Statistical Inference Course Project Part 2

Sheida Vanouei

July 24, 2018

Overview

In this paper, a sample dataset about Tooth Growth is being investigated with exploratory data analysis. The ToolthGrowth dataset contains observation for the effect of vitamin C on tooth growth in Guinea Pigs, which corresponds to the response is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).

Tooth Growth Dataset

Summary

```
data("ToothGrowth")


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 summary, we know that 30 observations are related to orange juice (supp = “OJ”), and other 30 are with ascorbic acid (supp = “VC”). Since the dose levels are fixed at 0.5, or 1, or 2 mg level, the median and mean or even quantile values do not carry much meaning in our analysis, we should focus on the length observation in the dataset, which the summary shows the values are between 4.20 and 33.90 with mean value at 28.81 and median at 19.25.

Structure

```
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

```
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
```

Unique

```
unique(ToothGrowth$len)
```

```
##  [1]  4.2 11.5  7.3  5.8  6.4 10.0 11.2  5.2  7.0 16.5 15.2 17.3 22.5 13.6
## [15] 14.5 18.8 15.5 23.6 18.5 33.9 25.5 26.4 32.5 26.7 21.5 23.3 29.5 17.6
## [29]  9.7  8.2  9.4 19.7 20.0 25.2 25.8 21.2 27.3 22.4 24.5 24.8 30.9 29.4
## [43] 23.0
```

```
unique(ToothGrowth$supp)
```

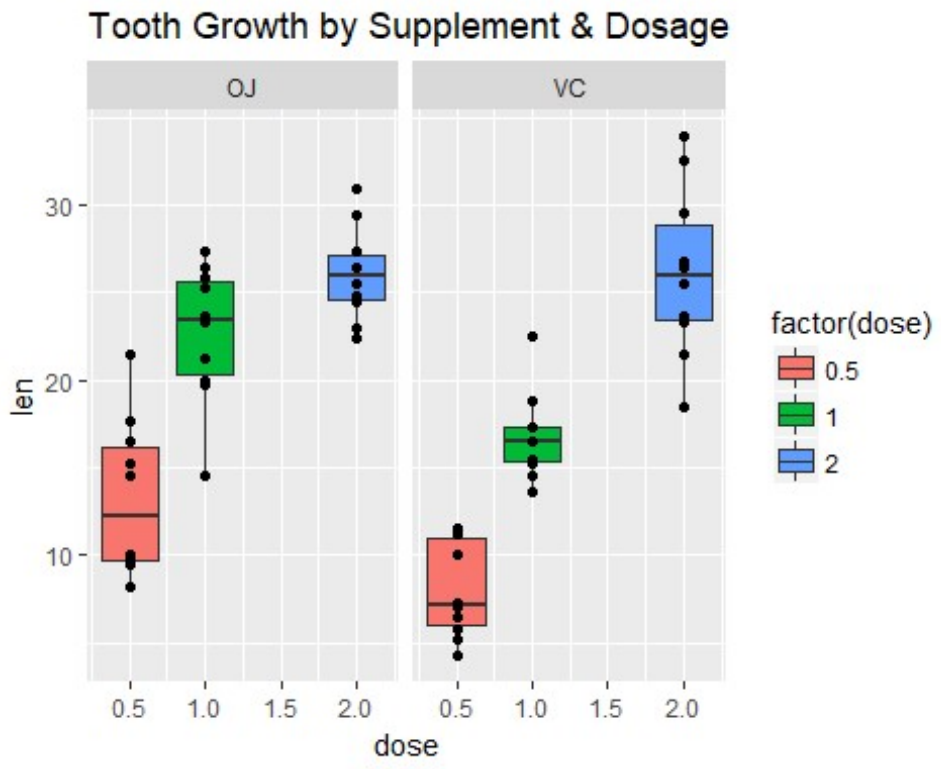
```
## [1] VC OJ
## Levels: OJ VC
```

```
unique(ToothGrowth$dose)
```

```
## [1] 0.5 1.0 2.0
```

Boxplot

```
library(ggplot2)
p <- ggplot(ToothGrowth, aes(x=dose, y=len))
p <- p + geom_boxplot(aes(fill=factor(dose)))
p <- p + geom_point()
p <- p + facet_grid(.~supp)
p <- p + ggtitle("Tooth Growth by Supplement & Dosage")
p
```




```
Based on the above chart, it appears that when the dosage is high at 2 mg,
the mean value of tooth growth appears to be similar between OJ and VC,
however, when the dosage is 0.5 mg or 1 mg, the chart definitely shows that
OJ has a obvious positive impact on tooth growth compared to VC.
```

Hypothesis Testing

Application methods have no impact on tooth growth.

H0: Both group have the same mean. HA: Means are different.

```
t.test(len ~ supp, data=ToothGrowth)

##
##  Welch Two Sample t-test
##
## data:  len by 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
```

Conclusion

The test returns a p-value of 0.06 and the confidence interval contains zero. The p-value is greater than the significance level of 0.05 and the confidence interval contains zero so that we can say that there is not enough evidence to reject the null hypothesis. Therefore, application methods have *no impact* on tooth growth

By Dose Level

Higher doses of vitamin C cause less tooth growth. H0: Mean of level 2 is equal than level 0.5. HA: Mean of level 2 is greater than level 0.5.

```
dose1 <- subset(ToothGrowth, dose %in% c(0.5, 1.0))
dose2 <- subset(ToothGrowth, dose %in% c(0.5, 2.0))
dose3 <- subset(ToothGrowth, dose %in% c(1.0, 2.0))
t.test(len ~ dose, paired = F, var.equal = F, data = dose1)

##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735

t.test(len ~ dose, paired = F, var.equal = F, data = dose2)

##
##  Welch Two Sample t-test
##
## data:  len by dose
## 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 in group 0.5   mean in group 2
##            10.605            26.100

t.test(len ~ dose, paired = F, var.equal = F, data = dose3)

##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

Conclusions

Supplement type has no effect on tooth growth.

Inreasing the dose level leads to increased tooth growth.