STAT 514 Homework#4
(Due Thursday 09/24/15 BEFORE CLASS)

1. An experiment is conducted to study the effect of hormones injected into test rats. There are two distinct hormones (A,B) each with two distinct levels. For purposes here, we will consider this to be four different treatments labeled {A,a,B,b}. Each treatment is applied to six rats with the response being the amount of glycogen (in mg) in the liver.

Treatment	Responses					
A	106	101	120	86	132	97
a	51	98	85	50	111	72
B	103	84	100	83	110	91
b	50	66	61	72	85	60

Three contrasts are of interest. They are:

Comparison	A	a	B	b
Hormone A vs Hormone B	1	1	-1	-1
Low level vs High level	1	-1	1	-1
Equivalence of level effect	1	-1	-1	1

(a) Compute the one-way ANOVA (using `PROC GLM`) to see if there are any treatment differences.

(b) Are these contrasts orthogonal? Why or why not?

(c) Compute the single degree of freedom sum of squares and test each null hypothesis. Interpret the results. (NOTE: Be careful if you use a character string variable to denote the treatment levels. The order of the treatments SAS uses in the contrast statement is different than A, a, B, b.)

2. Suppose you performed an ANOVA with $a = 4$ treatments and $n = 5$ observations per treatment. If the MS_E=16 and $\alpha = .05$, what would the minimum difference have to be between any two means for you to conclude they were significantly different if

(a) You performed the LSD comparison procedure?

(b) You performed the Bonferroni comparison procedure?

(c) You performed Tukey's multiple comparison procedure (use Table VII)?

(d) You performed Scheffe's procedure?

(e) Explain the relationship between power and the minimum difference. Also state which of the above four is the most powerful and least powerful comparison procedure.

3. The following data are the treatment means from an experiment where each treatment was randomly and equally allocated to a total of 27 experimental units.

Trt	1	2	3
Mean	3	1	0

Suppose the experimenter planned to test the following three hypotheses and assume $MS_{\mathrm{E}} = 5$.

a) $H_0: \quad 3\mu_2 = \mu_3 + 2\mu_1$
b) $H_0: \quad \mu_1 = \mu_3$
c) $H_0: \quad \mu = 3$

(a) Using $\alpha = .05$, test each of the three hypotheses (two-sided).

(b) Which of the three linear combinations of means are contrasts? Why?

(c) Are any pairs of combinations orthogonal? Which ones and why?

(d) Suppose the experimenter was also interested in testing $H_0 : \mu_1 = 0$. Since the sample mean of the third group is zero, it appears this is the same test as hypothesis b). Is it? Why or why not?

4. A factor with three levels was studied in an experiment. The data is given as follows, in which the first column includes the treatments and the second column includes the responses. You can download the data, `hw4.dat`, from the class website.

```
1  2.23
1  3.04
.  ...
3  8.12
```

(a) Test the hypothesis that there is no difference across the treatments (use $\alpha = .05$).

(b) Use proper plots to check whether the constant variance assumption is valid. Can you use a formal test to support your conclusion?

(c) Generate the log s_i vs. log $\bar{y}_{i\cdot}$ plot (`trans.sas`) and estimate the possible transformation for variance stabilization.

(d) Use the formal Box-Cox procedure to identify the optimal transformation. You need use `trans1.sas` for this data set and generate proper output and plot to make the choice.

(e) Repeat (a) and (b) for the transformed data. You need to use some sas functions in the data step to generate the new data.

5. Four different designs for a digital computer circuit are being studied to compare the amount of defects. The following data have been obtained (`defects.dat` on the class website):

```
design defect
1   7
1   2
......
4   7
```

(a) Is the amount of defects present the same for all four designs? Use $\alpha = 0.05$.

(b) Check the model assumptions. In particular, how do you think about the normality assumption? Can you use any formal test to support your conclusion?

(c) Use the Kruskal-Wallis test for the data and compare the results with (a).

6. Ten needles were randomly selected from a branch of a loblolly pine tree. The stomata (microscopic breathing holes) are arranged in rows. On each needle, four rows are randomly selected and the number of stomata per centimeter for each of the rows was determined. The data below is in the file named `stomata.dat`.

Needle									
1	2	3	4	5	6	7	8	9	10
149	136	143	121	148	129	127	134	117	129
143	139	142	133	121	134	130	137	128	132
138	129	124	126	124	127	123	119	117	131
131	143	134	130	128	113	125	130	118	137

(a) Why is the random effects model appropriate here?

(b) Estimate all relevant variance components.

(c) What percentage of the overall variation in stomata number per centimeter is due to the needle?

(d) Construct a 95% CI for this ratio.

(e) Compute a 95% confidence interval for the average number of stomata per centimeter (HINT: What is the experimental unit for this question or use results from HW#1 problem 2 to compute the proper standard deviation).