# Statistical Inference Project

Marcelo Nobrega

The project consists of two parts:

1. Basic inferential data analysis.

# Part 2: Basic Inferential Data Analysis Instructions

## 1. Provide a basic summary of the data.

We're going to analyze the ToothGrowth data in the R datasets package.

Loading libraries, the ToothGrowth data and performing some basic exploratory data analyses.

library(ggplot2)  
data("ToothGrowth")  
# Some basic exploratory data analyses  
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

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

tail(ToothGrowth)

## len supp dose  
## 55 24.8 OJ 2  
## 56 30.9 OJ 2  
## 57 26.4 OJ 2  
## 58 27.3 OJ 2  
## 59 29.4 OJ 2  
## 60 23.0 OJ 2

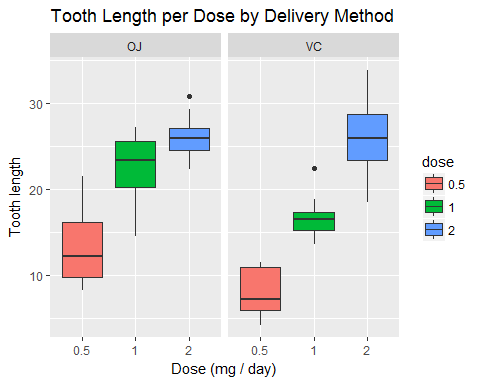
str(ToothGrowth)

## '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 ...

The ToothGrowth data consists in the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, (orange juice or ascorbic acid (a form of vitamin C and coded as VC).

60 observations and 3 variables. len: Tooth length Supp: Supplement type (VC or OJ) dose: Dose in miligrams/day

ToothGrowth$dose <- as.factor(ToothGrowth$dose)  
ggplot(ToothGrowth, aes(x = dose, y = len))+  
 geom\_boxplot(aes(fill = dose))+  
 facet\_grid(~ supp)+  
 xlab("Dose (mg / day)")+  
 ylab("Tooth length")+  
 ggtitle("Tooth Length per Dose by Delivery Method")



The larger is the dosage for both delivery methods longre is th tooth.

## 2. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)

unique(ToothGrowth$dose)

## [1] 0.5 1 2   
## Levels: 0.5 1 2

There are 3 levels which can be understood as 3 doseage groups, 0.1, 1.0 and 2.0 mg / day. I will compare the tooth growth by each doseage group and delivery method.

Comparing tooth groth by Delivery method using t-test.

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

##   
## 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

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

##   
## 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

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

##   
## 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

## 3. State your conclusions and the assumptions needed for your conclusions.

It was assumed that the sample data was paired.

Conclusions: Dosage equal to 0.5 mg/day: OJ had higher mean tooth length than VC, 13.23 vs. 7.98. P-value is 0.006 (< 0.05), so for the same dosage the different delivery methods have effect on the tooth length.

Dosage equal to 1.0 mg/day: OJ had higher mean tooth length than VC, 22.70 vs. 16.77. P-value is 0.001 (< 0.05). The different delivery methods have effect on the tooth length.

Dosage equal to 2.0 mg/day: Both delivery methods, OJ and VC, had similar average tooth length, 26.06 vs. 26.14. P-value is 0.963 (> 0.05), and the confidence interval conains zero, so the null hypothesis can't be rejected and the different delivery methods have no effect on the tooth length.