

# Week 12 Discussion

Hanying Jiang

## Exercise 1

### Question

Consider an experiment where we are interested in determining the effectiveness of several antibiotics applied at the same concentration. Treatments are randomly applied to cultured bacteria, and the percent reductions in bacterial population after a specified amount of time are given below:

Trt 1: 13.3, 13.7, 11.6, 11.9, 12.0, 12.9, 12.3, 12.1, 12.3, 12.3

Trt 2: 12.5, 13.9, 12.4, 14.0, 12.8

Trt 3: 16.8, 16.6, 17.1, 15.8, 17.4, 16.5

Trt 4: 13.2, 14.2, 13.5, 13.2, 14.3, 12.7, 14.9, 13.3, 13.5

- a. Analyze the data graphically. What properties from the graphical summaries are noteworthy for considering an ANOVA analysis?

```
# Data
trt1 <- c(13.3, 13.7, 11.6, 11.9, 12, 12.9, 12.3, 12.1, 12.3, 12.3)
trt2 <- c(12.5, 13.9, 12.4, 14, 12.8)
trt3 <- c(16.8, 16.6, 17.1, 15.8, 17.4, 16.5)
trt4 <- c(13.2, 14.2, 13.5, 13.2, 14.3, 12.7, 14.9, 13.3, 13.5)

# Create vectors of all observations and treatments
q1_vals <- c(trt1, trt2, trt3, trt4)
q1_trt <- c(rep("Trt1", 10), rep("Trt2", 5), rep("Trt3", 6), rep("Trt4", 9))
```

- b. Summarize the data numerically (your values should match the table below). What properties from the numeric summaries are noteworthy for considering an ANOVA analysis?

Group	Sample Mean	Sample SD	Sample Size
1	12.44	0.6586181	10
2	13.12	0.7726578	5
3	16.700	0.5513620	6
4	13.644	0.6857680	9

- c. Create an ANOVA table by hand, and check your work with the `aov` function in R.

Source	SS	df	MS	F	pvalue
Treat					
Error					
Tot					

- d. Decide whether we have sufficient evidence to reject the null after checking that the normality and equal variances assumptions are reasonably met.
- e. Conduct pairwise comparisons for all treatment pairs using 95% CI based on Fisher's method by hand. Make a table to summarize your findings of which groups are significantly different.

Summary Measures:

Group	Sample Mean	Sample SD	Sample Size
1	12.44	0.6586181	10
2	13.12	0.7726578	5
3	16.700	0.5513620	6
4	13.644	0.6857680	9

Treat	Mean	Group
Trt 3	16.70	
Trt 4	13.644	
Trt 2	13.12	
Trt 1	12.44	

- f. Explain why a researcher may instead choose to construct Tukey-Kramer adjusted CIs. How do these CIs differ from those computed in (e)? Look at the Tukey multiplier and 95% Tukey CIs below and resummairize the groupings found in (e) as necessary.

```
# Find Tukey multiplier
qtukey(0.95, nmeans = 4, df = 26)/sqrt(2)
```

```
## [1] 2.74332
```

Tukey CIs:

```
μ1 − μ2 : (−1.683, 0.323)
μ1 − μ3 : (−5.205, −3.315)
μ1 − μ4 : (−2.045, −0.363)
μ2 − μ3 : (−4.688, −2.472)
μ2 − μ4 : (−1.545, 0.496)
μ3 − μ4 : (2.091, 4.020)
```

Treat	Mean	Group
Trt 1	12.44	
Trt 2	13.12	
Trt 4	13.644	

Treat	Mean	Group
Trt 3	16.70	

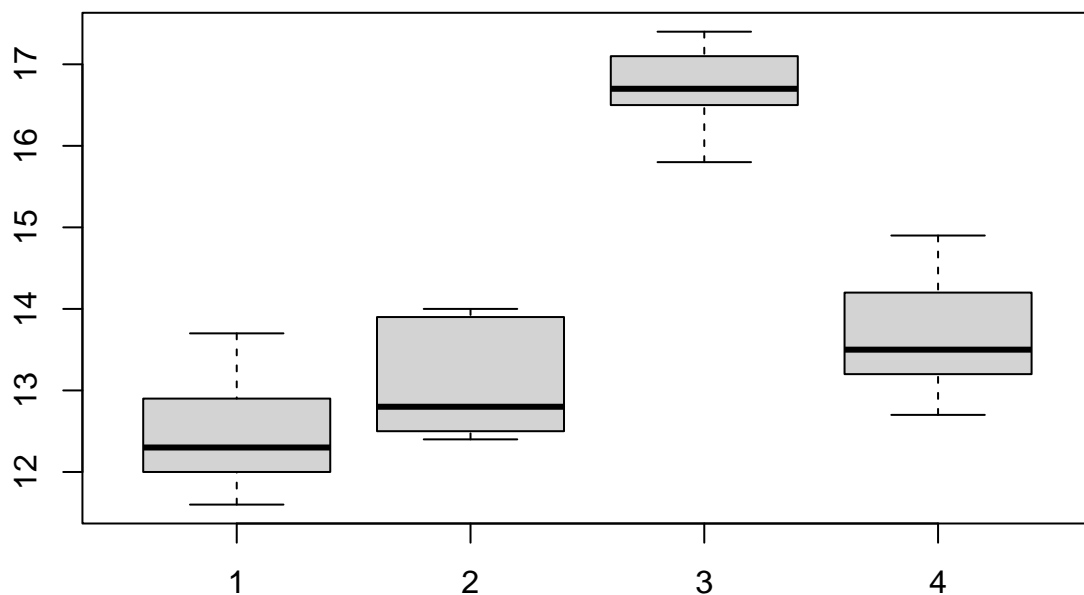
- g. A researcher could instead choose to build Bonferroni CIs to bound the family-wise error rate. What would the Bonferroni multiplier be in this case? How would Bonferroni CIs compare to the Tukey CIs? (you do not need to actually make the CIs).

## Solution

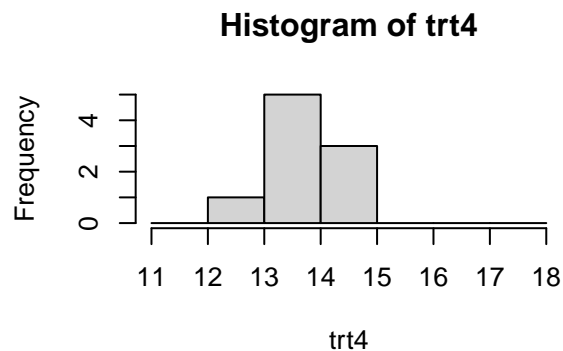
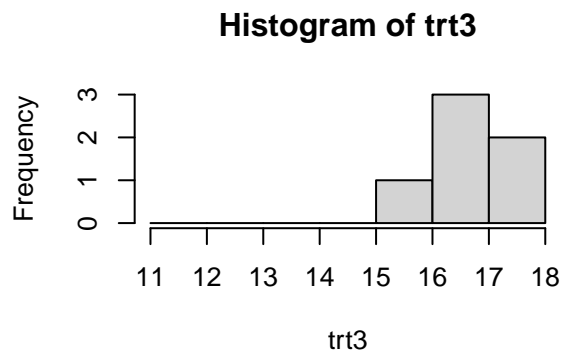
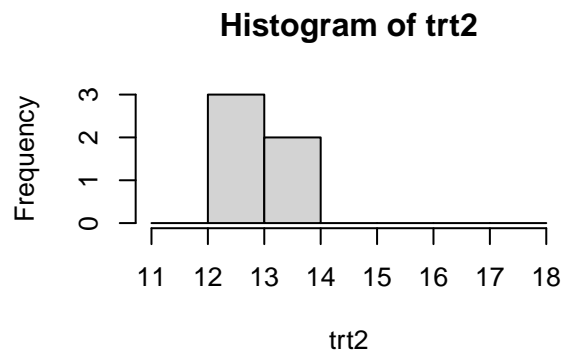
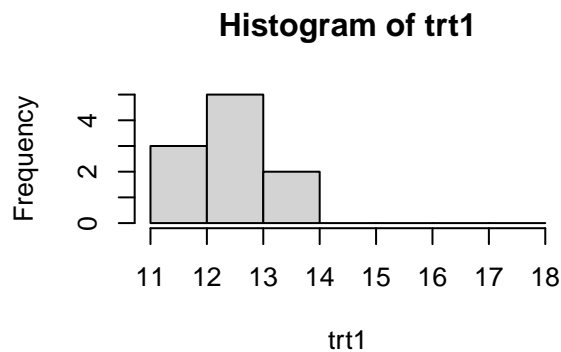
(a)

Let's check the boxplots and histograms.

```
# Data
trt1 <- c(13.3, 13.7, 11.6, 11.9, 12, 12.9, 12.3, 12.1, 12.3, 12.3)
trt2 <- c(12.5, 13.9, 12.4, 14, 12.8)
trt3 <- c(16.8, 16.6, 17.1, 15.8, 17.4, 16.5)
trt4 <- c(13.2, 14.2, 13.5, 13.2, 14.3, 12.7, 14.9, 13.3, 13.5)
# Create vectors of all observations and treatments
q1_vals <- c(trt1, trt2, trt3, trt4)
q1_trt <- c(rep("Trt1", 10), rep("Trt2", 5), rep("Trt3", 6), rep("Trt4", 9))
# Boxplot
boxplot(trt1, trt2, trt3, trt4, names = c("1", "2", "3", "4"))
```



```
# Histograms
par(mfrow = c(2, 2))
hist(trt1, breaks = seq(11, 18, 1))
hist(trt2, breaks = seq(11, 18, 1))
hist(trt3, breaks = seq(11, 18, 1))
hist(trt4, breaks = seq(11, 18, 1))
```



```
par(mfrow = c(1, 1))
```

The 4 group boxplots and histograms shows that there looks to be a difference in means between the groups. The groups also have similar variability.

The normality of residuals assumption will be evaluated after the ANOVA model is fit

(b)

Summarize the data numerically (your values should match the table below). What properties from the numeric summaries are noteworthy for considering an ANOVA analysis?

Use R to compute the quantities:

```
ybar1 <- mean(trt1)
s1 <- sd(trt1)
n1 <- length(trt1)
ybar2 <- mean(trt2)
```

```

s2 <- sd(trt2)
n2 <- length(trt2)
ybar3 <- mean(trt3)
s3 <- sd(trt3)
n3 <- length(trt3)
ybar4 <- mean(trt4)
s4 <- sd(trt4)
n4 <- length(trt4)
ybar <- mean(q1_vals)
s <- sd(q1_vals)
N <- length(q1_vals)

```

Group	Sample Mean	Sample SD	Sample Size
1	12.44	0.6586181	10
2	13.12	0.7726578	5
3	16.700	0.5513620	6
4	13.644	0.6857680	9
Overall	13.767	1.692	30

(c)

Create an ANOVA table by hand, and check your work with the `aov` function in R.

**Trt**

$SS_{\text{trt}}$ :

$$\begin{aligned}
 SS_{\text{trt}} &= \sum_{j=1}^4 n_j * (\bar{y}_j - \bar{y})^2 \\
 &= 10 * (12.44 - 13.767)^2 + 5 * (13.12 - 13.767)^2 + 6 * (16.700 - 13.767)^2 + 9 * (13.644 - 13.767)^2 \\
 &= 71.453.
 \end{aligned}$$

- $n_j$ : sample size for each group.
- $\bar{y}_j$ : sample mean for each group.
- $\bar{y}$ : total sample mean.

As we have four groups:  $df_{\text{trt}} = 4 - 1 = 3$ .

$$MS_{\text{trt}} = \frac{SS_{\text{trt}}}{df_{\text{trt}}} = \frac{71.453}{3} = 23.818$$

**Error**

$SS_E$ :

- $S_j$ : sample SD of the  $j$ -th group.

$$\begin{aligned}
 SS_E &= \sum_{j=1}^4 (n_j - 1) * S_j^2 \\
 &= (10 - 1) * 0.6586181^2 + (5 - 1) * 0.7726578^2 + (6 - 1) * 0.5513620^2 + (9 - 1) * 0.6857680^2 \\
 &= 11.574.
 \end{aligned}$$

$$df_E = n - 4 = 30 - 4 = 26.$$

$$MS_E = \frac{SS_E}{df_E} = \frac{11.574}{26} = 0.445.$$

**Total**

$$SS_{\text{Total}} = (n - 1) * S^2 = (30 - 1) * 1.692^2 = 83.023.$$

With the results above, we can also compute it in this way:

$$SS_{\text{Total}} = SS_{\text{trt}} + SS_E = 71.453 + 11.574 = 83.027.$$

We obtained the same results (although the rounding causes slight difference).

$$df_{\text{Total}} = n - 1 = 29.$$

**F**

$$F = \frac{MS_{\text{trt}}}{MS_E} = \frac{23.818}{0.445} = 53.524.$$

P-value:

```
1 - pf(53.524, df1 = 3, df2 = 26)
```

```
## [1] 2.918732e-11
```

Source	SS	df	MS	F	pvalue
Treat	71.453	3	23.818	53.524	2.918732e-11
Error	11.574	26	0.445		
Tot	83.027	29			

Check with R:

```
q1_mod <- aov(q1_vals ~ q1_trt)
summary(q1_mod)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## q1_trt         3   71.45   23.817    53.5 2.93e-11 ***
## Residuals    26   11.57    0.445
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

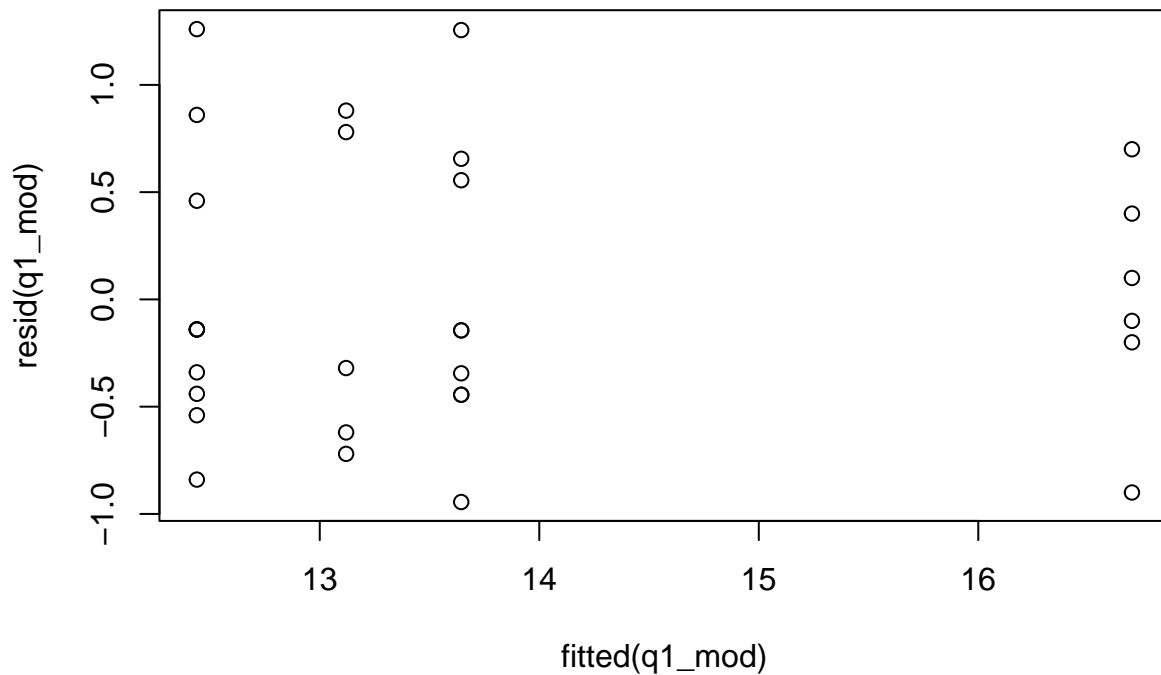
(d)

Decide whether we have sufficient evidence to reject the null after checking that the normality and equal variances assumptions are reasonably met.

**Check assumption:**

*Equal variances*

```
plot(fitted(q1_mod), resid(q1_mod))
```



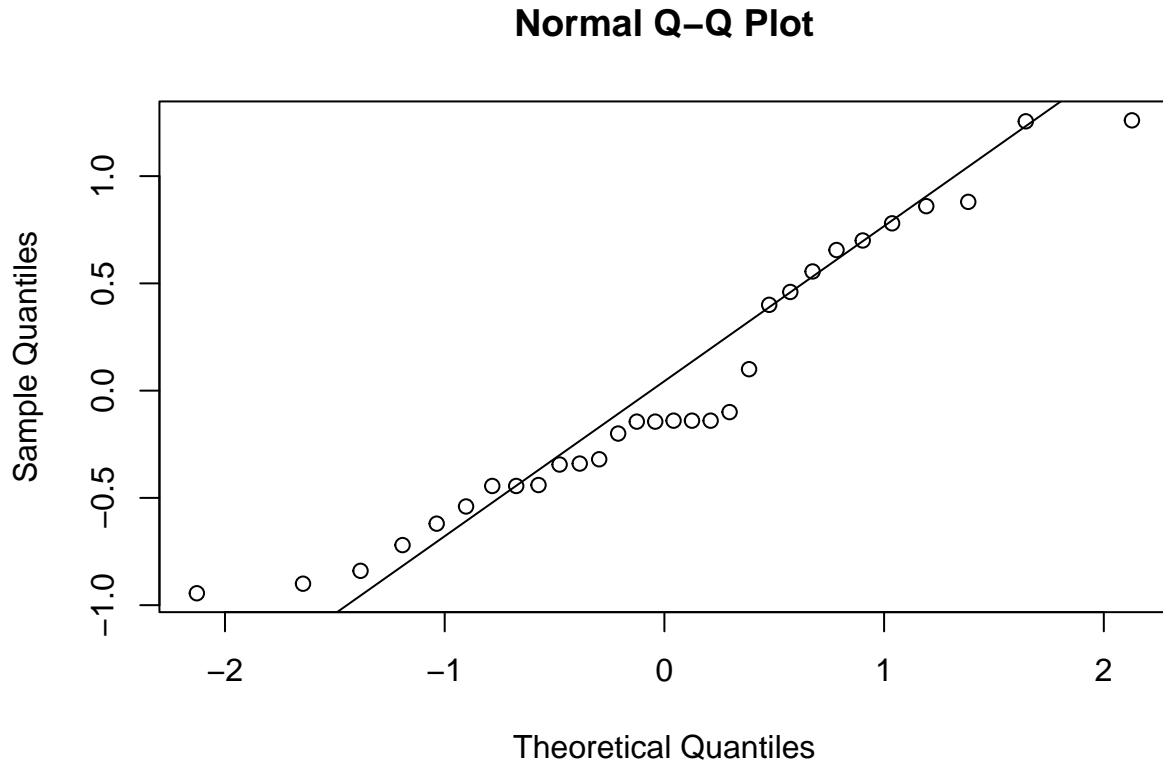
In the figure, the x-axis is the fitted value (the estimated population mean for each group). The y-axis is the residuals of each observation (the difference between the observed value and the fitted value.)

From our residual plots, we see roughly equal sample standard deviations between the 4 sample samples so equal variance in the populations seems reasonable.

#### *Normality*

We use QQ plot to check the normality of the residuals:

```
qqnorm(resid(q1_mod))  
qqline(resid(q1_mod))
```



We can see the points falls into a straight line roughly, which gives us no evidence that the populations of errors is nonnorm. So, the normality assumption is reasonable.

#### Conclusion:

From the results in (c), we have an extremely small p-value (p-value =  $2.918732e - 11 \approx 0$ ). This is very strong evidence against the null so we reject the null that all means are equal.

(e)

Conduct pairwise comparisons for all treatment pairs using 95% CI based on Fisher's method by hand. Make a table to summarize your findings of which groups are significantly different.

Summary Measures:

Group	Sample Mean	Sample SD	Sample Size
1	12.44	0.6586181	10
2	13.12	0.7726578	5
3	16.700	0.5513620	6
4	13.644	0.6857680	9

The confidence level is 0.95, thus  $\alpha = 1 - 0.95 = 0.05$ . The critical value:  $t_{1-\frac{\alpha}{2}, df_E} = t_{0.975, 26} = 2.056$ .

```
qt(1 - 0.05/2, 26)
```

```
## [1] 2.055529
```



$$\hat{\sigma}^2 = MS_{Error} = 0.4452.$$

CI for group  $i$  and group  $j$  ( $i \neq j$ ):

$$\bar{y}_i - \bar{y}_j \pm 2.056 * \sqrt{0.4452 * (\frac{1}{n_i} + \frac{1}{n_j})}.$$

CI for each pairs:

$$\begin{aligned}\mu_1 - \mu_2 : 12.44 - 13.12 \pm 2.056 * \sqrt{0.4452 * (\frac{1}{10} + \frac{1}{5})} &= (-1.431, 0.0714) \\ \mu_1 - \mu_3 : 12.44 - 16.7 \pm 2.056 * \sqrt{0.4452 * (\frac{1}{10} + \frac{1}{6})} &= (-4.968, -3.552) \\ \mu_1 - \mu_4 : 12.44 - 13.644 \pm 2.056 * \sqrt{0.4452 * (\frac{1}{10} + \frac{1}{8})} &= (-1.834, -0.574) \\ \mu_2 - \mu_3 : 13.12 - 16.7 \pm 2.056 * \sqrt{0.4452 * (\frac{1}{5} + \frac{1}{6})} &= (-4.411, -2.749) \\ \mu_2 - \mu_4 : 13.12 - 13.644 \pm 2.056 * \sqrt{0.4452 * (\frac{1}{5} + \frac{1}{8})} &= (-1.289, 0.241) \\ \mu_3 - \mu_4 : 16.7 - 13.644 \pm 2.056 * \sqrt{0.4452 * (\frac{1}{6} + \frac{1}{8})} &= (2.333, 3.779)\end{aligned}$$

Trt 1 has the smallest sample mean. Let Trt 1 to be group A. Then only the CI of  $\mu_1 - \mu_2$  contains 0. Thus, only Trt 2 belongs to group A, too. Similarly, we can define group B and C.

Treat	Mean	Group
Trt 3	16.70	C
Trt 4	13.644	B
Trt 2	13.12	AB
Trt 1	12.44	A

(f)

Explain why a researcher may instead choose to construct Tukey-Kramer adjusted CIs. How do these CIs differ from those computed in (e)? Look at the Tukey multiplier and 95% Tukey CIs below and resummairize the groupings found in (e) as necessary.

The Tukey-Kramer adjusted CIs will be wider than the Fisher CIs because they are doing an adjustment, as many pairwise comparisons (6) are being done in this same experiment and we want to control the overall (family-wide) error rate.

It uses a multiplier from the Q: Studentized Range Distribution(divided by sqrt(2)): 2.74332 will be the standard error multiplier (critical value) instead of 2.07. TukeyHSD code below gives Tukey CIs. We again see significant differences between groups 3 and 1, 4 and 1, 3 and 2, and 3 and 4.

```
# Find Tukey multiplier
qtukey(0.95, nmeans = 4, df = 26)/sqrt(2)
```

```
## [1] 2.74332
```

Tukey CIs:

$$\begin{aligned}\mu_1 - \mu_2 : 12.44 - 13.12 \pm 2.74332 * \sqrt{0.4452 * (\frac{1}{10} + \frac{1}{5})} &= (-1.683, 0.323) \\ \mu_1 - \mu_3 : 12.44 - 16.7 \pm 2.74332 * \sqrt{0.4452 * (\frac{1}{10} + \frac{1}{6})} &= (-5.205, -3.315) \\ \mu_1 - \mu_4 : 12.44 - 13.644 \pm 2.74332 * \sqrt{0.4452 * (\frac{1}{10} + \frac{1}{8})} &= (-2.045, -0.363) \\ \mu_2 - \mu_3 : 13.12 - 16.7 \pm 2.74332 * \sqrt{0.4452 * (\frac{1}{5} + \frac{1}{6})} &= (-4.688, -2.472) \\ \mu_2 - \mu_4 : 13.12 - 13.644 \pm 2.74332 * \sqrt{0.4452 * (\frac{1}{5} + \frac{1}{8})} &= (-1.545, 0.496) \\ \mu_3 - \mu_4 : 16.7 - 13.644 \pm 2.74332 * \sqrt{0.4452 * (\frac{1}{6} + \frac{1}{8})} &= (2.091, 4.020)\end{aligned}$$

Treat	Mean	Group
Trt 1	12.44	A
Trt 2	13.12	AB
Trt 4	13.644	B
Trt 3	16.70	C

(g)

A researcher could instead choose to build Bonferroni CIs to bound the family-wise error rate. What would the Bonferroni multiplier be in this case? How would Bonferroni CIs compare to the Tukey CIs? (you do not need to actually make the CIs).

```
# Find Bonferroni multiplier
qt(1 - 0.05/2/6, df = 26) # as we have 6 pairs for comparison.
```

```
## [1] 2.855521
```

Since our Bonferroni multiplier is 2.856 which is greater than our Tukey multiplier of 2.743, the intervals would all be slightly wider (more conservative). This suggests that the Tukey method is more powerful in this case, and would be more likely to detect a difference in means.