

Statistical Inference Part 2

Jency magdalene J

14/05/2020

Overview

The purpose of the this data analysis is to analyze the ToothGrowth data set by comparing the guinea tooth growth by supplement and dose. Let us do exploratory data analysis on the data set and provide a basic summary of the data.

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)
```

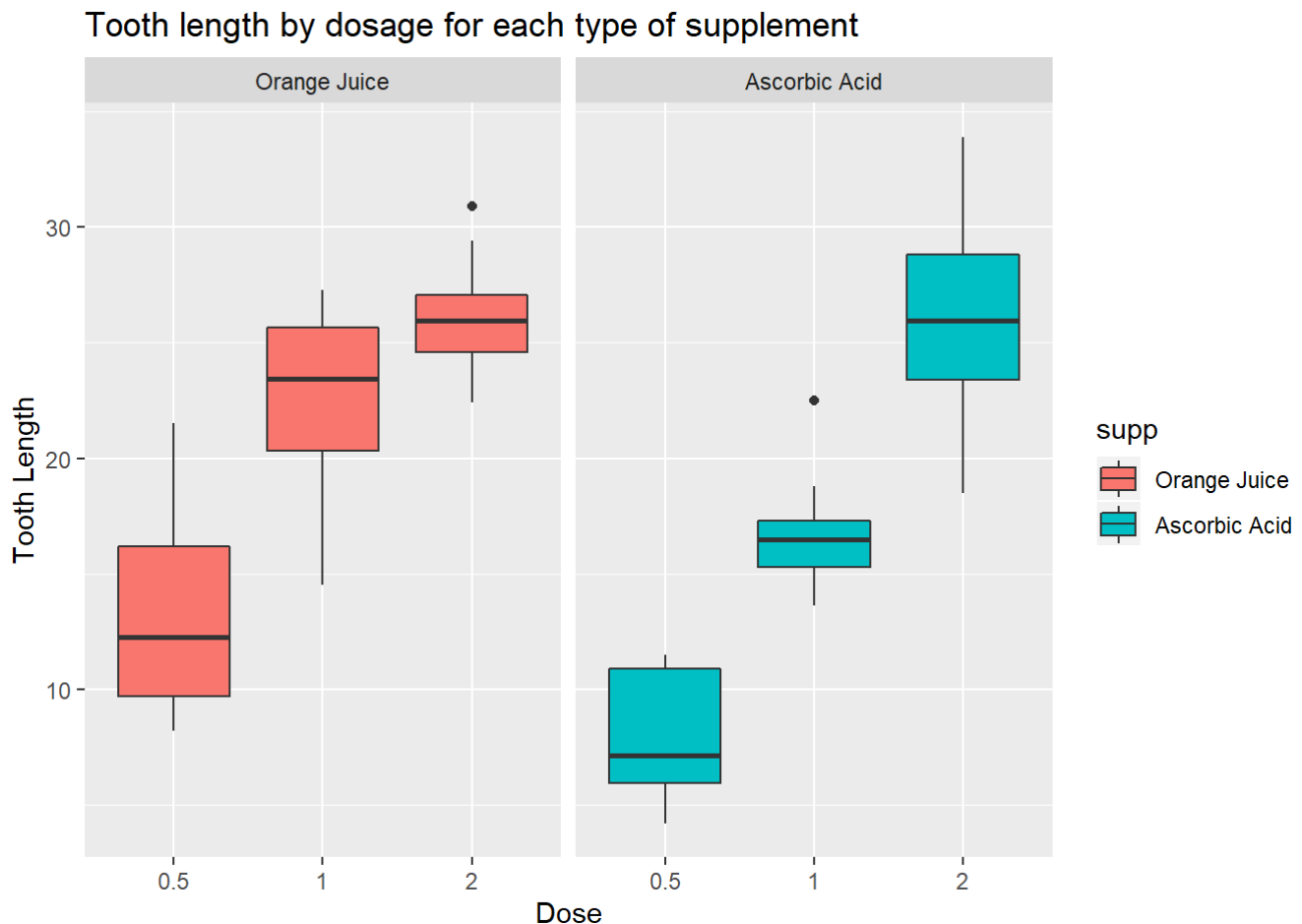
```
##   len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
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
```

Exploratory analysis

```
library(ggplot2)
levels(ToothGrowth$supp) <- c("Orange Juice", "Ascorbic Acid")
ggplot(ToothGrowth, aes(x=factor(dose), y=len)) +
  facet_grid(.~supp) +
  geom_boxplot(aes(fill = supp)) +
  labs(title="Tooth length by dosage for each type of supplement",
       x="Dose",
       y="Tooth Length")
```



Summary

The box plots seem to show, increasing the dosage increases the tooth growth. Orange juice is more effective than ascorbic acid for tooth growth when the dosage is .5 to 1.0 milligrams per day. Both types of supplements are equally as effective when the dosage is 2.0 milligrams per day.

Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose

T-test

Now we will compare tooth growth by supplement using a t-test.

```
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 Orange Juice mean in group Ascorbic Acid
##                20.66333                16.96333
```

The p-value of this test was 0.06. Since the p-value is greater than 0.05 and the confidence interval of the test contains zero we can say that supplement types seems to have no impact on Tooth growth based on this test. Now we'll compare tooth growth by dose, looking at the different pairs of dose values.

test 1

```
Tsub <- subset(ToothGrowth, ToothGrowth$dose %in% c(1.0,0.5))
t.test(len~dose,data=Tsub)
```

```
##
## 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
```

test 2

```
Tsub <- subset(ToothGrowth, ToothGrowth$dose %in% c(0.5,2.0))
t.test(len~dose,data=Tsub)
```

```
##
## 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
```

test 3

```
Tsub <- subset(ToothGrowth, ToothGrowth$dose %in% c(1.0,2.0))
t.test(len~dose,data=Tsub)
```

```
##
## 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
```

The p-value of each test was essentially zero and the confidence interval of each test does not cross over zero.

Thus we can assume that the average tooth length increases with an increasing dose, and therefore the null hypothesis can be rejected.

Conclusion

From the given data we found that the sample is representative of the population and the distribution of the sample means follows the Central Limit Theorem (CLT). In reviewing our t-test analysis from above, we can conclude that supplement delivery method has no effect on tooth growth/length, however increased dosages do result in increased tooth length.