Inferential Analysis of Tooth Growth in Guinea Pigs

Tyler Peterson

10/29/2016

Overview

In this project, we analyze the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods - orange juice (coded as OJ) or ascorbic acid (a form of vitamin C and coded as VC).

Exploring the data:

```
# Exploring the data
  str(df)
```

```
## '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(df)
  #df
```

From examining the dataset, we see that the structure of the ToothGrowth dataframe contains 2 levels for the supplementation column (supp), which is a factor variable, but an unknown number of levels for the dose and length (len) columns. In our analyses, we expect the len variable to be the dependent variable, so we are less concerned with the number of levels it contains, but need to know the number of levels the dose variable has. We aggregate the length based on dose size and and see the number of unique tooth lengths per dose size For consistency, we aggregate the length based on supplement type to see the number of unique tooth lengths supplement types. From this we see that both dose size and supplement size are discrete variables and will be treated as such in the analysis.

```
  aggregate(len ~ dose, df, function(x) length(unique(x)))
```

```
##   dose len
## 1  0.5  17
## 2  1.0  17
## 3  2.0  17
```

```
  aggregate(len ~ supp, df, function(x) length(unique(x)))
```

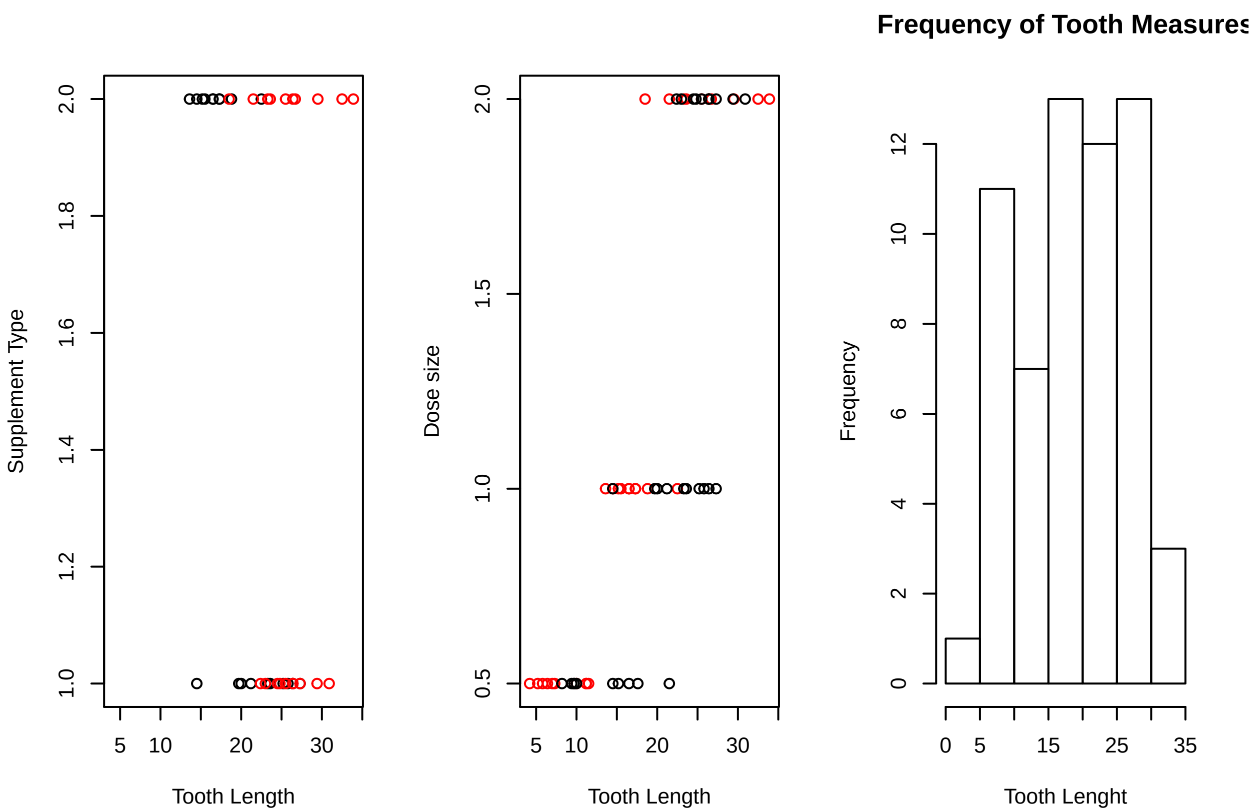
```
##   supp len
## 1   OJ  25
## 2   VC  27
```

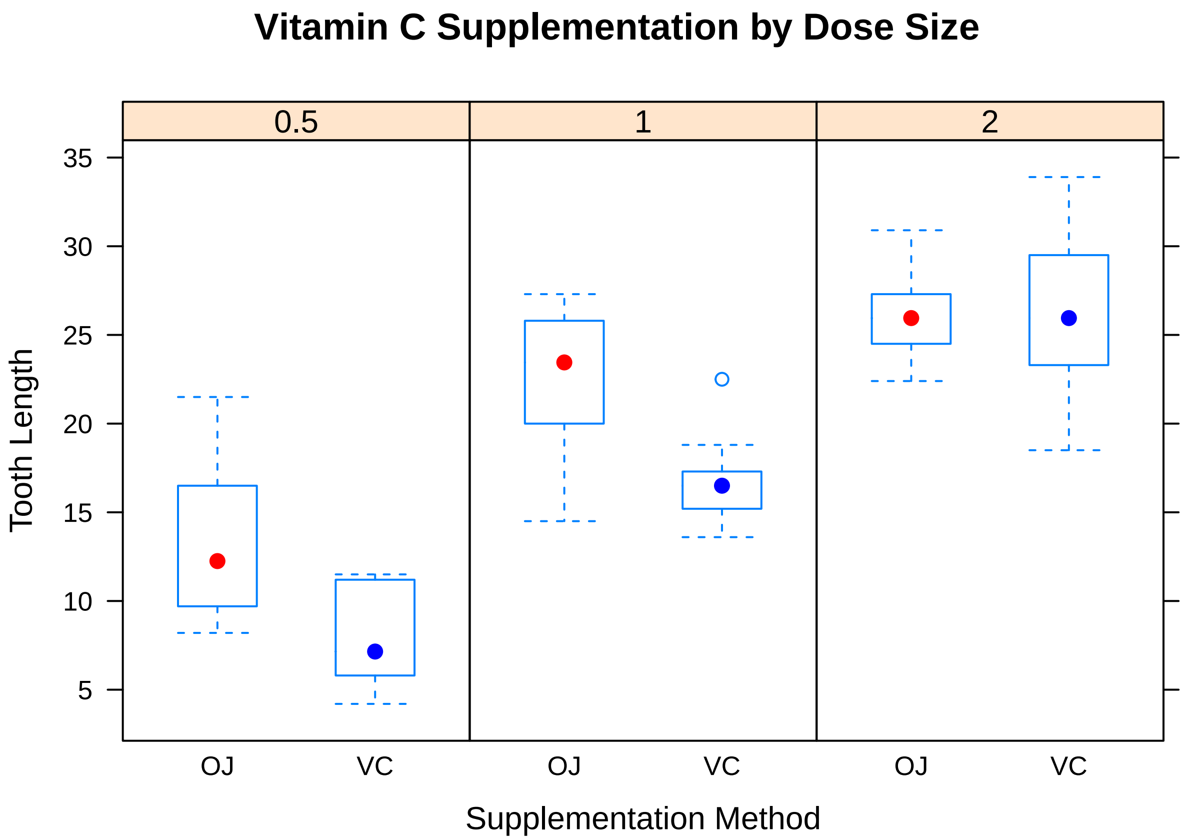
We create several plots to attempt to visually determine a pattern and see what to expect when performing the analysis. We first plot tooth length against supplement type with coloration per dose size. This plot is not very helpful in establishing a distinct relationship between supplement type and tooth length. We will not be using this plot in our analysis.

In our second plot, we plot tooth length against dose size with coloration per supplement type. In this plot, it appears that there is a correlation between tooth length and dose size; larger dose sizes lead to longer tooth lengths.

The third plot is a histogram of the tooth lengths. We include this plot to simply illustrate that there does not appear to be any outliers in the tooth lengths and we can include all tooth lengths in our analysis.

The final plot is a panel box plot segmenting the tooth length by dose size, then illustrating the supplement type per dose size. The correlation between tooth length and dose size is clearly visible in this graph, but again we're left with no clear distinction between supplement delivery type. There doesn't appear to be a strong relationship between delivering vitamin C through orange juice or ascorbic acid.





Assumptions

Before beginning our analysis, we assume that there is no dependency between the vitamin C delivery type and the dosage size. Furthermore, we assume that the guinea pig's tooth lengths are iid resulting from a random selection from within the population.

Analyzing the data

Once we visually inspect the data, we begin our hypothesis testing. In this testing, we use a single tailed Sudent's T Test and set our confidence level at 95%. Furthermore, we will be performing hypothesis testing on both the tooth length compared to the dose size and the tooth length compared to the supplement type.

First test is checking to see if the dose size affects the tooth length

- H0 = Tooth length is not affected by dose size
- Ha = Tooth length is affected by dose size

```
# Separate dose size into distinct subsets:
doseSub1 <- subset(df, df$dose %in% c(0.5, 1.0))
doseSub2 <- subset(df, df$dose %in% c(0.5, 2.0))
doseSub3 <- subset(df, df$dose %in% c(1.0, 2.0))

# Perform single tailed T Tests on the subsetted data
doseTTest1 <- t.test(len ~ dose, data = doseSub1)
```

```r
doseTTest2 <- t.test(len ~ dose, data = doseSub2)
doseTTest3 <- t.test(len ~ dose, data = doseSub3)

doseTTest1
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

```r
doseTTest2
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean in group 0.5   mean in group 2
##            10.605            26.100
```

```r
doseTTest3
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

Second test is checking to see if the supplement type affects tooth length

- H0 = Tooth length is not affected by supplement type
- Ha = Tooth length is affected by supplement type

```r
# Since there only exists 2 types of supplementation, we will not subset the dataset
suppTTest <- t.test(len~supp, data = df)
suppTTest
```

```
##
##  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
```

Conclusion:

Since the p-value in the len/dose t tests were $<<< 0.5$, we can reject the null hypothesis that the dose of vitamin C does not affect tooth length. However, since the p-value in the supplement test is > 0.5, we fail to reject the null hypothesis that the tooth length is not affected by the supplement type. We conclude that vitamin C is a valuable supplement to help rodent's teeth grow longer, it does not matter if the supplement is delivered through orange juice or ascorbic acid.