STAT 6910: HW 4

David Angeles

```
## Warning: package 'ggplot2' was built under R version 3.4.4

## Warning: package 'emmeans' was built under R version 3.4.4

## NOTE: As of emmeans versions > 1.2.3,
##       The 'cld' function will be deprecated in favor of 'CLD'.
##       You may use 'cld' only if you have package:multcomp attached.
```

Problem 1 (3a)

The meat cooking experiment was described in Exercise 14 of Chap. 3, and the data were given in Table 3.14, p. 68.

(a) Compare the effects of the six treatments, pairwise, using Scheffe's method of multiple comparisons and a 95% overall confidence level.

The data displayed below is prepresented by the codes 1, 2, 3, 4, 5, and 6 which denote the frying fat content at 10%, 15%, and 20% and the grilling fat content at 10%, 15% and 20% respectively, for the post-cooking weight data (in grams) for the meat cooking experiment.

```
head(Post_grams)
```

```
##      [,1] [,2]
## [1,]   81    1
## [2,]   88    1
## [3,]   85    1
## [4,]   84    1
## [5,]   84    1
## [6,]   85    2
```

```
colnames(Post_grams) <- c( "Weight", "Code")
Post_grams <- data.frame(Post_grams)
Post_grams$Code <- factor(Post_grams$Code)
summary(Post_grams)
```

```
##      Weight      Code
##  Min.   :71.00   1:5
##  1st Qu.:80.00   2:5
##  Median :82.00   3:5
##  Mean   :81.67   4:5
##  3rd Qu.:84.75   5:5
##  Max.   :88.00   6:5
```

We can see that the contrasts 1 - 2, 1 - 3, 1 - 4, 1 - 5, 1 - 6, 2 - 3, 2 - 4, 2 - 5, 2 - 6, 3 - 4, 3 - 5, 3 - 6, 4 - 5, 4 - 6, and 5 - 6 have the following 95% overall confidence intervals respectively.

```r
Post_weight_model <- aov(Weight ~ Code , data = Post_grams)
weight_means <- emmeans(object = Post_weight_model, specs = "Code")
estimates <- contrast(weight_means,method = "pairwise", adjust = "bonferroni")
alpha <- 0.05
v = 6
n = 30
wS <- sqrt((v-1) * qf(1-alpha,v-1,n-v))
cbind(summary(estimates)[,"estimate"] - wS* summary(estimates)[,"SE"],
      summary(estimates)[,"estimate"] + wS* summary(estimates)[,"SE"])
```

```
##                  [,1]        [,2]
##  [1,]   -3.5714537   8.771454
##  [2,]    2.2285463  14.571454
##  [3,]   -5.1714537   7.171454
##  [4,]   -7.7714537   4.571454
##  [5,]   -0.1714537  12.171454
##  [6,]   -0.3714537  11.971454
##  [7,]   -7.7714537   4.571454
##  [8,]  -10.3714537   1.971454
##  [9,]   -2.7714537   9.571454
## [10,]  -13.5714537  -1.228546
## [11,]  -16.1714537  -3.828546
## [12,]   -8.5714537   3.771454
## [13,]   -8.7714537   3.571454
## [14,]   -1.1714537  11.171454
## [15,]    1.4285463  13.771454
```

Problem 2 (5)

Exercise 15 of Chap. 3 (p. 67) concerns a study of the effects of four levels of sulfamerazine (0, 5, 10, 15 g per 100 lb of fish) on the hemoglobin content of trout blood. An analysis of variance test rejected the hypothesis that the four treatment effects are the same at significance level $\alpha = 0.01$.

(a) Compare the four treatments using Tukey's method of pairwise comparisons and a 99% overall confidence level.

The Tukey's method of pairwise comparisons and 99% overall confidence level are displyed below.

```r
Hemoglobin_2 <- matrix(c( rep(1, each = 10), rep(2, each = 10),
                          rep(3, each = 10), rep(4, each = 10),
            6.7, 7.8, 5.5, 8.4 ,7.0, 7.8, 8.6,7.4 ,5.8, 7.0,
```

```
            9.9, 8.4 ,10.4, 9.3 ,10.7 ,11.9, 7.1, 6.4 ,8.6, 10.6,
            10.4,8.1 ,10.6, 8.7, 10.7, 9.1, 8.8,8.1, 7.8, 8.0,
            9.3, 9.3 ,7.2, 7.8, 9.3, 10.2, 8.7, 8.6, 9.3, 7.2),
            byrow = FALSE, ncol = 2 )
Hemoglobin_2 <- data.frame(Hemoglobin_2)
Hemoglobin_2$X1 <- as.factor(Hemoglobin_2$X1)
colnames(Hemoglobin_2) <- c("Codes", "Hemo")

hem<- aov(Hemo ~ Codes, data = Hemoglobin_2)
hemoglobin_means<- emmeans(object = hem, specs = "Codes")

hemo.tukey <- contrast(hemoglobin_means,method = "pairwise", adjust = "tukey")
summary(hemo.tukey,infer = c(TRUE,TRUE), level = .99)
```

```
## contrast estimate          SE df  lower.CL     upper.CL t.ratio p.value
## 1 - 2        -2.13 0.5601141 36 -4.003146 -0.25685355  -3.803  0.0029
## 1 - 3        -1.83 0.5601141 36 -3.703146  0.04314645  -3.267  0.0122
## 1 - 4        -1.49 0.5601141 36 -3.363146  0.38314645  -2.660  0.0539
## 2 - 3         0.30 0.5601141 36 -1.573146  2.17314645   0.536  0.9498
## 2 - 4         0.64 0.5601141 36 -1.233146  2.51314645   1.143  0.6660
## 3 - 4         0.34 0.5601141 36 -1.533146  2.21314645   0.607  0.9292
##
## Confidence level used: 0.99
## Conf-level adjustment: tukey method for comparing a family of 4 estimates
## P value adjustment: tukey method for comparing a family of 4 estimates
```

(b) Compare the effect of no sulfamerazine on the hemoglobin content of trout blood with
the average effect of the other three levels. The overall confidence level of all intervals
in parts (a) and (b) should be at least 98%.

Using a FWER $\leq$ .02 and the code below, our data do not provide evidence that no
sulfamerazine on the hemoglobin content of trout blood with the average effect of the other
three levels.

```
contrast(hemoglobin_means,method = list(c(3,-1,-1,-1)))
```

```
## contrast          estimate       SE df t.ratio p.value
## c(3, -1, -1, -1)     -5.45 1.371994 36  -3.972  0.0003
```

Problem 3 (7a and b)

The soap experiment was described in Sect. 2.5.1, p. 20, and an analysis was given in Sect.
3.7.2, p. 50.

(a) Suppose that the experimenter had been interested only in the contrast $\tau_1 - \frac{1}{2}(\tau_2 + \tau_3)$,
which compares the weight loss for the regular soap with the average weight loss for

3

the other two soaps. Calculate a confidence interval for this single contrast.

Since it is being assumed that the experimenter migh have been intersted only in one contrast, there isn't a need to use a family-eise error rate. Therefore, using an α-level of $\alpha = .05$ and results on pg. 51, we can calcualte a confidence interval for this single contrast as follow:

$$\sum_{i=1}^{v} b_i \hat{Y}_{i\bullet} \pm t_{n-v,\alpha/2} \sqrt{MSE \sum_{i=1}^{v} \frac{b_i^2}{r_i}}$$

$$= -0.035 - \frac{1}{2}(2.7) - \frac{1}{2}(1.9925) \pm t_{9,.025} \sqrt{0.0772 \sum_{i=1}^{v} \frac{b_i^2}{r_i}}$$

$$= -2.38125 \pm 2.262 \sqrt{0.0772 \left(\frac{1}{4} + \frac{1/4}{4} + \frac{1/4}{4} \right)}$$

$$= -2.38125 \pm 2.262 \sqrt{0.0772 \left(\frac{3}{8} \right)}$$

$$= -2.38125 \pm .38487$$

Therefore the 95% confidence interval for the specific constrast is (-2.76612,-1.99638).

(b) Test the hypothesis that the regular soap has the same average weight loss as the average of the other two soaps. Do this via your confidence interval in part (a) and also via (4.3.13) and (4.3.15).

For this test, we will use the following:

The Null Hypothesis of No Difference Between Regular Soap and Others $H_0 : \tau_1 = \frac{1}{2}(\tau_2 + \tau_4)$ vs. The Alternative Hypothesis of Some Difference Between Regular Soap and Others
$$H_A : \tau_1 \neq \frac{1}{2}(\tau_2 + \tau_4)$$

Using the confidence interval, we see that 0 is not in the interval and therefore we can reject the null hypothesis at the $\alpha = .05$ level in favor of the alternative hypothesis that the regular soap does not have the same average weight loss as the average of the other two soaps. Next, using Equation 4.3.13 we get

$$\left| \frac{\sum c_i \bar{y}_i}{\sqrt{msE \sum c_i^2 / r_i}} \right| = \left| \frac{-2.38125}{\sqrt{0.0772 \left(\frac{3}{8} \right)}} \right| = 13.99525 > 2.201.$$

So we can reject the null hypothesis again. Lastly, Using Equation 4.3.15 we get

$$\left| \frac{\frac{\left(\sum c_i \bar{y}_i \right)^2}{\sum c_i^2 / r_i}}{msE} \right| = \left| \frac{\frac{5.67035}{3/8}}{.0772} \right|$$

$$= 195.8671 > 4.75.$$

```r
qf(.95, 1, 12)
```

```
## [1] 4.747225
```

Problem 4 (10)

Consider again the trout experiment in Exercise 15 of Chap. 3.

(a) Suppose the experiment were to be repeated. Suggest the largest likely value for the error mean square msE.

(b) How many observations should be taken on each treatment so that the length of each interval in a set of simultaneous 95% confidence intervals for pairwise comparisons should be at most 2 g per 100 ml?