Basic Inferential Data Analysis

Victor

September 21, 2016

## ToothGrowth analysis

In this analysis we are looking to make some inference based on the ToothGrowth data set provided in the datasets package. The ToothGrowth data shows the effect of Vitamin C on tooth growth. We observe to method to deliver the vitamin C (supplement 1: Orange Juice and supplement 2: ascorbic acid). The study was made at various level of delivered Vitamin C (.5,1 and 2 mg). We observe in our analysis that at higher levels (2mg) of supplement the delivery method doesn't have an impact. At lower levels (.5 and 1mg) we observe a significant difference between the impact of Vitamin C from orange juice versus ascorbic acid.

## The data and exploratory analysis

We want to understand more about our data and its structure. To do so we use the str() and summary() commands in r.

library(datasets)  
TG = ToothGrowth  
str(TG)

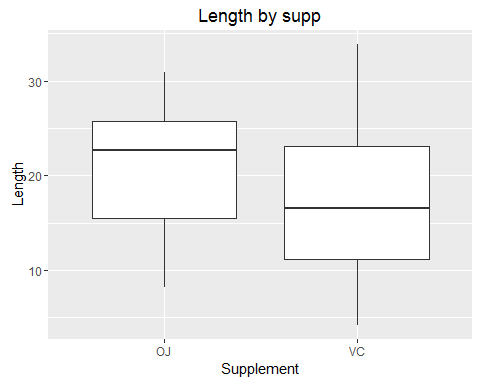
## 'data.frame': 60 obs. of 3 variables:  
## $ len : num 4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...  
## $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...  
## $ dose: num 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...

summary(TG)

## 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

We can see that this data set is made up of 60 observations of 3 variables. The data structure is 2 numerical variables and one factor variable. we see from the summary of the data set that the factor level are: OJ and AC, the first numerical variable len ranges from 4.2 to 33.9 and the second numerical variable (dose) ranges from 0.5 to 2.

In this first part of the analysis we are going to see if there is a significant difference in the mean length growth for subjects taking orange juice versus subjects taking ascorbic acid



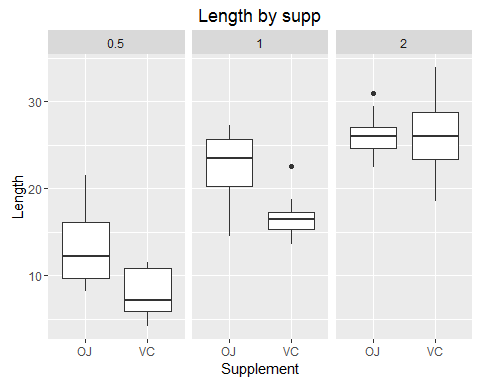
From the plot it is unclear if there is a significant difference is mean between the 2 groups. We will proceed to perform a t-test to see if the observe difference is significant.

##   
## Welch Two Sample t-test  
##   
## data: len by supp  
## 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 in group OJ mean in group VC   
## 20.66333 16.96333

Our test fails to reject the null hypothesis. Therefore we can say there is not enough evidence to show that the mean tooth lengths for subjects taking orange juice is different to subjects taking ascorbic acid. Although there isn't enough evidence to dissociate the 2 groups, we found our p-value to be 0.06063 which is really close to the rejection region at alpha = 0.95.

We then look at the sub-groupings between subjects taking orange juice and ascorbic acid. these 2 groups can be divide into subjects taknig 0.5,1 and 2mg of vitamin C through one of the two delivery method.

Our preliminary grahpical analysis shows the break down of each dosage levels by subject groups:



At first glanced it looks like at dosage level 0.5 and 1 mg of Vitamin C the delivery method has a significant impact on tooth growth. At 2mg it is unclear that the delivery impacts tooth growth.

We will perform t-tests in order to see if there is enough evidence to back-up our claims from the graphical analysis.

## 0.5mg Vitamin C 1.0mg Vitamin C 2.0mg Vitamin C  
## t-stat 3.169732784 4.032769634 -0.0461361  
## p-val 0.006358607 0.001038376 0.9638516  
## CI (lower bound) 1.719057271 2.802148249 -3.7980705  
## CI (upper bound) 8.780942729 9.057851751 3.6380705  
## mean (OJ) 13.230000000 22.700000000 26.0600000  
## mean (VC) 7.980000000 16.770000000 26.1400000

## Conclusion

We see from the table that at 0.5 and 1 mg dosage, the delivery method significantly impacts the tooth growth rate. We can make this claim because the p-values are .006 and .001 respectively. We then reject the null hypothesis and conclude that there is a significant difference on tooth growth based on the delivery method.

We also can observe that at 2 mg dosage, the delivery method doesn't significantly impact the tooth growth. This means that at a certain amount of Vitamin C intake tooth growth will not be impacted by the delivery method. We can make this claim because the p-value is .964 and thus we fail to reject the null, stating that there is not enough evidence to show that there is a significant difference in tooth growth from taking vitamin C through orange juice or ascorbic acid at 2 mg of dosage.