Part 2 of the Statistical Inference Project

Andy Tse

August 7, 2016

Overview

In the second part of the project, we will start to analyze the tooth growth data in the R datasets package. At the start of part 2 of the Statistical Inference Assessment, we will load the tooth growth data and analyze the summary there. There is going to be an analysis of different samples of the dose and supp sizes on the tooth growth.

Load the Datasets Data on Tooth Growth

```
data(ToothGrowth)
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
```

```
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 ...
```

```
dim(ToothGrowth)
```

```
[1] 60  3
```

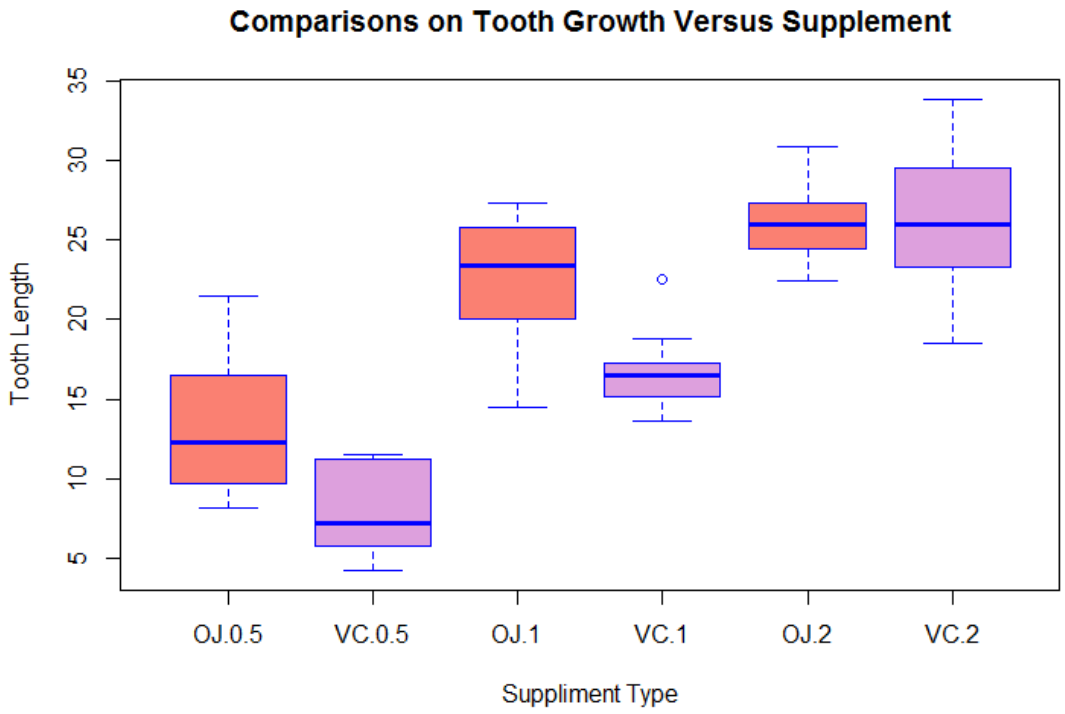
```
summary(ToothGrowth)
```

```
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
```

Summary of Data

```
library(ggplot2)
boxplot(len~supp*dose, data=ToothGrowth,
```

```
        col=(c("salmon","plum")), border="blue",
        main="Comparisons on Tooth Growth Versus Supplement",
xlab="Suppliment Type", ylab="Tooth Length")
```



Confidence Intervals to Test the Data

For this section, we will start to test the hypotheses on the tooth growth analysis. In addition, there is going to also be a comparison between the variences also. We are going to compare all the dosages from 0.5mg, 1.0mg, and 2.0mg for all levels.

```
library(ToothGrowth); data(ToothGrowth)
t.test(len~supp, paired=FALSE, var.equal=FALSE, data=ToothGrowth)


    Welch Two Sample t-test

data:  len by 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
```

T Test Growth in 0.5 Interval

```
t.test(len~supp, ToothGrowth[ToothGrowth$dose == 0.5,])

data:  len by 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
```

T Test Growth in 1.0 Interval

```
t.test(len~supp, ToothGrowth[ToothGrowth$dose == 1.0,])

    Welch Two Sample t-test

data:  len by 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
```

T Test Growth in 2.0 Interval

```
t.test(len~supp, ToothGrowth[ToothGrowth$dose == 2.0,])

    Welch Two Sample t-test

data:  len by 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

Based on all three dosages (0.5, 1.0, and 2.0mg) on the tooth growth, it has determined that for the degrees of freedom, they are almost close to equivalent to one to another. In the graph for 2.0mg dosage, it has shown that there is insuffient evidence to conclude that orange juice promote tooth growth. However, for the 0.5mg and 1.0mg dosages, there is a significant difference with the p-values. It also show that there is a higher growth with the orange juice for tooth growth between the 0.5mg and 1.0mg respectively.