Statistical Inference Project Part 2

Bryan Willauer

4/25/2020

Data Summary

The data comes from an experiment on the effect of vitamin C on tooth growth in Guinea Pigs. The data is formatted as a data frame with 60 observations of three variables.

	Name	Type	Description
1.	len	numeric	tooth length
2.	supp	factor	supplement type (vitamin C VC or orange Juice OJ)
3.	dose	numeric	dose in mg/day

```
library(tidyverse)
data("ToothGrowth")
str(ToothGrowth)
```

```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
head(ToothGrowth)
```

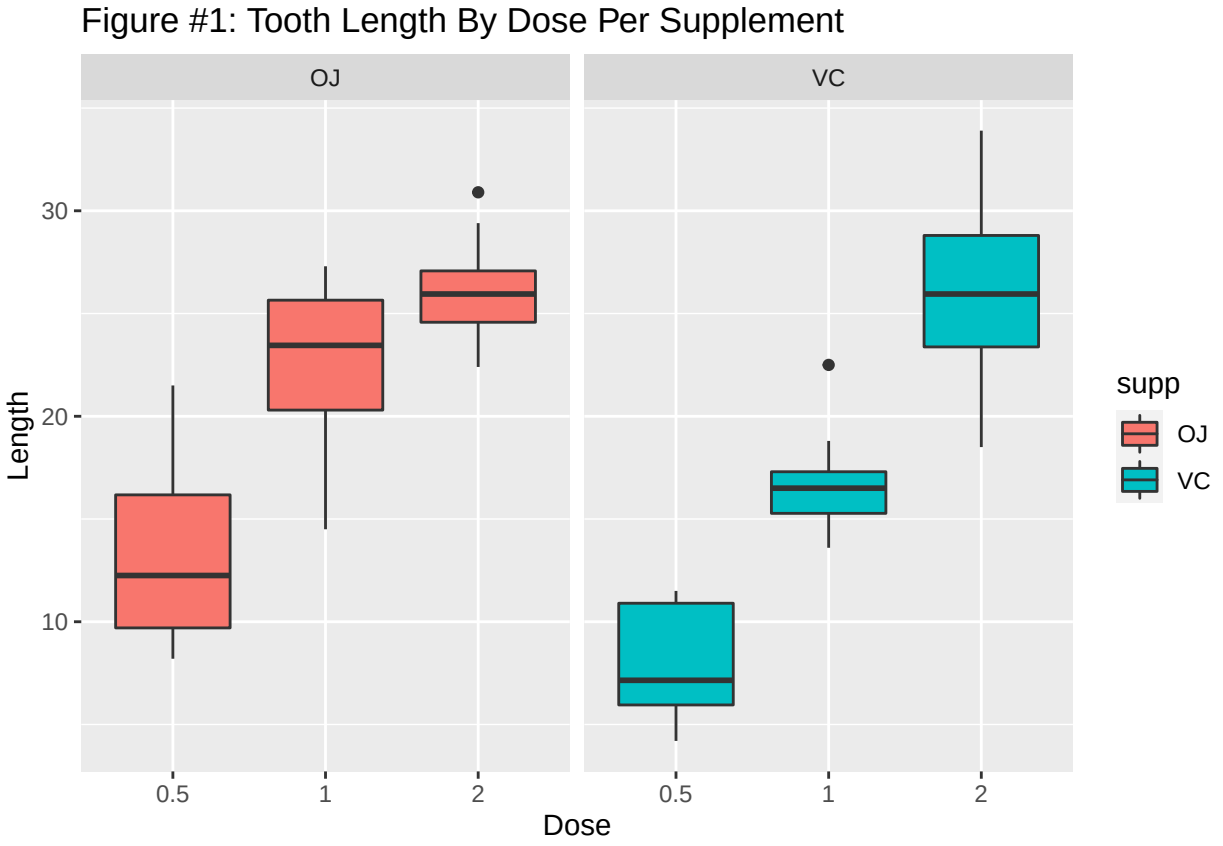
```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
summary(ToothGrowth)
```

```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

Exploratory Data Analysis

```
ToothGrowth %>% ggplot(aes(as.factor(dose),len, fill = supp)) +
        facet_grid(.~ supp) +
        geom_boxplot() +
        labs(x = "Dose", y = "Length") +
        ggtitle("Figure #1: Tooth Length By Dose Per Supplement")
```



Preliminary analysis seems to show that tooth length increases with increasing dosages of both supplements. However, lower dosages of Orange Juice look to do better than the lower dosages of Vitamin C.

Hypothesis Testing

The following t-test tests the null hypothesis that there is a difference in length due to choice of supplement.

```
t.test(ToothGrowth$len ~ ToothGrowth$supp)
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len by ToothGrowth$supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

The following t-tests test whether there is a difference between dosages of each supplement. The data was split into three groups, each group contains the individual dosages so we can compare how each supplement at each dosage compares. The first test tests the 0.5mg/day dosage.

```
dose0.5 <- subset(ToothGrowth, ToothGrowth$dose==c(0.5))
dose2 <- subset(ToothGrowth, ToothGrowth$dose==c(2))
dose1   <- subset(ToothGrowth, ToothGrowth$dose==c(1))
t.test(dose0.5$len ~ dose0.5$supp, paired = FALSE, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose0.5$len by dose0.5$supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

The second test tests the 1mg/day dosage.

```
t.test(dose1$len ~ dose1$supp, paired = FALSE, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose1$len by dose1$supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

the third test tests the 2mg/day dosage.

```
t.test(dose2$len ~ dose2$supp, paired = FALSE, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose2$len by dose2$supp
```

```
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

Conclusion

The first t-test, where we tested the null hypothesis of no difference between using either supplement resulted in not being able to reject the null hypothesis. The p-value was above .05 and the 95% confidence interval contained zero.

The tests related to the dosage of the supplement resulted in being able to reject the null hypothesis of no difference in supplements at the same dosage for the two lower dosages. Both tests for the two lower dosages had low p-values, high t values, and 95% confidence intervals not containg zero. Thus the smaller dosages showed significant tooth length difference when orange juice was supplemented rather than vitamin C. We could not reject the null hypothesis at the higher dosage. The p-value was high and the 95% confidence interval contained zero. So there was no significant difference between the supplements at 2mg/day.