Statistical Inference Part 2

Viktor M

Thursday, April 24, 2015

My github page

Synopsis

In this research we will analyze the ToothGrowth data in the R datasets package.

Here we assign all variables and prepare data set to be used in our research

```
data(ToothGrowth)
ToothGrowth$dose<-as.factor(ToothGrowth$dose)
head(ToothGrowth,3)
```

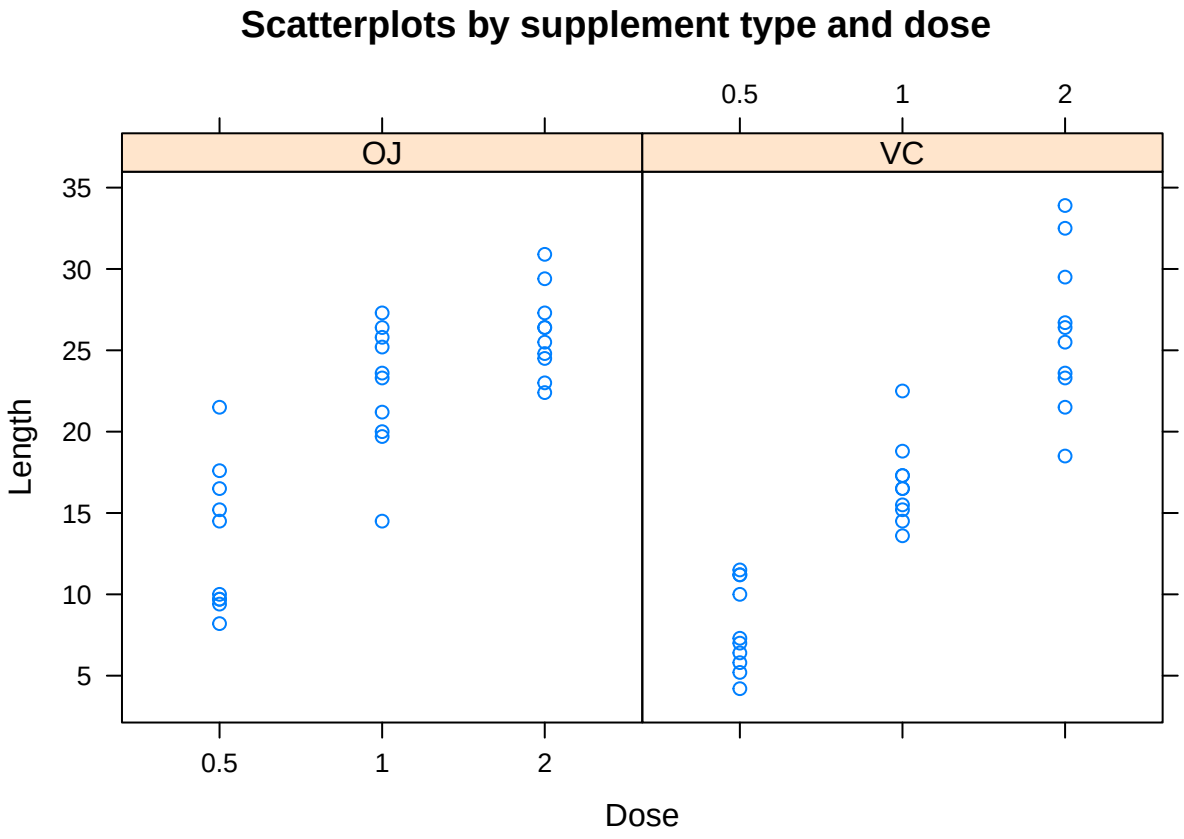
```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
```

```
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 ...
```

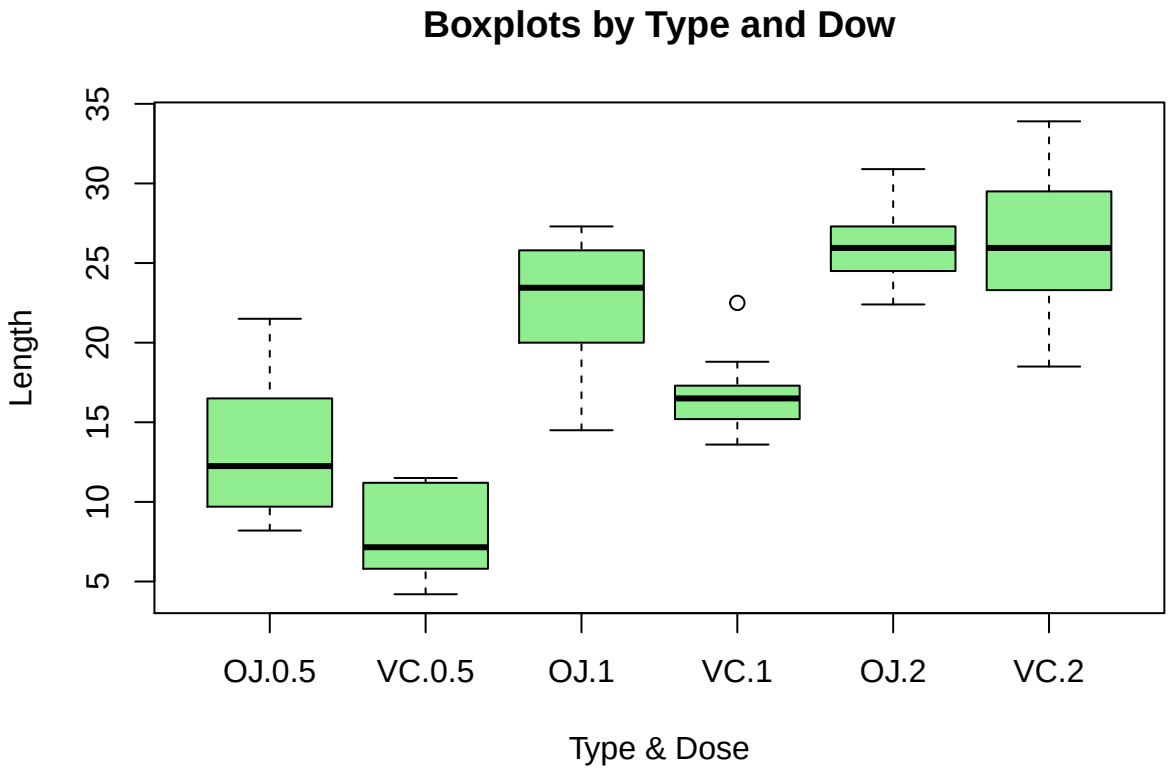
Create basic plot using lattice package

```
library(lattice)
xyplot(len~dose|supp, ToothGrowth,
       main="Scatterplots by supplement type and dose",
       ylab="Length", xlab="Dose")
```



Create boxplot

```
boxplot(len ~ supp * dose,ToothGrowth,
        col="lightgreen", ylab="Length", xlab="Type & Dose",main="Boxplots by Type and Dow")
```



As we can observe OJ promotes better tooghgrowth comparing to VC.

Show basic structrue and summary for all data

```
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
```

And show it for each of categories we researched

```
summary(ToothGrowth[ToothGrowth$supp=="OJ",])
```

```
##       len        supp     dose
##  Min.   : 8.20   OJ:30   0.5:10
##  1st Qu.:15.53   VC: 0   1  :10
##  Median :22.70           2  :10
##  Mean   :20.66
##  3rd Qu.:25.73
##  Max.   :30.90
```

And show it for each of categories we researched

```
summary(ToothGrowth[ToothGrowth$supp=="VC",])
```

```
##       len        supp     dose
##  Min.   : 4.20   OJ: 0   0.5:10
##  1st Qu.:11.20   VC:30   1  :10
##  Median :16.50           2  :10
##  Mean   :16.96
##  3rd Qu.:23.10
##  Max.   :33.90
```

Now we will compare growth dynamics by categories using Confidence intervals and Hypythesis

First for supp :

```
t.test(len ~supp,ToothGrowth[ToothGrowth$dose==0.5,],paired=FALSE, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 18, p-value = 0.005304
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.770262 8.729738
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

Using t.test with dose=0.5 we observe that OJ performs bettter then VC.

```
t.test(len ~supp,ToothGrowth[ToothGrowth$dose==1,],paired=FALSE, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 18, p-value = 0.0007807
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.840692 9.019308
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

Using t.test with dose=1 we observe that results are opposite.

```
t.test(len ~supp,ToothGrowth[ToothGrowth$dose==2,],paired=FALSE, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = -0.0461, df = 18, p-value = 0.9637
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.722999  3.562999
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

And for dose=2 there are not significant difference.

And for len

```
t.test(len ~dose,ToothGrowth[ToothGrowth$dose!=0.5,],paired=FALSE, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by dose
## 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:
##  -8.994387 -3.735613
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

```
t.test(len ~dose,ToothGrowth[ToothGrowth$dose!=2,],paired=FALSE, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by dose
## 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:
##  -11.983748  -6.276252
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

Conclusion

Assuming that each Guinea pig treated with suppliment type and dosage randomly assigned and 60 head cound can represent of all its population we ca state that: 1. Supplement type doesn't have any effect on tooth growth. 2. Increasing the dose level promotes tooth growth.