# ToothGrowth data Analysis

*Author: Jianlei (John) Sun*

*November 2, 2016*

# 1. Overview

In this project, we're going to analyze the ToothGrowth data in the R datasets package.

The response is 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 or ascorbic acid (a form of vitamin C and coded as VC).

- len : Tooth Length is numeric
- supp: Supplement is Factor consisting of OJ and VC
- dose: Dosage is numeric. As it is just 0.5, 1 and 2, we will have to convert it to factor so that it is easier for analysis

# 2. Load the ToothGrowth data and perform some basic exploratory data analyses

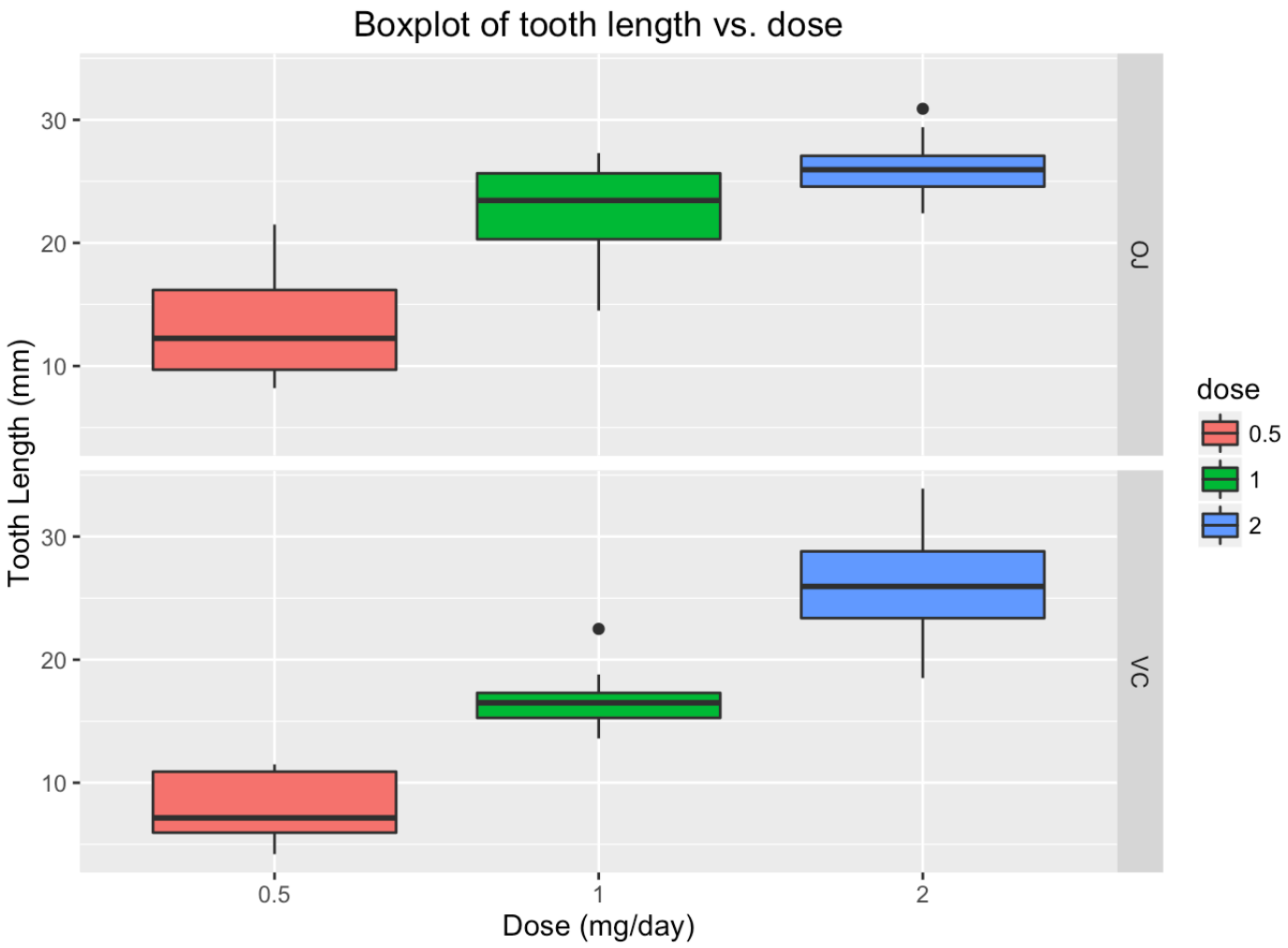
```
library(datasets)
library(ggplot2)

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 ...
```

Display a boxplot for the dataset:

```
qplot(dose, len, data = ToothGrowth, facets = supp ~ ., geom = "boxplot", fill = dose
, xlab = "Dose (mg/day)", ylab = "Tooth Length (mm)", main = "Boxplot of tooth length
vs. dose")
```




# 3. Provide a basic summary of the data.

```
summary(ToothGrowth)
```

```
##        len        supp     dose
##  Min.   : 4.20   OJ:30   0.5:20
##  1st Qu.:13.07   VC:30   1  :20
##  Median :19.25           2  :20
##  Mean   :18.81
##  3rd Qu.:25.27
##  Max.   :33.90
```

# 4. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)

# 4.1 Hypothesis Testing 1 - the impact of dosage on tooth growth

- Null Hypothesis: there is no effect of dosage on tooth growth.
- Alternate Hypothesis: higer doses produce greater tooth growth.

## Compare between dose of 0.5mg and 1mg:

```
t.test(len ~ dose, data = subset(ToothGrowth, dose %in% c(0.5, 1.0)))
```

```
##
##  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
```

Because the P value 1.268e-07 is less than 0.05, it is highly significant and we can reject the null hypothesis.

## Compare between dose of 1mg and 2mg:

```
t.test(len ~ dose, data = subset(ToothGrowth, dose %in% c(1.0, 2.0)))
```

```
##
##  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
```

Because the P value 1.906e-05 is less than 0.05, it is highly significant and we can reject the null hypothesis.

## Compare between dose of 0.5mg and 2mg:

```
t.test(len ~ dose, data = subset(ToothGrowth, dose %in% c(0.5, 2.0)))
```

```
##
##  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
```

Because the P value 4.398e-14 is less than 0.05, it is highly significant and we can reject the null hypothesis.

Therefore, we reject the null hypothesis, and accept the alternative, i.e., the higer doses produce greater tooth growth.

# 4.2 Hypothesis Testing 2 - differences between supplements by dosage

- Null Hypothesis: there is no effect between supplement OJ and VC
- Alternate Hypothesis: the OJ is better on toothgrowth than VC

## For 0.5mg:

```
t.test(len ~ supp, data = subset(ToothGrowth, dose == 0.5))
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

Because the p-value is below 0.05, we accept the alternative, i.e., OJ is a better supplement than VC at dosage of 0.5mg.

## For 1mg:

```
t.test(len ~ supp, data = subset(ToothGrowth, dose == 1))
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

Because the p-value is below 0.05, we accept the alternative, i.e., OJ is a better supplement than VC at dosage of 1mg.

### For 2mg:

```
t.test(len ~ supp, data = subset(ToothGrowth, dose == 2))
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

Because the p-value is higher than 0.05, we cannot reject the null hypothesis. There is no significant difference between the two supplement at dosage of 1mg.

# 5. State your conclusions and the assumptions needed for your conclusions.

- Higher dosages have a positive influence on toothgrowth.
- The p-value indicators for doses of 0.5 and 1.0 are less than 5% and the confidence intervals of the test do not contain 0. However, the p-value for the dose of 2.0 is greater than 5% and the confidence test contains 0. Therefore, we conclude that OJ is better than VC in promoting tooth growth for dosages of 0.5mg and 1mg. However, there is no significant difference at dosage of 2mg.