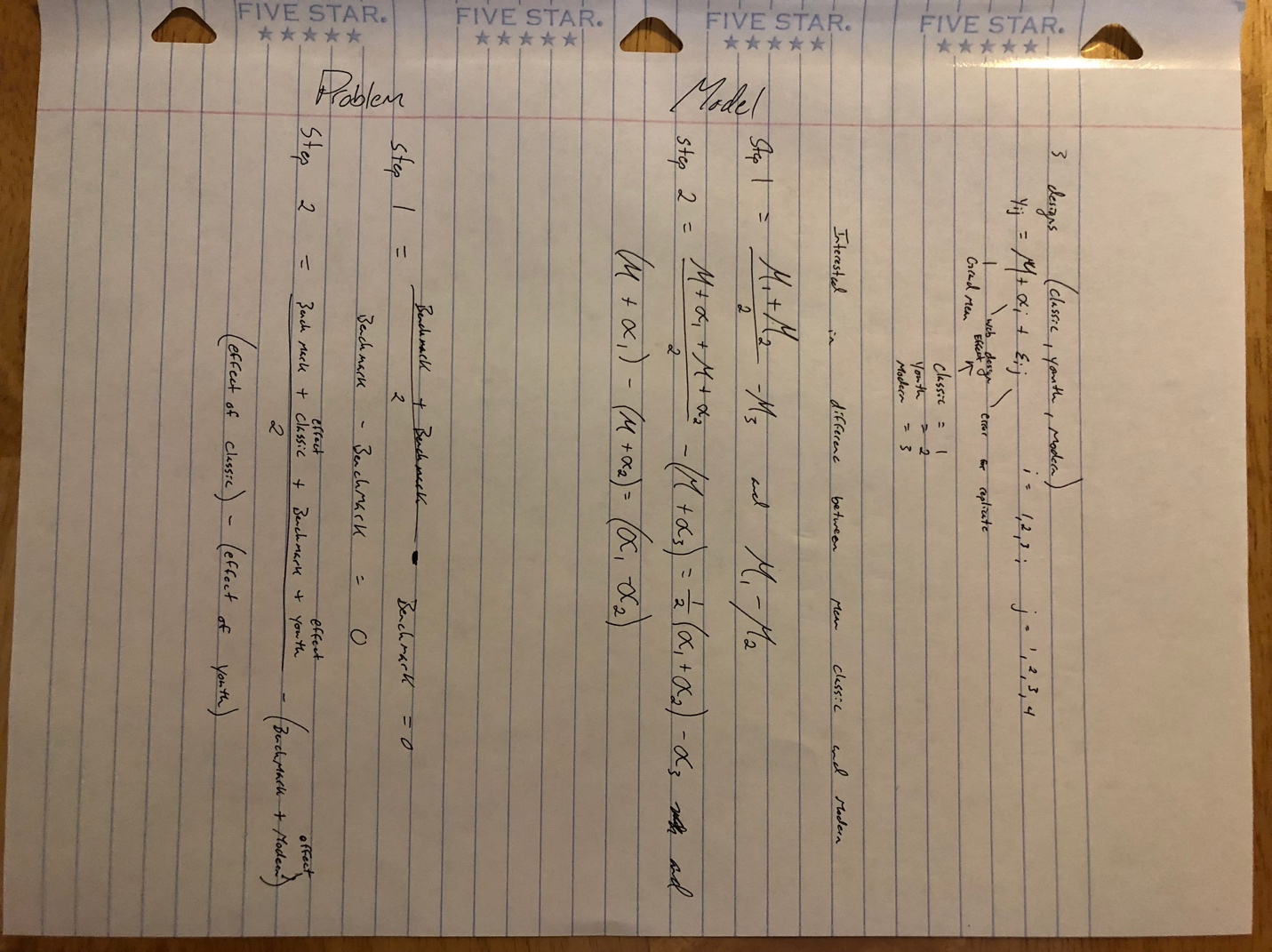
2. (5 pts.) You assume that the online marketing company runs an experiment involving now just the 3 different website designs (classic, youth, and modern). They are interested in the sales associated with each possible online experience and randomly assign each of 12 similar days (all were Tuesdays) to one of the 3 possible treatments. They wish to assess the effects for website design. At the end of each day, the total sales are recorded.

Let the model for these data be written:

yij = µ + αi + εij; i = 1, 2, 3; j = 1, 2, 3, 4;

where µ is the grand mean (benchmark), αi is the effect for the ith level of website design, and εij is the error term for the jth replicate in cell (i). Note that for website design, “classic” is level i = 1, “youth” is level i = 2, and “modern” is level i = 3, which is how the factors levels are put in an R dataset.

We are interested in the difference between the mean sales for classic website design and the mean sales for the modern website design. Write the appropriate contrast IN TERMS OF the model parameters (µ, αi, and εij). Hint: The final result for this problem is the second step of three steps when we worked on contrasts in the last unit.



3. (3 pts.) An agronomist is interested in comparing five varieties of beans to see which has the largest yield. The farm used in the experiment is laid out in 25 plots as shown below, and the plan is to plant each variety in five of the 25 plots. However, there is some concern about lack of uniformity in the 25 plots. Specifically, the farmland slopes upward from south to north (which affects water availability) and the plots in the east tend to have higher concentrations of silt.



Assuming that (1) there are no interactions among variety, water availability, and silt concentration, and

(2) the **only factor of interest is variety**. Which is the most appropriate design?

(Circle one)

(a) BF[1]

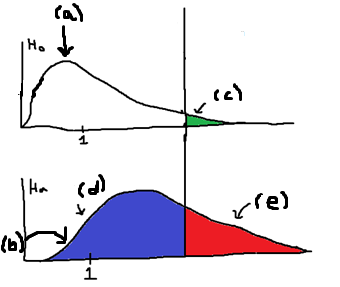
(b) BF[2]

(c) BF[3]

(d) CB[1]

(e) LS[1]

4. The figure below illustrates issues related to the power of hypothesis tests in ANOVA. The top picture refers to the null hypothesis and the bottom picture refers to the alternative hypothesis.



1. (1 pt) What is the name of the **distribution** labeled (a)?
   1. F distribution
2. (1 pt) What is the name of the **distribution** labeled (b)?
   1. Normal Distribution
3. (1 pt) What is the name for the **area** marked as (c)?
   1. Type 1 error (alpha)
4. (1 pt) What is the name for the **area** marked as (d)?
   1. Beta
5. (1 pt) What is the name for the **area** marked as (e)?
   1. Power

5. Kudzu is a plant that was imported to the United States from Japan and now covers over seven million acres in the South. The plant contains chemicals called isoflavones that have been shown to have beneficial effects on bones. In one study, 45 rats were randomly put into three groups, the control group, low dose of isoflavones or high dose of isoflavones. Each group had 15 rats in their group. One of the outcomes examined was the bone mineral density in the femur (in grams per square centimeter).The mean density for each group is given below.

* Control: 0.21887
* Low Dose: 0.21593
* High Dose: 0.23507

(a) (5 points) Fill in the missing values (underlined) in the ANOVA table below:

Source df

Mean \_\_1\_\_\_\_

Treatment \_\_2\_\_\_\_

Residuals (Error) \_\_42\_\_\_\_

Total \_\_45\_\_\_\_

(b) (2 point) For the F-statistic used to test if Treatment is significant, what are the degrees of freedom for the F distribution used to calculate a p-value?

Numerator = 2

Denominator = 42

6. (4 points) What are the two main reasons we use randomization when assigning subjects to treatments?

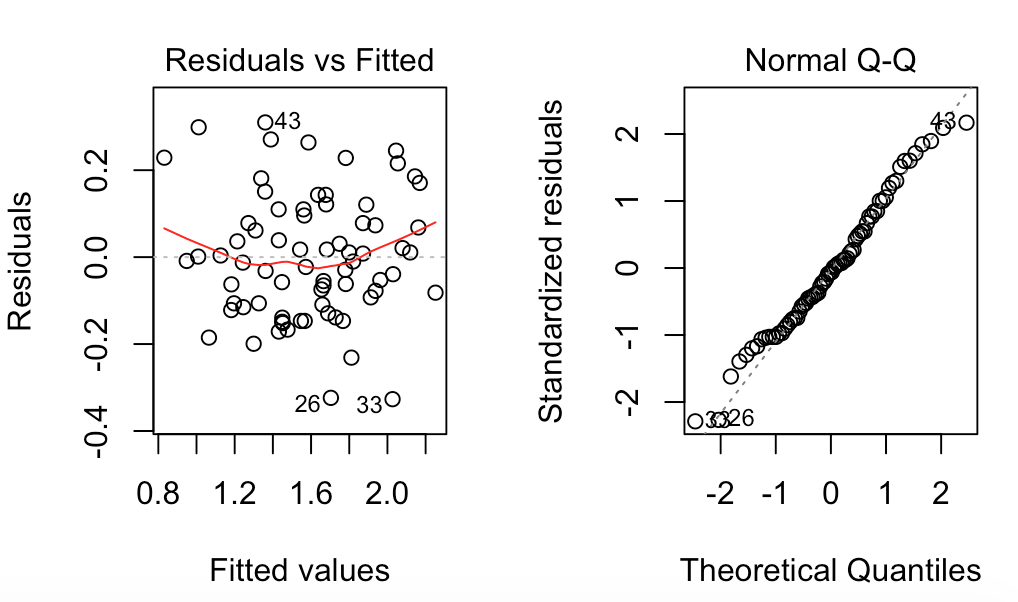
(i) Protects us against bias

(ii) Allows us to use p-value and sampling distributions when analyzing data

7. An experiment was performed where the test engineer wishes take 72 batteries to see how much the temperature (three levels) and the type of electrolyte (four levels) and would affect the activated lives in hours and would use a **Split Plot Design** for the design and analysis. The design was set up where 24 batteries were randomly assigned to each temperature, which is the **whole plot factor**. There were six replicates for each temperature.

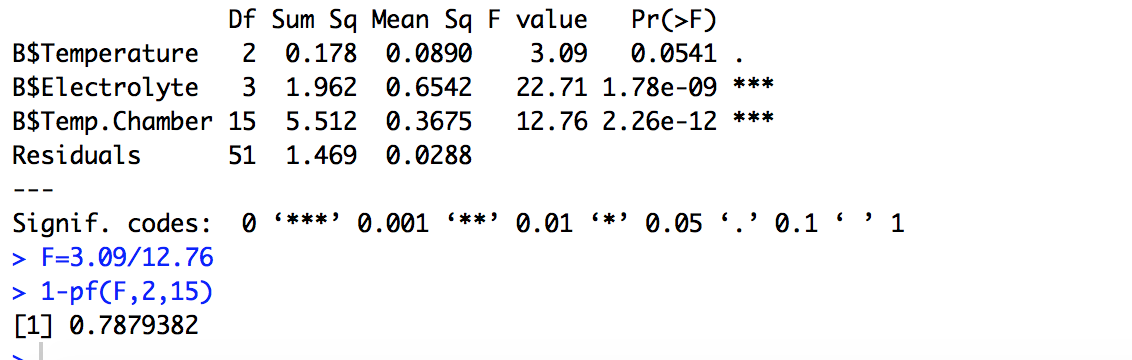
Four batteries were put in a temperature chamber at a time with one temperature and each of the four batteries received one level of electrolyte where all four electrolytes were represented, where electrolyte is the **subplot factor**. Thus, for each replicate the test temperature is maintained, and the four levels of electrolyte are tested simultaneously at the same temperature. The data is in ActiveBattery. **Do not forget to check to see if you need to do as.factor** (12 points).

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



These two graphical summaries teach us a lot about the data itself. First, the variance appears to have some concerns. It tails on both sides and seems to be more concentrated in the middle suggesting that the variance is not constant. Secondly, the normality of the data is questionable. Although most of the data follows the straight and narrow path, the tails seem to go off track and curve. This suggests to us that the data is also not normally distributed. For purposes of the exam, the process will move onward.

1. (5 points) Get an ANOVA table using **software**

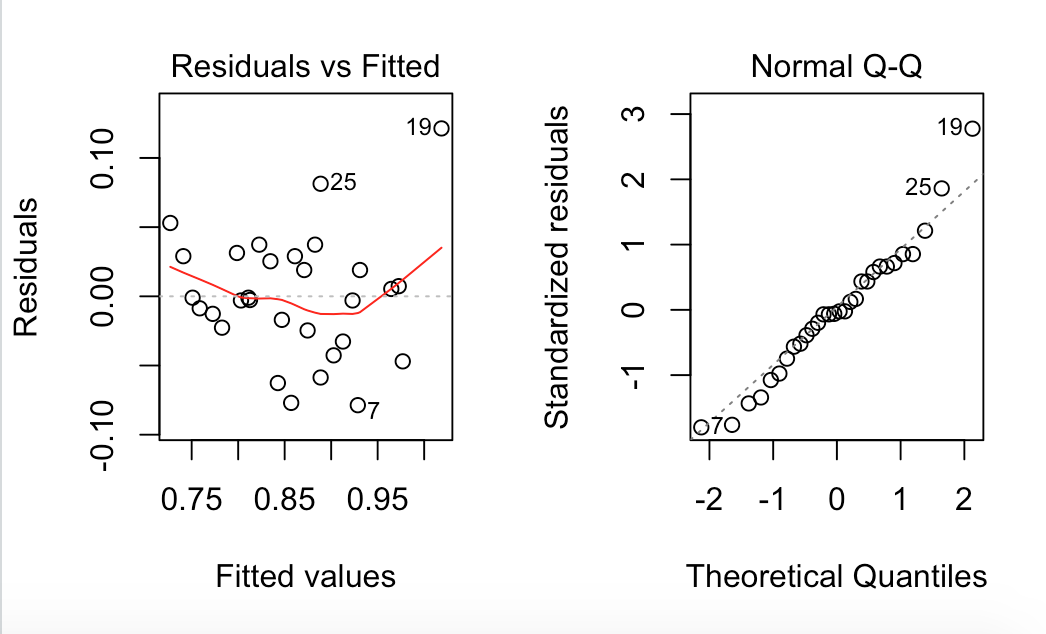


This table shows us a lot as well. We are able to see levels of significance and make conclusions based on this table. At the same time, because we are running a split plot analysis, the B$Temperature value for F-statistic and p-value is inaccurate. We therefore have to recalculate it using the given equation. This then gives us another measurement which draws very plain conclusions.

1. (5 points) For the **whole-plot factor, subplot factor, and interaction**: 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.
   1. Whole Plot
      1. Null = The population means of the different temperatures are equal. Alternative = At least one of the population means is different.
      2. .242
      3. num = 2 den = 15
      4. 0.7879382
      5. We fail to reject the null hypothesis
      6. We fail to reject the null hypothesis that the populations means of the different temperatures are equal because of the high p-value.
   2. Subplot
      1. Null = The population means of the different electrolyte groups are equal Alternative = At least one of the population means is different
      2. 22.71
      3. num = 3 den = 51
      4. 1.78e-09
      5. We reject the null hypothesis
      6. We reject the null hypothesis that the groups are the same because of the very low p-value.
   3. Interaction
      1. Null = There is no interaction Alternative = There is an interaction
      2. 12.76
      3. num = 15 den = 51
      4. 2.26e-12
      5. We reject the null hypothesis that there is no interaction
      6. Because of the low p-value we are able to conclude that there is an interaction.

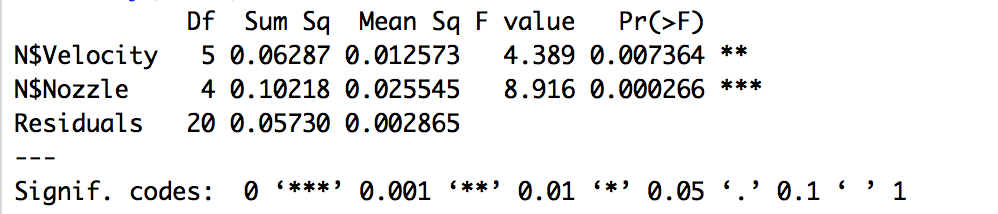
8. An article in the Fire Safety Journal (“The Effect of Nozzle Design on the Stability and Performance of Turbulent Water Jets,” Vol.4 August 1981) describes an experiment in which a shape factor was determined for several different nozzle designs at six levels of jet efflux velocity (m/s). Interest focused on potential differences between nozzle designs, with velocity considered as a nuisance variable. The researchers used a **Complete Block Design (CB[1])**. The data is called NozzleDesign. **Do not forget to check to see if you need to do as.factor** (12 points)

1. Check the assumption of residuals being normally distributed and constant variance using **software** (3 points).



These two graphical summaries teach us about the appropriateness of using our analysis. The first suggest that the variance is not constant. The tailing ends, and dip in the middle suggests that a transformation my be needed, but at least that the variance is not constant. The second graphic show us that the normality is also questionable because it shows that the upper end tails. It appears that the bottom end levels out, but the top follows an upward trend suggesting a skew.

1. Get an ANOVA table using software (3 points)

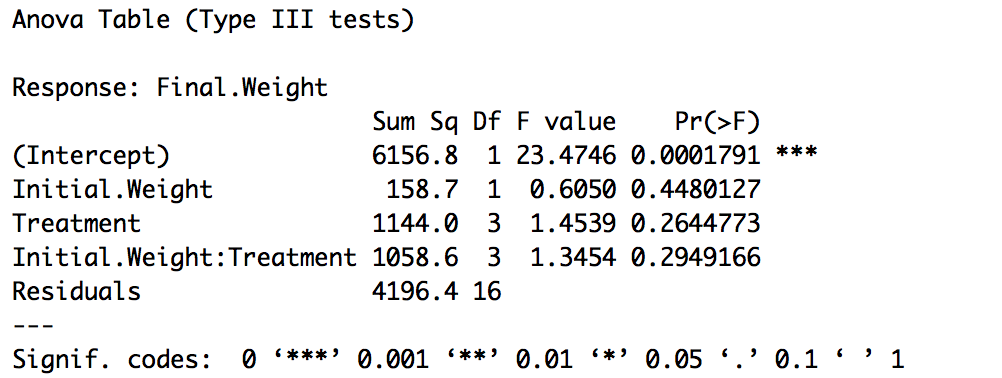


The table shows us a lot about being able to use the analysis and make conclusions, eventually making assumptions for the future.

1. For effect nozzle design: 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. (3 points)
   * 1. Null = The population means of the nozzle size are equal. Alternative = At least one of the population means is different.
     2. 8.916
     3. num = 4 den = 20
     4. 0.000266
     5. We have sufficient evidence to reject the null hypothesis.
     6. We are able to safely conclude that at least one of the population means of the nozzles is different.
2. Does the blocking factor turn out to be an important source of variability(3 points)?
   1. The blocking appears to be an important source of variability. We can tell this by seeing that the blocking factor had a significant p-value.

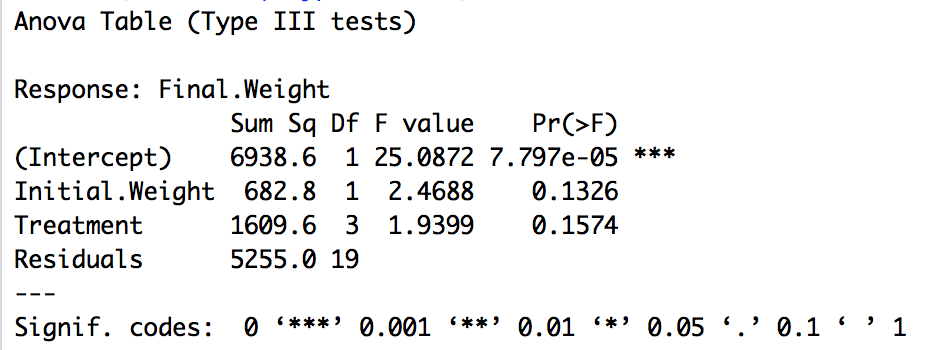
9. An experiment was conducted in which the weight gains of pigs for 4 different feeds were compared. The initial weight of the pigs was used as a nuisance variable or a covariate, so the experimenters used **ANCOVA**. Pigs were assigned to feeds completely at random. The data is called pigfeed. **Do not forget to check to see if you need to do as.factor for the feed factor.**

1. Check if there is an interaction between the feed and initial weight (Show the table)(4 points).



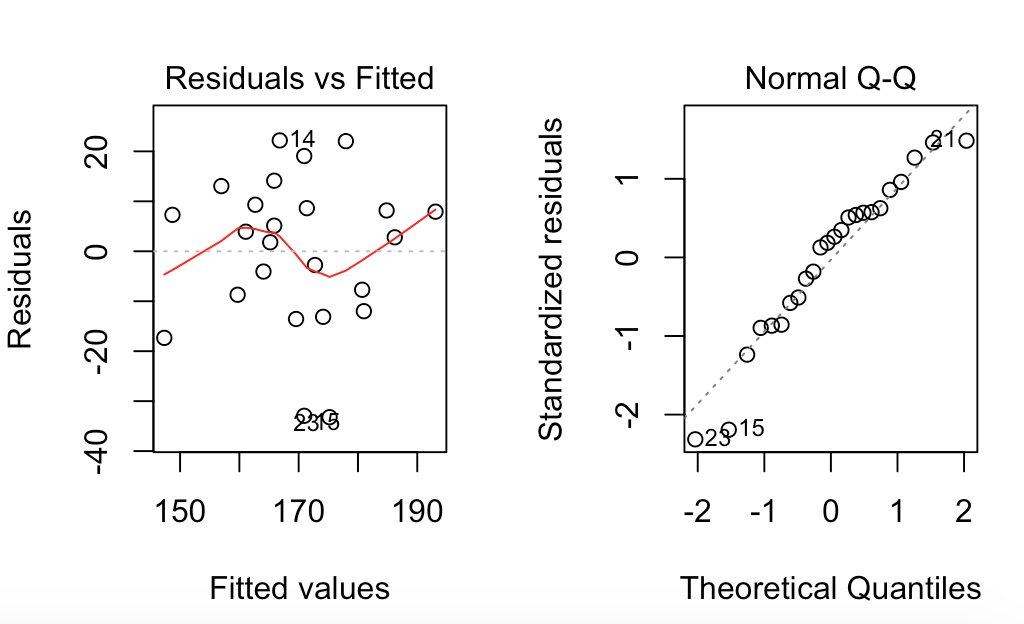
It does not appear that there is any significance in our ancova analysis. This includes the target factor as well as the potential interactions. We will now look to see how it looks without the interaction.

1. Get the ANCOVA model (show the table (3 points).



Looking at this new table, we are able to see that the interaction did pull some weight off of the other variables, but it did not pull enough weight to have changed the significance of the other factors.

1. Check the assumption of residuals being normally distributed and constant variance using **software** (4 points).



These graphical summaries also show us that the data is not normal and that the variance is not constant.

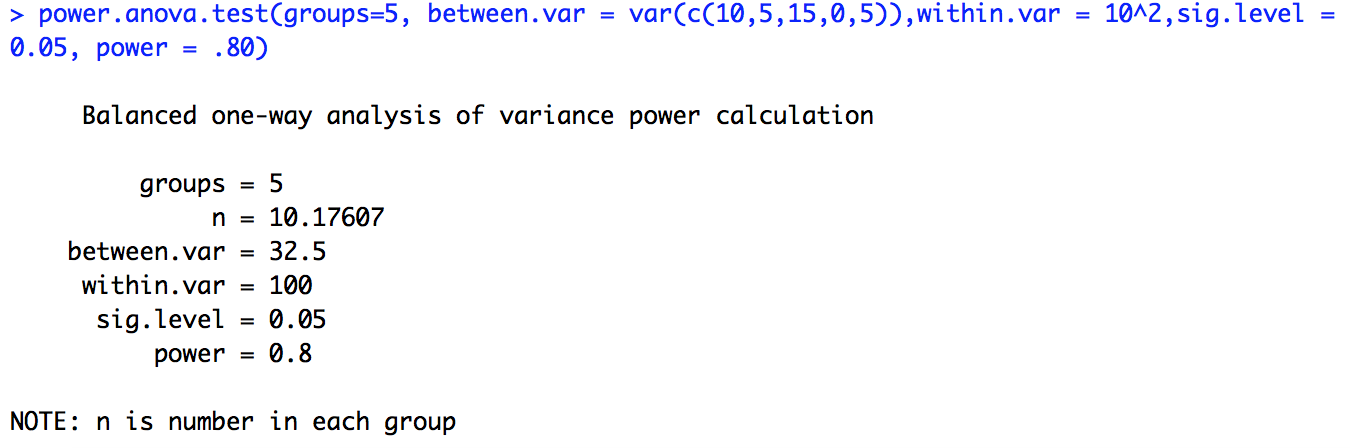
1. For the effect of interest (Different Feeds): i) state the null and alternative hypotheses, ii) give the test statistic, iii) give the degrees of freedom, iv) state the p-value, v) determine whether you should reject or not reject the null hypothesis, and vi) write a sentence which gives an appropriate conclusion. (3 points).
   * 1. Null = There is no difference in the population means Alternative = At least one of the groups is different
     2. 1.4 with interaction 1.9 without
     3. With Interaction Num = 3 Den = 16 With No interaction num = 3 den = 19
     4. With interaction 0.264 Without interaction 0.157
     5. We fail to reject the null hypothesis that the population means are equal
     6. There is insufficient evidence to reject the null hypothesis because of the high p-value.

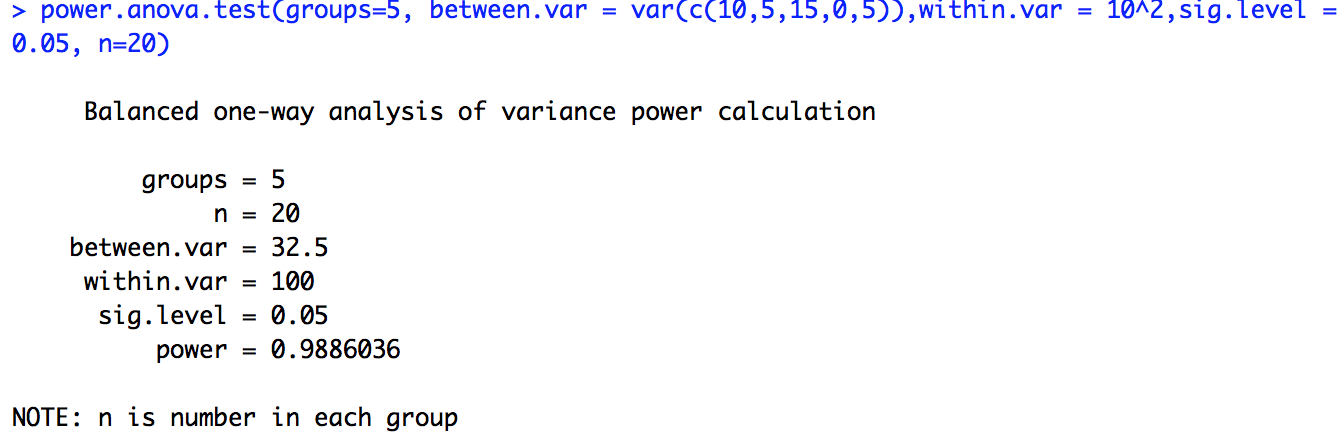
10. We are interested in comparing 5 different methods when training for a marathon:

* Method A: Long runs only
* Method B: Long runs with regular interval training
* Method C: No long runs with regular interval training
* Method D: Long runs with intense interval training
* Method E: No long runs with intense interval training

You are interested in assessing the power of the F test (in ANOVA) for detecting differences in preparation method means when the significance level is α = 0.05. Suppose that marathon times in minutes have a standard deviation of 10, and suppose we would like to evaluate the possibility that the group means are , , , and

(b) (3 points) What is the smallest value for the group size (n) that gives 80% power?

  
We need to have at least 11 samples in each group

(c) (3 points) With a sample size of 20, what is the power of the test?

With 20 samples per group we end up with a power of 98.8%.

11. Brother Cromar would like to improve the Experimental Design Class with two factors, assessments and textbooks. He would like to compare two different types of unit assessments, traditional assessments and case study assessments. Also, he is considering five different types of textbooks from five different authors (Authors A, B, C, D, and E).

In the class, there are twenty students where four groups of five students are created. Each of the four groups was randomly assigned to receive either traditional unit assessments or the case study unit assessments **(unit assessment – whole plot factor)**, where 2 groups got one type of unit assessment and the remaining two got the other.

Within each participating group, each student was randomly assigned a textbook (**textbook – sub plot factor**), where each student would randomly receive one of the five textbooks.

Other “blinded faculty members” who are evaluating the results did not know which groups got which unit assessment and which students got which textbooks. These other faculty members were chosen to evaluate the results of the experiment. The students were all given the same final exam at the end of the semester to evaluate the unit assessment factor, the textbook factor, and the interaction of unit assessment type and textbook.

(a) (2 pts.) What is the response variable?

Final Exam at the end of the semester

(b) (3 pts.) What are the experimental factors, and what are the levels of each factor?

Whole plot = Unit assessments 2

Block = groups 4

Within plot = Textbook 5

(c) (2 pts.) Was blocking used in this study? If so, what was the block?

Blocking was used in this study by randomly creating four groups of five students. Each group then randomly received a textbook assigned to each member.

(d) (2 pts.) What is the experimental unit for evaluating the effect of unit assessment?

The mean score of the group. The block is the experimental unit of the whole plot.

(e) (2 pts.) What is meant by “a blinded faculty members” and what is the purpose of blinding here?

Blinded faculty member means that they do not know about the experiment/which students or groups received which treatment. This is done to prevent bias.

(f) (2 pts.) Write down the statistical model for the observed values, defining all symbols used (symbols used include y, α, β, γ, ε, i, j, and k).

y\_ijk = Mu + alpha\_i + Beta\_j(i) + Gamma\_k + ((alpha)(Gamma))\_ik + epsilon\_ijk

i = 1, ...., I; j = 1, ....., J; k = 1, ...., K

Mu = Benchmark

alpha\_i = effect of the i^th unit assessment – group factor level

Beta\_j(i) = effect of the j^th group level nested in the i^th unit – group factor level

Gamma\_k = effect of the k^th textbook – group factor level

((alpha)(Gamma))\_ik = effect of the ik^th level of the interaction

epsilon\_ijk = residual error for the observed value Gamma\_ijk

(g) (2 pts.) Using your notation for the model in part (f), write down the null and alternative hypotheses for testing the effect for the textbook.

H\_0 = The mean for each level of the unit assessment – group factor is the same

H\_0 = the mean for each level of the textbook – textbook factor is the same

H\_0 = there is no interaction

H\_a = At least one of the means is different

H\_a = At least one of the means is different

H\_a = There is an interaction

(h) (8 pts.) Give a partial ANOVA table below. That is, list all sources and the degrees of freedom (df) for each. (Hint: the first line should be “Mean (benchmark)” and the final line should be “Total”.)

**Source df**

**Grand Mean 1**

**Unit Assesment 1**

**Group 3**

**Textbook 4**

**Interaction 4**

**Residuals 7**

**Total 20**

(i) (2 pts.) What are the degrees of freedom for the F distribution used to test if the factor unit assessment is significant?

Num = 1 Den = 3

