Statistical Inference Assignment, Part B

Teja Kodali

22 March 2015

Overview

Here, we will investigate the ToothGrowth data in the R datasets package. We will load the ToothGrowth data, perform some basic data analysis and provide a summary of the data. We will use confidence intervals and hypotheses tests to compare tooth growth by supp and dose.

Loading the data

```
library(ggplot2)
library(datasets)
data(ToothGrowth)
str(ToothGrowth)
```

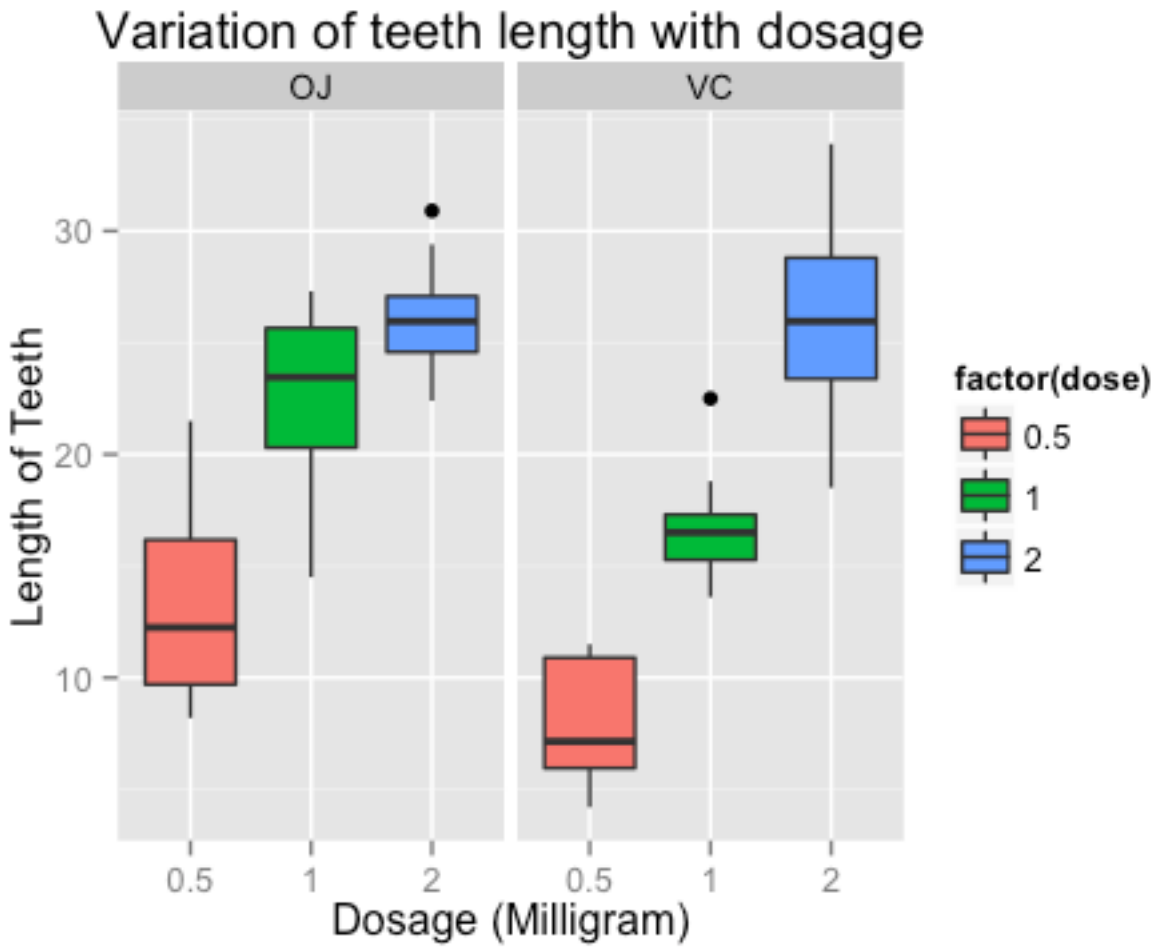
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
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
```

Plotting

```
plot <- ggplot(ToothGrowth,
               aes(x=factor(dose),y=len,fill=factor(dose)))
plot + geom_boxplot(notch=F) + facet_grid(.~supp) +
     scale_x_discrete("Dosage (Milligram)") +
     scale_y_continuous("Length of Teeth") +
     ggtitle("Variation of teeth length with dosage")
```



As we can see, the length of the teeth is higher when dosage is higher, regardless of whether it was Orange Juice of Vitamin C.

Summary of data

```
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
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
```

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

```
##
##      0.5  1  2
```

```
##   OJ  10 10 10
##   VC  10 10 10
```

Confidence intervals and hypotheses tests

```
supp.t1 <- t.test(len~supp, paired=F, var.equal=T, data=ToothGrowth)
supp.t2 <- t.test(len~supp, paired=F, var.equal=F, data=ToothGrowth)
supp.result <- data.frame("p-value"=c(supp.t1$p.value,
supp.t2$p.value),
                          "Conf-
Low"=c(supp.t1$conf[1],supp.t2$conf[1]),
                          "Conf-
High"=c(supp.t1$conf[2],supp.t2$conf[2]),
                          row.names=c("Equal Var","Unequal Var"))
supp.result
```

```
##                 p.value   Conf.Low Conf.High
## Equal Var   0.06039337 -0.1670064  7.567006
## Unequal Var 0.06063451 -0.1710156  7.571016
```

Conclusions

Based on the analysis above, we can conclude that

1. The 2mg dose has larger impact on tooth growth than 1mg and 0.5mg, while 1mg dose has more impact than 0.5mg dose. So there is a different in the growth of the tooth while the doses are larger.
2. There is no doubt that Orange Juice and Vitamin C have obvious different impact on tooth growth.

Assumptions

It is assumed that

1. There are no other supplements consumed by the subjects.
2. Orange Juice and Vitamin C are the only two factors influencing the growth of the teeth.