Study Of The Effect of Vitamin C on Tooth Growth in Guinea Pigs

Tony Lin

Overview :

Analyze the ToothGrowth data in the R datasets package

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

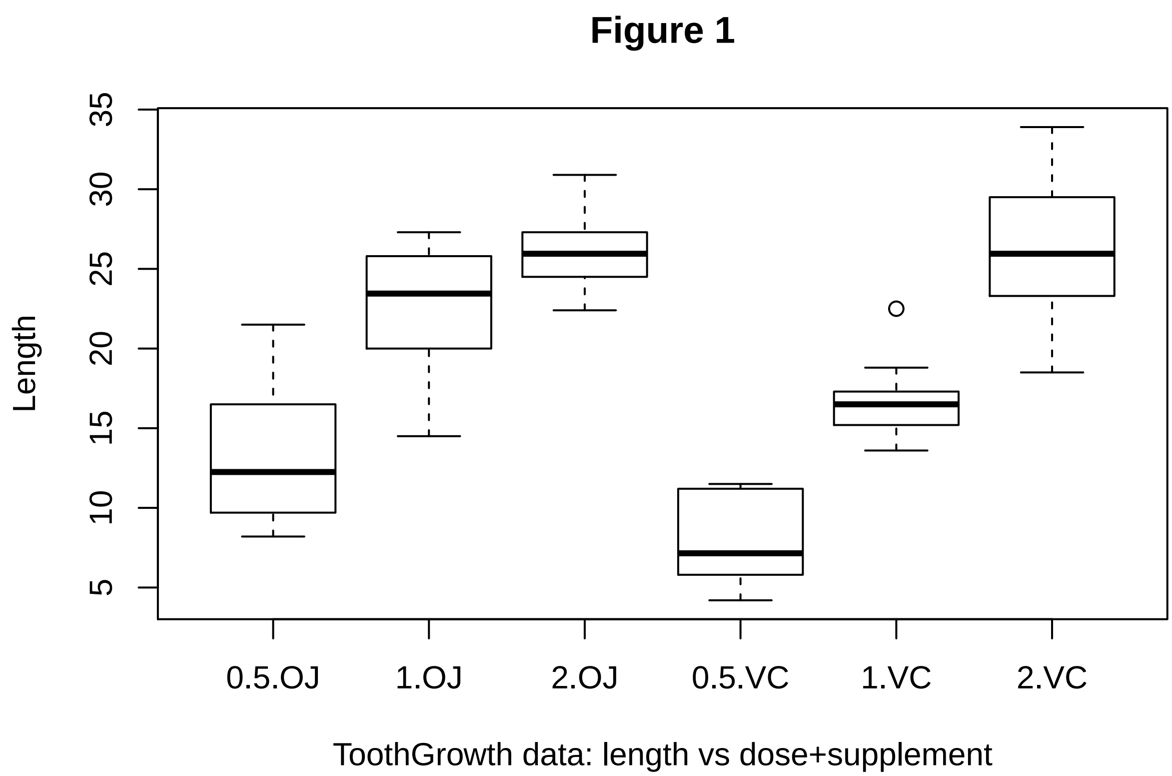
```
library(datasets)
library(lattice)
dim(ToothGrowth)
```

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

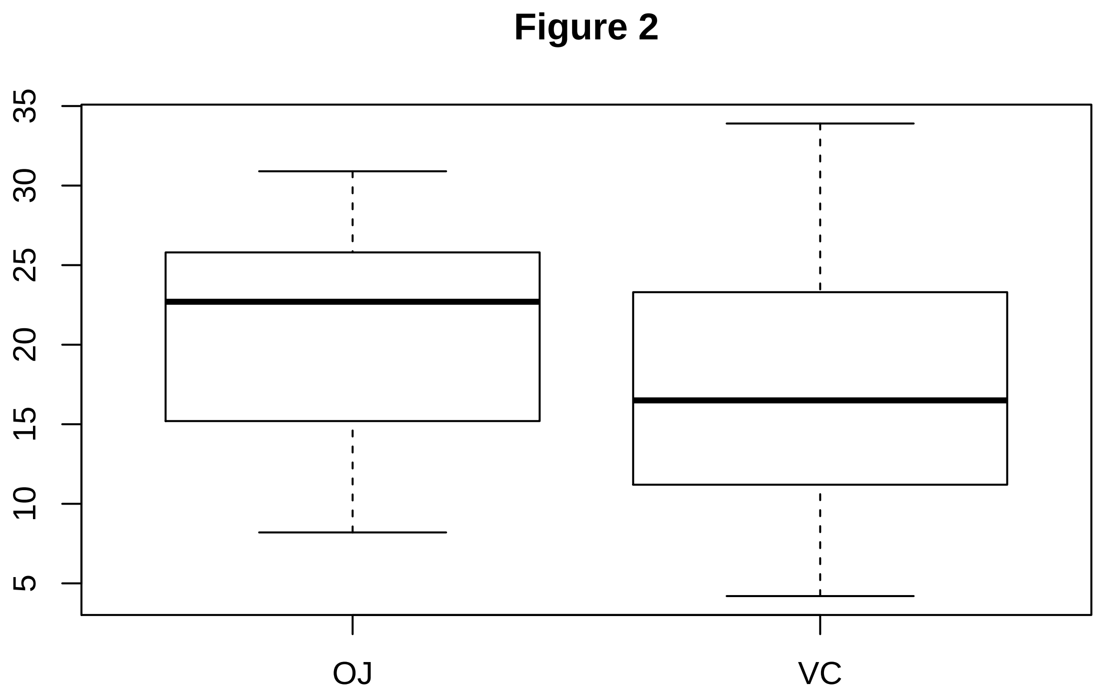
```
summary(ToothGrowth)
```

```
##       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
```

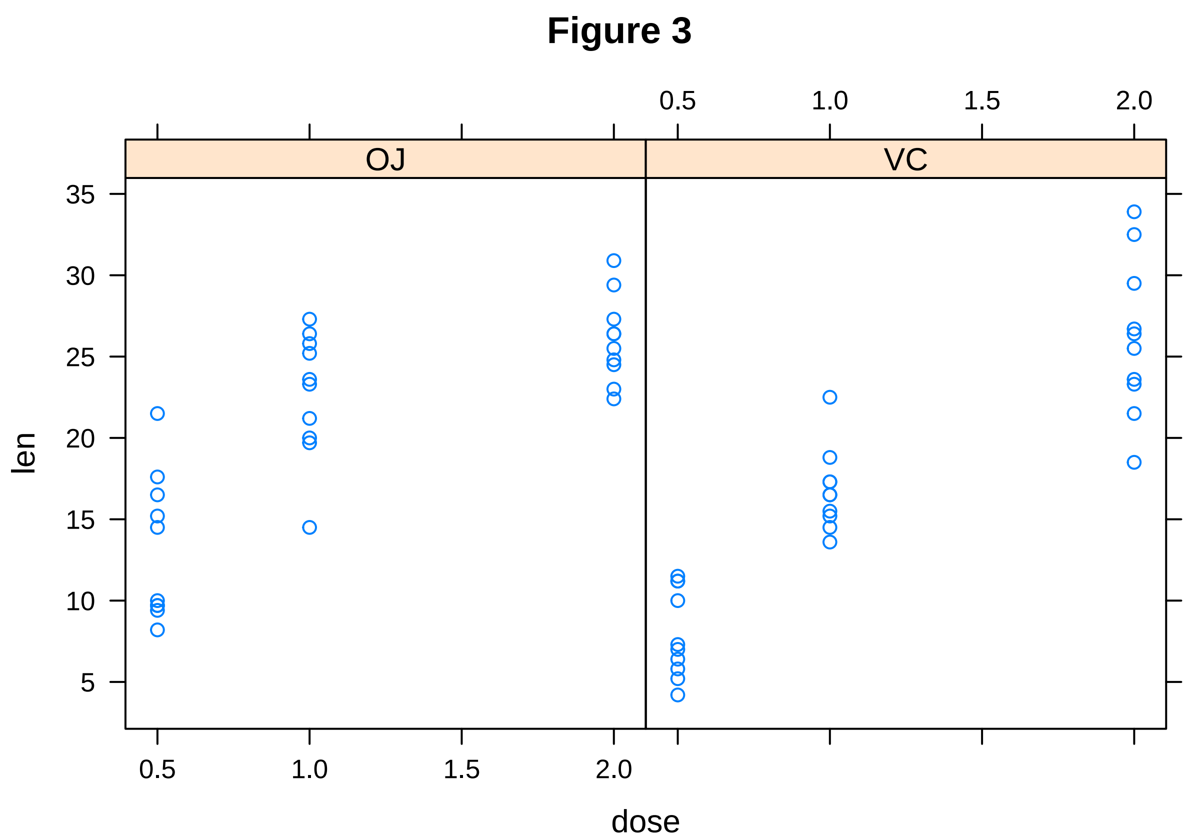
```
boxplot(len~dose+supp, data=ToothGrowth,main="Figure 1", ylab="Length", xlab =
    "ToothGrowth data: length vs dose+supplement")
```



```r
boxplot(len~supp, data=ToothGrowth, main="Figure 2")
```



```r
xyplot(len~dose|supp, data=ToothGrowth, main="Figure 3")
```



2. Basic summary of the data.

The dataset is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).

From the length distribution plot (figure 1,3), we see the trend that length increase when dose increased from 0.5 to 2.0, regardless of what supplement type is. Also, figure 2 tell us the median of length from OJ observations are larger that VC observations.

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

First hypothesis : Is there significant difference two supplement type (OJ vs VC).

H0 : OJ Mean = VC Mean. Ha : OJ Mean > VC Mean

We will do a t test for this hypothesis. Assume the data is not paired. Using one sided and 95% confidence level. Assume the vaiance is equal for both sets of observation.

```r
oj_data <- ToothGrowth[ToothGrowth$supp == "OJ",]$len
vc_data <- ToothGrowth[ToothGrowth$supp == "VC",]$len
t.test(oj_data, vc_data, var.equal=TRUE, paired=FALSE, conf.level = .95, alternative = "greater")
```

```
##
##  Two Sample t-test
##
## data:  oj_data and vc_data
## t = 1.9153, df = 58, p-value = 0.0302
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4708204       Inf
## sample estimates:
## mean of x mean of y
##  20.66333  16.96333
```

Therefor, we reject the H0 and conclude that the average length of teeth by OJ is larger than by VC with 95% confident level.

We will test the below hypotheis using VC samples only. The same test can be applied to OJ samples too.

Second Hypothesis : For VC samples, more dose will result larger teeth length.

Let's do the test by comparing the mean difference between dose=0.5 and dose=1.0. ###H0 : m05 = m10 Ha : m10 > m05 ** Let's do this one sided, non paired, same variance, confident level=.95 t test.

```r
vc_05 <- ToothGrowth[ToothGrowth$supp=="VC" & ToothGrowth$dose == .5,]$len
vc_10 <- ToothGrowth[ToothGrowth$supp=="VC" & ToothGrowth$dose == 1,]$len
t.test(vc_10, vc_05, var.equal=TRUE, paired=FALSE, conf.level = .95, alternative = "greater")
```

```
##
##  Two Sample t-test
##
## data:  vc_10 and vc_05
## t = 7.4634, df = 18, p-value = 3.246e-07
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.747719      Inf
## sample estimates:
## mean of x mean of y
##     16.77      7.98
```

p-value is extremely small and we have great confidence that H0 can be rejected.

We may do the same t test between does=1 and dose=2 samples

```
vc_20 <- ToothGrowth[ToothGrowth$supp=="VC" & ToothGrowth$dose == 2,]$len
vc_10 <- ToothGrowth[ToothGrowth$supp=="VC" & ToothGrowth$dose == 1,]$len
t.test(vc_20, vc_10, var.equal=TRUE, paired=FALSE, conf.level = .95, alternative = "greater")
```

```
##
##  Two Sample t-test
##
## data:  vc_20 and vc_10
## t = 5.4698, df = 18, p-value = 1.699e-05
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.399483      Inf
## sample estimates:
## mean of x mean of y
##     26.14     16.77
```

p-value is extremely small and we reject H0 again.

4. Conclusions and the assumptions

Assumptions :
. The observations are iid.
. The variance are constant
Conclusion :
. OJ supplement result larger teeth length on average than VC supplement.
. mean of len(dose=0.5) < mean of len(dose=1) < mean of len(does=2)