Statistical Inference: Project 1 Part 2

JR Sturino

November 17, 2015

Overview

Using the ToothGrow data set perform basic data analysis, summarize the data, use condifence intervals to compare tooth growth by supplement and and dose, and give a conclusion about the data in general.

Initial Analysis of Data

```
library(datasets)
library(ggplot2)
library(dplyr)

toothData <- ToothGrowth
head(toothData)

##     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
```

Look at the general summary of the data

```
summary(toothData)

##       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
```

Note that there are only two supplements in this data set, OJ and VC.

Lets quickly examine the data inside of the dataset by looking at the distribution across supplments of tooth length and dosage:

```
ggplot(toothData, aes(x=as.factor(dose), y=len, color=as.factor(dose))) +
  geom_boxplot() + xlab("Supplement") + ylab("Tooth Length") +
facet_grid(.~supp)
```
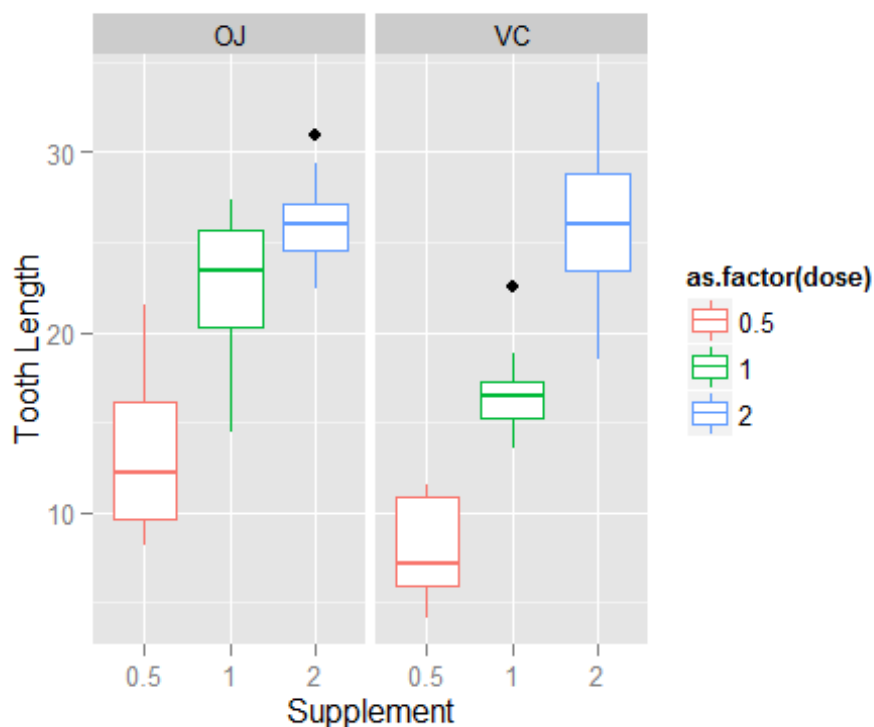
Summary

As we look further into the data, we can see that tooth length is greater for OJ than VC in the dosages of 0.5 and 1.0, but as the dosage reaches 2.0 tooth length looks very similar for the two supplements.

```r
#only use the complete cases of the data
isValidData <- complete.cases(toothData)
toothData <- toothData[isValidData,]
head(toothData)

##     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

toothInfo <- group_by(toothData, dose, supp) %>%
  summarize(toothTotal = sum(len), toothMean = mean(len))

ggplot(toothInfo, aes(x=as.factor(dose), y=toothMean, group=supp, fill=supp))
+
  geom_bar(stat="identity") +xlab("Dosage") + ylab("Average Tooth Length") +
  ggtitle("Tooth Length by Dosage and Supplement Type") +
```
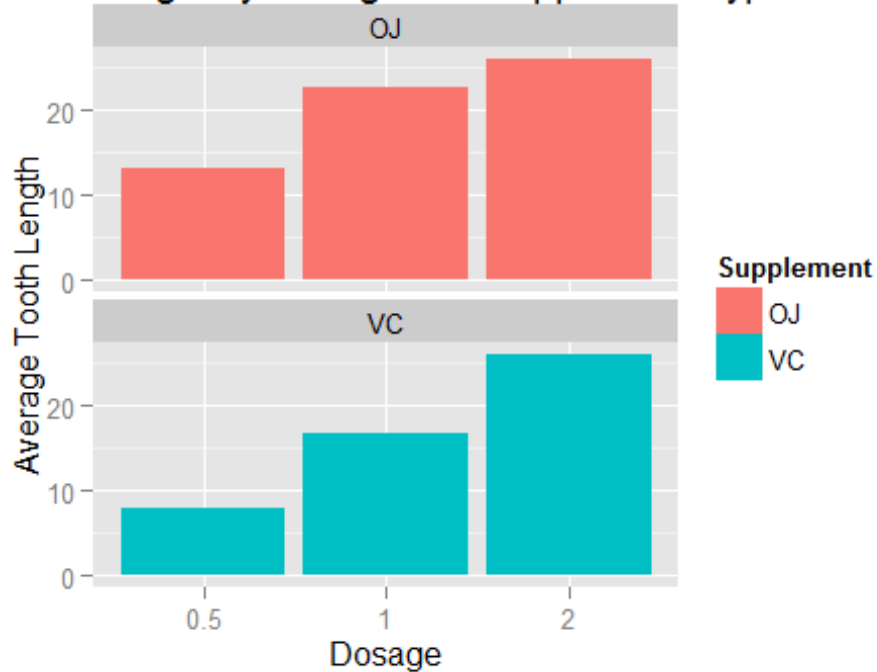
```
    facet_wrap(~supp, ncol=1) +
    guides(fill=guide_legend(title="Supplement"))
```

Tooth Length by Dosage and Supplement Type

Confidence intervals

0.5mg Dosage Test

```
dose.5 <- t.test(len ~ supp, var.equal=T,
data=toothData[toothData$dose==.5,])
ci.5 <-dose.5$conf[1:2]
pv.5 <- dose.5$p.value
```

The Confidence Interval of the 0.5mg Dosage

```
## [1] 1.770262 8.729738
```

The pvalue of the 0.5mg Dosage

```
## [1] 0.005303661
```

1mg Dosage Test

```
dose1 <- t.test(len ~ supp, var.equal=T, data=toothData[toothData$dose==1,])
ci1 <- dose1$conf[1:2]
pv1 <- dose1$p.value
```

The Confidence Interval of the 1mg Dosage

```
## [1] 2.840692 9.019308
```

The pvalue of the 1mg Dosage

```
## [1] 0.0007807262
```

2mg Dosage Test

```
dose2 <- t.test(len ~ supp, var.equal=T, data=toothData[toothData$dose==2,])
ci2 <- dose2$conf[1:2]
pv2 <- dose2$p.value
```

The Confidence Interval of the 2mg Dosage

```
## [1] -3.722999  3.562999
```

The pvalue of the 2mg Dosage

```
## [1] 0.9637098
```

Conclusion

After looking at the data provided multiple conclusions can be drawn.

1. The difference of tooth growth when using the supplement OJ over the supplement VC is much greater when the dosage ranger are .5 and 1.0.
2. When the 2mg dosage is used the effect of tooth growth is nearly the same when using the supplements OJ and VC.