MTH 3220 Lab 7

Due Thu., Oct. 31

1 Part A: Two-Factor ANOVA With More Than One Observation Per Group

1.1 Memory Data Set

The file **memory.txt** contains data from an experiment to test the effects of two factors, level of **reinforcement** (*none* or *verbal*) and time of **isolation** (20, 40, or 60 minutes), on students' abilities to recall information. The response variable is a score on a test that assesses a student's ability to recall items.

1. Use `read.table()` to read the data into a data frame in R.
2. Carry out a ***two-factor ANOVA*** using the **full model**. An interaction term is specified using a colon (`:`), for example:

   ```
   my.anova <- aov(Score ~ Reinforcement + Isolation + Reinforcement:Isolation,
                   data = my.data)
   ```

   Then look at the ANOVA table using:

   ```
   summary(my.anova)
   ```

3. Make a ***histogram*** and ***normal probability plot*** of the **residuals** to check the normality assumption:

   ```
   hist(my.anova$residuals)
   qqnorm(my.anova$residuals)
   qqline(my.anova$residuals)
   ```

4. Make a plot of the **residuals** (y-axis) versus the **fitted values** (x-axis) to check the constant standard deviation assumption:

   ```
   plot(x = my.anova$fitted.values, y = my.anova$residuals)
   abline(h = 0)
   ```

5. The function `interaction.plot()` will produce an interaction plot. Among its arguments are:

`x.factor` a factor whose levels will form the x axis.

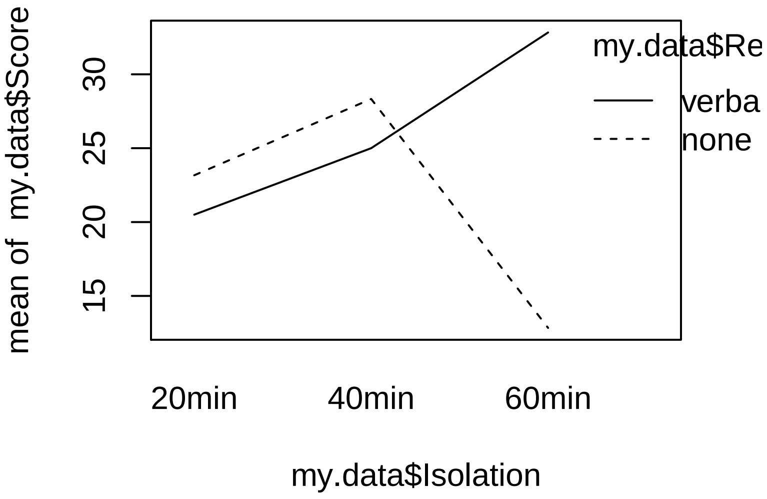
`trace.factor` another factor whose levels will form the traces (lines).

`response` a numeric variable giving the response.

Make an interaction plot, with the time of **isolation** on the x-axis and the levels of **reinforcement** represented by separate lines, for example by typing:

```
interaction.plot(x.factor = my.data$Isolation,
                 trace.factor = my.data$Reinforcement,
                 response = my.data$Score)
```

You plot should look similar to the one below:



6. Because the interaction effect is significant, it **doesn't make sense to try to interpret the results of the tests for main effects**.

   Instead, we'll compare **group means** using ***Tukey's multiple comparisons procedure*** to decide **which *group means*** differ. Carry out **Tukey's procedure** by typing:

   ```
   TukeyHSD(my.anova)
   ```

2 Part B: Summarizing and Graphing Data from Two-Factor Studies

2.1 Memory Data Set (Cont'd)

In this problem we'll use the data from the file **memory.txt** (again).

1. Compute the sample **grand mean** $\bar{X}_{...}$ using:

   ```
   mean(my.data$Score)
   ```

2. The function `aggregate()` will compute summary statistics separately for each level of a factor and also separately for each group in a two-factor study.

   a) Use `aggregate()` to compute the ***level means*** $\bar{X}_{i..}$ for the two levels of **reinforcement**:

   ```
   aggregate(Score ~ Reinforcement, data = my.data, FUN = mean)
   ```

   b) Use `aggregate()` to compute the ***level means*** $\bar{X}_{.j.}$ for the three levels of **isolation**:

   ```
   aggregate(Score ~ Isolation, data = my.data, FUN = mean)
   ```

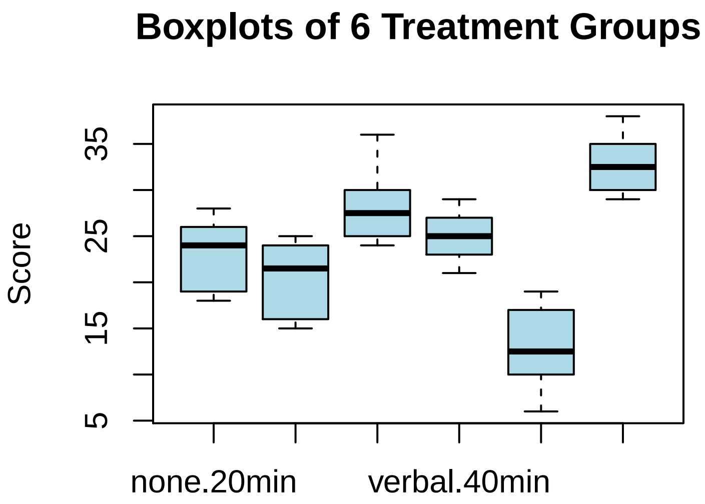
   c) Use `aggregate()` to compute the ***group means*** $\bar{X}_{ij.}$ for the six **groups**:

   ```
   aggregate(Score ~ Reinforcement + Isolation, data = my.data, FUN = mean)
   ```

3. We can pass a *formula* to `boxplot()` along with our *data frame* to make ***side-by-side boxplots*** of the scores for the six **groups**. Make the **boxplots**, for example by typing:

   ```
   boxplot(Score ~ Reinforcement + Isolation, data = my.data,
           ylab = "Score", col = "lightblue",
           main = "Boxplots of Six Treatment Groups")
   ```

   The result should by similar to:



3 Part C: Three-Factor ANOVA

3.1 Electrical Stimulation of Muscles Data Set

Muscles that have lost connections to nerve fibers are said to be *denervated.* Motor nerve transmissions to denervated muscles are decreased or nonexistent, leading loss of muscle tissue.

A number of experiments have suggested that electrical stimulation may be helpful in preventing the loss of tissue in denervated muscles.

A three-factor experiment on rats was conducted to learn something about the best type of **current** and the most effective **method of treatment**. The factors and their levels are shown below:

Factor A	**Factor B**	
Number of treatment	Length of treatment	**Factor C**
periods daily	(minutes)	Type of current
1	1	Galvanic
3	2	Faradic
6	3	60 cycle alternating
	5	25 cycle alternating

Treatments were started on the third day after denervation and continued for 11 consecutive days. There were 48 different combinations of methods of treatments ($IJK = 3 \times 4 \times 4 = 48$), each of which was applied to two rats, using 96 rats in all.

The rats' muscles were denervated by the removal of a small part of the sciatic nerve from the gastroecnemius-soleus (calf) muscle on one side. The weights of the denervated muscles were measured at the end of the experiment. Since this depends on the size of the animal, the weight of the corresponding unaffected muscle on the other side was also measured for use as a reference.

Data from the experiment are contained in the file **muscles.txt**. The variable `RelativeDiff` contains the relative difference in muscle weights, i.e.

$$\text{Relative Difference} = \frac{\text{Normal Wt} - \text{Denervated Wt}}{\text{Normal Wt}}.$$

1. Use `read.table()` to read the data into a data frame in R.

2. We want to check for a ***three-factor interaction*** graphically by making **four** separate **interaction plots** of `NumberOfTrts` and `TrtLength`, one plot for each `TypeOfCurrent`.

   To make the four **interaction plots**, we'll first create four new data sets, one for each `TypeOfCurrent`.

   The `subset()` function extracts a subset of rows from a data frame. Among its arguments are:

`x`	a data frame.
`subset`	a `"logical"` expression indicating rows to keep.

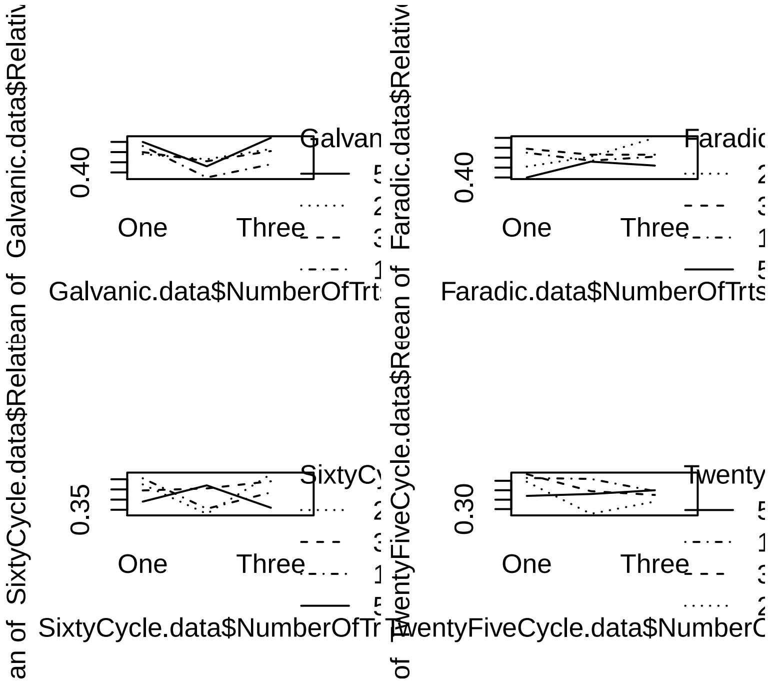
   Use `subset()` to create the four new data sets, for example by typing:

```
Galvanic.data <- subset(x = my.data, subset = my.data$TypeOfCurrent == "Galvanic")
Faradic.data <- subset(x = my.data, subset = my.data$TypeOfCurrent == "Faradic")
SixtyCycle.data <- subset(x = my.data, subset = my.data$TypeOfCurrent == "60Cycle")
TwentyFiveCycle.data <- subset(x = my.data, subset = my.data$TypeOfCurrent == "25Cycle")
```

3. Use `interaction.plot()` to make four separate interaction plots, one for each `TypeOfCurrent`, with `NumberOfTrts` on the horizontal axis and and `TrtLength` represented by separate lines, by typing:

```
interaction.plot(x.factor = Galvanic.data$NumberOfTrts,
  trace.factor = Galvanic.data$TrtLength, response = Galvanic.data$RelativeDiff)
interaction.plot(x.factor = Faradic.data$NumberOfTrts,
  trace.factor = Faradic.data$TrtLength, response = Faradic.data$RelativeDiff)
interaction.plot(x.factor = SixtyCycle.data$NumberOfTrts,
  trace.factor = SixtyCycle.data$TrtLength, response = SixtyCycle.data$RelativeDiff)
interaction.plot(x.factor = TwentyFiveCycle.data$NumberOfTrts,
  trace.factor = TwentyFiveCycle.data$TrtLength,
  response = TwentyFiveCycle.data$RelativeDiff)
```

   Your graphs should look like this:



4. Carry out a three-factor ANOVA, with `RelativeDiff` as the response and `NumberOfTrts`, `TrtLength`, and `TypeOfCurrent` as the factors:

```
my.anova <- aov(RelativeDiff ~ NumberOfTrts + TrtLength + TypeOfCurrent
                                + NumberOfTrts:TrtLength
                                + NumberOfTrts:TypeOfCurrent
                                + TrtLength:TypeOfCurrent
                                + NumberOfTrts:TrtLength:TypeOfCurrent,
                data = my.data)

summary(my.anova)
```

5. Check the **normality assumption** by making a ***histogram*** and ***normal probability plot*** of the **residuals**:

```
hist(my.anova$residuals)

qqnorm(my.anova$residuals)
qqline(my.anova$residuals)
```

6. Check the **constant standard deviation assumption** by making a plot of the **residuals** versus the **fitted values**:

```
plot(x = my.anova$fitted.values, y = my.anova$residuals)
abline(h = 0)
```