One-way ANOVA Homework

PSYCH 710

Before answering the questions in this handout, you should read the handout that provides an example of how to conduct a one-way ANOVAs in R, which can be obtained here: http://www.psychology.mcmaster.ca/bennett/psy710/notes/Anova_example.pdf.

Michaelson-Morley Speed of Light Data

Load the data from the Michaelson-Morley experiments that measured the speed of light:

> lightspeed <- morley # load it into R

morley is one of R’s built-in data sets; we’ve saved a copy of the data in our variable, lightspeed. You can read more about the data with the help command

> ?morley # read about the data file

and get a complete list of built-in data sets with the command

> data()

Next, we need to convert the variable Expt from an vector of integers into a qualitative grouping variable, or factor:

> names(lightspeed)

[1] "Expt"  "Run"  "Speed"

> class(lightspeed$Expt)

[1] "integer"

> lightspeed$Expt <- as.factor(lightspeed$Expt)

> class(lightspeed$Expt)

[1] "factor"

Finally, if you have not done so already, please set up R to use the sum-to-zero definition for effects:

> options(contrasts=c("contr.sum","contr.poly"))

Answer all of the following questions.

1. Inspect the data:

   (a) Use tapply to calculate the mean and standard deviation for the dependent variable, Speed, for each experiment.

       > with(lightspeed,tapply(Speed,Expt,mean)) # mean
Figure 1: Speed of light measured in different experiments.

泡 0 3 4 5 909.0 856.0 845.0 820.5 831.5
> with(lightspeed,tapply(Speed,Expt,sd)) # stan dev
1 2 3 4 5 104.92604 61.16414 79.10686 60.04165 54.21934
(b) Use boxplots to examine the distribution of Speed for each experiment. **Answer:** The following command was used to create Figure 1:
> boxplot(Speed~Expt,data=lightspeed,main="lightspeed")

2. Use ANOVA to investigate the proposition that the measured speed of light varied significantly across experiments. Clearly state the null and alternative hypotheses, as well as your conclusion.

> lm.02 <- lm(Speed~Expt,data=lightspeed) # linear model
> anova(lm.02) # anova table

Analysis of Variance Table

<table>
<thead>
<tr>
<th>Response: Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Df</td>
</tr>
<tr>
<td>Expt</td>
</tr>
<tr>
<td>Residuals</td>
</tr>
</tbody>
</table>

---

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
**Answer:** The null hypothesis is that the average light speed was the same in all experiments. The alternative hypothesis is that average light speed was not the same in all experiments. The effect of experiment was significant \((F(4, 95) = 4.28, p = 0.0031)\), and therefore the null hypothesis is rejected in favor of the alternative that average speed varied across experiments.

3. Calculate Cohen’s \(f\).

\[
\text{summary(lm.02)}
\]

Call:
```
lm(formula = Speed ~ Expt, data = lightspeed)
```

Residuals:
```
  Min 1Q Median 3Q Max
-259.00 -42.62  2.25 41.75 161.00
```

Coefficients:
```
  Estimate Std. Error t value Pr(>|t|)
(Intercept)  852.400    7.423  114.827  < 2e-16 ***
Expt1        56.600    14.847   3.812  0.000245 ***
Expt2         3.600    14.847   0.242   0.808933
Expt3        -7.400    14.847  -0.498   0.619336
Expt4        -31.900    14.847  -2.149  0.034207 *
```

---

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 74.23 on 95 degrees of freedom
Multiple R-squared: 0.1529,      Adjusted R-squared: 0.1173
F-statistic: 4.288 on 4 and 95 DF, p-value: 0.003114

\[
\text{adj.r.squared} <- .1173;
\]
\[
\text{omega.squared} <- \text{adj.r.squared}
\]
\[
\text{cohen.f} <- \sqrt{\text{adj.r.squared}/(1-\text{adj.r.squared})}
\]

\[
[1] 0.3645377
\]

4. List the coefficients for the full model. Show how these coefficients are related to the group means.

**Answer:** Note: the following answer assumes that you have used the `options` function to set up R to use the sum-to-zero constraint to define the effects of different levels within a factor (see above). Here are the coefficients of the best-fitting (least-squares) model. The first value is the intercept and the next four values represent \(\alpha_j\) for groups \(j = 1 \cdots 4\):

\[
\text{coef(lm.02)}
\]

```
(Intercept)  Expt1  Expt2  Expt3  Expt4
     852.4     56.6     3.6    -7.4   -31.9
```

Using the sum-to-zero constraint, \(\alpha_{j=5}\) can be calculated from the other \(\alpha\)'s:

\[
0 - 56.6 - 3.6 - (-7.4) - (-31.9)
\]

\[
[1] -20.9
\]
The intercept corresponds to the unweighted mean of the group means. In other words, it is the mean of the group means ignoring the number of observations per group:

```r
> (Yu <- mean(with(lightspeed,tapply(Speed,Expt,mean)) ) ) # unweighted mean
[1] 852.4
```

The value of $\alpha_j$ corresponds to the difference between the mean of group $j$ and the unweighted mean:

```r
> group.means <- with(lightspeed,tapply(Speed,Expt,mean))
> (alphas <- group.means - Yu )
```

```
1   2   3   4   5
56.6 3.6 -7.4 -31.9 -20.9
```