STAT 502 - Homework 5

Due date: Monday, 22nd November. Total points: 20.

1. **(7 Points)** (Analyze the data from Exercise 6.2 in Oehlert.) Bacteria in solution are often counted by a method known as serial dilution plating. Petri dishes with a nutrient agar are inoculated with a measured amount of solution. After 3 days of growth, an individual bacterium will have grown into a small colony that can be seen with the naked eye. Counting original bacteria in the inoculum is then done by counting the colonies on the plate. Trouble arises because we don't know how much solution to add. If we get too many bacteria in the inoculum, the petri dish will be covered with a lawn of bacterial growth and we won't be able to identify the colonies. If we get too few bacteria in the inoculum, there may be no colonies to count. The resolution is to make several dilutions of the original solution (1:1, 10:1, 100:1, and so on) and make a plate for each of these dilutions. One of the dilutions should produce a plate with 10 to 100 colonies on it, and that is the one we use. The count in the original sample is obtained by multiplying by the dilution factor. Suppose that we are trying to compare three different Pasteurization treatments for milk. Fifteen samples of milk are randomly assigned to the three treatments, and we determine the bacterial load in each sample after treat- ment via serial dilution plating. The following table gives the counts.

Treatment 1	26×10^2	29×10^2	20×10^2	22×10^2	32×10^2
Treatment 2	35×10^3	23×10^3	20×10^3	30×10^3	27×10^3
Treatment 3	29×10^5	23×10^5	17×10^5	29×10^5	20×10^5

Test the null hypothesis that the three treatments have the same effect on bacterial concentration.

`R` **hints:**
Construct a `data.frame` with two columns, one for the factor and one for the response.

```
dat <- data.frame(responses = c(2600, 2900, 2000, 2200,3200,
                      35000,23000,20000,30000,27000,
                      2900000,2300000,1700000,2900000,2000000),
        treatment = as.factor(c(rep("1",5),rep("2",5),rep("3",5))))
```

   (a) **(1Pt)** Based on the data, do you expect the linear model ANOVA assumptions to be satistfied?

   (b) **(2Pts)** Perform the ANOVA and look at the residuals and their diagnostic plots. What conclussions can you draw?

   (c) **(2Pts)** Would a BoxCox transformation of the response be reasonable in this case? How do you suggest to transform the response?

   (d) **(2Pts)** After transforming the data, perform an ANOVA again and look at the new residuals and their diagnostic plots. What can you conclude about this model.

2. **(7 Points)** (Source: R.G. Peterson, *Agricultural Field experiments - Design and Analysis*, 1994, p. 113) A plant scientist wants to test the effect of a new herbicide on lentils. He considers the following treatments:

1.	Control (without weeding, no fertilizer)	`TR` = 1
2.	weeding by hand	`TR` = 2
3.	spraying with herbicide before	`TR` = 3
4.	spraying with herbicide after	`TR` = 4
5.	weeding by hand + fertilizer	`TR` = 5
6.	spraying with herbicide before + fertilizer	`TR` = 6
7.	spraying with herbicide after + fertilizer	`TR` = 7

The data ("lentil.dat") is available on the course Canvas page. In this exercise, we only consider the variable `TR` (treatment) and as response we consider the variable `Y` (the harvesting weight).

(a) **(2Pts)** Perform an ANOVA wih a 0.002-level test for any differences between treatments. In addition, check the model assumptions by analyzing the residuals (histogram, Tukey-Anscombe (fitted values vs. residuals) plot, QQ-plot).

`R` **hints:**

- `d.len <- read.table("lentil.dat", header = TRUE)`
- `d.len$TR <- factor(d.len$TR)`
- Use the function `stripchart()` to plot the data (`TR` on the x-axis and `Y` on th y-axis).

(b) In order to detect existing differences between treatments, we consider the following contrasts:

	Treatment						
Contrast	k_1	k_2	k_3	k_4	k_5	k_6	k_7
C1	-6	$+1$	$+1$	$+1$	$+1$	$+1$	$+1$
C2	0	-1	-1	-1	$+1$	$+1$	$+1$
C3	0	$+2$	-1	-1	$+2$	-1	-1
C4	0	0	-1	$+1$	0	-1	$+1$
C5	0	-2	$+1$	$+1$	$+2$	-1	-1
C6	0	0	$+1$	-1	0	-1	$+1$

i. **(1Pt)** Are these contrasts orthogonal?

ii. **(2Pts)** Which questions do contrasts C_1, C_2, C_3 and C_4 address (Give your interpretation of what each of these contrast is testing)?

(c) **(2Pts)** Test all contrasts in 2b. Use the Bonferroni correction so that the family wise (exeperiment wise) error rate is smaller than 0.048 (Remember that we already performed an F-test at 0.002 level).

`R` **hints:**

- Contrasts: Generate a matrix `mat.contr` with the contrasts C1 to C6 in the **rows** (either use `rbind()` or `matrix()`). Then use the following `R` code:
  ```
  library(multcomp)
  fit.mc <- glht(fit.len, linfct = mcp(TR = mat.contr))
  ```
- ANOVA tables for contrasts:
  ```
  summary(fit.mc, test = adjusted("none"))
  ```

3. **(6 Points)** The response time (in milliseconds) was determined for three types of electrical circuits. The results were:

Type 1	9	12	10	8	15
Type 2	20	21	23	17	30
Type 3	6	5	8	16	7

(a) **(2Pts)** Test the hypothesis that all types have the same expected response time at the 0.01-level (and check the assumptions of your model).

`R` **hints:**
We want to construct a `data.frame` with two columns, one for the factor and one for the response.

```
y <- c(9, 12, 10, 8, 15,
       20, 21, 23, 17, 30,
       6, 5, 8, 16, 7)
type <- c(...)
circ <- data.frame(Type = type, Y = y)
circ$Type <- as.factor(circ$Type)
```

(b) **(2Pts)** Construct a set of orthogonal contrasts, such that one of the contrasts compares circuit Type 2 with the other two ("Type 2 vs. Type 1 and Type 3").

(c) **(2Pts)** Test the contrasts using the Bonferroni-Holm criterion correction.

The Bonferroni-Holm procedure for performing k hypothesis tests with the family wise (experiment wise) error rate of α is as follows:

- Sort p-values of the k tests from smallest to largest: $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(k)}$.
- Let $j \in \{1, \ldots, k\}$ be the minimal index for which $p_{(j)} > \frac{\alpha}{k-j+1}$.
- Reject the null hypotheses $H_{0(i)}$, for $i < j$ and do not reject the rest. (If $p_{(j)} \leq \frac{\alpha}{k-j+1}$, for all $j \in \{1, \ldots, k\}$, then reject all $H_{0(j)}$.)

This criterion is less conservative and hence (uniformly) more powerful than the Bonferroni criterion. See `p.adjust` in `R` for this method and more.

`R` **hints:**

```
library(multcomp)
mat.contr <- rbind(...)
circ.mc <- glht(circ.fit, linfct = mcp(Type = mat.contr))
summary(circ.mc, test = adjusted("holm"))
```