Assignment-Part2

Stephanie Lum

January 31, 2016

Overview

This report conducts analyses on the ToothGrowth data in the R datasets package.

This report will show:
1. A basic summary of the data
2. Basic exploratory data analyses
3. The use of confidence intervals and/or hypothesis tests to compare tooth growth by supplement and dosage
4. Assumptions and conclusions drawn from the tests and the data

```
# Load packages
library(datasets)
library(ggplot2)

# Load data
data("ToothGrowth")
```

Basic Summary of the Data

```
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 ...
```

This data frame has the measurements of 60 guinea pig subjects (rows) on 3 variables (columns).
The variables are described as follow (from the R Documentation file):
1. len = Tooth length, measured in length of odontoblasts
2. supp = Supplement type, ascorbic acid (VC) or orange juice (OJ)
3. dose = Dose in milligrams/day

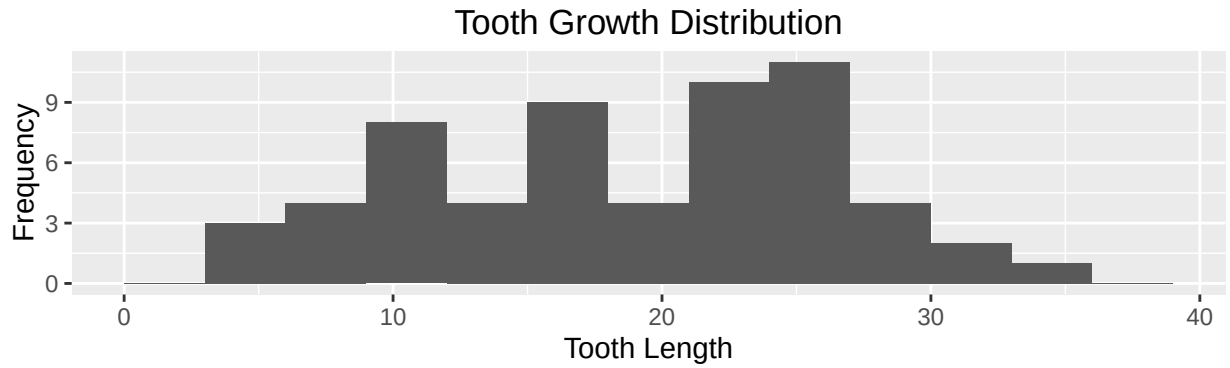
```
head(ToothGrowth)
```

```
##    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
```

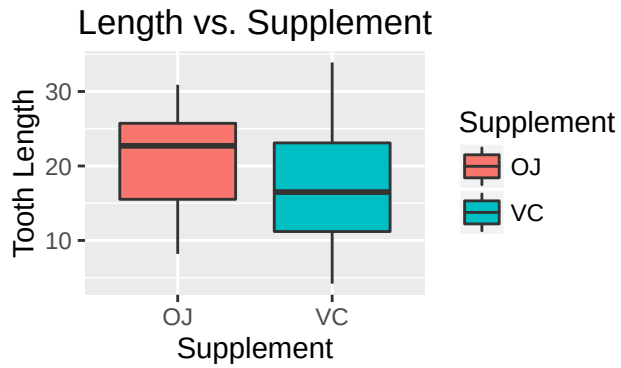
```
summary(ToothGrowth)
```

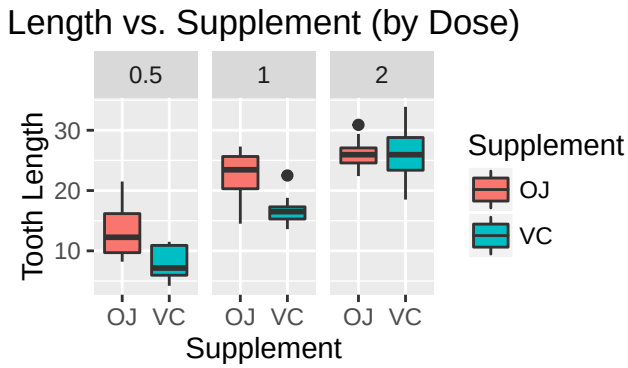
```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

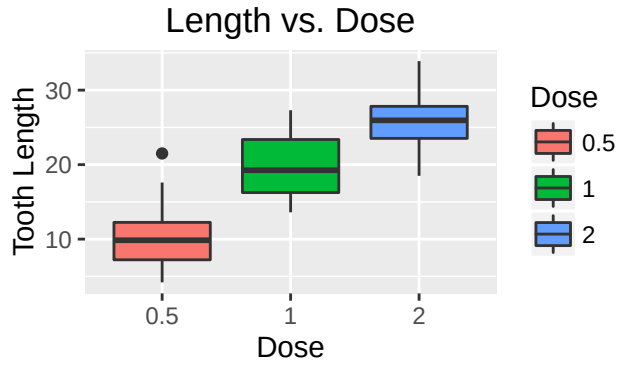
This histogram shows a distribution of the tooth lengths between all the subjects.

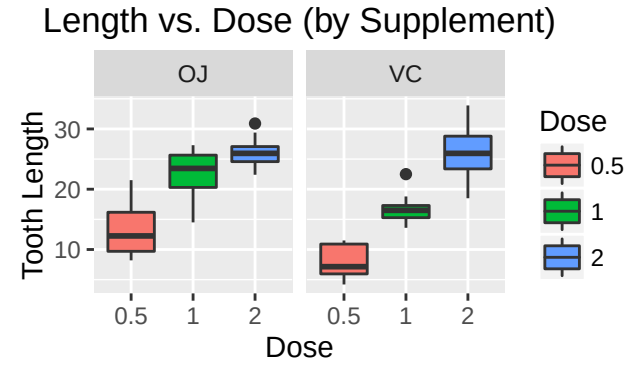


Basic Exploratory Analysis









Based on the box plots, it can be initially inferred that a higher dosage of Vitamin C will promote more tooth growth. The supplement types do not seem to affect the tooth length as much as the dosage does. The calculations of confidence intervals and hypothesis testing will confirm or reject this inference.

Confidence Intervals and Hypothesis Testing

The null hypothesis for this experiment is that the change in dosage and/or supplement will have no effect on the tooth length.

The alternative hypothesis for this experiement is that a change of supplement and/or a change in dosage will have an effect on the tooth length.

Comparing Supplement Types

The first test will be to conduct a T-Test between the tooth length and the supplement type. It is assumed there are unequal variances between the two groups. The output for the T-Test can be found in the Appendix.

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

The reported p-value is 0.06 and the confidence interval contains 0. Therefore, the null hypothesis that supplement type will affect tooth length should not be rejected. However, this test compared the supplement across all dosages. Further investigation between the dosages of different supplements and tooth length will need to be conducted. The output for the T-Tests can be found in the Appendix.

```
# Subset the data into different dosages
ToothGrowth_0.5 <- subset(ToothGrowth, dose == 0.5)
ToothGrowth_1.0 <- subset(ToothGrowth, dose == 1.0)
ToothGrowth_2.0 <- subset(ToothGrowth, dose == 2.0)

# T-Test to Compare tooth length and supplement type at a dose of 0.5 mg/day
t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = ToothGrowth_0.5)

# T-Test to compare tooth length and supplement type at a dose of 1.0 mg/day
t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = ToothGrowth_1.0)

# T-Test to compare tooth length and supplement type at a dose of 2.0 mg/day
t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = ToothGrowth_2.0)
```

The p-values for the first two T-Tests are well below 0.05, which means the null hypothesis that supplement type does not affect tooth length can be rejected at doses of 0.5 and 1.0 milligrams/day. From the means, it is evident that OJ is the better supplement at those doses. However, at 2.0 mg/day, both supplements are almost equally potent at promoting tooth growth.

Comparing Dosage Levels

The next three tests will be to conduct T-Tests between the tooth length and different dosages of the supplements. Once again, it is assumed there are unequal variances between the groups. The output for the T-Tests can be found in the Appendix.

```r
# Subset the dosages into paired groupings based on supplement
ToothGrowth_0.5_1.0 <- subset(ToothGrowth, dose == c(0.5, 1.0))
ToothGrowth_1.0_2.0 <- subset(ToothGrowth, dose == c(1.0, 2.0))
ToothGrowth_0.5_2.0 <- subset(ToothGrowth, dose == c(0.5, 2.0))

# T-Test to compare a dose 0.5 mg/day to 1.0 mg/day
t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = ToothGrowth_0.5_1.0)

# T-Test to compare a dose of 1.0 mg/day to 2.0 mg/day
t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = ToothGrowth_1.0_2.0)

# T-Test to compare a dose of 0.5 mg/day to 2.0 mg/day
t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = ToothGrowth_0.5_2.0)
```

The p-values for all three tests are well below 0.05, indicating that the null hypothesis that the dose level does not affect the tooth length can be rejected.

Conclusion

From the graphs and the tests conducted, it is clear that there is a correlation between the dose of Vitamin C given to the guinea pigs and their tooth length. What is less clear is the effect of the supplement type on the tooth length. However, there is indication that at the two lower dosages of 0.5 mg/day and 1.0 mg/day, orange juice (OJ) is more effective than ascorbic acid (VC) at promoting tooth growth.

Several assumptions were made during the course of this report:
1. The guinea pig subjects in this experiment represent the whole population of guinea pigs and that this is an unbiased sample with random assignment of dosages and supplement types.
2. The T-Test was the correct model to use for analysis of the p-value and confidence intervals.
3. The variances for the T-Tests were unequal.

Appendix

T-Tests

```r
# T-Test to compare tooth length to supplement type
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
```

```r
# T-Test to Compare tooth length and supplement type at a dose of 0.5 mg/day
t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = ToothGrowth_0.5)
```

```
##
##  Welch Two Sample t-test
##
## 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
```

```r
# T-Test to compare tooth length and supplement type at a dose of 1.0 mg/day
t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = ToothGrowth_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
```

```r
# T-Test to compare tooth length and supplement type at a dose of 2.0 mg/day
t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = ToothGrowth_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
```

```r
# T-Test to compare a dose 0.5 mg/day to 1.0 mg/day
t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = ToothGrowth_0.5_1.0)
```

```
##
##  Welch Two Sample t-test
##
```

```
## data:  len by dose
## t = -4.4725, df = 17.976, p-value = 0.0002952
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -14.43327  -5.20673
## sample estimates:
## mean in group 0.5   mean in group 1
##             10.63             20.45
```

```
# T-Test to compare a dose of 1.0 mg/day to 2.0 mg/day
t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = ToothGrowth_1.0_2.0)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -3.6827, df = 17.949, p-value = 0.00171
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -10.899993  -2.980007
## sample estimates:
## mean in group 1 mean in group 2
##           19.02           25.96
```

```
# T-Test to compare a dose of 0.5 mg/day to 2.0 mg/day
t.test(len ~ dose, paired = FALSE, var.equal = FALSE, data = ToothGrowth_0.5_2.0)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -7.3335, df = 17.635, p-value = 9.362e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -19.72833 -10.93167
## sample estimates:
## mean in group 0.5   mean in group 2
##             10.63             25.96
```