Lab Exercise 5 - Simple ANOVA

There is no answer sheet this week. Hand in your work on a separate sheet, typed! No e-mails! You may copy and paste stuff from R if you wish, but if you do, SET THE FONT TO A MONOSPACED FONT (such as courier or courier new). I will not be pleased to see R output that is not lined up properly because it is in a proportionally spaced font! Please CLEARLY label and explain what you are giving me. You were given the name of the dataset you should use in class. (If you didn't get one, see me.) Here is an example of the kind of thing you should do along with some helpful pointers.

----------------------------------------------------------------------------------------------------------------------------
William King
Psyc 480 - exercise 5

Dataset: Firemen.csv (### first tell me what dataset you were using; I explained this one in class ###)

(### all stuff copied from R in a monospaced font--this is courier 10 point ###)

> firemen = read.csv("http://ww2.coastal.edu/kingw/psyc480/labdata/Firemen.csv")

Variables: The IV was risk rating assigned to the fireman by his superior. The DV was the fireman's score on the Rotter Locus of Control scale. On this scale, low scores indicate internal locus of control.

Design: This was a quasi-experimental design using intact groups.

Hypotheses:
   null - there is no difference in mean locus of control scores by risk rating
   alternative - there is a difference in mean locus of control scores by risk rating

Summary of the dataset:

> summary(firemen)

Rotter  Area     Risk
Min.   : 4.00  Charleston:25 A:26
1st Qu.: 8.00   Horry    :25 B:33
Median :10.00   NYC      :25 C:16
Mean   : 9.96
3rd Qu.:12.00
Max.   :16.00

(### you do not have to show the R commands, and I especially don't want to see the mistakes you made in R; notice my dataset contains variables I am not using; yours may or may not ###)

> with(firemen, tapply(Rotter, Risk, mean))  # these are the group means
A           B           C
8.961538    10.212121   11.062500

> with(firemen, tapply(Rotter, Risk, var))   # these are the group variances
A           B           C
9.958462    7.297348    6.862500

(### once again, you don't have to show me the R commands, but make sure it is clear what the numbers are that you are showing me; one way to do that is to comment your R commands ###)
Testing assumptions of ANOVA:
  homogeneity of variance - met (the variances appear to be reasonably similar)
  normality - met (the qqnorm plots are reasonably straight)

(### FYI: the commands I used to draw these plots were...
  > par(mfrow=c(1,3))
  > with(firemen, tapply(Rotter, Risk, qqnorm))
  ...I set mfrow equal to one row and three columns, because I was drawing three graphs, one for each
  level of the IV; if you have more than three levels, you might consider drawing them in two rows ###)

Graphics:

Note: points plotted as filled circles are group means.

(### make sure I know what you're showing me here; the graph was resized to fit on the page; FYI, the
commands I used to draw this graph were...
  > boxplot(Rotter ~ Risk, data=firemen)
  > points(x=1:3, y=c(8.96,10.21,11.06), pch=16, cex=1.5)
  > title(main="Rotter Scores by Risk Ratings")
  > title(xlab="Risk Ratings")
  > title(ylab="Rotter Scores")  ...###)

2
ANOVA summary table and decision regarding the hypotheses:

```r
> aov.out = aov(Rotter ~ Risk, data=firemen)
> summary(aov.out)

Df  Sum Sq  Mean Sq  F value Pr(>F)
Risk         2   47.5  23.733   2.919 0.0604 .
Residuals   72 585.4   8.131
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Since \( p > .05 \), I cannot reject the null hypothesis and cannot say there were any significant differences among the sample means. I find no evidence that Rotter scores are related to Risk ratings.

Effect size:

```r
> 47.5 / (47.5 + 585.4)   # this is eta-squared
[1] 0.07505135
```

The effect size was small.

Post hoc tests:

(### I want you to do post hoc tests whether or not you rejected the null; ordinarily you would not do them if the null hypothesis was not rejected; you can use any of the three tests we discussed in class, provided it is appropriate--Fisher LSD is not appropriate if you are doing too many comparisons ###)

Pairwise post hoc comparisons were computed using the Fisher LSD procedure.

<table>
<thead>
<tr>
<th>comparison</th>
<th>mean diff.</th>
<th>LSD</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk A vs. Risk B</td>
<td>1.25</td>
<td>1.495</td>
<td>not significant</td>
</tr>
<tr>
<td>Risk A vs. Risk C</td>
<td>2.10</td>
<td>1.812</td>
<td>significant</td>
</tr>
<tr>
<td>Risk B vs. Risk C</td>
<td>0.85</td>
<td>1.737</td>
<td>not significant</td>
</tr>
</tbody>
</table>

(### here is another way to compute this; instead of computing three t-tests, compute a least significant difference using the t-test formula, and then compare the obtained mean differences to the computed values; this will not work if you are going to use the Fisher calculations as the basis of a Bonferroni test, however ###)

\[
t = \frac{\text{mean1} - \text{mean2}}{sp \cdot \sqrt{\frac{1}{n1} + \frac{1}{n2}}}
\]

\[
\text{LSD.AB} = 2.000 \times 2.851 \times \sqrt{\frac{1}{26} + \frac{1}{33}} = 1.495
\]
\[
\text{LSD.AC} = 2.000 \times 2.851 \times \sqrt{\frac{1}{26} + \frac{1}{16}} = 1.812
\]
\[
\text{LSD.BC} = 2.000 \times 2.851 \times \sqrt{\frac{1}{33} + \frac{1}{16}} = 1.737
\]