

Exploring ToothData

ToothData Problem 2

Load the ToothGrowth data and perform some basic exploratory data analyses

Provide a basic summary of the data.

Below is the loading of the data along with some basic exploration and summarization

```
data("ToothGrowth")
summary(ToothGrowth)
```

```
##      len      supp      dose
##  Min.   : 4.2    OJ:30    Min.   :0.50
##  1st Qu.:13.1    VC:30    1st Qu.:0.50
##  Median :19.2                Median :1.00
##  Mean   :18.8                Mean   :1.17
##  3rd Qu.:25.3                3rd Qu.:2.00
##  Max.   :33.9                Max.   :2.00
```

```
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 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

The len column shows the length of teeth in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) in the dose column with each of two delivery methods (orange juice or ascorbic acid) in the supp column ###Use confidence intervals and hypothesis tests to compare tooth growth by supp and dose. (Use the techniques from class even if there's other approaches worth considering)

```
t5 <- ToothGrowth[ToothGrowth$dose == 0.5,]
t10 <- ToothGrowth[ToothGrowth$dose == 1.0,]
t20 <- ToothGrowth[ToothGrowth$dose == 2.0,]

t.test(t5$len[t5$supp == "OJ"], t5$len[t5$supp == "VC"], paired=TRUE, alternative="greater")
```

```
##
## Paired t-test
##
## data:  t5$len[t5$supp == "OJ"] and t5$len[t5$supp == "VC"]
## t = 2.979, df = 9, p-value = 0.007736
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  2.02 Inf
## sample estimates:
## mean of the differences
##                5.25
```

```
t.test(t10$len[t10$supp == "OJ"], t5$len[t10$supp == "VC"], paired=TRUE, alternative="greater")
```

```
##
## Paired t-test
##
## data: t10$len[t10$supp == "OJ"] and t5$len[t10$supp == "VC"]
## t = 12.22, df = 9, p-value = 3.287e-07
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
## 12.51 Inf
## sample estimates:
## mean of the differences
## 14.72
```

```
t.test(t20$len[t20$supp == "OJ"], t5$len[t20$supp == "VC"], paired=TRUE, alternative="greater")
```

```
##
## Paired t-test
##
## data: t20$len[t20$supp == "OJ"] and t5$len[t20$supp == "VC"]
## t = 17.9, df = 9, p-value = 1.204e-08
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
## 16.23 Inf
## sample estimates:
## mean of the differences
## 18.08
```

4. State your conclusions and the assumptions needed for your conclusions.

The results show a low p-value for each dose level, and since the p-value is very low, we reject the null hypothesis. There is strong evidence of a mean increase in tooth growth between Ascorbic Acid and Orange Juice.

This assumes that the guinea pigs were randomly selected for supplement type, and appropriate time was given to reset the pigs tooth growth. Further, we're dosage levels were also separately done and appropriate time was given to reset the pigs in between switching dosages.