Statistical Inference Project, Part 2

Nannette Spear

July 23, 2018

Overview

Part II of the assignment invovles analyzing the ToothGrowth data in the R datasets package. The analysis involves providing basic summary of the data, exploratory data analysis, and comparing tooth growth by supplements and dosages. Conclusions will be drawn from the analysis.

Load knitr package

```
## Warning: package 'knitr' was built under R version 3.4.4
```

Set the Working Directory

Load the ToothGrowth Data

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

Basic Summary of Data

```
dim(ToothGrowth)
```

```
## [1] 60  3
```

```
names(ToothGrowth)
```

```
## [1] "len"  "supp" "dose"
```

```
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 ...
```

Data Exploration

```
aggregate(len ~ supp, ToothGrowth, mean)
```

```
##   supp      len
## 1   OJ 20.66333
## 2   VC 16.96333
```

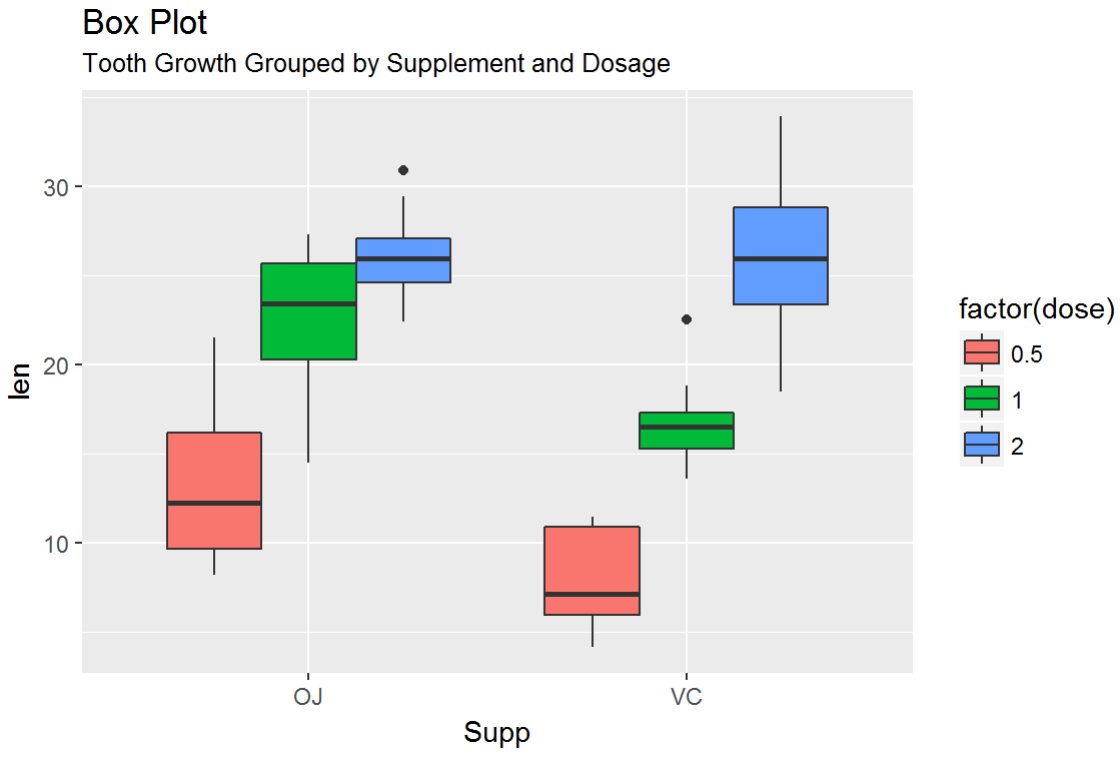
```
aggregate(len ~ dose, ToothGrowth, mean)
```

```
##   dose    len
## 1  0.5 10.605
## 2  1.0 19.735
## 3  2.0 26.100
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.4.4
```

```
d <- ggplot(ToothGrowth, aes(supp, len))
d + geom_boxplot(aes(fill = factor(dose))) +
  labs(title = "Box Plot",
       subtitle = "Tooth Growth Grouped by Supplement and Dosage ",
       x = "Supp",
       Y = "Length")
```



Assumptions - Comparing tooth growth by supplements and dosages.

1. Null Hypothesis: The difference in the average tooth length is zero between the supplements.

```
t.test(ToothGrowth$len ~ ToothGrowth$supp, paired = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len by ToothGrowth$supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

- The t value of 1.9153 falls within the confidence interval, therefore we fail to reject the null hypothesis that the difference in means is equal to 0.

2. Test the Null Hypothesis that the difference in the average tooth length is zero amongst dosages.

- There are three comparisons: .5 to 1 mg, 1 to 2 mg, and .5 to 2 mg.
- One is shown below:

- Compare dosages .5mg and 2mg

```
dose.5b <- subset(ToothGrowth, ToothGrowth$dose == .5, 1)
dose2b <- subset(ToothGrowth, ToothGrowth$dose == 2, 1)
dosetest3 <- t.test(dose.5b, dose2b)
dosetest3
```

```
##
##  Welch Two Sample t-test
##
## data:  dose.5b and dose2b
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean of x mean of y
##    10.605    26.100
```

Conclusion comparing tooth length to dosage.

For each test comparing doseages, the p-value is much smaller compared to the alpha of 0.05. Hence we reject the Null Hypothesis that the difference in means amongst dosage is zero.