STAT 461: Midterm 2 (Take-Home)

This is a take-home exam. You are allowed to use any non-human sources (internet, books, notes, etc), but you are NOT allowed to receive help from or work with any other person. If the instructor feels that cheating may have happened, an oral component will be added to the exam, with students each individually explaining their work. Any cheating will be addressed in accordance with Penn State's Academic Integrity policies.

For each experiment below, conduct a full analysis of the experimental data, and answer any additional questions as stated in each problem. In all cases, you should check model assumptions and make transformations to the response variable as needed. Your answer should contain all R code used, and you should describe the results of all important hypothesis tests you conduct.

1 (10pts)

The experiment was run to compare the effects of auditory and visual cues on speed of response of a human subject. A personal computer was used to present a "stimulus" to a subject, and the reaction time required for the subject to press a key was monitored. The subject was warned that the stimulus was forthcoming by means of an auditory or a visual cue. The experimenters were interested in the effects on the subjects' reaction time of the auditory and visual cues and also in different elapsed times between cue and stimulus. Thus, there were two different treatment factors: "cue stimulus" at two levels "auditory" or "visual," and "elapsed time between cue and stimulus" at three levels "five," "ten," or "fifteen" seconds.

The data are as follows:

```
cue=c(rep("auditory",9),rep("visual",9))
elapsed.time=rep(rep(1:3,each=3),2)
reaction.time=c(204,170,181,167,182,187,202,198,236,257,279,269,283,235,260,256,281,258)
```

1.1 Give a plot of either the response variable (reaction time), or the mean response variable, versus the two treatment factors: cue and elapsed time. You plot or plots should make it clear which treatments correspond to which response variables.

1.2 Give a complete analysis of this experimental data, under a 2-way complete model. You should show all R code used, and explain all important choices and results in your analysis. Interpret the results in the context of the experiment.

2 (10pts)

An experiment was conducted to compare melting times in seconds for three different brands of margarine (coded 1–3) and one brand of butter (coded 4). The butter was used for comparison purposes. The sizes and shapes of the initial margarine/butter pats were as similar as possible, and these were melted one by one in a clean frying pan over a constant heat.

```
Brand<-c(rep(1,10), rep(2, 10), rep(3,10), rep(4, 10))
Times<-c(167, 171, 178, 175, 184, 176, 185, 172, 178, 178,
         231, 233, 236, 252, 233, 225, 241, 248, 239, 248,
         176, 168, 171, 172, 178, 176, 169, 164, 169, 171,
         201, 199, 196, 211, 209, 223, 209, 219, 212, 210)
```

Give a complete analysis of this experimental data, under a 1-way complete model. You should show all R code used, and explain all important choices and results in your analysis. Interpret the results in the context of the experiment.

3 (10pts)

An analysis is conducted to determine if different species and types of wood influence the nitrogen content in standing trees. Trees are divided into two kinds of wood: hard wood (oak, ash, and maple), and soft wood (pine, spruce, and fir). A random selection of 4 trees of each kind (24 trees total) was chosen from all trees in the State Game Lands, and the nitrogen content was measured.

```
wood=read.table("wood.csv",header=TRUE)
wood
```

```
##        Type Species Nconc
## 1  softwood    pine    12
## 2  softwood    pine    13
## 3  softwood    pine    11
## 4  softwood    pine    12
## 5  softwood  spruce    15
## 6  softwood  spruce    19
## 7  softwood  spruce    17
## 8  softwood  spruce    17
## 9  softwood     fir    10
## 10 softwood     fir    12
## 11 softwood     fir    11
## 12 softwood     fir    17
## 13 hardwood   maple    18
## 14 hardwood   maple    20
## 15 hardwood   maple    21
## 16 hardwood   maple    16
## 17 hardwood     oak    20
## 18 hardwood     oak    14
## 19 hardwood     oak    17
## 20 hardwood     oak    15
## 21 hardwood     ash    19
## 22 hardwood     ash    22
## 23 hardwood     ash    21
## 24 hardwood     ash    21
```

3.1 Is this experiment a completely randomized design? Why or why not?

3.2 Give a complete analysis of this data, under a 2-way nested model, with species nested in type. Show all R code used, and explain all important choices and results in your analysis. Interpret the results in the context of the experiment.

4 (10pts)

An experiment was conducted to determine the best recipe for different kinds of canned beans. Beans are divided into four different crocks (i=1,2,3,4). The beans are soaked before cooking for either a long or a short time (j=short,long). Two of the crocks are randomly chosen to soak for a short time, and the other two crocks are allowed to soak for a long time. After soaking, the beans from each crock are divided into three jars, and are used to make baked beans using one of three recipes (k=Original, Barbecue, or Refried). Finally, beans from each jar are fed to people, and the average taste rating of for each jar is recorded.

```
beans=read.table("Beans.csv",header=TRUE)
beans
```

```
##    Crock SoakTim   Recipe Jar Rating
```

```
## 1      1    Long Original   1     45
## 2      1    Long Barbecue   2     50
## 3      1    Long  Refried   3     44
## 4      2   Short Original   4     33
## 5      2   Short Barbecue   5     40
## 6      2   Short  Refried   6     40
## 7      3    Long Original   7     46
## 8      3    Long Barbecue   8     49
## 9      3    Long  Refried   9     45
## 10     4   Short Original  10     32
## 11     4   Short Barbecue  11     41
## 12     4   Short  Refried  12     41
```

4.1 Explain why Jar is not treated as a factor in this experiment.

4.2 Give a complete analysis of this data, under the following model.

$$Y_{ijk} = \mu + \alpha_j + \beta_{i(j)} + \gamma_k + (\alpha\gamma)_{jk} + \epsilon_{ijk}, \quad \epsilon_{ijk} \sim N(0, \sigma^2)$$

$$\beta_{i(j)} \sim N(0, \sigma^2_{crock})$$

Show all R code used, and explain all important choices and results in your analysis. Interpret the results in the context of the experiment.