

Statistical Inference Assignment, Part B

Fabio Fornari

December 26, 2015

Now in the second portion of the class, we're going to analyze the ToothGrowth data in the R datasets package. It is an analysis of the publically available ToothGrowth data set to study the effect of Vitamin C on tooth growth in Guinea pigs.

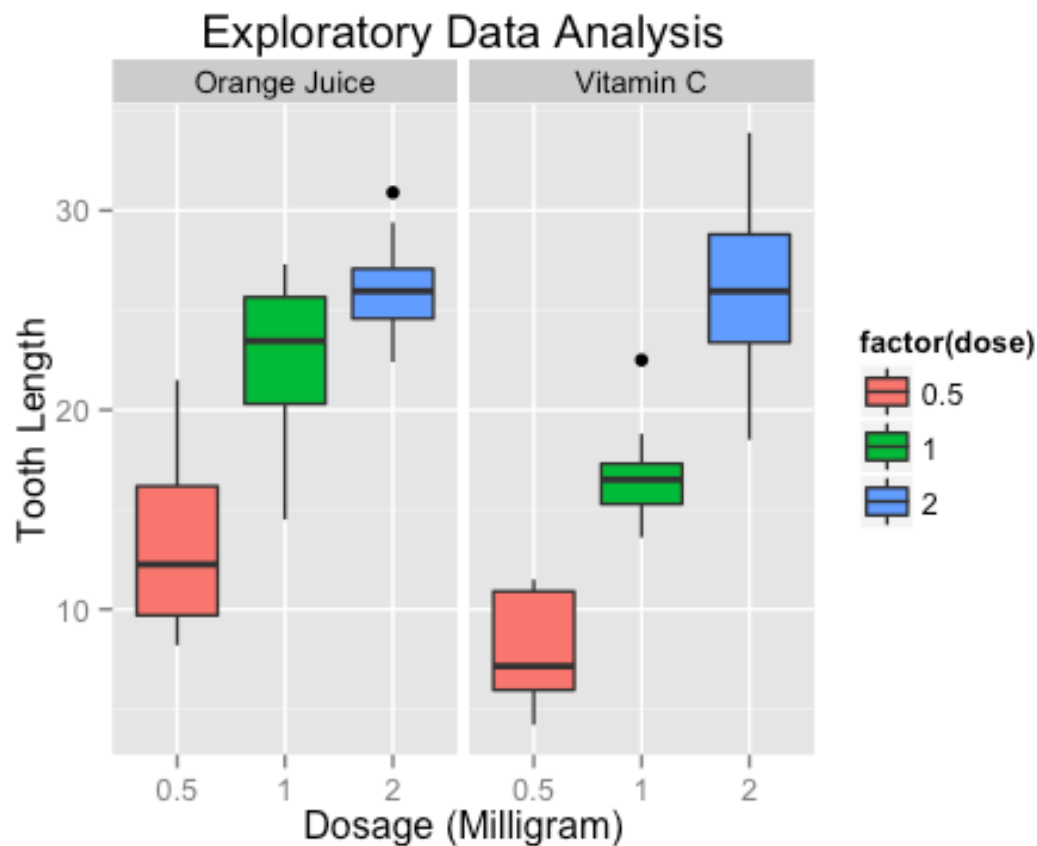
Questions 1

Load the ToothGrowth data and perform some basic exploratory data analyses

```
library(ggplot2)
library(datasets)
library(xtable)
data(ToothGrowth)
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
supplements <- c('Orange Juice', 'Vitamin C')
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: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 1 ...

g <- ggplot(ToothGrowth, aes(x=factor(dose), y=len, fill=factor(dose)))
g + geom_boxplot(notch=F) + facet_grid(.~supp, labeller=function(var, value)
supplements[value]) +
  scale_x_discrete('Dosage (Milligram)') +
  scale_y_continuous('Tooth Length') +
  ggtitle('Exploratory Data Analysis')
```



Questions 2

Provide a basic summary of the data.

```
summary(ToothGrowth)
```

```
##      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
```

Question 3

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

```
supp.t1 <- t.test(len~supp, paired=F, var.equal=T, data=ToothGrowth)
supp.t2 <- t.test(len~supp, paired=F, var.equal=F, data=ToothGrowth)
supp.result <- data.frame("p-value"=c(supp.t1$p.value, supp.t2$p.value),
                          "Conf-Low"=c(supp.t1$conf[1],supp.t2$conf[1]),
                          "Conf-High"=c(supp.t1$conf[2],supp.t2$conf[2]),
                          row.names=c("Equal Var","Unequal Var"))
supp.result
```

##		p.value	Conf.Low	Conf.High
##	Equal Var	0.06039337	-0.1670064	7.567006
##	Unequal Var	0.06063451	-0.1710156	7.571016

Question 4

State your conclusions and the assumptions needed for your conclusions.

In all cases the t distribution was used, the subjects were assumed to be independent, there was assumed to be no interaction between the supplementation regime and the dosage.

The hypothesis tested was that there is a difference in efficacy between the Orange Juice and Vitamin C supplement regimes. In this case the null hypothesis was that there was no difference between the two regimes. In the results the confidence interval includes the value of zero meaning it isn't possible to rule out the null hypothesis that there was no difference between the two regimes.

The 2mg dose has larger impact on tooth growth than 1mg and 0.5mg, while 1mg dose has more impact than 0.5mg dose. So there is a difference in the tooth growth while the doses are larger.