Statistical Inference Course Project

Part 2: Basic Inferential Data Analysis

Artem M.

November 4, 2017

Overview

We performed analysis of the ToothGrowth data in the R datasets package. Our analysis consisted of two parts: Basic Exploratory analysis and Comparison of the tooth growth by supplement type(orange juice/ascorbic acid) and dose of the vitamin C(milligrams/day).

Basic Exploratory analysis

Dataset

The dataset is part of the packaged datasets that come with the R language software. The name of the dataset is "The Effect of Vitamin C on Tooth Growth in Guinea Pigs". The dataset provides observations of the response is 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). The dataset comes in a dataframe format with 60 observations of the following 3 variables:

- **len** - Tooth length, numeric;
- **supp** - Supplement type (VC-ascorbic acid or OJ - orange juice), factor;
- **dose** - Dose in milligrams/day, numeric.

Data loading

We loaded the ToothGroth data from the R provided datasets:

```r
library(datasets)
data("ToothGrowth")
```

Basic exploratory analysis and data summary.

We started our analysis by understanding the dataset via R summary function:

```r
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
```

From the data summary the range of the tooth lengh was 4.20-33.90; 30 of the guinea pigs were given the ascorbic acid(VC) and 30 were given orange juice(OJ). Dose of the supplement was 0.5, 1, and 2.

To further explore the relationiship between tooth growth and factors such as supplement type and does we used the following box plots (Fig.1 and Fig.2):

```
library(ggplot2)
ggplot(ToothGrowth, aes(x=supp, y=len)) + geom_boxplot()+
        labs(title="Fig.1 Tooth length vs supplement type:" )+
        labs(x="Supplement type", y="Tooth Length")
```

Fig.1 Tooth length vs supplement type:
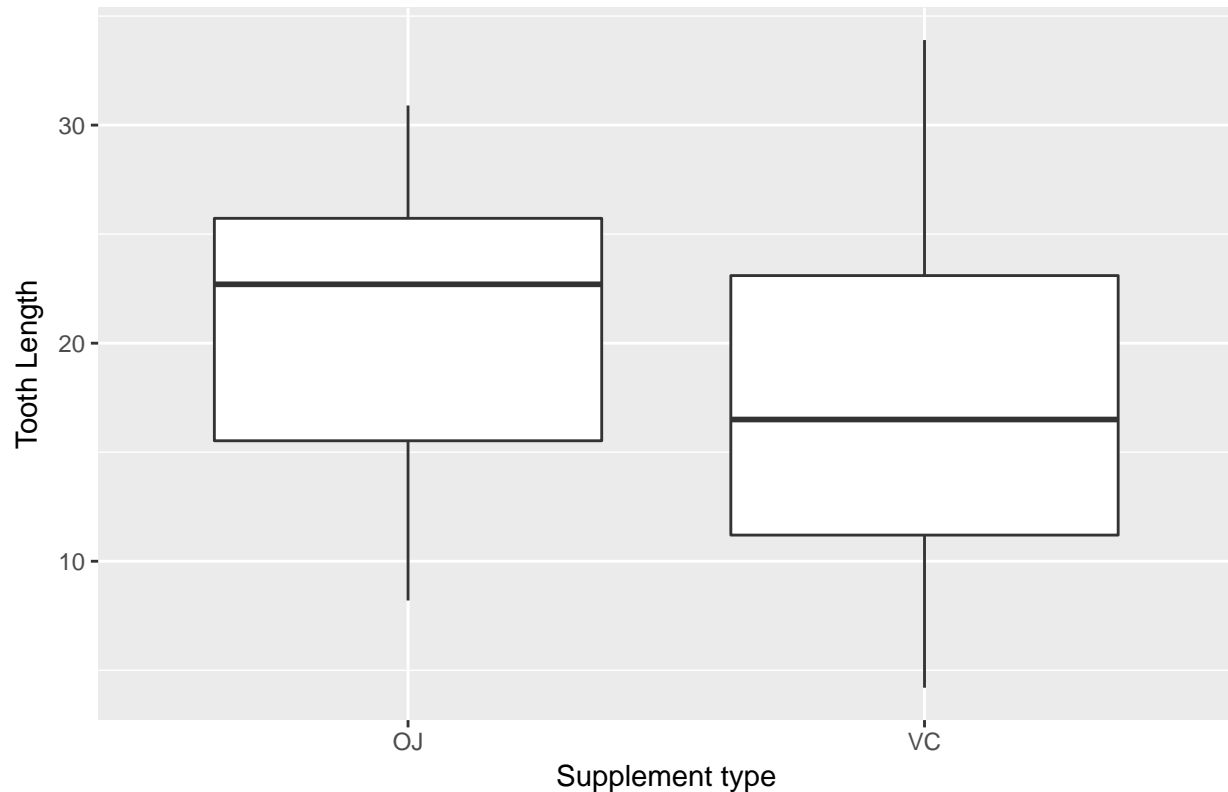
From Fig.1 we observed that the tooth length is on average larger for Orange juice supplement than for the ascorbic acid:

```
# Average tooth lenght for orange juice supplement

mean(ToothGrowth[ToothGrowth$supp=="OJ",1])
```

```
## [1] 20.66333
```

```
# Average tooth lenght for ascorbic acid supplement
mean(ToothGrowth[ToothGrowth$supp=="VC",1])
```
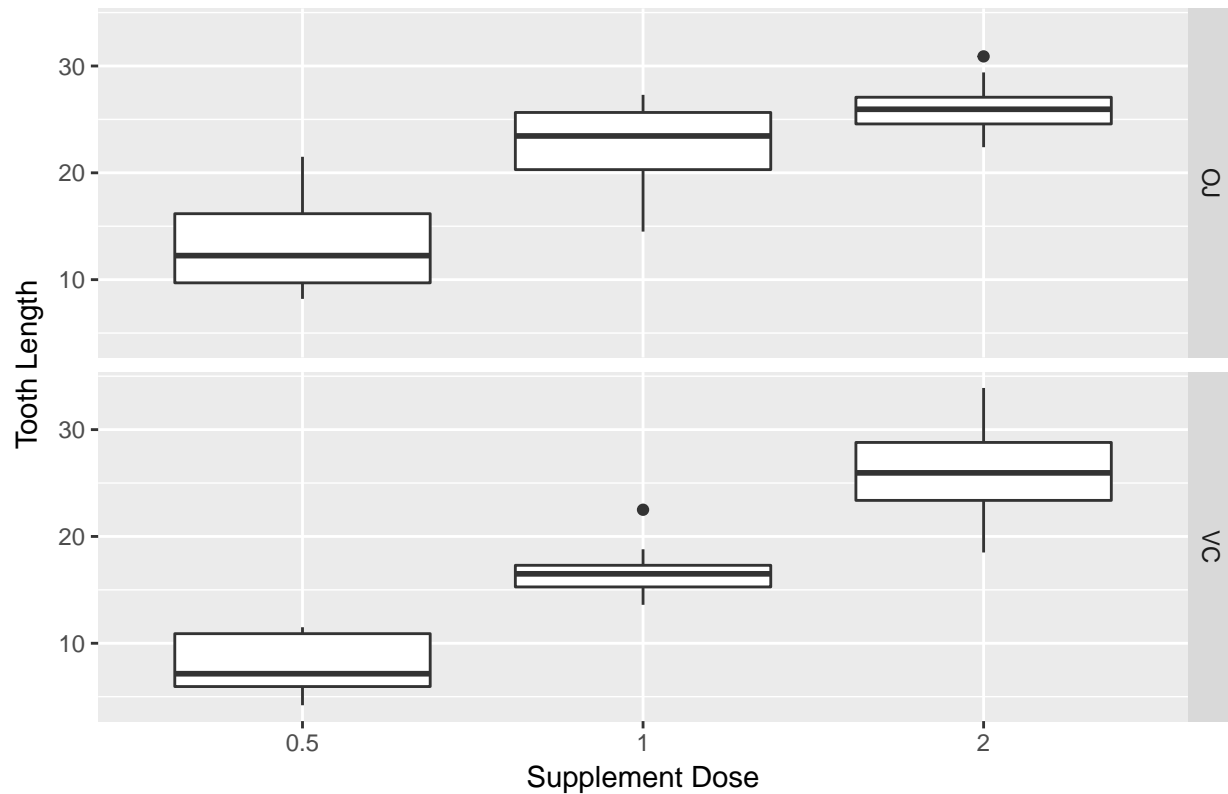
```
## [1] 16.96333
```

The following boxplot shows relationiships of the tooth length with the dose:

```
ggplot(ToothGrowth, aes(x=factor(dose), y=len)) +
        geom_boxplot()+
        facet_grid(supp ~ .)+
```

```
    labs(title="Fig.2 Tooth length vs supplement dose" )+
    labs(x="Supplement Dose", y="Tooth Length")
```

Fig.2 Tooth length vs supplement dose

From Fig.2 we observed that the tooth length is on average becomes larger when increasing the amount of the supplements regardless of the type of the supplement. We also observed that with the amount of supplement reaching 2 mg/day the average tooth growth were very close regardless of the type of the supplement:

Average tooth length for the orange juice supplement with a dose 2 mg/day

```
mean(ToothGrowth[ToothGrowth$supp=="OJ" & ToothGrowth$dose==2,1])
```

```
## [1] 26.06
```

Average tooth length for the ascorbic acid supplement with a dose 2 mg/day

```
mean(ToothGrowth[ToothGrowth$supp=="VC" & ToothGrowth$dose==2,1])
```

```
## [1] 26.14
```

Comparison of the tooth growth by supplement type and dose

We would like to explore the following questions:

1. Is there convincing evidence that the average tooth growth for the guinea pigs taking orange juice is different than the average tooth growth for the guinea pigs taking the ascorbic acid?

2. Is there convincing evidence that the average tooth growth for the guinea pigs taking 2 mg/day of vitamin C is the same regardless of the type of the supplement?

Since our sample was quite small we performed a T test to validate our hypothesis:

For the first question our null hypothesis H_0 was that the average tooth length is equal for both of the supplements:

```
t.test(ToothGrowth[ToothGrowth$supp=="OJ",1]-ToothGrowth[ToothGrowth$supp=="VC",1])
```

```
##
##  One Sample t-test
##
## data:  ToothGrowth[ToothGrowth$supp == "OJ", 1] - ToothGrowth[ToothGrowth$supp ==    "VC", 1]
## t = 3.3026, df = 29, p-value = 0.00255
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.408659 5.991341
## sample estimates:
## mean of x
##       3.7
```

Based on the T test, we observerd that the 95% confidence interval was [1.4,5.99] and did not include 0, additionaly p-value was roughly 0.2%, which gives us ability to reject null hypothesis; meaning there is indeed difference between average tooth growth that is dependent on the type of the vitamin C supplement.

For the second question we first isolated the observations with supplement intake that is equal to 2 mg/day:

```
dose2<-ToothGrowth[ToothGrowth$dose == 2, ]
```

Our null hypothesis H_0 was that the average tooth length is equal for both of the supplements when the dose is equal to 2 mg/day:

```
t.test(dose2[dose2$supp=="OJ",1]-dose2[dose2$supp=="VC",1])
```

```
##
##  One Sample t-test
##
## data:  dose2[dose2$supp == "OJ", 1] - dose2[dose2$supp == "VC", 1]
## t = -0.042592, df = 9, p-value = 0.967
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -4.328976  4.168976
## sample estimates:
## mean of x
##     -0.08
```

Results of the test showed that the 95% confidence interval was [-4.3,4.2] and did include 0; p-value was roughly 97%. Based on these results we fail to reject our null hypothesis, meaning the average tooth growth is the same when the supplement intake is 2 mg/day regardless of the type of the supplement.

Conclusion

Our analysis showed that on average the tooth growth is larger for the guinea pigs taking orange juice supplement. However when dose of the supplement is 2mg/day the average tooth growth is similar for both types of the supplements.