Statistical Inference Project Part 2

Umair Ahmed Khan

24 July 2016

Overview

The purpose of this document is to analyse the ToothGrowth dataset by comparing the Guinea tooth growth against supplement and dose. First part contains the exploratory data analysis of the dataset. Then the comparison with confidence intervals is shown in order to make conclusions about the tooth growth.

Load the ToothGrowth data and perform exploratory data analyses

```
library(datasets)
data(ToothGrowth)
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 ...
```

```
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
```

```
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
```

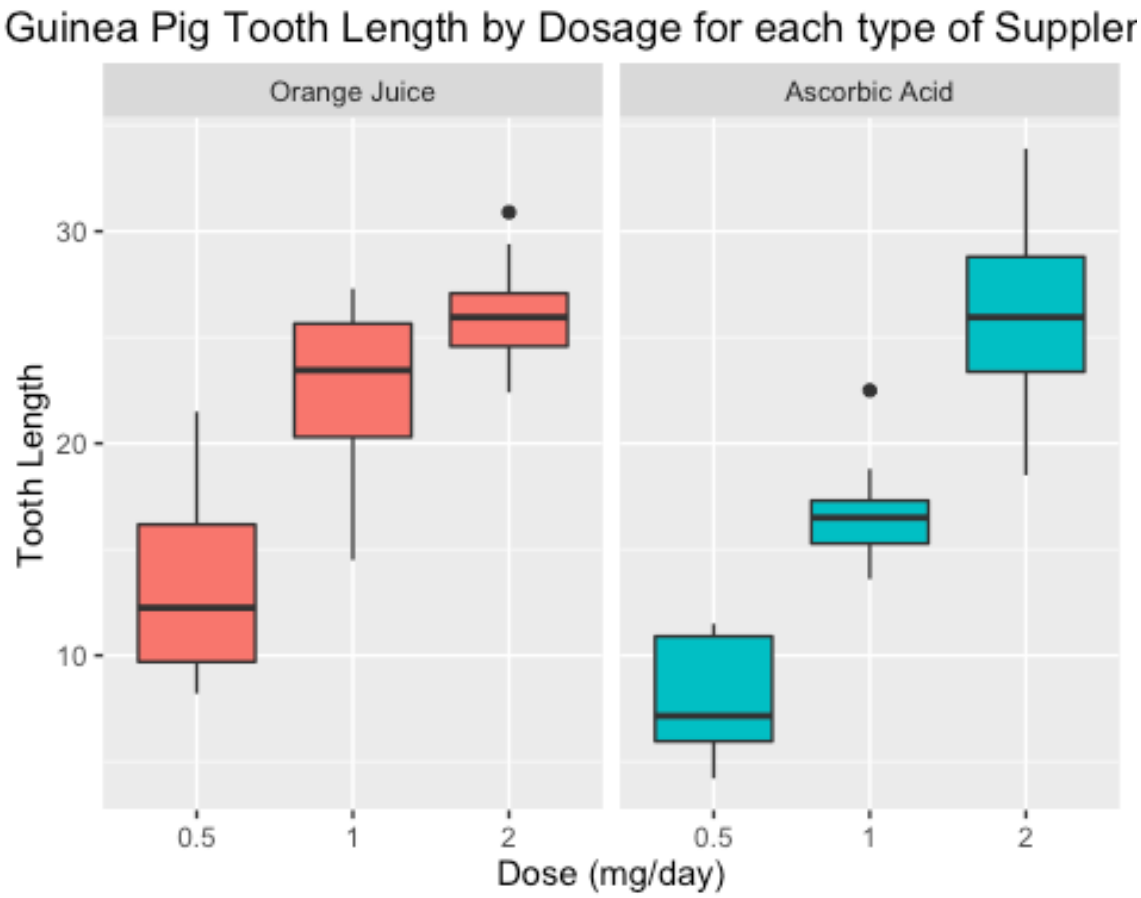
```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.2.4
```

```
t = ToothGrowth
levels(t$supp) <- c("Orange Juice", "Ascorbic Acid")
```

```
ggplot(t, aes(x=factor(dose), y=len)) +
  facet_grid(.~supp) +
  geom_boxplot(aes(fill = supp), show_guide = FALSE) +
  labs(title="Guinea Pig Tooth Length by Dosage for each type of
Supplement",
    x="Dose (mg/day)",
    y="Tooth Length")

## Warning: `show_guide` has been deprecated. Please use `show.legend`
## instead.
```



Basic summary of data

The box plots is suggesting that the dosage is directly proportional to the tooth growth. Orange juice is more effective than ascorbic acid for tooth growth when the dosage is 0.5 to 1.0 mg/day. Both type of supplements are equally effective when the dosage is 2.0 mg/day.

Using Confidence Intervals & Hypothesis Tests to compare tooth growth by supplement and dose.

Hypothesis 1

Orange Juice & Ascorbic Acid deliver the same tooth growth throughout the dataset.

```
hypoth1<-t.test(len ~ supp, data = t)
hypoth1$conf.int
```

```
## [1] -0.1710156  7.5710156
## attr(,"conf.level")
## [1] 0.95
```

```
hypoth1$p.value
```

```
## [1] 0.06063451
```

Results suggests that the confidence intervals includes 0 and the p-value is greater than the threshold of 0.05. Therefore the null hypothesis couldn't be rejected.

Hypothesis 2

For the dosage of 0.5 mg/day, the two supplements deliver the same tooth growth.

```
hypoth2<-t.test(len ~ supp, data = subset(t, dose == 0.5))
hypoth2$conf.int
```

```
## [1] 1.719057 8.780943
## attr(,"conf.level")
## [1] 0.95
```

```
hypoth2$p.value
```

```
## [1] 0.006358607
```

Results suggests that the confidence interval does not include 0 and the p-value is below the 0.05 threshold. Therefore the null hypothesis can be rejected. The alternative hypothesis that 0.5 mg/day dosage of orange juice delivers more tooth growth than ascorbic acid is accepted.

Hypothesis 3

For the dosage of 1 mg/day, the two supplements deliver the same tooth growth

```
hypoth3<-t.test(len ~ supp, data = subset(t, dose == 1))
hypoth3$conf.int
```

```
## [1] 2.802148 9.057852
## attr(,"conf.level")
## [1] 0.95
```

```
hypoth3$p.value

## [1] 0.001038376
```

Results suggests that the confidence interval does not include 0 and the p-value is smaller than the 0.05 threshold. Therefore the null hypothesis can be rejected. The alternative hypothesis that 1 mg/day dosage of orange juice delivers more tooth growth than ascorbic acid is accepted.

Hypothesis 4

For the dosage of 2 mg/day, the two supplements deliver the same tooth growth

```
hypoth4<-t.test(len ~ supp, data = subset(t, dose == 2))
hypoth4$conf.int

## [1] -3.79807  3.63807
## attr(,"conf.level")
## [1] 0.95

hypoth4$p.value

## [1] 0.9638516
```

Results suggests that the confidence interval does include 0 and the p-value is larger than the 0.05 threshold. Therefore null hypothesis cannot be rejected.

Conclusions & Assumptions

Orange juice delivers more tooth growth than ascorbic acid for dosages 0.5 & 1.0. Orange juice and Ascorbic Acid deliver the same amount of tooth growth for dose amount 2.0 mg/day. For the entire dataset we cannot conclude that Orange Juice is more effective that Ascorbic Acid.

Assumptions

- Normal distribution of the tooth lengths
- No other unmeasured factors are affecting tooth length