**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 one handwritten formula sheet that you prepare (8.5x11 – both sides). Once you start reading the test, you cannot update your notes. **Attach a picture of your 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 items in I-Learn by the posted due date for full credit. **This includes the test, notes, graphs, and R output.** Remember, I will grade the R output not the code itself; though sometimes having the code may be helpful for partial credit.
* Once you have taken the exam, you will not talk to anyone about the contents of the exam (except Brother Palmer).
* **For problems #7-10 you will be using R.**
* You may use the following: a calculator, your one page of notes, scratch paper, R data files provided for you in I-Learn in Test 2.
* By submitting your exam you certify that 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.

Start Date/Time: \_\_\_\_\_\_\_1:00 pm\_\_\_\_\_\_ End Date/Time: \_\_\_\_\_\_2:20 pm\_\_\_\_\_

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

2. (2 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 down the vector of weights (or coefficients) for the desired contrast.

1 = 1

2 = 0

3 = -1

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]

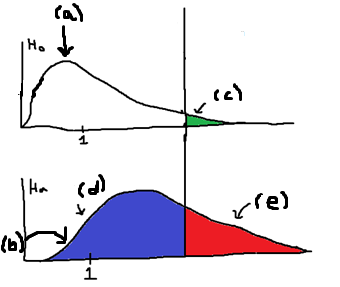
(b) BF[2]

(c) BF[3]

(d) CB[1]

(e) LS[1] 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.



(a) (1 pt) What is the name of the **distribution** labeled (a)? Null Hypothesis

(b) (1 pt) What is the name of the **distribution** labeled (b)? Alternative Hypothesis

(c) (1 pt) What is the name for the **area** marked as (c)? Alpha Level

(d) (1 pt) What is the name for the **area** marked as (d)? Critical Value

(e) (1 pt) What is the name for the **area** marked as (e)? 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) \_\_\_\_\_41\_

Total \_\_\_\_44\_\_

(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?

2,44

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

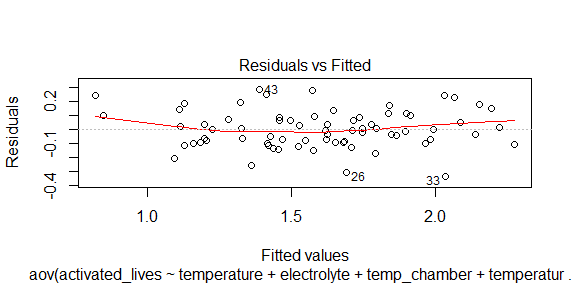
(i) Eliminate any possible bias.

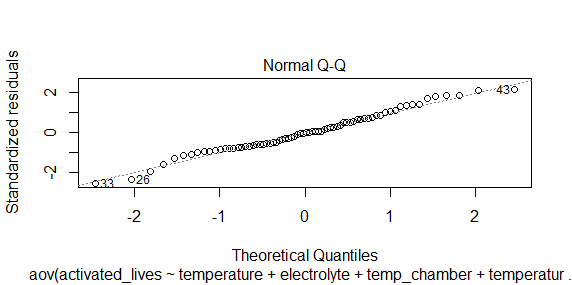
(ii)  Homogeneous treatment groups

7. An experiment was performed where the test engineer measures 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. The engineer will 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) Determine whether the assumption of residuals being normally distributed and constant variance are met using **software**. Explain/show how you made that determination.





The Variance looks constant and based on the QQPlot the residuals follow the straight line and look normally distributed.

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

Error: electrolyte

Df Sum Sq Mean Sq

temperature:electrolyte 3 1.962 0.6542

Error: Within

Df Sum Sq Mean Sq F value Pr(>F)

temperature 2 0.178 0.0890 3.183 0.051 .

temp\_chamber 15 5.512 0.3675 13.139 8.92e-12 \*\*\*

temperature:electrolyte 6 0.211 0.0351 1.255 0.297

Residuals 45 1.259 0.0280

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

1. (6 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.

H0 = α1 = α2 = α3= α4 = α5 = 0

Ha = At least one is Different

F-Value2,45 =  3.183

DF = 2,45

P-Value = 0.051

We will fail to reject the null hypothesis, but we could reject the null with the same p-value where it is in a gray area.

H0 = y1 = y2 = y3=y4 =yα5 = 0

Ha = At least one is Different

F-Value15,45 = 13.139

DF = 15,45

P-Value = 8.98e-12

We have sufficient evidence to reject the null hypothesis.

H0 = αy = αy = αy= αy = αy = 0

Ha =There is an interaction

F-Value6,45 = 1.255

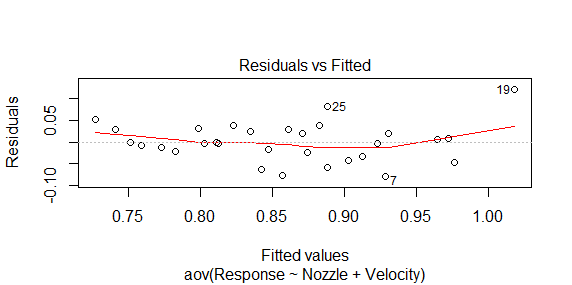
DF = 6,45

P-Value = 0.297

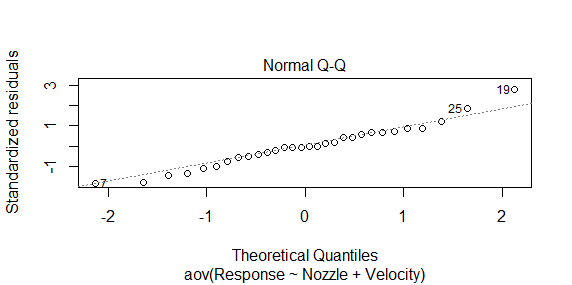
We have insufficient evidence to reject the null hypothesis.

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. Determine whether the assumption of residuals being normally distributed and constant variance are met using **software** (3 points). Explain/show how you made that determination.



The constant variance looks constant based on this plot.



The residuals look normally distributed based on this plot following the straight line

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

Analysis of Variance Table

Response: Response

Df Sum Sq Mean Sq F value Pr(>F)

Nozzle 4 0.102180 0.025545 8.9162 0.0002655 \*\*\*

Velocity 5 0.062867 0.012573 4.3886 0.0073642 \*\*

Residuals 20 0.057300 0.002865

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

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)

H0 = α1 = α2 = α3= α4 = α5 = 0

Ha = At least one is Different

F-Value 4,20 = 8.9162

DF = 4,20

P-Value = 0.0002655

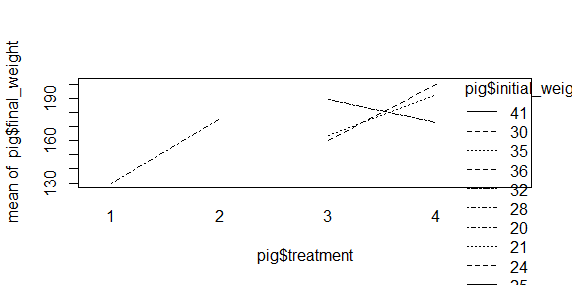
We have sufficient evidence to reject the null hypothesis.

1. Does the blocking factor turn out to be an important source of variability (3 points)?

Yes it does based on the p-value of 0.0073642

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).



Yes, there is an interaction.

1. Get the ANCOVA model (show the table (3 points). (if the interaction term tested above is not significant remove it)

Anova Table (Type III tests)

Response: final\_weight

Sum Sq Df F value Pr(>F)

(Intercept) 6938.6 1 25.0872 7.797e-05 \*\*\*

initial\_weight 682.8 1 2.4688 0.1326

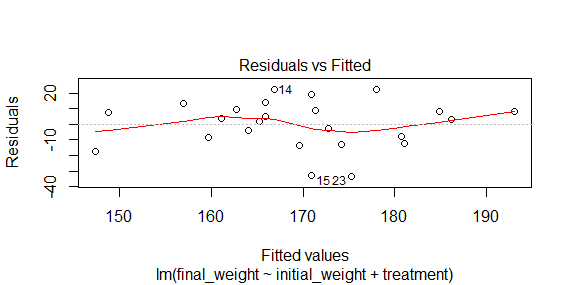
treatment 1609.6 3 1.9399 0.1574

Residuals 5255.0 19

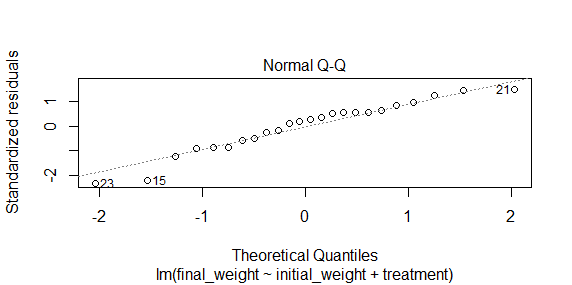
---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

1. Determine whether the assumption of residuals being normally distributed and constant variance are met using **software** (4 points). Explain/show how you made that determination.



The variance looks constant. It could be a bit better but we will say it looks constant and use it.



The Residuals look to be normally distributed by following the straight line and not deviating to much from the line.

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).

Ho = H0 = α1 = α2 = α3= α4 = 0

Ha = At least one is Different

F-Value3,19 = 1.9399

DF= 3,19

P-Value = 0.1574

We have insufficient evidence to reject the null hypothesis.

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

(a) (3 points) What is the smallest sample size per group (n) that gives 80% power?

22

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

0.7698023

(c) (1 point) You are interested in comparing the means of methods with long runs against the methods without. Write down the vector of weights (or coefficients) for the desired contrast.

A= 1/3

B= 1/3

C= -1/2

D= 1/3

E= -1/2

(d) (1 point) You are interested in comparing the means of methods with regular interval training against the methods with intense interval training. Write down the vector of weights (or coefficients) for the desired contrast.

A= 0

B= 1/2

C= 1/2

D= -1/2

E= -1/2

(e) (1 point) Are the contrasts in parts c and d orthogonal? Show the calculations that prove your answer.

1/3 1/3 -1/2 1/3 -1/2

0 ½ ½ -1/2 -1/2

1/3 1/5 0 -.17 -1/4

Not Orthogonal

11. Brother Palmer 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?

Improvement of students within the course

Final Exam Grade

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

Assessments and Textbooks

Textbooks is A,B,C,D,E

Assessments is traditional unit assessment or case study

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

Block on Student

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

Final Exam Grade

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

The faculty member was not told what the students got before- hand.

(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).

Textbook = α

Assessment = β

Y = αijklm + βij + εijklm

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

Ho: α1+ α2 +α3 + α4 + α5 = 0

Ha : AT least one is different

Ho: β1 + β2 = 0

Ha: At least one is different.

(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**

**Mean**  1

Treatment 2

Interaction 2

Residual 14

Total 19

1. (2 pts.) What are the degrees of freedom for the F distribution used to test if the factor unit assessment is significant?

DF = 2,19