

The Effect of Vitamin C on Tooth Growth in Guinea Pigs

Mark E. Drummond

2015-05-23

Contents

Overview	1
Requirements	1
Exploratory Data Analysis	1
Basic Summary	2
Hypothesis Testing	4
Dosage vs. Tooth Growth	4
Delivery Method vs. Tooth Growth	5

Overview

We analyze the response in guinea pig tooth growth to vitamin C supplementation. From the data set description:

The response is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).

The ‘ascorbic acid’ delivery method refers to vitamin C tablets.

Requirements

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Provide a basic summary of the data.
3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose.
4. State your conclusions and the assumptions needed for your conclusions.

Exploratory Data Analysis

First we clean up our environment and set the seed for our random number generator to ensure consistent results during testing.

```
# Clean up the working environment.
rm(list = ls())
# We set the seed for consistent results. Comment this line to allow for a
# random seed each time the code is run.
set.seed(19647)
```

Next we load the data and review the basic structure and summary statistics.

```
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 ...
```

```
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
```

```
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
```

The dosage is actually a factor so let's convert it to such:

```
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
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: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 1 ...
```

Basic Summary

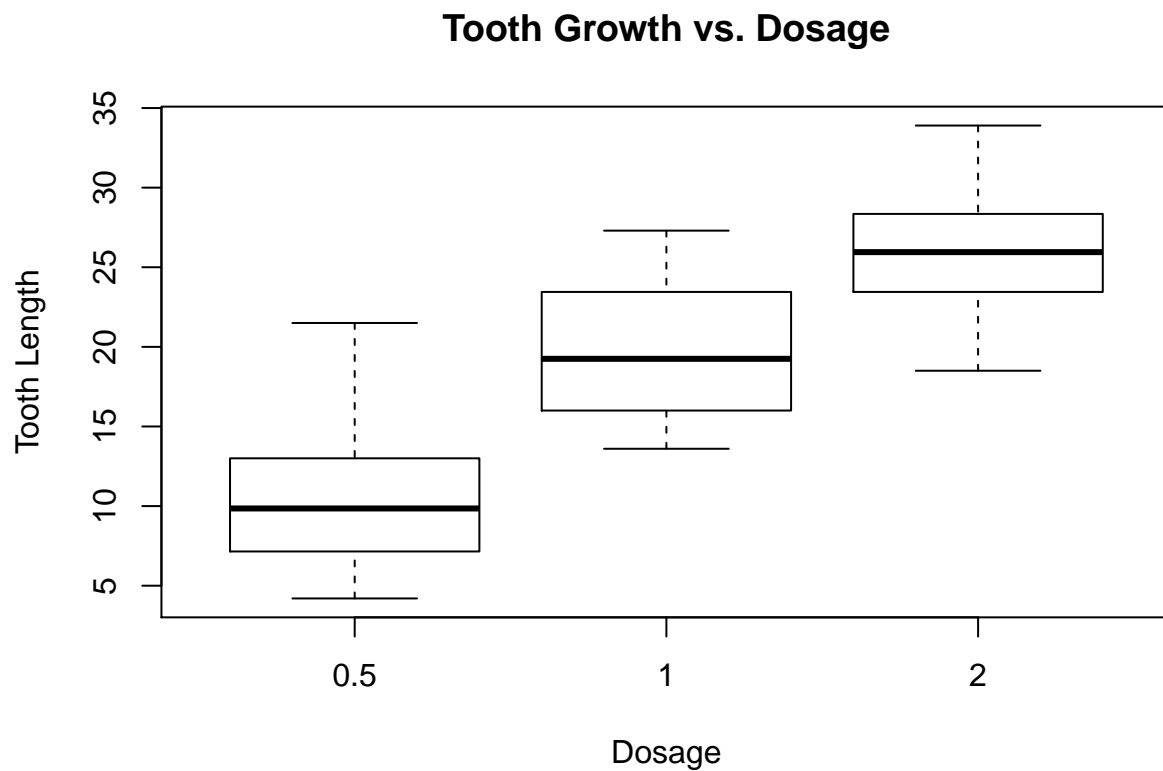
Each dosage / delivery method pair involved 10 guinea pigs.

```
table(ToothGrowth$dose, ToothGrowth$supp)
```

```
##  
##      OJ VC  
## 0.5 10 10  
## 1   10 10  
## 2   10 10
```

Comparing tooth growth against dosage:

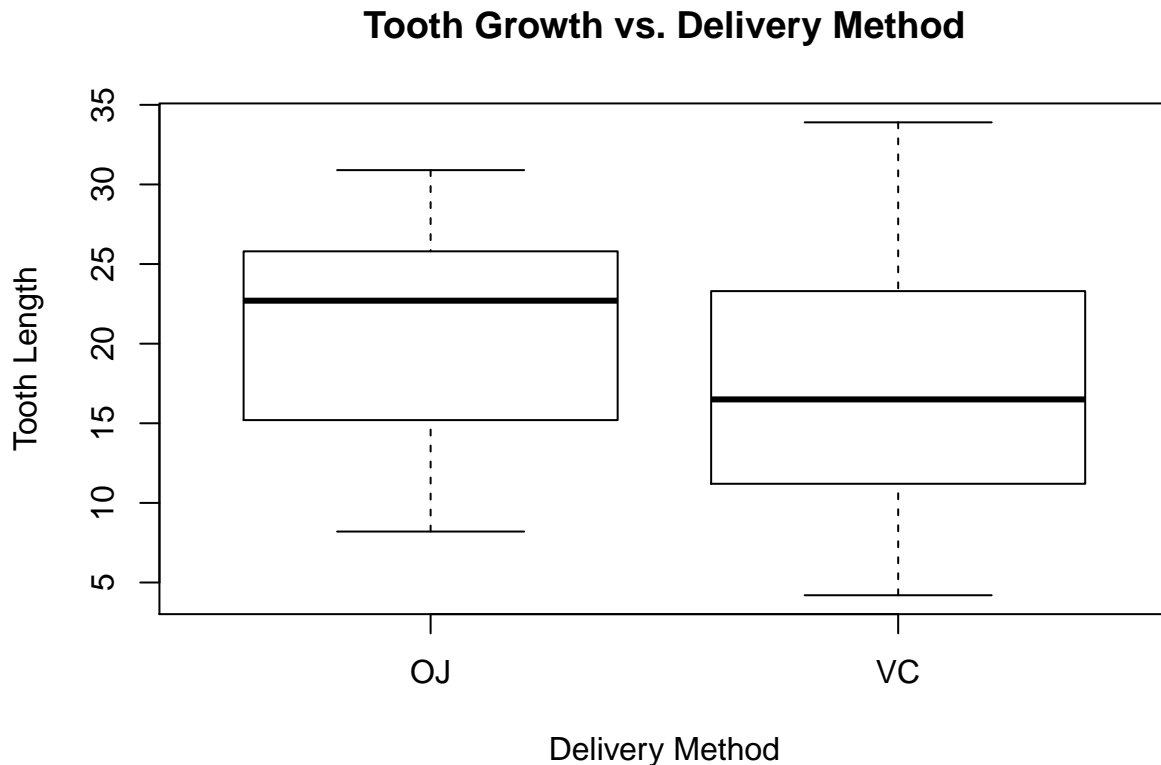
```
boxplot(ToothGrowth$len ~ ToothGrowth$dose,  
        main = "Tooth Growth vs. Dosage",  
        xlab = "Dosage",  
        ylab = "Tooth Length")
```



Based on the plot, it seems reasonable to assume a relationship between dosage and tooth growth.

Comparing tooth growth against vitamin C delivery method:

```
boxplot(ToothGrowth$len ~ ToothGrowth$supp,  
        main = "Tooth Growth vs. Delivery Method",  
        xlab = "Delivery Method",  
        ylab = "Tooth Length")
```



The relationship here is less clear. While the bulk of the data related to orange juice as a delivery method is associated with increased tooth growth, the dispersion of the data for vitamin C tablets as a delivery method makes the situation less certain.

Hypothesis Testing

Based on the results above, our intuition leads us to believe that higher dosage is related to increased tooth growth, while the delivery method is unclear. We wish to test our intuition in the first case, and see what a statistical analysis can tell us about the second by answering the following questions:

1. Does a higher dosage yield more tooth growth?
2. With respect to increased tooth growth, is orange juice a more effective vitamin C delivery method vs. tablets?

Dosage vs. Tooth Growth

- Null hypothesis: Higher dosage *does not* yield more tooth growth.
- Alt hypothesis: Higher dosage *does* yield more tooth growth.

Calculate the mean and standard deviation of tooth length aggregated by dosage, and the standard error of the mean:

```

n <- 20 # There are 20 guinea pigs per dosage group
mean_growth_by_dose <- tapply(ToothGrowth$len, ToothGrowth$dose, mean)
sd_growth_by_dose <- tapply(ToothGrowth$len, ToothGrowth$dose, sd)
std_error_of_mean <- sd_growth_by_dose / sqrt(n)

df <- data.frame(dose = unique(ToothGrowth$dose),
                 mean_tooth_length = as.numeric(mean_growth_by_dose),
                 sd_tooth_length = as.numeric(sd_growth_by_dose),
                 std_error_of_mean = as.numeric(std_error_of_mean))

Z <- 1.96
for(i in 1 : nrow(df)) {
  print(df[i, c("mean_tooth_length")] + c(-1, 1) * Z * df[i, c("std_error_of_mean")])
}

## [1] 8.632892 12.577108
## [1] 17.79985 21.67015
## [1] 24.44591 27.75409

```

Conclusion

Based on the results above, we reject the null hypothesis and state with 95% confidence that higher dosage does yield increased tooth growth.

Delivery Method vs. Tooth Growth

- Null hypothesis: Orange juice *is not* a more effective delivery method (does not yield more tooth growth vs. tablets).
- Alt hypothesis: Orange juice *is* a more effective delivery method (yields more tooth growth vs. tablets).

```

OJ <- ToothGrowth[ToothGrowth$supp == "OJ",]
VC <- ToothGrowth[ToothGrowth$supp == "VC",]

# Two-sided test first
t.test(OJ$len, VC$len)

##
## Welch Two Sample t-test
##
## data: OJ$len and VC$len
## 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 of x mean of y
## 20.66333 16.96333

```

Conclusion

Based on the results of the t-test, we cannot reject the null hypothesis. There may be no statistically significant difference in the vitamin C delivery method.