Statistical Inference project part 2

Abdelrahman Elsehaily

September 29, 2017

1.Introduction

The goal of this document is to provide basic summary of "ToothGrowth" dataset in R and usin t-test and confidence intervals try to answer two main questions:
1.Is there significant difference in the lenght between the two delivery methods?
2.Is there significant difference in the lenght between different doses?

2.Analysis

Loading ggplot for plotting the data

```
library("ggplot2")
```

```
## Warning: package 'ggplot2' was built under R version 3.4.1
```

2.1.Summary of the data

The column of the 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
```

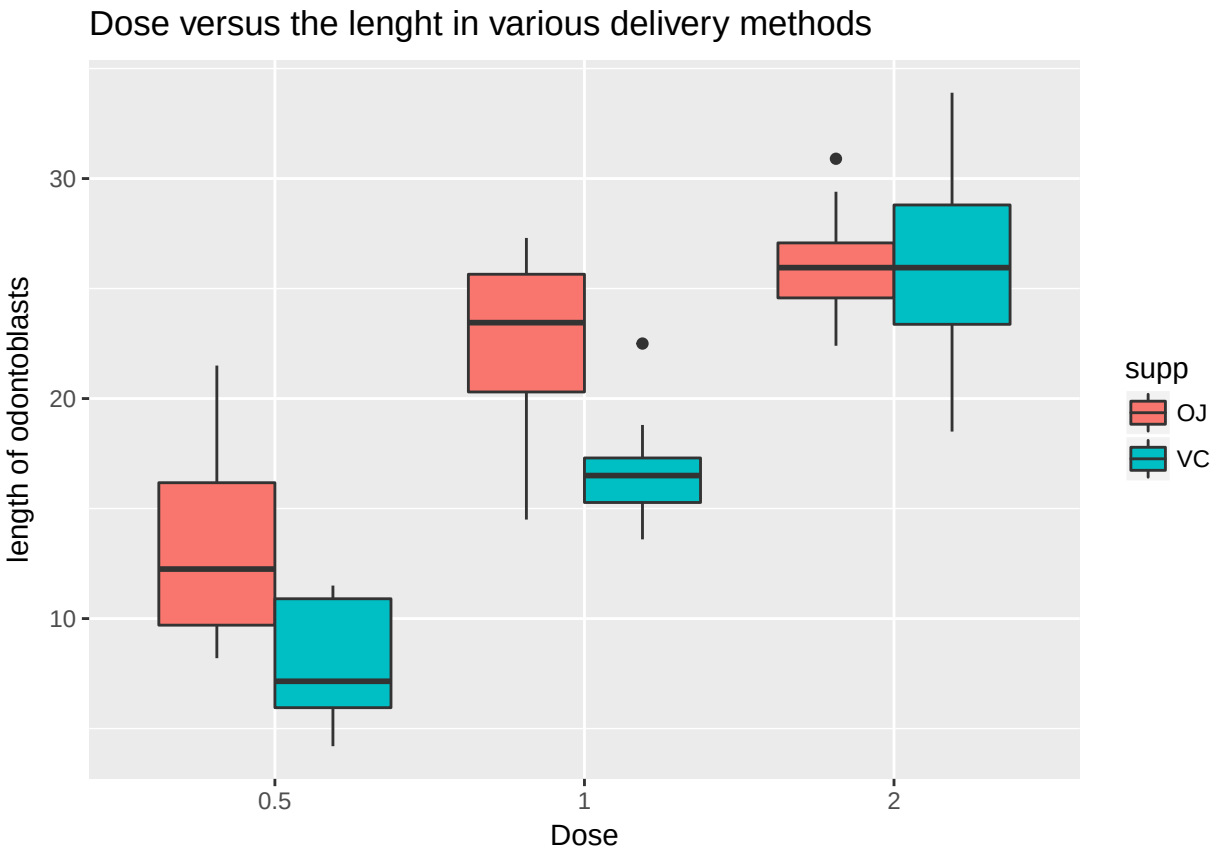
The first rows of the data set

```
head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

Using boxplot I plotted the dataset as y-axis is the lenght of odontoblasts and x-axis is the doses.

```
ggplot(data=ToothGrowth,aes(x=as.factor(dose),y=len,fill=supp))+geom_boxplot()+ggtitle("Dose versus the
```



The plot shows that the length is higher when using orange juice(OJ) as the delivery method that when using th ascorbic acid (VC), and obviously the lenght increase when the dose increase.

2.2.what is the effect of supply method(orange juice vs ascorbic acid)?

We wanna see if there is significant difference between the two delivery methods(supp) so our hypothesis I think is the following:
H0: M0-Ma=0
Ha: M0-Ma!=0
as M0,Ma are the means of the deliveries method.

```
t.test(len~supp,data=ToothGrowth)
```

```
##
##  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
```

The p-value is not significant as the p-value >0.025 (considering the two sided) We fail to reject the null hypothesis H0.

2.3.Does the dose affet the length?

We wanna see if there is significant difference between the different doses so and our hypothesis I think is the following(here I impelemented the one side test as we want to know if the bigger does gives higher lenght or the same):
H0: M0-Ma=0
Ha: M0-Ma>0
as M0,Ma are the means of lenght given two doses. ### 0.5 dose versus 1.0 dose

```
t.test(len~dose,data=ToothGrowth[ToothGrowth$dose ==0.5 | ToothGrowth$dose ==1,])
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

as the p-value is lower thean 0.05, so the there is siginficant difference of the length. ### 1.0 dose versus 2.0 dose

```
t.test(len~dose,data=ToothGrowth[ToothGrowth$dose ==1 | ToothGrowth$dose ==2,])
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

as the p-value is lower thean 0.05, so the there is siginficant difference of the length so we reject the null hypothesis.

3.conslusion

There's no significant differnce in the lenght of odontoblasts with the different delivery methods.
Ther's significance increase in the length as the dose increase.