Stat 441/541 Homework #2 Due: Monday, 01/30/17. Name: _____

1. You will be sent a file with SAS code that will produce an analysis for the data in Problem 3.28, page 135. An experiment was performed to investigate the effectiveness of five insulating materials. Four samples were tested at an elevated voltage level to accelerate the time to failure. The failure times (in minutes) are:

Insulating Material				
1	2	3	4	5
110	1	880	495	7
157	2	1256	7040	5
194	4	5276	5307	29
178	18	4355	10050	2

You are asked to test if the means across the five insulating materials are all equal.

(a) (1.5pt) What are the ANOVA values of SS_E, MS_E, SS_{trt}, MS_{trt}, the F-statistic, and its p-value?

(b) (0.5pt) State the null and alternative hypotheses in terms of means for the F-test.

(c) (0.5pt) What are the five sample variances s_i^2 for $i = 1, 2, 3, 4, 5$?

(d) (2pt) Look at the residuals vs the predicted values plot, the normal probability plot, and the histogram. Comment on what you observe and the implication with respect to the ANOVA F-test.

(e) (2pt) In summary, what (if anything) would you conclude from this analysis? You can use $\alpha = .05$ to draw any conclusions. Your answer should use complete sentences and be in the context of the study.

2. Consider the design and data from Problem 3.12, page 132. A pharmaceutical manufacturer wants to investigate the bioactivity of a new drug. A completely randomized single-factor experiment was conducted with three dosage levels, and the following results were obtained.

Dosage	Observations			
20 g	24	28	37	30
30 g	37	44	31	35
40 g	42	47	52	38

(a) (1.5pt) Assuming the constraint $\tau_2 = 0$, write out the response vector y and matrices X, $X'X$, $(X'X)^{-1}$, and $X'y$ (using either the a or $a + 1$ parameter form for X).

(b) (.5pt) What is the parameter vector θ corresponding to your choice of X in (a)?

(c) (1pt) Assuming the constraint $\tau_2 = 0$, determine the least squares model parameter estimates using $(X'X)^{-1}X'y$.

(d) (1.5pt) Assuming the constraint $\sum \tau_i = 0$, write out the X, $X'X$, $(X'X)^{-1}$, and $X'y$ matrices (using either the a or $a + 1$ parameter form for X).

(e) (.5pt) What is the parameter vector θ corresponding to your choice of X in (d)?

(f) (1pt) Assuming the constraint $\sum \tau_i = 0$, determine the least squares model parameter estimates using $(X'X)^{-1}X'y$.

3. You are given partial SAS output for three multiple comparison procedures (MCPs). The experiment had $a = 6$ treatments (called groups). The results are based using $\alpha = .05$ for each MCP.

   (a) (2.5pt) For each of the fifteen tests, circle the null hypotheses that are rejected.

   For Fisher's Least Significance Difference (LSD) test

$H_0 : \mu_1 = \mu_2$	$H_0 : \mu_1 = \mu_3$	$H_0 : \mu_1 = \mu_4$	$H_0 : \mu_1 = \mu_5$	$H_0 : \mu_1 = \mu_6$
$H_0 : \mu_2 = \mu_3$	$H_0 : \mu_2 = \mu_4$	$H_0 : \mu_2 = \mu_5$	$H_0 : \mu_2 = \mu_6$	$H_0 : \mu_3 = \mu_4$
$H_0 : \mu_3 = \mu_5$	$H_0 : \mu_3 = \mu_6$	$H_0 : \mu_4 = \mu_5$	$H_0 : \mu_4 = \mu_6$	$H_0 : \mu_5 = \mu_6$

   For Tukey's studentized range test

$H_0 : \mu_1 = \mu_2$	$H_0 : \mu_1 = \mu_3$	$H_0 : \mu_1 = \mu_4$	$H_0 : \mu_1 = \mu_5$	$H_0 : \mu_1 = \mu_6$
$H_0 : \mu_2 = \mu_3$	$H_0 : \mu_2 = \mu_4$	$H_0 : \mu_2 = \mu_5$	$H_0 : \mu_2 = \mu_6$	$H_0 : \mu_3 = \mu_4$
$H_0 : \mu_3 = \mu_5$	$H_0 : \mu_3 = \mu_6$	$H_0 : \mu_4 = \mu_5$	$H_0 : \mu_4 = \mu_6$	$H_0 : \mu_5 = \mu_6$

   For Bonferroni's test

$H_0 : \mu_1 = \mu_2$	$H_0 : \mu_1 = \mu_3$	$H_0 : \mu_1 = \mu_4$	$H_0 : \mu_1 = \mu_5$	$H_0 : \mu_1 = \mu_6$
$H_0 : \mu_2 = \mu_3$	$H_0 : \mu_2 = \mu_4$	$H_0 : \mu_2 = \mu_5$	$H_0 : \mu_2 = \mu_6$	$H_0 : \mu_3 = \mu_4$
$H_0 : \mu_3 = \mu_5$	$H_0 : \mu_3 = \mu_6$	$H_0 : \mu_4 = \mu_5$	$H_0 : \mu_4 = \mu_6$	$H_0 : \mu_5 = \mu_6$

   (b) (2pt) On the SAS output, draw letters or lines vertically that correspond to your results in (a).

   (c) (1pt) If $\alpha = .05$ was changed to $\alpha = .10$, would it be possible to see more significant differences or fewer significant differences in the Bonferroni MCP? Briefly justify your answer.

For Stat 541 Students

4. Suppose the last response ($y_{34} = 38$) in Problem 2 is missing.

   (a) (1pt) Assuming the constraint $\tau_2 = 0$, write out the response vector y and matrix X using the $a + 1$ parameter form for X.

   (b) (1pt) Assuming the constraint $\sum n_i \tau_i = 0$, write out the response vector y and matrix X using the $a + 1$ parameter form for X.

5. (1pt) Verify $\dfrac{-\sum_{i=1}^{a-1} n_i(\overline{y}_{i\cdot} - \overline{y}_{\cdot\cdot})}{n_a} = \overline{y}_a - \overline{y}_{\cdot\cdot}$.

Multiple Comparison Test Results

t Tests (LSD) for y

Error Mean Square	7.200758
Critical Value of t	2.06390
Least Significant Difference	3.5027

t Grouping	Mean	N	group
	29.3	5	6
	24.0	5	5
	23.7	5	4
	20.0	5	3
	16.4	5	2
	15.4	5	1

Tukey's Studentized Range (HSD) Test for y

Error Mean Square	7.200758
Critical Value of Studentized Range	4.37265
Minimum Significant Difference	5.2475

Tukey Grouping	Mean	N	group
	29.3	5	6
	24.0	5	5
	23.7	5	4
	20.0	5	3
	16.4	5	2
	15.4	5	1

Bonferroni (Dunn) t Tests for y

Error Mean Square	7.200758
Critical Value of t	3.25838
Minimum Significant Difference	5.5299

Bon Grouping	Mean	N	group
	29.3	5	6
	24.0	5	5
	23.7	5	4
	20.0	5	3
	16.4	5	2
	15.4	5	1