Tooth Growth, Exploratory Data Analysis and Comparison

Sebastian Cruz

December 25, 2015

Synopsis

To compare the impact of two supplements of Vitamin C at three dossages (0.5mg, 1.0mg, 2.0mg) on tooth growth in guinea pigs. The length of teeth in six groups of ten guinea pigs were measured in the experiment. Each guinea pig received one of the three doses by one of the supplements. Supplement type and dosage do impact guinea pig tooth growth, with Orange Juice resulting in a higher mean tooth growth at 0.5mg and 1.0mg dosages. No statistically significant difference was found at the 2.0mg dose between supplement types. Overall and regardless of dose, Orange Juice and Ascorbic Acid resulted in about the same mean tooth growth.

Data Summary and Exploration

The ToothGrowth dataset contains 60 observations and 3 variables (Tooth Length, Supplement Type, and Dosage). The ToothGrowth dataset is converted to a data table and relabled for clarity. Dosage and Supplement are made into factors as we will be comparing these groups.

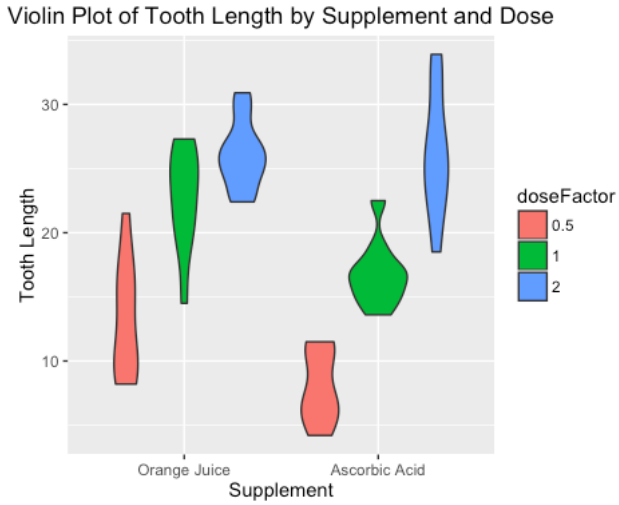
Summary

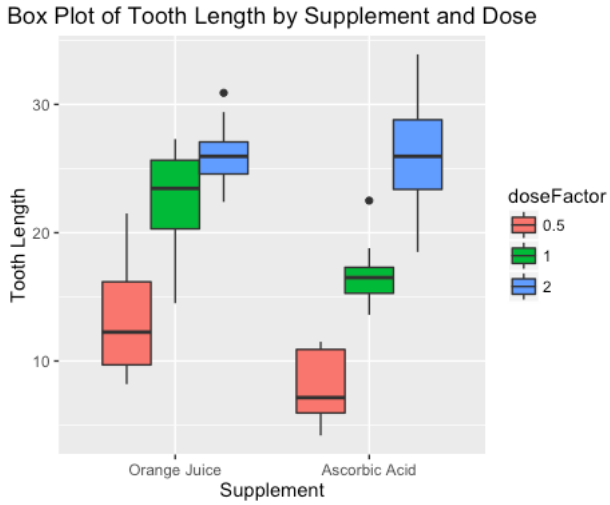
Summary metrics by Supplement and Dose

```
##       supplement dose  n  Min     Q1 Median  Mean     Q3  Max   StdDev
## 1: Ascorbic Acid  0.5 10  4.2  5.950   7.15  7.98 10.900 11.5 2.746634
## 2: Ascorbic Acid  1.0 10 13.6 15.275  16.50 16.77 17.300 22.5 2.515309
## 3: Ascorbic Acid  2.0 10 18.5 23.375  25.95 26.14 28.800 33.9 4.797731
## 4:  Orange Juice  0.5 10  8.2  9.700  12.25 13.23 16.175 21.5 4.459709
## 5:  Orange Juice  1.0 10 14.5 20.300  23.45 22.70 25.650 27.3 3.910953
## 6:  Orange Juice  2.0 10 22.4 24.575  25.95 26.06 27.075 30.9 2.655058
```

Exploratory Analysis

The violin plots show that the distribution of teeth length tends to be shifted higher and tend to be longer than the equivalent dose with Ascorbic Acid with the exception of the 2.0mg dose.





The box plots show that the distribution of teeth length values tend to have wider distributions that are larger and with higher means than the equivalent dose with Ascorbic Acid with the exception of the 2.0mg dose.

Confidence Intervals and Hypothesis Testing

The following will test our hypothesis to determine the significance of any differences between the mean tooth length by supplement and dose. The confidence interval will be set at 95% and hypothesis testing will utilize $\alpha = 0.05$.

Assumptions

T-test are conducted with the assumption that the length data follows a normal distribution, that the guinea pigs were given their dose and supplement assignments at random, thus making the groups independent. The confidence intervals will be unpaired.

Tooth Length by Supplement Mean Comparison

The Null Hypothesis is that the mean tooth length values for Orange Juice and Ascorbic Acid are not different in a statistically significant way. The alternative hypothesis is that the mean differences are statistically significant.

```
## t = 1.9153, df = 55.309, p-value = 0.06063
## 95 percent confidence interval: -0.1710156  7.5710156
##  mean in group Orange Juice mean in group Ascorbic Acid
##                    20.66333                    16.96333
```

As the 95% confidence interval contains zero and the p-value is greater than $\alpha = 0.05$ at 0.06063 we cannot reject H_0 and conclude there is no statistically significant difference in tooth length regardless of dose between either supplement.

0.5mg Dose Tooth Lenth by Supplement Mean Comparison

The Null Hypothesis is that the mean tooth length values for Orange Juice and Ascorbic Acid at 0.5mg are not different in a statistically significant way.

```
## t = 3.1697, df = 14.969, p-value = 0.006359
## 95 percent confidence interval: 1.719057 8.780943
##  mean in group Orange Juice mean in group Ascorbic Acid
##                       13.23                        7.98
```

As the 95% confidence interval is above zero and the p-value is less than $\alpha = 0.05$ at 0.006359 we can reject H_0 and conclude there is a statistically significant difference in tooth length and that Orange Juice results in longer teeth than Ascorbic Acid.

1.0mg Dose Tooth Lenth by Supplement Mean Comparison

The Null Hypothesis is that the mean tooth length values for Orange Juice and Ascorbic Acid at 1.0mg are not different in a statistically significant way.

```
## t = 4.0328, df = 15.358, p-value = 0.001038
## 95 percent confidence interval: 2.802148 9.057852
##  mean in group Orange Juice mean in group Ascorbic Acid
##                       22.70                       16.77
```

As the 95% confidence interval is above zero and the p-value is less than $\alpha = 0.05$ at 0.001038 we can reject H_0 and conclude there is a statistically significant difference in tooth length and that Orange Juice results in longer teeth than Ascorbic Acid.

2.0mg Dose Tooth Lenth by Supplement Mean Comparison

The Null Hypothesis is that the mean tooth length values for Orange Juice and Ascorbic Acid at 2.0mg are not different in a statistically significant way.

```
## t = -0.046136, df = 14.04, p-value = 0.9639
## 95 percent confidence interval: -3.79807  3.63807
##  mean in group Orange Juice mean in group Ascorbic Acid
##                       26.06                       26.14
```

As the 95% confidence interval contains zero and the p-value is greater than $\alpha = 0.05$ at 0.9639 we cannot reject H_0 and conclude there is no statistically significant difference in tooth length between either supplement.

Conclusion

Supplement type and dosage do have a statistically significant impact on tooth growth in guinea pigs. Orange Juice results in a statistically significant higher mean tooth length than Ascorbic Acid at 0.5mg and 1.0mg dosages but was no different than Ascorbic Acid at 2.0mg dosage. However, regardless of dose and just comparing by supplement type, no statistically significant difference was found between supplement types when examining the mean tooth lengths.

Appendix

```r
#Required Libraries
library(datasets)
library(data.table)
library(ggplot2)

#Load and Process Tooth Growth Dataset
dtToothGrowth = data.table(ToothGrowth) #Convert to data.table
names(dtToothGrowth) <- c("length", "supplement", "dose") #
levels(dtToothGrowth$supplement) = c('Orange Juice','Ascorbic Acid')
dtToothGrowth$doseFactor = factor(dtToothGrowth$dose) #Dose as a Factor for u
se in ggplot

#Summerize Data by Supplement and Dose
dtToothGrowthSummary =
  dtToothGrowth[,
    list(
      n = nrow(.SD)
      ,Min = min(length)
      ,Q1 = quantile(length, probs = 0.25)
      ,Median = median(length)
      ,Mean = mean(length)
      ,Q3 = quantile(length, probs = 0.75)
      ,Max = max(length)
      ,StdDev = sd(length)
    ),
    by = list(supplement, dose)
  ]

#ToothGrowth Summary by Supplement and Dose
dtToothGrowthSummary

#Tooth Length by Supplement and Dose Violin Plot
ggplot(dtToothGrowth, aes(x=supplement, y=length)) +
  geom_violin(aes(fill=doseFactor)) +
  ggtitle("Violin Plot of Tooth Length by Supplement and Dose") +
  xlab("Supplement") +
  ylab("Tooth Length")

#Tooth Length by Supplement and Dose Box Plot
ggplot(dtToothGrowth, aes(x=supplement, y=length)) +
  geom_boxplot(aes(fill=doseFactor)) +
  ggtitle("Box Plot of Tooth Length by Supplement and Dose") +
  xlab("Supplement") +
  ylab("Tooth Length")

#Welch Two Sample t-test 0.5mg Dose
t.test(
    length ~ supplement
    ,data = dtToothGrowth[dose == 0.5]
```

```
    ,paired = FALSE
    ,var.equal = FALSE
)

#Welch Two Sample t-test 1.0mg Dose
t.test(
    length ~ supplement
    ,data = dtToothGrowth[dose == 1.0]
    ,paired = FALSE
    ,var.equal = FALSE
)

#Welch Two Sample t-test 2.0mg Dose
t.test(
  length ~ supplement
  ,data = dtToothGrowth[dose == 2.0]
  ,paired = FALSE
  ,var.equal = FALSE
)
```