Tooth Growth in Guinea Pigs and the Effect of Vitamin C

Wayne Hockensmith

January 31, 2017

In this project we will analyze the ToothGrowth data in the R datasets package. The data was collected while giving guinea pigs vitamin-c from orange juice and ascorbic acid. Doses were given in the amounts of .5, 1.0, and 2.0 millgrams per day.

Four things we were ask to do:

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.

1. Load the ToothGrowth data and perform some basic exploratory data analyses.

```
data(ToothGrowth) # Load the library "ToothGrowth":
str(ToothGrowth)  # Check the data structure:
```

```
## '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(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
```

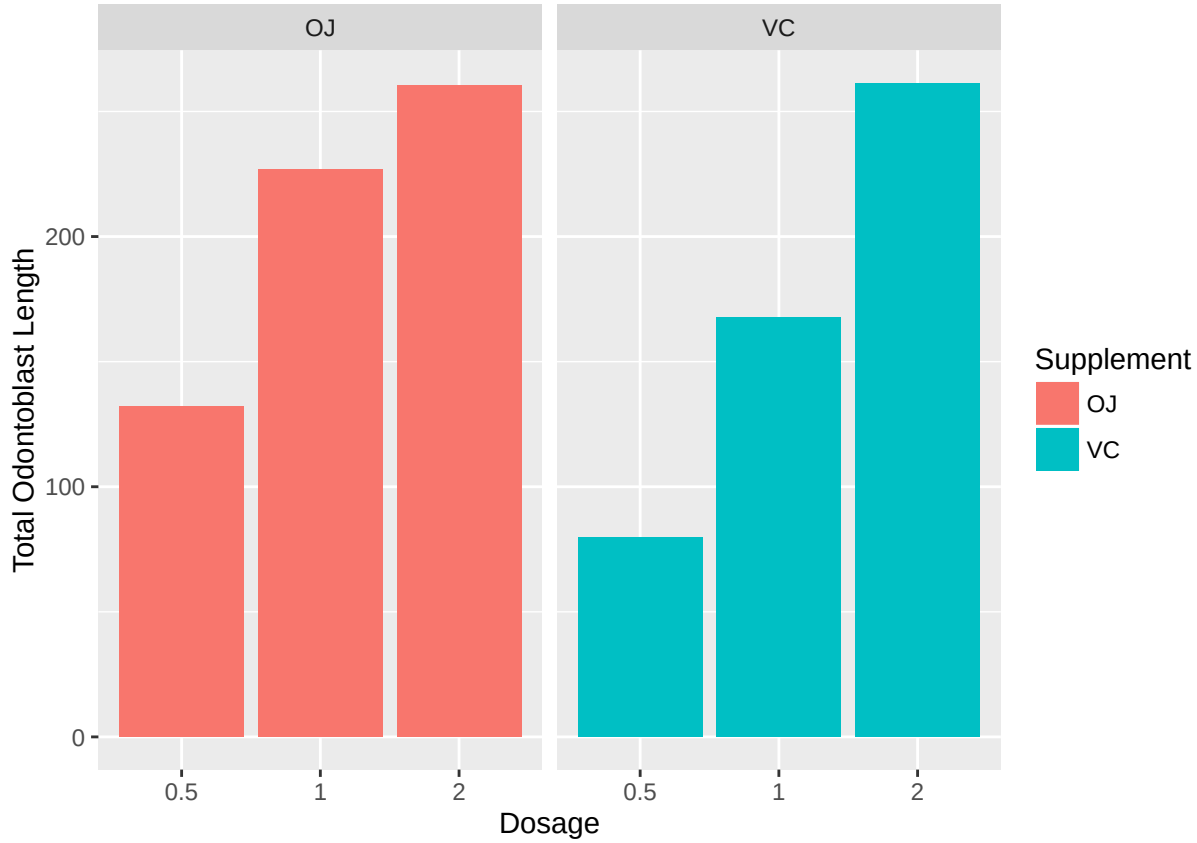
2. Provide a basic summary of the data.

There are sixty observations with 3 variables (columns).
The first column is len(tooth length).
The second column is supp(suppliment).
The third column is dose(dossage).

3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose.

First we use a simple ggplot to compair the data.

```
# Add a needed library first:
library(ggplot2)
# Build plot
ggplot(data=ToothGrowth, aes(x=as.factor(dose), y=len, fill=supp)) +
    geom_bar(stat="identity",) +
    facet_grid(. ~ supp) +
    xlab("Dosage") +
    ylab("Total Odontoblast Length") +
    guides(fill=guide_legend(title="Supplement"))
```



At first glance it would appear that orange juice has more of an effect then asorbic acid dose for 0.5 and 1

milligrams. However at 2 milligrams there is no ifference. Column and bar charts don't always tell the whole story or a true story. So we will move onto some hypothesis testing to determine if the column chart is correct.

To further investigate this lets use the null hypothesis that each supplement has the same effect at a certain dosage on the tooth.
1. H_0: $\mu_{OJ0.5} = \mu_{VC0.5}$
2. H_0: $\mu_{OJ1.0} = \mu_{VC1.0}$
3. H_0: $\mu_{OJ2.0} = \mu_{VC2.0}$

```
# Split the data into groups by the dosage ammounts:
dos05 <- subset(ToothGrowth, dose == 0.5)
dos10 <- subset(ToothGrowth, dose == 1.0)
dos20 <- subset(ToothGrowth, dose == 2.0)

# A t-test is conducted between the supplements:
t05 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = dos05)
t05$p.value; t05$conf[1]
```

```
## [1] 0.006358607
```

```
## [1] 1.719057
```

```
t10 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = dos10)
t10$p.value; t10$conf[1]
```

```
## [1] 0.001038376
```

```
## [1] 2.802148
```

```
t20 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = dos20)
t20$p.value; t20$conf[1]
```

```
## [1] 0.9638516
```

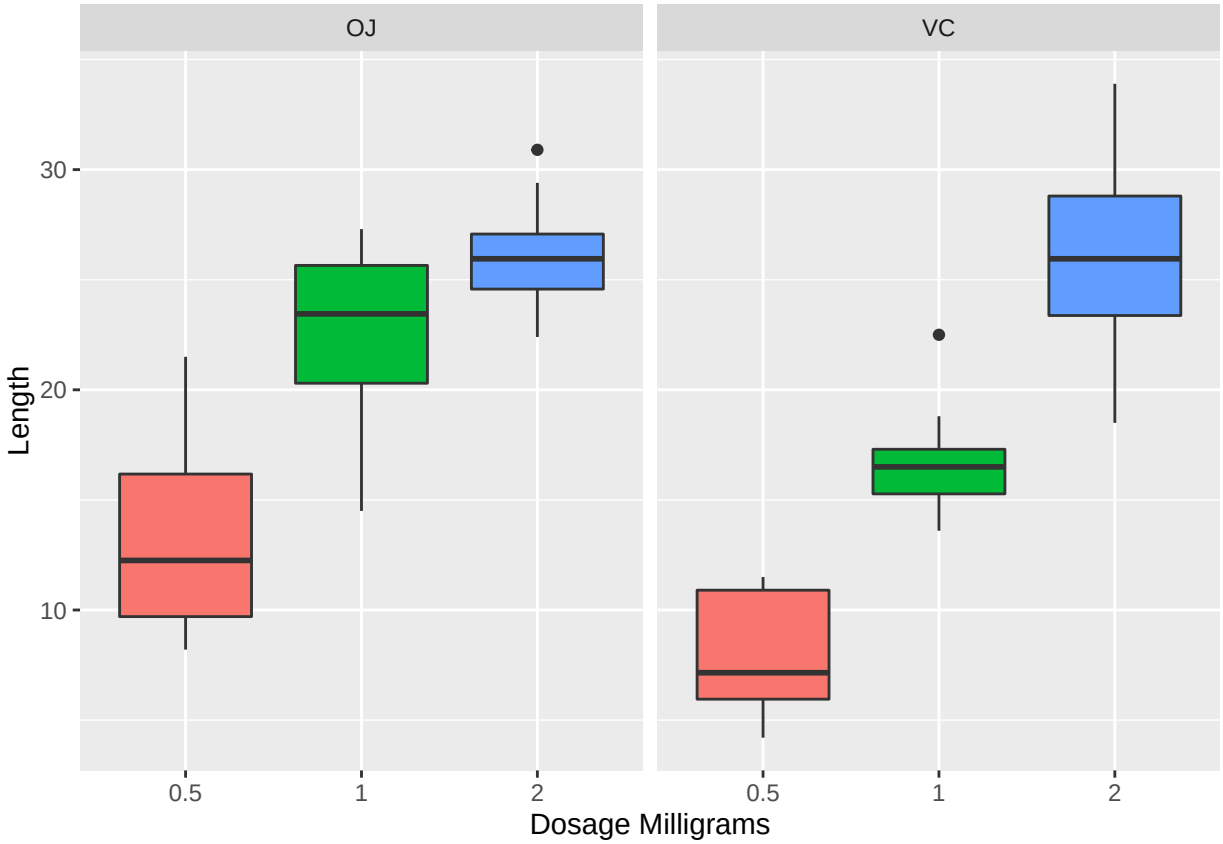
```
## [1] -3.79807
```

The 1.0 dosage has a confidence interval of 1.719-8.781 while dosage 2.0 has a confidence interval of 2.802-9.058. The 1.0 dosage has a p-value of 0.006358607 and the 1.5 dosage has a p-value of 0.001038. This would indicate that there is a signifigant difference in mean values between the two supplements at these dosages.

The 3.0 dosage has a p-value of 0.9639 and a confidence interval below zero -3.798-3.638, indicates that there is no significance between the supplements at this dosage.

Lets do one more validation using a box plot.

```
par(mfrow = c(1,2))
p1 <- ggplot(ToothGrowth, aes(x = factor(dose), y = len, fill = factor(dose))) +
  ylab("Length") +
  xlab("Dosage Milligrams")
p1 + geom_boxplot() + guides(fill=FALSE) + facet_grid(. ~ supp)
```



The box plot shows that orange juice still promotes more tooth growth with dosages of 0.5 and 1 millingrams. However with a dosage of 2 milligrams the mean tooth growth is the same (25.95) but ascorbic acid has a wider range of results. Ascorbic acid at the dose of 2 milligrams had a range from 18.5 to 33.9 where as orange juice had a range only of 22.4 to 30.9. ascorbic acid does have two observations over orange juice but the not enough to offset the fact toat fifty percent of the ascorbic acid lies less then the meadian growth of those treated with orange juice.

4. conclusions and the assumptions

My conclusion for this project is orange juice is a better supplement to promote tooth growth, atleast with Guinea Pigs. At lower dosages orange juice significantly out performed ascorbic aacid and at the higher dosage the orange juice was more consistant.
This conclusion is based on the assumtions that all Guinea pigs:
1. The same age.
2. The same size.
3. Had the same diet.
4. Had no deseases or had the same ones.
5. All came from the same genetic stock.
6. All had the same living conditions.