# Effect of Vitamin C on Tooth Growth

*R. Handsfield*

*February 19, 2016*

## Abstract

In this project, we analyze the `datasets::ToothGrowth` data set to determine the effect Vitamin C supplement and dosage on the average tooth length of guinea pigs. We use exploratory analysis and hypothesis testing to achieve our results.

## Overview

The `ToothGrowth` dataset reports the effect of Vitamin C on tooth growth in guinea pigs. The data frame contains 3 columns representing, *tooth length*, *supplement type*, and *dose (mg)*.

## Exploratory Analysis

```
library("datasets")
data(ToothGrowth)
```

```
dftg <- ToothGrowth
head(dftg, 3L)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
```

```
str(dftg)
```

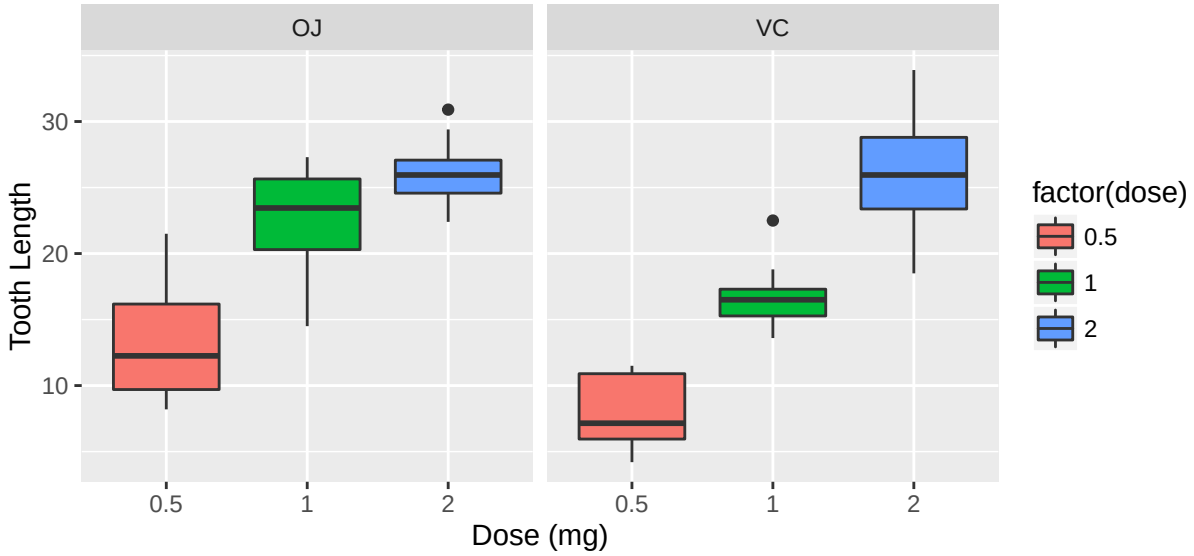
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
table(dftg[, c('supp','dose')])
```

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

There are two treatments (OJ and VC), each administered at 3 different dosages to independant groups of 10 guinea pigs.

```
g <- ggplot(dftg) + aes(x=factor(dose), y=len, fill=factor(dose)) + geom_boxplot()
g <- g + facet_grid(.~supp) + xlab("Dose (mg)") + ylab("Tooth Length")
g
```




## Comparing Treatments

- Null hypothesis: $H_0 : \mu_1 = \mu_2$
- Alternative hypothesis: $H_a : \mu_1 \neq \mu_2$

Summary of Mean Tooth Lengths

| Group | Mean Length |
|---|---|
| OJ | 20.663 |
| VC | 16.963 |
| 0.5 mg | 10.605 |
| 1.0 mg | 19.735 |
| 2.0 mg | 26.1 |

The means of the supplement and dosage groups are all distinct, but are they different enough to suggest that each experimental parameter had an effect?

### T-Testing

Comparing standard errors of each experimental group tells us whether we should use equal variance t-tests.

```
##    group std.error
## 1     OJ 1.2060049
## 2     VC 1.5091635
## 3 0.5 mg 1.0061776
## 4   1 mg 0.9873216
## 5   2 mg 0.8439257
```

Neither the treatment nor the dosage groups have similar variances, therefore we use unequal variance t-tests.

```
t <- t.test(len~supp, data=dftg, paired=FALSE, var.equal=FALSE)
data.frame(t$estimate, t$p.value)
```

```
##                   t.estimate  t.p.value
## mean in group OJ   20.66333 0.06063451
## mean in group VC   16.96333 0.06063451
```

```
t <- t.test(len~dose, data=rbind(df_half, df_one), paired=FALSE, var.equal=FALSE)
data.frame(t$estimate, t$p.value)
```

```
##                    t.estimate    t.p.value
## mean in group 0.5      10.605 1.268301e-07
## mean in group 1        19.735 1.268301e-07
```

```
t <- t.test(len~dose, data=rbind(df_two, df_one), paired=FALSE, var.equal=FALSE)
data.frame(t$estimate, t$p.value)
```

```
##                  t.estimate   t.p.value
## mean in group 1      19.735 1.90643e-05
## mean in group 2      26.100 1.90643e-05
```

The means of every group-pair are different, with p-values of 0.06 or better.

## Conclusions

Every experimental parameter caused some change in the tooth growth of a guinea pig population. From t-testing group pairs, we obtain the following results.

| Group | Mean Tooth Length | P-Value |
|---|---|---|
| OJ | 20.663 | 0.06 |
| VC | 16.963 | 0.06 |
| | | |
| 0.5 mg | 10.605 | $1.27 \times 10^{-7}$ |
| 1.0 mg | 19.735 | $1.27 \times 10^{-7}$ |
| | | |
| 1.0 mg | 19.735 | $1.91 \times 10^{-5}$ |
| 2.0 mg | 26.1 | $1.91 \times 10^{-5}$ |

To obtain these results, we assume that each guinea pig group is experimentally independant, and has a variance unequal to that of any other group.

## Appendix 1: Subsetting Groups

```
# subset individual groups
df_oj <- dftg[dftg$supp=='OJ',]
df_vc <- dftg[dftg$supp=='VC',]

df_half <- dftg[dftg$dose==0.5,]
df_one <- dftg[dftg$dose==1,]
df_two <- dftg[dftg$dose==2,]
```

## Appendix 2: Standard Errors of Groups

```
data.frame(group=c('OJ', 'VC', '0.5 mg', '1 mg', '2 mg'),
       std.error= c(sd(df_oj$len)/sqrt(30),
            sd(df_vc$len)/sqrt(30),
            sd(df_half$len)/sqrt(20),
            sd(df_one$len)/sqrt(20),
            sd(df_two$len)/sqrt(20)
            )
    )
```

```
##    group std.error
## 1     OJ 1.2060049
## 2     VC 1.5091635
## 3 0.5 mg 1.0061776
## 4   1 mg 0.9873216
## 5   2 mg 0.8439257
```

## Appendix 3: T-Test Results

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

```
##
##  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(len~dose, data=rbind(df_half, df_one), paired=FALSE, var.equal=FALSE)
```

```
##
##  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
t.test(len~dose, data=rbind(df_two, df_one), paired=FALSE, var.equal=FALSE)
```

```
##
##  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
```