

## ANALYSIS OF TOOTH GROWTH DATA

By: Piyush Neupane

### SYNOPSIS

ToothGrowth data is part of R datasets package. It lists growth of teeth in test subjects based on two types of supplements, and 3 types of dosage for each supplement.

The purpose of this analysis is to perform confidence interval/hypothesis test to check whether teeth growth is different for different supplements and different dosages.

```
# set libraries and datasets
library(datasets)
library(ggplot2)
```

### EXPLORATORY ANALYSIS OF DATA

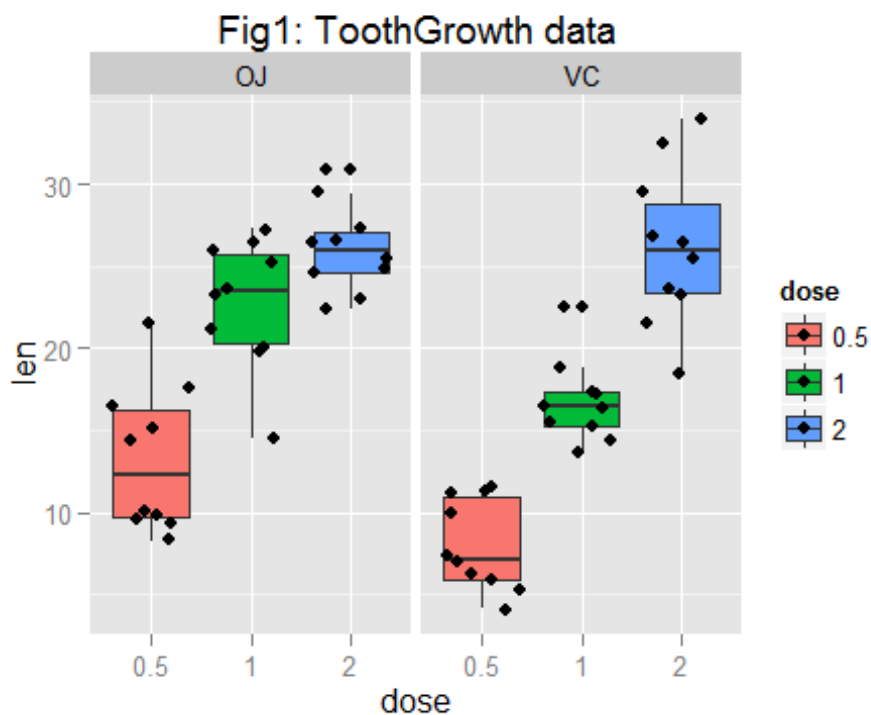
```
head(ToothGrowth)
```

```
##      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
```

```
# Change dose to factor
TG1 = ToothGrowth
TG1$dose = factor(TG1$dose)
summary(TG1)
```

```
##      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
```

```
# exploratory plots
ggplot(TG1, aes(x=dose, y=len, fill=dose))+geom_boxplot() +
geom_jitter()+facet_grid(.~supp)+ggtitle("Fig1: ToothGrowth data")
```



The plots shows that teeth growth with dose 0.5 and 1 for supplement OJ is higher than with supplement VC. However, with dose 2, both supplements have very high growth (compared to lower doses)

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

1a) Compare teeth growth between two supplements (regardless of the dosage) assuming equal variance and Paired test

```
supp1a <- t.test(len~supp, data=TG1, paired=T, var.equal=T)
supp1aTtest <- data.frame("p" = supp1a$p.value,
                          "95pctConfLower" = supp1a$conf[1],
                          "95pctConfupper" = supp1a$conf[2],
                          row.names = "1a) Equal Var/Paired Test"
)
```

1b) Compare teeth growth between two supplements (regardless of the dosage) assuming unequal variance and Paired test

```
supp1b <- t.test(len~supp, data=TG1, paired=T, var.equal=F)
supp1bTtest <- data.frame("p" = supp1b$p.value,
                          "95pctConfLower" = supp1b$conf[1],
                          "95pctConfupper" = supp1b$conf[2],
                          row.names = "1b) Unequal Var/Paired Test"
)
```

2a) Compare teeth growth between two supplements (regardless of the dosage) assuming equal variance and UnPaired test

```
supp2a <- t.test(len~supp, data=TG1, paired=F, var.equal=T)
supp2aTtest <- data.frame("p" = supp2a$p.value,
                           "95pctConfLower" = supp2a$conf[1],
                           "95pctConfupper" = supp2a$conf[2],
                           row.names      ="2a) Equal Var/UnPaired Test")
```

**2b) Compare teeth growth between two supplements (regardless of the dosage) assuming unequal variance and UnPaired test**

```
supp2b <- t.test(len~supp, data=TG1, paired=F, var.equal=F)
supp2bTtest <- data.frame("p" = supp2b$p.value,
                           "95pctConfLower" = supp2b$conf[1],
                           "95pctConfupper" = supp2b$conf[2],
                           row.names = "2b) Unequal Var/UnPaired Test")
)
```

```
# Combine all test results
```

```
rbind(supp1aTtest,supp1bTtest,supp2aTtest,supp2bTtest)
```

##		p	X95pctConfLower	X95pctConfupper
##	1a) Equal Var/Paired Test	0.002549842	1.4086586	5.991341
##	1b) Unequal Var/Paired Test	0.002549842	1.4086586	5.991341
##	2a) Equal Var/UnPaired Test	0.060393371	-0.1670064	7.567006
##	2b) Unequal Var/UnPaired Test	0.060634508	-0.1710156	7.571016

It seems that assuming unpaired Test, we fail to reject the null hypothesis because  $p\text{-value} > .05$  at the default significance level of .05.

### Perform hypothesis among each dose levels (Unpaired and Variance not same)

### Dose 0.5 vs 1

```
dose05_1<- subset(TG1, dose %in% c(0.5,1))  
t.dose05_1<- t.test(len~dose, data=dose05_1, paired=F, var.equal=F)  
dose05_1.Ttest <- data.frame("p" = t.dose05_1$p.value,  
                             "95pctConfLower" = t.dose05_1$conf[1],  
                             "95pctConfupper" = t.dose05_1$conf[2],  
                             row.names ="1) Dose .5 vs 1")
```

### Dose 0.5 vs 2

[illegible]

```

"95pctConfupper" = t.dose05_2$conf[2],
row.names        ="2) Dose .5 vs 2"
)

```

### Dose 1 vs 2

```

dose1_2<- subset(TG1, TG1$dose %in% c(1,2))
t.dose1_2<- t.test(len~dose, data=dose1_2, paired=F, var.equal=F)
dose1_2.Ttest <- data.frame("p"           = t.dose1_2$p.value,
                           "95pctConfLower" = t.dose1_2$conf[1],
                           "95pctConfupper" = t.dose1_2$conf[2],
                           row.names      ="3) Dose 1 vs 2"
)

rbind(dose05_1.Ttest, dose05_2.Ttest,dose1_2.Ttest)

```

```

##                p X95pctConfLower X95pctConfupper
## 1) Dose .5 vs 1 1.268301e-07      -11.983781      -6.276219
## 2) Dose .5 vs 2 4.397525e-14      -18.156167     -12.833833
## 3) Dose 1 vs 2 1.906430e-05       -8.996481      -3.733519

```

Looking at the 95% confidence intervals for different pairs of dosages, we see that none of them include 0. Also, the p-value is much smaller than the chosen significance level of .05. So, we can reject the null hypothesis.

## ASSUMPTIONS

For the comparison involving two supplements (OJ and VC), we looked at both Paired and Unpaired tests, with equal and unequal variances. However, the assumption is that the test is actually Unpaired with unequal variance.

For the comparison involving 3 different dosages (0.5mg, 1mg and 2mg), we assumed Unpaired test and unequal variances for each pair.

Also, the significance level was set to the default of .05.

## CONCLUSIONS

From the boxplot, we can see that as the amount of dosage increases for both supplements, the length of teeth increases.

**The confidence interval for Unpaired test of supplement includes 0. Also, the p-value >.05. So we conclude that we do not have enough evidence to reject the null hypothesis (Effect of supplement on the length of teeth is not significant).**

**However, the confidence interval for Unpaired test of dosages does not include 0. Also, the p-value <.05. So we conclude that we HAVE enough evidence to reject the null hypothesis, and conclude that as we increase the dosage, the length of teeth increases.**