ECHE 313 Exam 2

Write your name: Prof Rennec

Write the following statement, and sign your name acknowledging your agreement:

""I neither received nor gave external assistance on this exam "

_____

The exam layout is such that all problems add up to 100 points. This point total is meant to guide you in time allocation for this test, will be adjusted such that this exam is 25% of your total class grade. You may use your class materials, class notes, homework, book, and Minitab to aid in the completion of the exam, but you are not allowed to use any outside resources. The exam is designed to be completed in 1 hr and 15 min, but you will have between 10 AM 4/9 to 10 AM 4/10 to complete the exam. Note: Once you start the exam, you have 4 hours to complete the test. All of the necessary tables are provided in the back of your book. Please show your work and all steps necessary to arrive at your answer. It is expected you do this entire test by hand, using a calculator and the tables found in the back of your book. So while you are allowed to use Minitab if you wish to check your answers, it should not replace the work you have done by hand using the tables for your answers. Where applicable, it is best to write out the general formals you want to use first, then write in your subbed in values so it is clear. Also, when reporting numbers, keep the same number of decimals as the data you are working with unless there are specific directions on how many sig figs to report. Good luck!

Premise: You work as a process engineer for a company that makes eclipse glasses. What a great time to be in this business! These glasses work by filtering light, reducing light exposure to the eye. Unless otherwise stated, your company uses α=0.05.

Problem 1 (67 points total) Your team wants to change the concentration of materials that are used to block light in your product to save money. Material B is particularly expensive and is typically set at 5.00 wt% in the material. You want to know if the concentration of material B has an impact on the % transmittance of light, so you vary material B and test the amount of light that is transmitted through the material, and get the following data:

	% Transmittance			$\bar{y}_{i.}$	
	Replicate				
Material B (wt%)	1	2	3	Ave.	Standard Deviation
(typical) 5.00	0.0023	0.0023	0.0024	0.0023	0.000058
4.00	0.0023	0.0025	0.0028	0.0025	0.000252
3.00	0.0027	0.0024	0.0021	0.0024	0.000300
2.00	0.0032	0.0038	0.0035	0.0035	0.000300
1.00	0.0041	0.0039	0.0045	0.0042	0.000306

5
4
3
2
1

$$\bar{y}_{..} = \frac{0.0023 + 0.0025 + 0.0024 + 0.0035 + 0.0042}{5} = 0.00298$$

a. (1 points) List the *name* of the hypothesis test you would use

ANOVA

b. (2 points) Name the parameter of interest in the context of the problem statement

Multiple means, specifically % transmittance at different levels of material B

c. (2 points) Write out the correct null hypothesis for this test

$H_0: \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = 0$

d. (2 points) Write out the correct alternative hypothesis for this test

$H_1: \tau_i \neq 0$ for at least one i $(i = 1\text{-}5)$

e. (6 points) Write the formula of the test statistic, and other formulas needed to fully calculate the test statistic from the data above, and calculate its value given that SS_E is 0.00000068

$n = 3$

$a = 5$

$$F_0 = \frac{SS_{Treatments}/a-1}{SS_E/a(n-1)} = \frac{0.0000079/4}{0.00000068/10} = \boxed{29}$$

$$SS_{Treatments} = n\sum_{i=1}^{a}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = \text{treatment sum of squares}$$

$$= 3\sum_{i=1}^{5}\left[(0.0023 - 0.00298)^2 + (0.0025 - 0.00298)^2 + \ldots + (0.0042 - 0.00298)^2\right] = 0.0000079$$

f. (4 points) Describe what the test statistic represents (specifically the numerator and denominator of the test statistic as they related to sources of variation) and what it means if the test statistic is very large versus very small

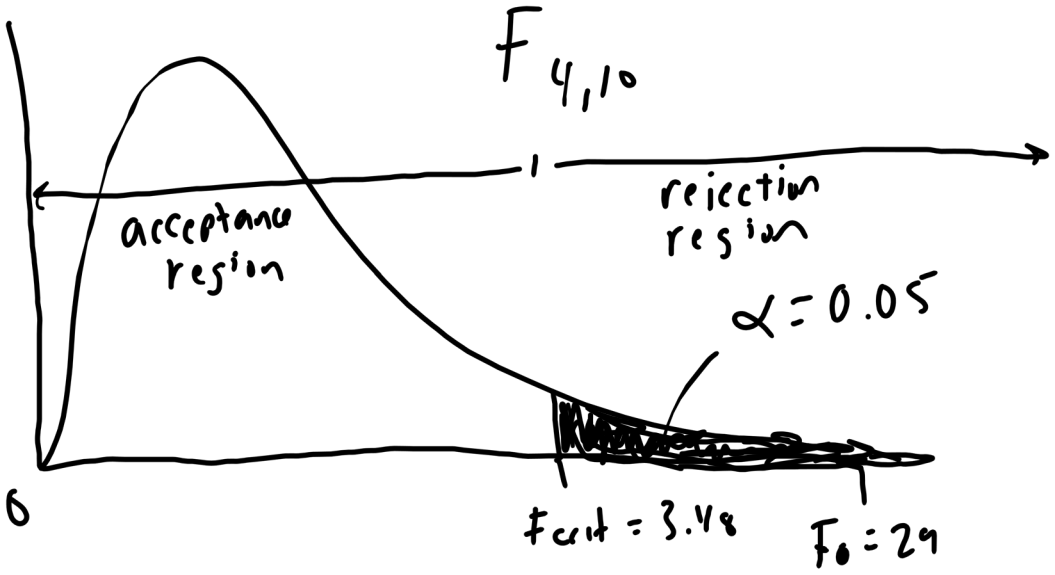
The numerator describes variance from treatments and denominator describes variance from random error. It gives "signal vs. noise". If F_0 is large, variance is coming from treatments (but hypothesis test will determine if significant). If it is small (unity) variance is comparable to random error.
-note that DOF impact the value

g. (4 points) Define your general rejection criteria using fixed significance level testing, sub in values and write out the critical value

Reject if $F_0 > F_{crit} = F_{\alpha, a-1, a(n-1)} = F_{0.05, 4, 10}$

$= 3.48$

h. (10 points) Sketch out the reference distribution for this problem and label: the type of reference distribution used and any defining characteristics (shape, DOF if applicable, and where zero is), the critical value(s), the test statistic, alpha, the acceptance region, and the rejection region.



i. (3 points) Perform any remaining calculations and state your conclusions. Remember to state them in the context of the problem statement, and phrase them appropriately

$F_0 > F_{crit}$ Reject the null hypothesis and accept the alternative.

The wt% of material B impacts % transmittance.

j. (7 points) Describe when it is appropriate to do post-hoc testing, and assuming it is appropriate in this case, conduct a Fisher's post hoc test and identify which groups are significantly different than the normal process setting (Material B at 5%).

A post hoc test is only appropriate after ANOVA H_0 is rejected

Declare sig. different if $|\bar{y}_{i.} - \bar{y}_{j.}| > LSD$

$$LSD = t_{\alpha/2,a(n-1)}\sqrt{\frac{2MS_E}{n}} = t_{0.025,10}^{(2.228)}\sqrt{\frac{2\left(\frac{0.000008868}{10}\right)}{3}} = 0.000474$$

5 vs 4: $|0.0023 - 0.0025| = 0.0002$

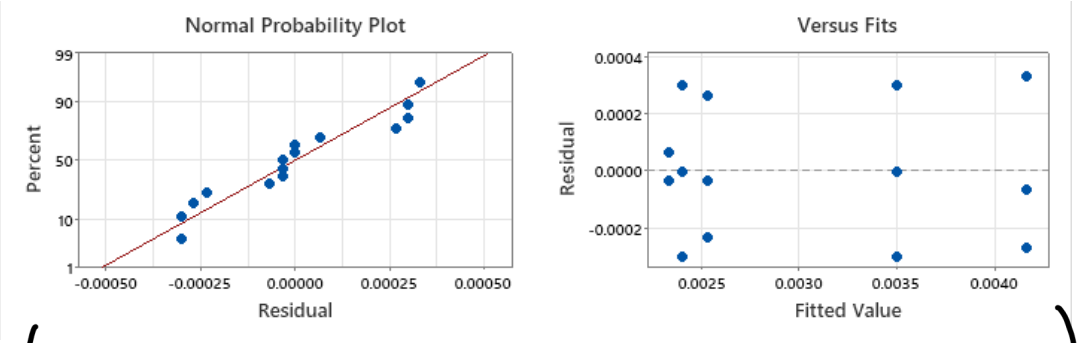
5 vs 3: $|0.0023 - 0.0024| = 0.0001$

5 vs 2: $|0.0023 - 0.0035| = 0.0012$ *

5 vs 1: $|0.0023 - 0.0042| = 0.0019$ *

1 and 2 wt % are sig. different than 5 wt %

k. (4 points) If these are the residual plots:



1. State what assumptions are being checked for each plot

That errors and observations are normally distributed

constant variance

2. Describe what you should do if the assumptions are violated

Transform the data and perform ANOVA again

l. You are concerned that the variance of the data for 5 wt% treatment is different than the other treatments.

1. (1 points) What is the name of the hypothesis test I should perform if I want to know if the variance from 5% treatment and 4% treatment are different?

2 variance test

2. (2 points) Write out the correct null and alternative hypotheses for this test

$H_0: \sigma_1^2 = \sigma_2^2$ $\quad H_1: \sigma_1^2 \neq \sigma_2^2$

5wt. 4wt%.

3. (5 points) Calculate the test statistic (report 3 significant figures)

Write out the general formula: Sub in values and calculate:

$$F_0 = \frac{S_1^2}{S_2^2} = \frac{0.000058^2}{0.000252^2} = 0.0526$$

4. (8 points) 1. Write out the rejection critical using p-value method and 2. determine the p-value from your test statistic (report 2 sig figs in your p-value). 3. Show the p-value graphically. In your sketch, you should have the reference distribution labeled with any defining characteristics (shape, zero, and degrees of freedom if applicable), the test statistic, and the p-value).

$F_{0.9,6,12}$

$= 1/F_{0.1,12,2}$

$= 1/9 = 0.11$

$\approx 0.05 = 1/2$ p-value

$F_{.95,12}$

$= 1/F_{0.05,12}$

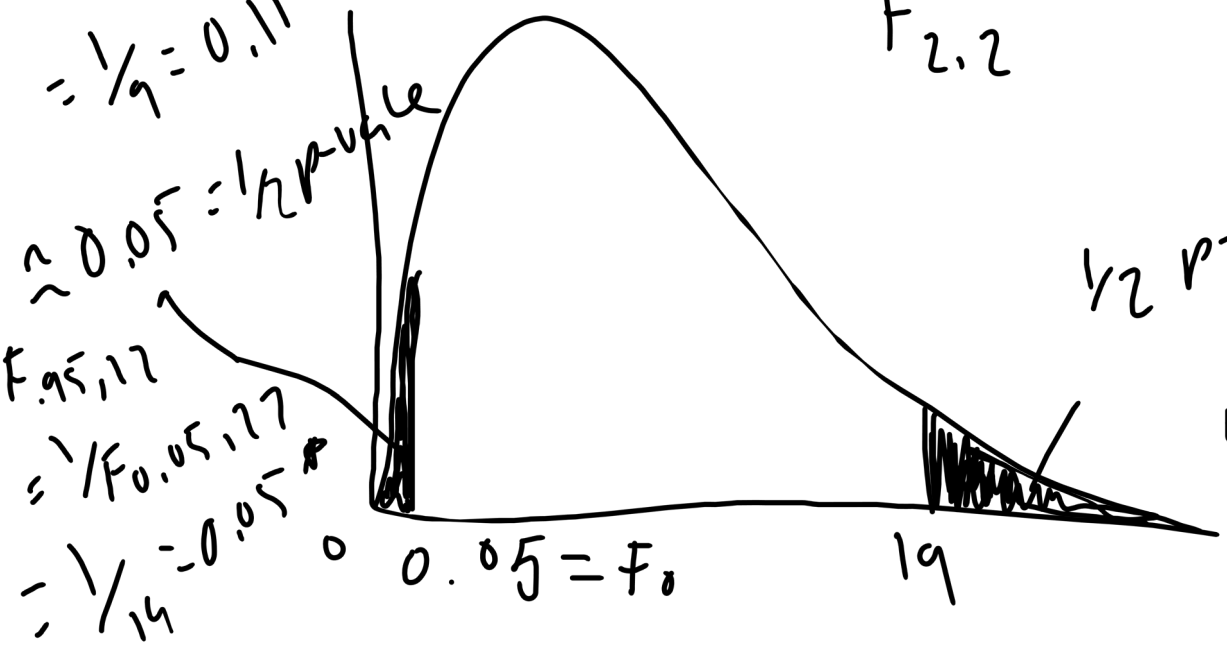
$= 1/14 = 0.05$

Reject H_0 if p-value $> \alpha$

$F_{2,2}$

1/2 p-value ≈ 0.05

p-value ≈ 0.1



5. (3 points) State your conclusions. Remember to state them in the context of the problem statement, and phrase them appropriately. Also use your conclusion to answer: is there cause for concern?

p-value > α : Fail to reject H_0. We do not have enough evidence to say that the variance at 5wt% is different than 4wt%. No cause for concern (constant variance still valid)

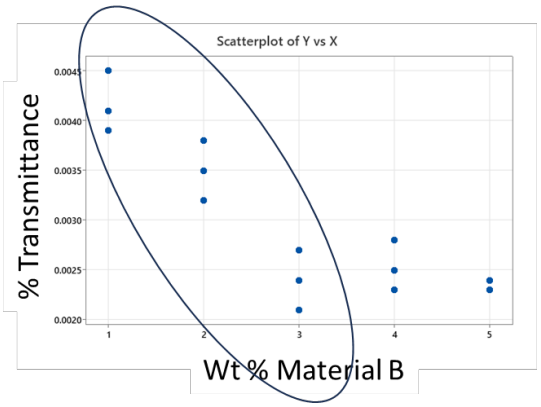
m. (3 points) Provide and answer to your boss who was wondering: "Why is the rejection criteria for ANOVA one sided, whereas sometimes an F-test can be one- or two-sided?"

ANOVA is looking to see if a "signal" is larger than "noise". We therefore only care if the variance of the "signal" is larger than random error, thus it is one sided.

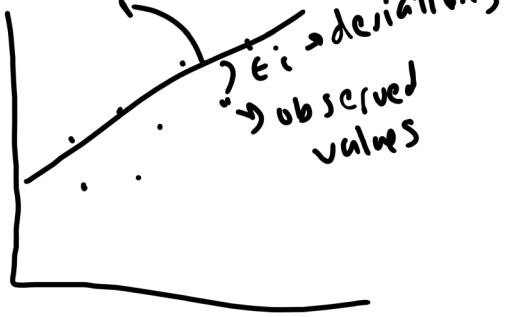
In an F-test we are determining if two variances are different, thus, we need a two sided test.

Problem 2. (22 points total) You suspect that the Material B wt% (from 1 to 3 wt% - predictor variable) is linearly correlated to % Transmittance (response variable). You want to perform a hypothesis test to see if the linear correlation is significant.

Performing linear regression on these data



a. (3 points) You perform the liner regression and your estimated coefficient for the slope is -0.00088 and the intercept is 0.0051. What is the name of the method that is used to estimate these coefficients, and describe generally how that method works. (Hint: I am NOT looking for how to calculate S_{xx} or S_{xy} here). You can use a diagram and general formulas for your answer if needed.



The method of least squares is used to estimate the coefficients in linear regression.

In this method, the vertical deviations from data and the regression line (ϵ_i's) are minimized.

$$\text{Minimize}: \sum_{i=1}^{n} \epsilon_i^2$$

b. This is your Minitab output from the simple linear regression:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression		0.00000468167		51.16	0.000
Error		0.00000064056			
Total					

(1) (2 points) State the degrees of freedom for the $SS_{Regression}$ and the SS_{Error}

SS_R DOF: 1

SS_E DOF: $n-2=7$

(2) (2 points) Write out the formula used and calculate the SS_{Total} using information given in the table above

$SS_T = SS_E + SS_R = 0.000006\,4056 + 0.000000\,568167$

$= 0.00000053$

(3) (2 points) Write out the formula for ~~and calculate~~ R^2

$$\frac{SS_R}{SS_T} = \frac{0.000000\,568167}{0.00000053} = 0.88$$

(4) (3 points) State your conclusions. Remember to state them in the context of the problem statement, and phrase them appropriately. Assume that no issues were noted with the residual plots.

P-value < α Reject H_0 and accept H_1.

The wt% of Material B from 1-3% is linearly correlated with % transmittance

c. (10 points) You think there might be another variable that could be useful in your model (wt% of another material, Material A). You add in this variable and run multiple linear regression to get the following output:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	0.000005	0.000002	26.21	0.001
Wt% B	1	0.000002	0.000002	19.92	0.004
Wt% A	1	0.000000	0.000000	1.03	0.349
Error	6	0.000001	0.000000		
Total	8	0.000005			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.0003018	89.73%	86.30%	76.89%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.005844	0.000760	7.69	0.000	
X	-0.001100	0.000246	-4.46	0.004	4.00
X2	-0.000289	0.000285	-1.02	0.349	4.00

(1) (3 points) The R^2 increased after adding the new variable. Should we keep the variable in the model? Why or why not?

No. R^2 is not a valid tool for determining the significance of a variable in regression. In addition, the partial F-test shows p-value $> 0.05 = \alpha$ for wt% A, so there isn't enough evidence to say it has a significant relationship with % composition

(2) (3 points) R^2 adjusted before adding Material A as a predictor was 86%. Explain why R^2 adjusted did not go up when adding the new variable. Use the formula in your answer.

$$R^2_{adj} = \frac{SSR/k}{SST/n-1}$$

R^2 adj. includes degrees of freedom which prevents R^2 from increasing simply by adding more regressor terms

(3) (4 points) State your conclusions appropriately in the context of the problem statement and write out the final model

Fail to reject H_0. We don't have enough evidence to say wt% A has a significant relationship with % transmittance. The final model is:

$$y = 0.0051 - 0.00088x$$

y: % transmittance; x: wt% material B

Problem 3. (11 points total) Your company wants to establish and Xbar control chart on the process for component C wt%. There are 22 preliminary samples, each of size 8, on the internal diameter of the seal. The summary data (in mm) are as follows:

$$\bar{R} = \frac{158}{22} = 7.18$$

$$\bar{\bar{x}} = \frac{283}{22} = 13.0$$

Sample #	Range (R)	Average (Xbar)
1	5	15
2	9	11
3	8	11
4	4	13
5	9	13
6	5	10
7	10	14
8	5	14
9	9	12
10	8	13
11	8	15
12	8	11
13	8	15
14	8	12
15	5	13
16	8	14
17	7	14
18	4	11
19	9	12
20	8	15
21	8	14
22	5	11
Sum	158	283

a. (8 Points) Write out the general formulas for and calculate the centerline as well as the three sigma control limits that should be used on the Xbar and R control chart

xbar: UCL: $\bar{\bar{x}} + A_2\bar{R} = 13.0 + 0.373(7.18) = 15.7 \sim 16$

centerline: $\bar{\bar{x}} = 13.0 \sim 13$

LCL: $\bar{\bar{x}} - A_2\bar{R} = 13.0 - 0.373(7.18) = 10.3 \sim 10$

R: UCL: $D_4\bar{R} = 1.864(7.18) = 13.4 \sim 13$

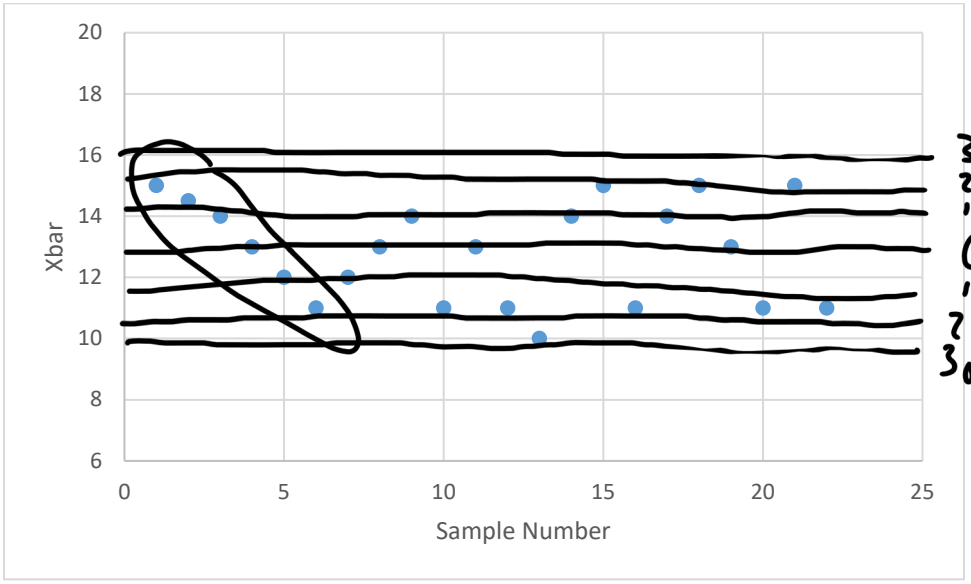
centerline: $\bar{R} = 7.18 \sim 7$

LCL: $D_3\bar{R} = 0.136(7.18) = 0.98 \sim 1$

b. (3 points) In the next month, you take an additional 20 samples in a similar manner as before. Sketch in your sigma control limits (UCL and LCL) and centerline from part A, as well as the 1 sigma and 2 sigma warning limits on this graph of your data. After analyzing this plot, what is your recommendation?

In control: does not break rules 1-4

BUT



Process is should be investigated (breaks sensitizing rule 5)