Statistical Inference part 2: Basic inferential data analysis

Aayush Garg

30 August 2017

# Instructions

This is a part 2 of final assignment,Stastistical Inference project.Our aim is to analyse and make sense of inbuilt R data "ToothGrowth".For more detailed instructions please visit: <https://www.coursera.org/learn/statistical-inference/peer/3k8j5/statistical-inference-course-project>

# Basic Inferential Data Analysis

### Loading data and some basic exploratory analysis

library(datasets); data("ToothGrowth");rbind(head(ToothGrowth,n=3 ),tail(ToothGrowth,n=3))

## len supp dose  
## 1 4.2 VC 0.5  
## 2 11.5 VC 0.5  
## 3 7.3 VC 0.5  
## 58 27.3 OJ 2.0  
## 59 29.4 OJ 2.0  
## 60 23.0 OJ 2.0

str(ToothGrowth) # Visualising structure of data

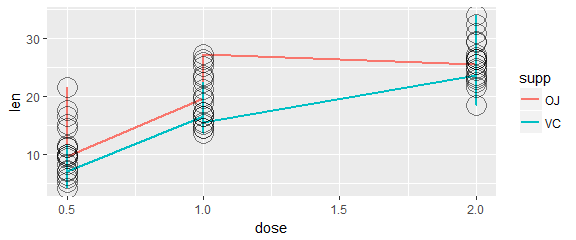
## 'data.frame': 60 obs. of 3 variables:  
## $ len : num 4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...  
## $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...  
## $ dose: num 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...

is.na(ToothGrowth$len|| ToothGrowth$supp || ToothGrowth$dose)# checking NAs

## [1] FALSE

For a better Visualisation of data and its all component , a plot is drawn between length and doses , based on supplement, for a visualised pre-decision factors:

library(ggplot2)  
ggplot(ToothGrowth, aes(x = dose, y = len, dose = factor(supp)) ) +  
geom\_line(size = 1,aes(colour = supp))+geom\_point(size =7,pch = 21,alpha = .5)



## More detailed Testing

len <- ToothGrowth$len;supp <- ToothGrowth$supp;dose <- ToothGrowth$dose #Subsetting individual columns  
# T test based on supplenets provided  
t.test(len[supp == "OJ"], len[supp == "VC"], paired = FALSE, var.equal = FALSE)[c("p.value","conf.int")]

## $p.value  
## [1] 0.06063451  
##   
## $conf.int  
## [1] -0.1710156 7.5710156  
## attr(,"conf.level")  
## [1] 0.95

# T test based on doses  
t.test(len[dose == 1], len[dose == 2], paired = FALSE, var.equal = TRUE)[c("p.value","conf.int")]

## $p.value  
## [1] 1.810829e-05  
##   
## $conf.int  
## [1] -8.994387 -3.735613  
## attr(,"conf.level")  
## [1] 0.95

By these results it is established that while different supplement provided doesn't afftect overall growth (CI contains 0, but p value 0.06 almost near alpha=0.05 value),but further exploration would be required.On the other hand, increase in doses have definite positive effects.

oj.test<-ToothGrowth[ToothGrowth$supp == "OJ",]; #subsetting based on OJ and testing for doses  
oj.0.5<-oj.test[oj.test$dose==0.5,]$len;oj.1<-oj.test[oj.test$dose==1,]$len  
t.test(oj.0.5,oj.1 , paired = FALSE, var.equal = TRUE)[c("p.value","conf.int")]#test p value and conf interval for doses

## $p.value  
## [1] 8.357559e-05  
##   
## $conf.int  
## [1] -13.410814 -5.529186  
## attr(,"conf.level")  
## [1] 0.95

vc.test<-ToothGrowth[ToothGrowth$supp == "VC",]# #subsetting based on VC and testing for doses  
vc.x<-vc.test[vc.test$dose==0.5,]$len;vc.y<-vc.test[vc.test$dose==1,]$len  
t.test(vc.x,vc.y , paired = FALSE, var.equal = TRUE)[c("p.value","conf.int")] #test p value and conf interval for doses

## $p.value  
## [1] 6.492265e-07  
##   
## $conf.int  
## [1] -11.264346 -6.315654  
## attr(,"conf.level")  
## [1] 0.95

Here doses comparision for(0.5,1), for OJ as well as VC, proves growth of tooth in length with higher dose amount.Although rate of increment of tooth growth for both supplement differs substantially.

#comparision of tooth growth for dose=0.5 in OJ and VC  
 vc.0.5<-ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose == 0.5,]  
 oj.0.5<-ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 0.5,]  
t.test(oj.0.5$len,vc.0.5$len , paired = FALSE, var.equal=F)[c("p.value","conf.int")]

## $p.value  
## [1] 0.006358607  
##   
## $conf.int  
## [1] 1.719057 8.780943  
## attr(,"conf.level")  
## [1] 0.95

#comparision of tooth growth for dose=2 in OJ and VC  
vc.2<-ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose == 2,]  
 oj.2<-ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 2,]  
t.test(oj.2$len,vc.2$len , paired = FALSE, var.equal=F)[c("p.value","conf.int")]

## $p.value  
## [1] 0.9638516  
##   
## $conf.int  
## [1] -3.79807 3.63807  
## attr(,"conf.level")  
## [1] 0.95

For different supplements and same dose, OJ is better for small doses,0.5 & 1, but with increase in dose difference become narrower.For dose=2, both have almost same effects and results are inconclusive.

# Conclusion

By using Confidence Intervals for Tooth Growth by supp and dose, we found out that the higher the dose, the longer the tooth grows in length. Comparing similar amounts of dosage for OJ (Orange Juice) and VC (Vitamin C), we notice that OJ is more effective with regards to Tooth Growth than VC in lower doses such as 0.5 and 1.0. With regards to 2.0, it's inconclusive or hard to tell.