

Statistical Inference Project Part 2

Ryan Wissman

Sunday, January 25, 2015

Coursera Statistical Inference Course Project

This is part two of the project for the class on statistical inference. For this project we will be analyzing the ToothGrowth dataset found in R. We will address the following questions using this data:

- 1) Load the ToothGrowth data and perform some basic exploratory data analyses
- 2) Provide a basic summary of the data.
- 3) 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)
- 4) State your conclusions and the assumptions needed for your conclusions.

Loading data and initial analysis (Q1)

Load the data.

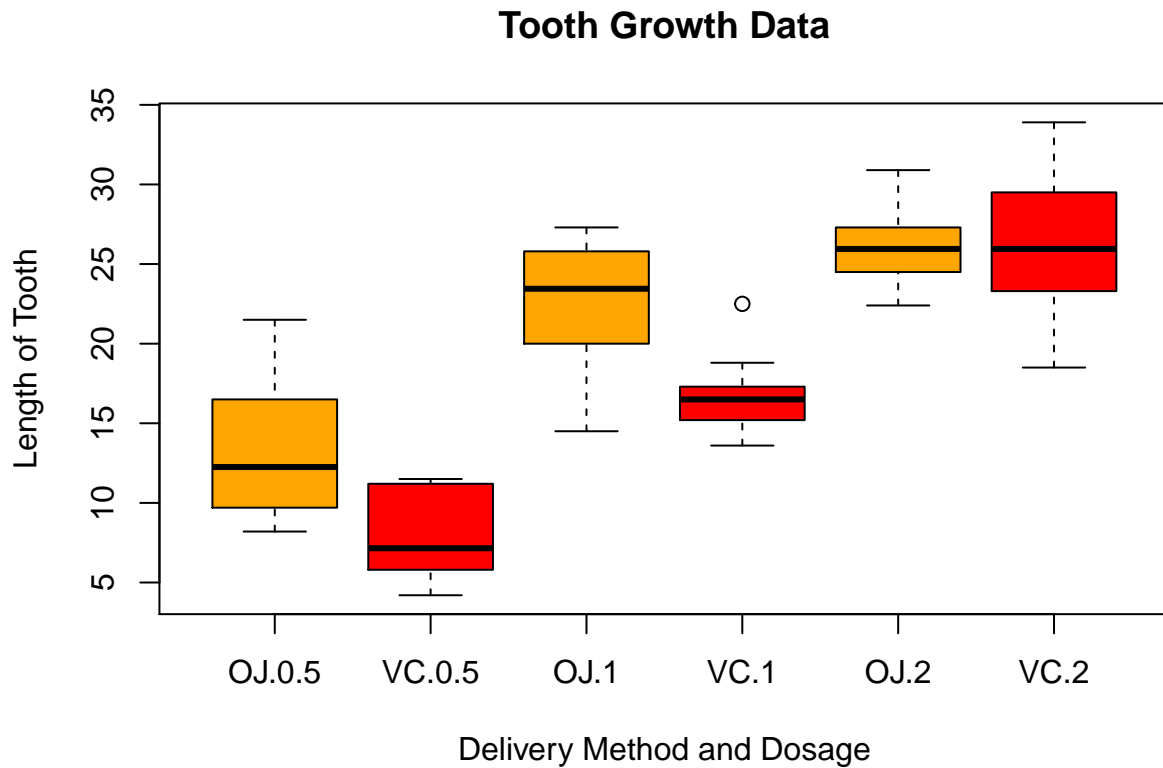
```
library(datasets)
data(ToothGrowth)

#Get the first few rows of the 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
```

The data shows the effect of vitamin c on tooth growth in guinea pigs. We know that the length of the tooth is given by (len), as well as the delivery method (OJ) or (VC), and the three dose levels (0.5mg, 1mg, and 2mg). A great way to quickly visualize this data would be to create a box and whisker plot illustrating the tooth growth by dosage level and delivery method.

```
boxplot(len ~ supp * dose, data=ToothGrowth, col=(c("orange", "red")), ylab="Length of Tooth",
        xlab="Delivery Method and Dosage", main="Tooth Growth Data")
```



After examining the above chart we could hypothesize that guinea pigs experienced more tooth growth when given orange juice when compared to ascorbic acid. Furthermore, tooth growth was further increased when exposed to higher dose levels.

Data Summary (Q2)

We will now perform a very basic summary of the data.

```
#set dose to factor to get count
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
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
```

We will also look at the mean for each dosage level and delivery method

```
with(ToothGrowth, tapply(len, list(supp,dose), mean))
```

##		0.5	1	2
##	OJ	13.23	22.70	26.06
##	VC	7.98	16.77	26.14

Look at the data this way furthers our hypothesis that orange juice is the superior delivery method to promote tooth growth.

Analysis, Hypothesis Test, and Conclusion (Q3 & Q4)

We will now test our hypothesis that orange juice is the superior vitamin C delivery method and results in better tooth growth among guinea pigs than ascorbic acid.