

# A simple analysis about the ToothGrowth data set

*Rick*

*September 27, 2015*

## Synopsis:

We used the ToothGrowth data to make a simple T-test analysis to compare tooth growth by supp and dose.

## Data Processing

First, we load the data.

```
data(ToothGrowth)
```

Let us see a brief summary of this dataset.

```
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
```

```
levels(ToothGrowth$supp)
```

```
## [1] "OJ" "VC"
```

```
levels(as.factor(ToothGrowth$dose))
```

```
## [1] "0.5" "1"  "2"
```

```
tapply(ToothGrowth$supp, ToothGrowth$supp, length)
```

```
## OJ VC
## 30 30
```

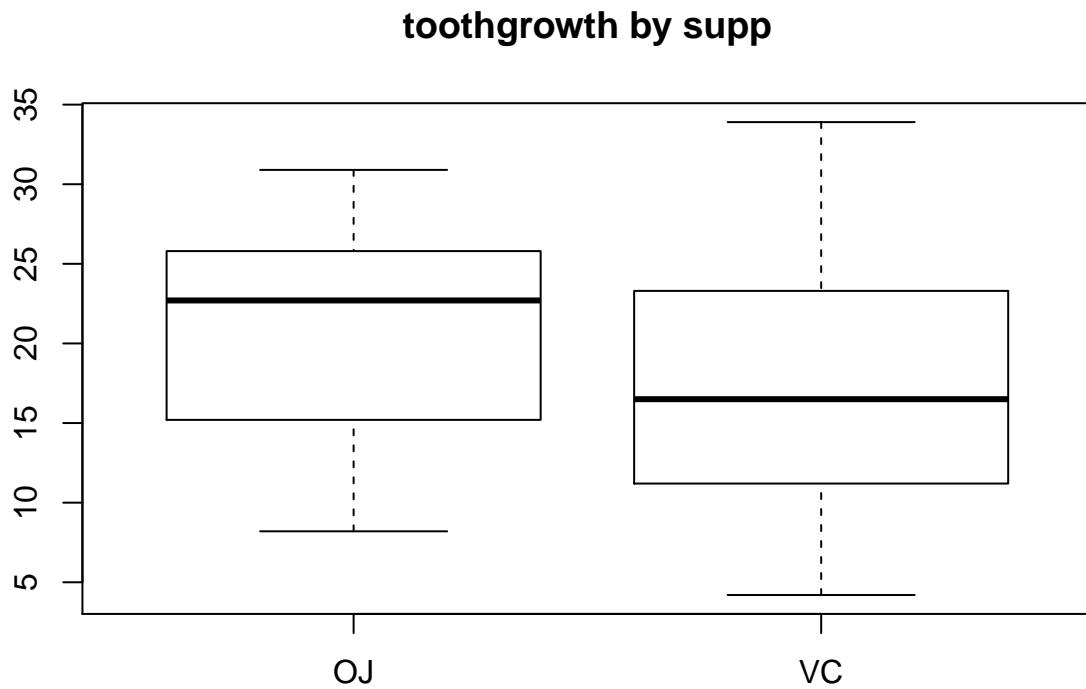
```
tapply(ToothGrowth$dose, ToothGrowth$dose, length)
```

```
## 0.5  1  2
## 20 20 20
```

To be more easy understanding, let us plot (boxplot) the Toothgrowth by dose and by supp, respectively.

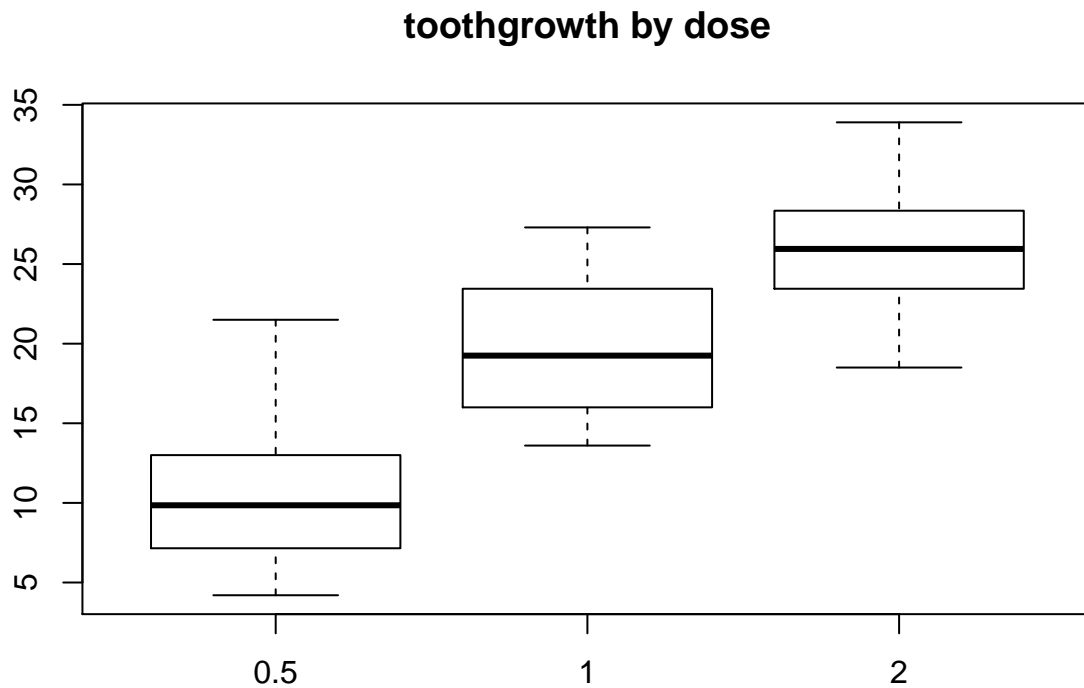
By supp:

```
with(ToothGrowth,plot(supp,len,type="p",main="toothgrowth by supp"))
```



By dose:

```
with(ToothGrowth,plot(as.factor(dose),len,type="p",main="toothgrowth by dose"))
```



## T-test

Then we make t test first by supp to test the hypothesis that:  $H_0$ : supp has no effect on the mean of tooth growth. And we assume that the variance of each group is the same.

```
t.test(len~supp,paired=TRUE,var.equal=TRUE,data=ToothGrowth)
```

```
##
## Two Sample t-test
##
## data: len by supp
## t = 1.9153, df = 58, p-value = 0.06039
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.1670064 7.5670064
## sample estimates:
## mean in group OJ mean in group VC
## 20.66333 16.96333
```

Hence, we fail to reject the hypothesis. We can conclude that using a 95 percent confidence interval, supp has no effect on the true mean of toothgrowth.

Then we make t test first by dose to test the hypothesis that:  $H_0$ : does has no effect on the mean of tooth growth.

Because there are three different doeses. We compare them two by two.

```
split(ToothGrowth, ToothGrowth$dose)[[1]][["dose"]][1]
```

```
## [1] 0.5
```

```
dose1<-split(ToothGrowth, ToothGrowth$dose)[[1]][["len"]]  
split(ToothGrowth, ToothGrowth$dose)[[2]][["dose"]][1]
```

```
## [1] 1
```

```
dose2<-split(ToothGrowth, ToothGrowth$dose)[[2]][["len"]]  
split(ToothGrowth, ToothGrowth$dose)[[3]][["dose"]][1]
```

```
## [1] 2
```

```
dose3<-split(ToothGrowth, ToothGrowth$dose)[[3]][["len"]]
```

And we assume that the variance of each group is the same for all the tests.

(1) The t test between dose 0.5 and 1

```
t.test(dose1,dose2,paired=TRUE,var.equal=TRUE,data=ToothGrowth)
```

```
##  
## Paired t-test  
##  
## data: dose1 and dose2  
## t = -6.9669, df = 19, p-value = 1.225e-06  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -11.872879 -6.387121  
## sample estimates:  
## mean of the differences  
## -9.13
```

(2) The t test between dose 1.0 and 2.0

```
t.test(dose2,dose3,paired=TRUE,var.equal=TRUE,data=ToothGrowth)
```

```
##  
## Paired t-test  
##  
## data: dose2 and dose3  
## t = -4.6046, df = 19, p-value = 0.0001934  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -9.258186 -3.471814  
## sample estimates:  
## mean of the differences  
## -6.365
```

(3) The t test between dose 0.5 and 2.0

```
t.test(dose1,dose3,paired=TRUE,var.equal=TRUE,data=ToothGrowth)
```

```
##  
## Paired t-test  
##  
## data: dose1 and dose3  
## t = -11.291, df = 19, p-value = 7.19e-10  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -18.3672 -12.6228  
## sample estimates:  
## mean of the differences  
## -15.495
```

For all the three tests, we reject the hypothesis under the 95 percent confidence interval. Hence, we can say that dose has very obvious effect on the tooth growth.

## Conclusion

Finally, we can conclude that under the assumption that the variance for all the compared groups are the same, and after applying t tests, using a 95 percent confidence interval, we can say that supp has no effect on toothgrowth and dose has large effect on it.