**Tooth Growth Data Analysis**

***Author: Yves E.***

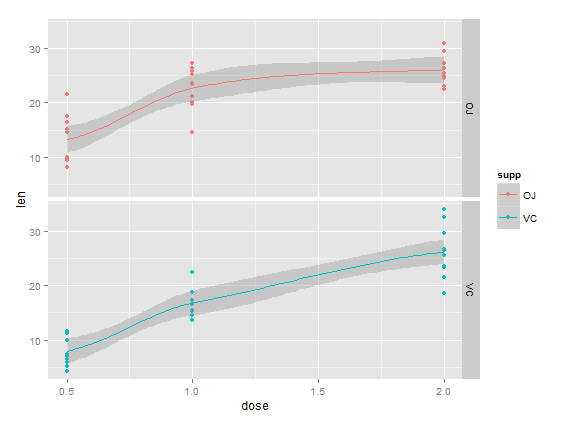
This report presents the analysis performed on Tooth growth dataset. This data set shows the effect of Vitamin C on Tooth Growth in Guinea Pigs.

**Analysis**

# Load data and perform a basic exploratory data analyses

question1 <- function(){  
 **library**(datasets)  
 d<- ToothGrowth  
 **library**(ggplot2)  
 g <- **qplot**(dose,len,data=d, facets=supp~.,

geom = **c**("point","smooth"),color=supp)  
 g  
}



**Comment:** On the basis of this graphic we can assert that the mean of teeth length increases with the Vitamine C dose. The graphics shows that each supplement (supp) has 3 groups by dose (0.5, 1 and 2):

# A basic summary of data:

> 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

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

1. T-tests for mean (of the length) difference by supplement type :

> t.test(len~supp, data=ToothGrowth, paired=FALSE)

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

**Comment**: According to the T-test above, **p-value=0.06** is **greater** than **α =0.05** (α for confidence interval of 95%) that means that we **fail to reject the null hypothesis**. The confidence interval **[-0.17, 7.60]** contains the 0, there is no effect on changing the supplement type.

1. T-tests for mean (of the length) difference by dose

On the basis of data exploratory performed on the first question, we have 3 subsets per dose. We are going to compare, for each dose :

* Split dataset by dose :

> s <- split(ToothGrowth,ToothGrowth$dose)

> summary(s)

Length Class Mode

0.5 3 data.frame list

1 3 data.frame list

2 3 data.frame list

* Compare OJ and VC for the first subset : dose = 0.5

> d0.5 <- s$`0.5`

> t.test(len~supp, data=d0.5, paired=FALSE)

Welch Two Sample t-test

data: len by supp

t = 3.1697, df = 14.969, p-value = 0.006359

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

1.719057 8.780943

sample estimates:

mean in group OJ mean in group VC

13.23 7.98

**Comment**: The p-value=0.006 **is less** than α=0.05 then we **reject the null hypothesis** for the Dose = 0.5.

* Compare OJ and VC for dose = 1

> d1 <- s$`1`

> t.test(len~supp, data=d1, paired=FALSE)

Welch Two Sample t-test

data: len by supp

t = 4.0328, df = 15.358, p-value = 0.001038

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

2.802148 9.057852

sample estimates:

mean in group OJ mean in group VC

22.70 16.77

**Comment**: The p-value=0.001 **is less** than α=0.05 then we **reject the null hypothesis** for the Dose = 1.

* Compare OJ and VC for dose = 2

> d2 <- s$`2`

> t.test(len~supp, data=d2, paired=FALSE)

Welch Two Sample t-test

data: len by supp

t = -0.046136, df = 14.04, p-value = 0.9639

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-3.79807 3.63807

sample estimates:

mean in group OJ mean in group VC

26.06 26.14

**Comment**: The p-value=0.96 **is greater** than α=0.05 then we **fail to** **reject the null hypothesis** for the Dose = 2.

# General conclusion

Based on different T-tests above, we can say that:

1. There is a limit effect on teeth length when changing the supplement in the whole sample of Guinea pigs.
2. The tests show that there are effects when the dose of vitamine C is 0.5 or 1. It’s not necessary to increase the Vitamine C dose to reach 2 because there no effect.