

Final Project, Part 2

A. Kramer

Wednesday, February 18, 2015

Assignment

Now in the second portion of the class, we're going to analyze the ToothGrowth data in the R datasets package.

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. (Only use
4. State your conclusions and the assumptions needed for your conclusions.

Overview

The data frame loaded contains 60 observations on 3 variables. The variables are:

1. Length of the tooth
2. Supplement type (VC or OJ)
3. Dosage in milligrams

The data set is interrogated to provide the basic summary of the data and compare tooth growth by supplement and dosage (using confidence intervals).

Solution

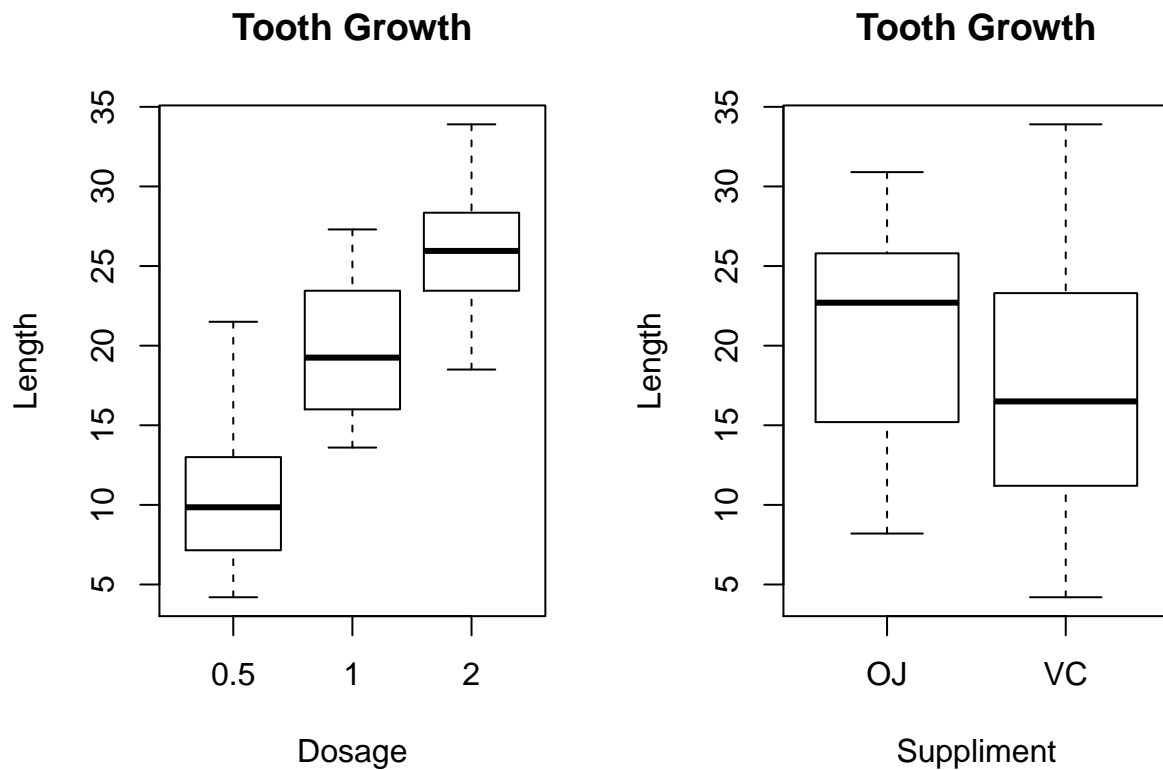
Question 1: Load the ToothGrowth data and perform some basic exploratory data analyses

```
library(datasets)
tmp <- ToothGrowth
str(tmp)
```

```
## '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 ...
## $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

Plotting tooth length vs. supplement and tooth length vs. dosage to visually investigate the relationships between the variables.

```
par(mfrow=c(1,2))
boxplot(len~dose, data=tmp, xlab = "Dosage", ylab = "Length", main = "Tooth Growth")
boxplot(len~supp, data=tmp, xlab = "Supplement", ylab = "Length", main = "Tooth Growth")
```



Question 2: Provide a basic summary of the data.

```
dim(tmp)
```

```
## [1] 60 3
```

```
tmp$dose <- as.factor(tmp$dose)
summary(tmp)
```

```
##      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
```

Displaying table of supplement and dosage usage in the data frame.

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

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

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

Calculating Confidence Intervals for equal and unequal variances.

```
t1 <- t.test(len~supp, paired=F, var.equal=T, data=tmp)
t2 <- t.test(len~supp, paired=F, var.equal=F, data=tmp)
results = cbind(c(t1$p.value, t2$p.value), c(t1$conf[1],t2$conf[1]), c(t1$conf[2],t2$conf[2]))
colnames(results) <- c("P-Value", "Confidence Low", "Confidence High")
rownames(results) <- c("Equal Variances","Unequal Variances")
results
```

```
##                P-Value Confidence Low Confidence High
## Equal Variances  0.06039337      -0.1670064      7.567006
## Unequal Variances 0.06063451      -0.1710156      7.571016
```

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

Based on the analysis performed we can conclude that the 2mg dosage provides greater impact on tooth growth than 1mg or 0.5mg dosage. We can also conclude that both, Vitamin C and Orange Juice impact tooth growth.