

# Homework 6 - Stats 230 One-way ANOVA and multiple comparisons: Knees and Sea Slugs

Dhyey Mavani

date

**PROBLEMS TO TURN IN: #5.6, #5.8, #5.16, #5.18, #5.38, #5.46, #Additional 1**

**Exercise 5.6 - Multiple Choice** SOLUTION: B

**Exercise 5.8 - Multiple Choice** SOLUTION: C

**Exercise 5.16**

5.16 part a:

SOLUTION: The explanatory variable is the dog food type and the response variable in this case is the average amount of time each dog sleeps per 24 hours for next week.

5.16 part b:

SOLUTION: This is a randomized experiment because we know that in this setting we randomly assign each dog to one of the three diets.

5.16 part c:

SOLUTION: Although dogs were not randomly selected from a population, the diet treatments were assigned to dogs in a randomized manner, so it is still appropriate to conduct an analysis of variance in this given situation.

**Exercise 5.18 (continues 5.16)**

5.18 part a: Experimental Units

SOLUTION: The Border Collies are the units in this experiment.

5.18 part b: Treatments

SOLUTION: The treatments are the type of food they are provided with which are salmon, chicken, and beef diets.

5.18 part c: Balanced Experiment

SOLUTION: They have to make sure that each treatment has equal number of units assigned to it. This makes the experiment balanced

```
data(Meniscus)
#Method is recorded numerically, we need it to be a factor
# FYI - The 3 methods are vertical suture, meniscus arrow, FastT-Fix
Meniscus <- mutate(Meniscus, Method = factor(Method)) #overwrites the variable
```

### Exercise 5.38

5.38 part a: Hypotheses

SOLUTION:

$H_0 : \alpha_{verticalsuture} = \alpha_{meniscusarrow} = \alpha_{FastT-Fix}.$

$H_A : \alpha_{verticalsuture} \neq \alpha_{meniscusarrow} \neq \alpha_{FastT-Fix}.$

OR in words ...

$H_0$ : There is no difference in mean stiffness among the three treatment groups namely vertical suture, meniscus arrow and FastT-Fix.

$H_A$ : There is a difference between at least one of the means of stiffness of the three groups namely vertical suture, meniscus arrow and FastT-Fix.

5.38 part b: Conditions

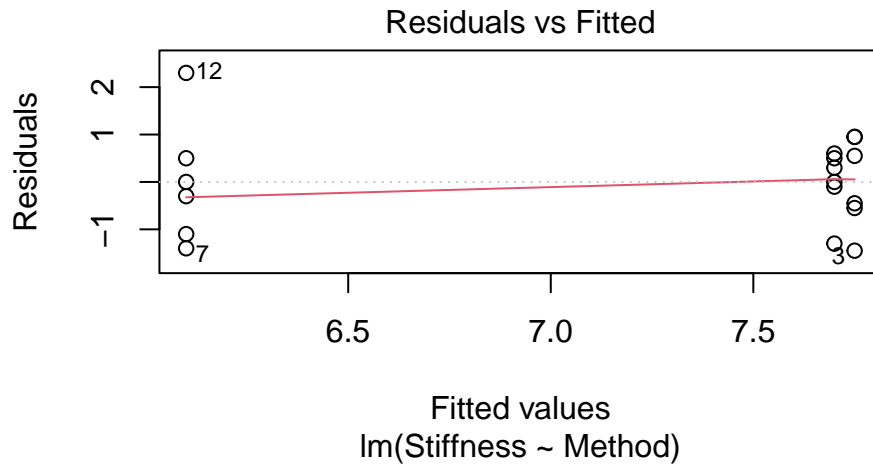
SOLUTION:

The majority of the points of the qq plot fall along the line, so we have normality of errors condition met in this case. We can say that based on the residuals vs. fitted plot, there is mostly equal variance among the three treatment groups. Since each specimen unit was randomly assigned to one of the three treatments, randomization is satisfied. We can also see that the errors are centered at 0. So we can essentially say that all conditions are satisfied.

```
favstats(Stiffness ~ Method, data = Meniscus)
```

##	Method	min	Q1	median	Q3	max	mean	sd	n	missing
## 1	1	6.3	7.225	7.80	8.600	8.7	7.75	0.971082	6	0
## 2	2	4.7	5.200	5.95	6.475	8.4	6.10	1.326650	6	0
## 3	3	6.4	7.625	7.85	8.150	8.3	7.70	0.692820	6	0

```
modNew <- lm(Stiffness ~ Method, data = Meniscus)
plot(modNew, which = 1)
```



5.38 part c: ANOVA

SOLUTION: Based on the t-test output we can see below, the p-value of 0.0219 ( $< 0.05$  which is significant), we can say that the data provides us with strong evidence that the mean value of stiffness significantly differs based on the type of meniscus repair between at least two treatment groups on average.

```
anova(modNew)
```

```
## Analysis of Variance Table
##
## Response: Stiffness
##          Df Sum Sq Mean Sq F value Pr(>F)
## Method    2  10.57   5.285   4.981 0.0219 *
## Residuals 15   15.91   1.061
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Exercise 5.46 (continues 5.38) SOLUTION:**

There are differences in stiffness between the treatments meniscus arrow - vertical suture, and FasT-Fix - meniscus arrow according to the p-values of Fisher's LSD.

We can see that since the confidence interval of meniscus arrow - vertical suture contains only negative values while the confidence interval of FasT-Fix - meniscus arrow contains only positive values, this implies that vertical suture results in less stiffness than meniscus arrow.

Hence, we can say Fast-Fix is better treatment than meniscus arrow.

On the other hand, there is no significant difference between FasT-Fix and vertical suture as we can see in the output below.

Thence, doctors should use FasT-Fix and vertical suture over meniscus arrow.

```
meniscus_aov <- aov(Stiffness ~ Method, data = Meniscus)
summary(meniscus_aov)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Method      2   10.6    5.29    4.98  0.022 *
## Residuals   15   15.9    1.06
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
PostHocTest(meniscus_aov, method = "lsd")
```

```
##
##   Posthoc multiple comparisons of means : Fisher LSD
##   95% family-wise confidence level
##
## $Method
##      diff   lwr.ci   upr.ci   pval
## 2-1 -1.65 -2.91757 -0.38243 0.0142 *
## 3-1 -0.05 -1.31757  1.21757 0.9341
## 3-2  1.60  0.33243  2.86757 0.0168 *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Additional 1 - More MCs - Based on 5.71 and 8.14** Read problem 5.71 for context. We are going to examine all three MC methods here and compare them. We will assume conditions for inference are met. If you read 8.14 as well, note there is a typo there - there are 6 time periods here.

```
data(SeaSlugs)
#Time is recorded numerically in the data set. We need it as a factor.
SeaSlugs <- mutate(SeaSlugs, TimeF = factor(Time))
#Use TimeF for your analysis as the categorical predictor
```

ADD1 part a: Fit the requested ANOVA from 5.71, and verify that the overall F is significant.

SOLUTION: Based on the t-test shown below, we can see that p-value of 0.000607 ( $< 0.05$  which is significant). This implies that we have strong evidence that the mean percentage of larvae that metamorphosed differs over the six time periods.

```
favstats(Percent ~ TimeF, data = SeaSlugs)
```

```
##   TimeF   min    Q1 median    Q3   max    mean    sd n missing
## 1     0 0.357 0.47525 0.5000 0.52475 0.857 0.535667 0.168786 6      0
## 2     5 0.125 0.26050 0.4000 0.45025 0.467 0.345500 0.142392 6      0
## 3    10 0.067 0.08350 0.1330 0.28300 0.333 0.177667 0.123888 6      0
## 4    15 0.000 0.05350 0.2405 0.28125 0.333 0.183333 0.147040 6      0
## 5    20 0.067 0.10775 0.2335 0.26700 0.437 0.219167 0.138391 6      0
## 6    25 0.000 0.08350 0.1330 0.23350 0.412 0.168667 0.148465 6      0
```

```
modMy <- lm(Percent ~ TimeF, data = SeaSlugs)
anova(modMy)
```

```
## Analysis of Variance Table
##
## Response: Percent
##           Df Sum Sq Mean Sq F value    Pr(>F)
## TimeF      5 0.6309 0.12618    5.965 0.000607 ***
## Residuals 30 0.6346 0.02115
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ADD1 part b: Perform Fisher's LSD and identify significant differences in mean percentage metamorphosed.

SOLUTION: Let's use the Fisher's LSD as done below. We can see that there are significant differences in mean percentage metamorphosed between time intervals 5 and 0, 10 and 0, 15 and 0, 20 and 0, 25 and 0, and 25 and 5.

```
seaslugs_aov <- aov(Percent ~ TimeF, data = SeaSlugs)
PostHocTest(seaslugs_aov, method = "lsd")
```

```
##
##   Posthoc multiple comparisons of means : Fisher LSD
##   95% family-wise confidence level
##
## $TimeF
##           diff      lwr.ci      upr.ci    pval
## 5-0   -0.19016667 -0.361663 -0.01867040 0.03092 *
## 10-0  -0.35800000 -0.529496 -0.18650374 0.00018 ***
## 15-0  -0.35233333 -0.523830 -0.18083707 0.00022 ***
## 20-0  -0.31650000 -0.487996 -0.14500374 0.00072 ***
## 25-0  -0.36700000 -0.538496 -0.19550374 0.00014 ***
## 10-5  -0.16783333 -0.339330  0.00366293 0.05478 .
## 15-5  -0.16216667 -0.333663  0.00932960 0.06296 .
## 20-5  -0.12633333 -0.297830  0.04516293 0.14292
## 25-5  -0.17683333 -0.348330 -0.00533707 0.04370 *
## 15-10  0.00566667 -0.165830  0.17716293 0.94665
## 20-10  0.04150000 -0.129996  0.21299626 0.62476
## 25-10 -0.00900000 -0.180496  0.16249626 0.91536
## 20-15  0.03583333 -0.135663  0.20732960 0.67263
## 25-15 -0.01466667 -0.186163  0.15682960 0.86252
## 25-20 -0.05050000 -0.221996  0.12099626 0.55210
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ADD1 part c: Perform Tukey's HSD and identify significant differences in mean percentage metamorphosed.

SOLUTION: Let's use the Tukey's HSD as done below. We can see that there are significant differences in mean percentage metamorphosed between time intervals 10 and 0, 15 and 0, 20 and 0, and 25 and 0.

```
PostHocTest(seaslugs_aov, method = "hsd")
```

```
##
## Posthoc multiple comparisons of means : Tukey HSD
## 95% family-wise confidence level
##
## $TimeF
##      diff      lwr.ci      upr.ci    pval
## 5-0    -0.19016667 -0.445579  0.0652459 0.2397
## 10-0   -0.35800000 -0.613413 -0.1025874 0.0023 **
## 15-0   -0.35233333 -0.607746 -0.0969208 0.0028 **
## 20-0   -0.31650000 -0.571913 -0.0610874 0.0085 **
## 25-0   -0.36700000 -0.622413 -0.1115874 0.0017 **
## 10-5   -0.16783333 -0.423246  0.0875792 0.3666
## 15-5   -0.16216667 -0.417579  0.0932459 0.4039
## 20-5   -0.12633333 -0.381746  0.1290792 0.6641
## 25-5   -0.17683333 -0.432246  0.0785792 0.3114
## 15-10  0.00566667 -0.249746  0.2610792 1.0000
## 20-10  0.04150000 -0.213913  0.2969126 0.9960
## 25-10 -0.00900000 -0.264413  0.2464126 1.0000
## 20-15  0.03583333 -0.219579  0.2912459 0.9980
## 25-15 -0.01466667 -0.270079  0.2407459 1.0000
## 25-20 -0.05050000 -0.305913  0.2049126 0.9901
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ADD1 part d: Perform Bonferroni's adjustment for the multiple comparisons and identify significant differences in mean percentage metamorphosed.

SOLUTION: Let's use Bonferroni's adjustment for the multiple comparisons as done below. We can see that there are significant differences in mean percentage metamorphosed between time intervals 10 and 0, 15 and 0, 20 and 0, and 25 and 0.

```
PostHocTest(seaslugs_aov, method = "bonferroni")
```

```
##
## Posthoc multiple comparisons of means : Bonferroni
## 95% family-wise confidence level
##
## $TimeF
##      diff      lwr.ci      upr.ci    pval
## 5-0    -0.19016667 -0.457941  0.0776077 0.4638
## 10-0   -0.35800000 -0.625774 -0.0902256 0.0028 **
## 15-0   -0.35233333 -0.620108 -0.0845589 0.0033 **
## 20-0   -0.31650000 -0.584274 -0.0487256 0.0108 *
## 25-0   -0.36700000 -0.634774 -0.0992256 0.0021 **
## 10-5   -0.16783333 -0.435608  0.0999411 0.8217
## 15-5   -0.16216667 -0.429941  0.1056077 0.9444
## 20-5   -0.12633333 -0.394108  0.1414411 1.0000
## 25-5   -0.17683333 -0.444608  0.0909411 0.6554
## 15-10  0.00566667 -0.262108  0.2734411 1.0000
## 20-10  0.04150000 -0.226274  0.3092744 1.0000
## 25-10 -0.00900000 -0.276774  0.2587744 1.0000
## 20-15  0.03583333 -0.231941  0.3036077 1.0000
```

```
## 25-15 -0.01466667 -0.282441 0.2531077 1.0000
## 25-20 -0.05050000 -0.318274 0.2172744 1.0000
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ADD1 part e: Compare the significant differences found by the three methods. (E.g. do they agree? etc.)

SOLUTION: We can see that Tukey's HSD and Bonferroni's adjustment found the same pairs of time intervals that had differences in mean percentage metamorphosed although they have slightly different significance levels. On the other hand, we can see that Fisher's LSD found additional differences between time intervals 5 and 0, and 25 and 5. This honestly makes sense to me because it is a property of Fisher's LSD that it finds the most potential differences while some of them might not be true ones.