Statistical inference: Project 2

Mrugank Akarte

27 February 2016

Overview

In this report, *ToolGrowth* data set was used to perform basic exploratory data analysis. Later with the help of confidence intervals and hypothesis testing, growth of tooth by supplement and dose was analyzed.

Setting up the data

```
library(datasets)
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
```

```
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 ...
```

Changing the variable *dose* to factor

```
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 ...
```

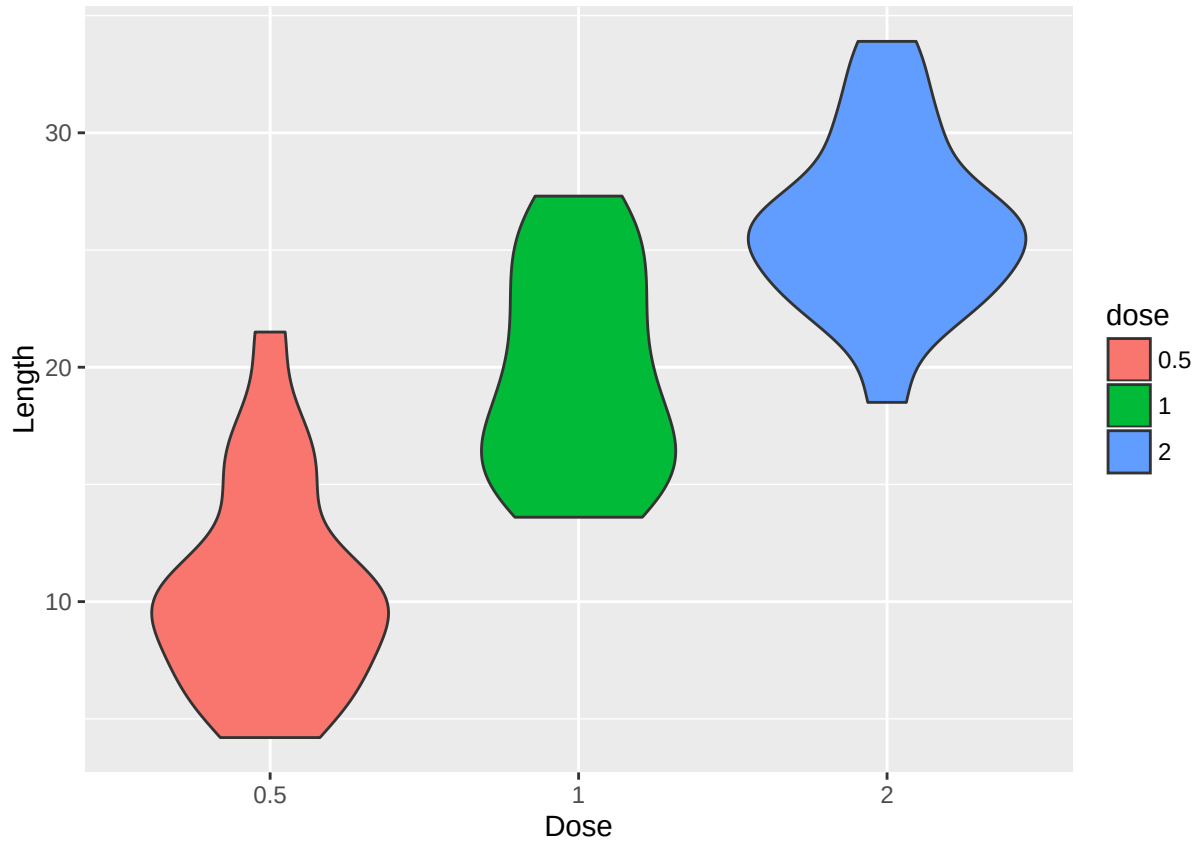
Exploratory data analysis

Checking the effect of dose on toothgrowth.

```
meandose<-split(ToothGrowth$len,f = ToothGrowth$dose)
sapply(meandose, mean)
```

```
##     0.5       1       2
## 10.605 19.735 26.100
```

```
library(ggplot2)
ggplot(data=ToothGrowth, aes(x=dose,y=len,group=dose))+geom_violin(aes(fill=dose))+xlab("Dose")+ylab("L
```



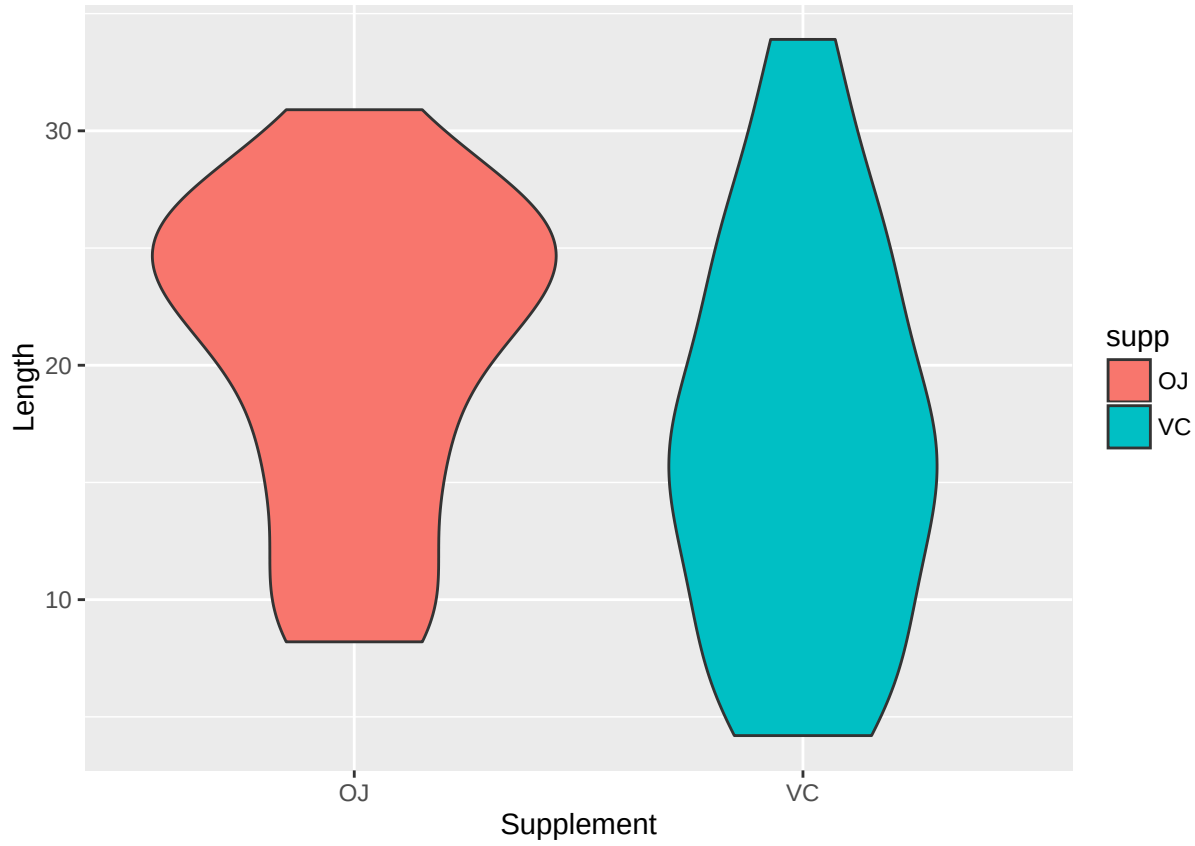
From above violin graph we can see significant increase in length of tooth with increase in dose.

Checking the effect of supplement on toothgrowth.

```
meansupp<-split(ToothGrowth$len, ToothGrowth$supp)
sapply(meansupp, mean)
```

```
##       OJ       VC
## 20.66333 16.96333
```

```
ggplot(data=ToothGrowth, aes(x=supp,y=len,group=supp))+geom_violin(aes(fill=supp))+xlab("Supplement")+yl
```



From above graph we can see increase in tooth growth for supplement type 'OJ' more than other supplement.

Confidence Interval and Hypothesis testing

```
t.test(ToothGrowth$len[ToothGrowth$supp=='OJ'],ToothGrowth$len[ToothGrowth$supp=='VC'],paired = F)
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "OJ"] and ToothGrowth$len[ToothGrowth$supp == "VC"]
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean of x mean of y
##  20.66333  16.96333
```

The p-value for the t-test is greater than 0.05, hence we cannot reject the null hypothesis. Earlier from the graph we stated that we can see growth in tooth due to supplement 'OJ' more than supplement'VC'. But due to lack of strong evidence we cannot confirm any relation between supplement and tooth growth.

Testing for relation between dose and toothgrowth

```
t.test(ToothGrowth$len[ToothGrowth$dose==1],ToothGrowth$len[ToothGrowth$dose==0.5],paired = F,var.equal
```

```
##
##  Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$dose == 1] and ToothGrowth$len[ToothGrowth$dose == 0.5]
## t = 6.4766, df = 38, p-value = 1.266e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   6.276252 11.983748
## sample estimates:
## mean of x mean of y
##    19.735    10.605
```

The p-value for t-test is very small, i.e it less than 0.05. Hence we can safely conclude that, with increase in dose from 0.5mg to 1mg there is growth in tooth. We can also see that zero is not included in the confidence interval.

```
t.test(ToothGrowth$len[ToothGrowth$dose==2],ToothGrowth$len[ToothGrowth$dose==1],paired = F,var.equal =
```

```
##
##  Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$dose == 2] and ToothGrowth$len[ToothGrowth$dose == 1]
## t = 4.9005, df = 38, p-value = 1.811e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  3.735613 8.994387
## sample estimates:
## mean of x mean of y
##    26.100    19.735
```

Again the p-value is approx. zero, hence we conclude that with increase in dose we observe growth in tooth. Again zero is not included in the confidence interval.

Conclusion

From above analysis we can conclude that:

*There is no relation between the supplement and the growth of tooth.

*With increase in dose we can see increase in growth of tooth.