

Tooth Growth

Gary Chamberlain

Saturday, October 25, 2014

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

```
library(datasets)
data(ToothGrowth); names(ToothGrowth) <- c("Length", "Supplement", "Dosage")
summary(ToothGrowth)
```

```
##      Length      Supplement      Dosage
##  Min.   : 4.2      OJ:30      Min.   :0.50
## 1st Qu.:13.1      VC:30      1st Qu.:0.50
##  Median :19.2                      Median :1.00
##   Mean  :18.8                      Mean   :1.17
## 3rd Qu.:25.3                      3rd Qu.:2.00
##   Max.  :33.9                      Max.   :2.00
```

2. Provide a basic summary of the data.

```
mean.Oj.0.5 <- mean(subset(ToothGrowth, Dosage=='0.5' & Supplement=='OJ')$Length)
mean.Oj.1 <- mean(subset(ToothGrowth, Dosage=='1' & Supplement=='OJ')$Length)
mean.Oj.2 <- mean(subset(ToothGrowth, Dosage=='2' & Supplement=='OJ')$Length)
mean.vc.0.5 <- mean(subset(ToothGrowth, Dosage=='0.5' & Supplement=='VC')$Length)
mean.vc.1 <- mean(subset(ToothGrowth, Dosage=='1' & Supplement=='VC')$Length)
mean.vc.2 <- mean(subset(ToothGrowth, Dosage=='2' & Supplement=='VC')$Length)
sd.Oj.0.5 <- sd(subset(ToothGrowth, Dosage=='0.5' & Supplement=='OJ')$Length)
sd.Oj.1 <- sd(subset(ToothGrowth, Dosage=='1' & Supplement=='OJ')$Length)
sd.Oj.2 <- sd(subset(ToothGrowth, Dosage=='2' & Supplement=='OJ')$Length)
sd.vc.0.5 <- sd(subset(ToothGrowth, Dosage=='0.5' & Supplement=='VC')$Length)
sd.vc.1 <- sd(subset(ToothGrowth, Dosage=='1' & Supplement=='VC')$Length)
sd.vc.2 <- sd(subset(ToothGrowth, Dosage=='2' & Supplement=='VC')$Length)
```

The means for Orange Juice are 13.23 for 0.5 mg, 22.7 for 1.0 mg and 26.06 for 2.0 mg. The standard deviations are 4.4597 for 0.5 mg, 3.911 for 1.0 mg and 2.6551 for 2.0 mg. The means for Vitimin C are 7.98 for 0.5 mg, 26.14 for 1.0 mg and 16.77 for 2.0 mg. The standard deviations are 2.7466 for 0.5 mg, 2.5153 for 1.0 mg and 4.7977 for 2.0 mg.

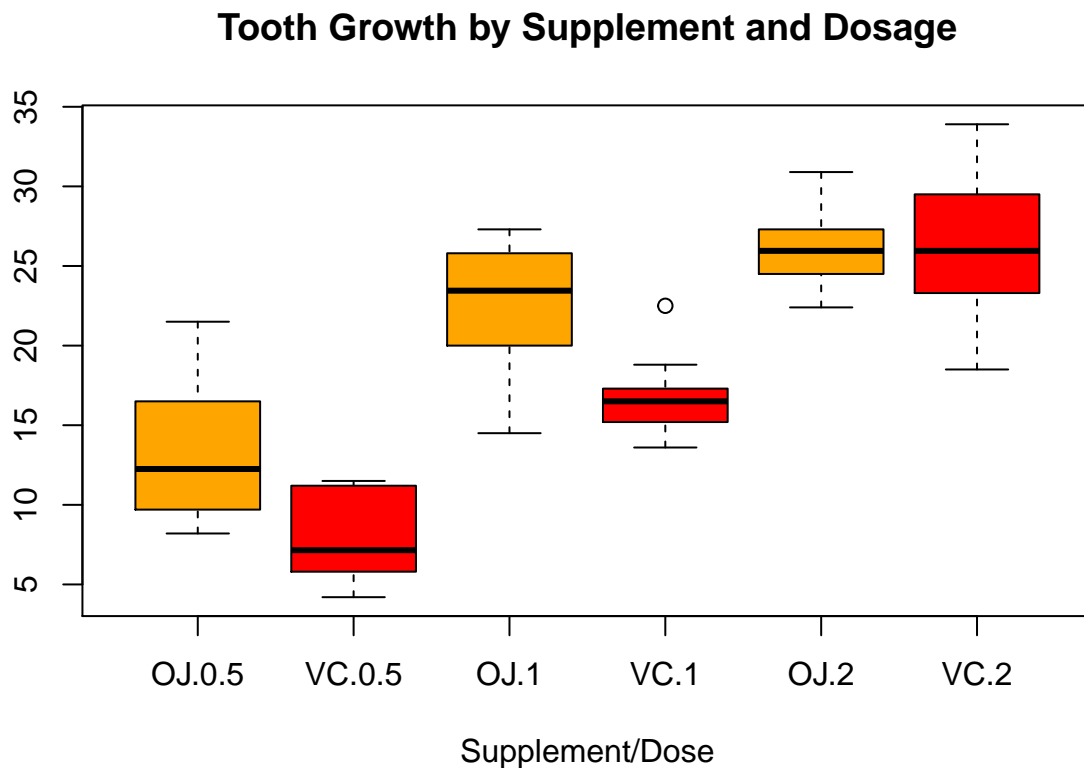
3. Use confidence intervals and hypothesis tests to compare tooth growth by supp and dose.

(Use the techniques from class even if there's other approaches worth considering)

Null hypothesis that means are of the two supplements are equal

```
t.test.0.5 <- t.test(Length ~ Supplement, var.equal=F,  
                    subset(ToothGrowth, Dosage==0.5), alternative = "g" )  
t.test.1 <- t.test(Length ~ Supplement, var.equal=F,  
                    subset(ToothGrowth, Dosage==1), alternative = "g" )  
t.test.2 <- t.test(Length ~ Supplement, var.equal=F,  
                    subset(ToothGrowth, Dosage==2), alternative = "g" )
```

P values are 0.0032 for 0.5 mg 5.1919×10^{-4} for 1.0 mg and 0.5181 for 2.0 mg Rejected for 0.5 and 1.0 mg doses but not rejected at 2.0mg dose as the high p.value indicates the means are similar. The plot below shows the means and confidence intervals for orange juice vs vitamin c at each dose. This illustrates what we know from the p values above.



4.State your conclusions and the assumptions needed for your conclusions.

In general conclusion we can say that Vitamin C increases tooth growth. Orange juice appears to work better than vitamin c supplements as shown by the null hypothesis being rejected above, except at the higher dose of 2mg where the means are similar and the null hypothesis was not rejected. Although the confidence in the supplement for the 2mg dose is wider.