Statistical Inference Course Project: PART 2

2. Basic Inferential Data Analysis

1. Basic Exploratory Analysis and Summary of Data

The Tooth Growth data contains 60 records including the dose, supplement given and length of tooth in a guinea pig for an experiement. There are two types of supplement; VC and OJ each with 30 trials. There are 3 types of doses; 0.5, 1.0 and 2.0mg, each with 20 trials.

Supplements Summary Table

```
##    supp meanLength sdLength
## 1:   VC      16.96    8.266
## 2:   OJ      20.66    6.606
```

Doses Summary Table

```
##    dose meanLength sdLength
## 1:  0.5      10.61    4.500
## 2:  1.0      19.73    4.415
## 3:  2.0      26.10    3.774
```

3. Use Confidence Intervals and Hypothesis Tests to compare tooth growth by supplement and dose

Comparing by Supp
Independent t confidence interval for OJ - VC

```
## [1] -0.17  7.57
```

Most of the time the OJ out performs the VC with 95% confidence

Comparing by Dose

Independent t confidence interval comparing doses 0.5 vs 1.0 milligrams

```
## [1]  6.28 11.98
```

Dose of 1.0 always out performs does of 0.5 in stimulating tooth growth with 95% confidence

Independent t confidence interval comparing doses 0.5 vs 2.0 milligrams

```
## [1] 12.84 18.15
```

Dose of 2.0 always out performs does of 0.5 in stimulating tooth growth with 95% confidence

Independent t confidence interval comparing doses 2.0 vs 1.0 milligrams

```
## [1] 3.74 8.99
```

Dose of 2.0 always out performs does of 1.0 in stimulating tooth growth with 95% confidence

b) Hypothesis Test

Comparing by Supp

```
##
##  Welch Two Sample t-test
##
## data:  dt[supp == "OJ"]$len and dt[supp == "VC"]$len
## t = 1.915, df = 55.31, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.171  7.571
## sample estimates:
## mean of x mean of y
##     20.66     16.96
```

Comparing by Dose

```
##
##  Welch Two Sample t-test
##
## data:  dt[dose == 0.5]$len and dt[dose == 1]$len
## t = -6.477, df = 37.99, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.984  -6.276
## sample estimates:
## mean of x mean of y
##     10.61     19.73
```

```
##
##  Welch Two Sample t-test
##
## data:  dt[dose == 0.5]$len and dt[dose == 2]$len
## t = -11.8, df = 36.88, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.16 -12.83
## sample estimates:
## mean of x mean of y
##     10.61     26.10
```

```
##
##  Welch Two Sample t-test
##
## data:  dt[dose == 1]$len and dt[dose == 2]$len
## t = -4.901, df = 37.1, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996 -3.734
## sample estimates:
## mean of x mean of y
##     19.73     26.10
```

4. Conclusions and Assumptions

Conclusions

1. Most of the time the OJ out performs the VC with 95% confidence
2. Dose of 1.0 always out performs does of 0.5 in stimulating tooth growth with 95% confidence
3. Dose of 2.0 always out performs does of 0.5 in stimulating tooth growth with 95% confidence
4. Dose of 2.0 always out performs does of 1.0 in stimulating tooth growth with 95% confidence

Assumptions

1. T distribution is used as data may not be normally distributed
2. Assume unequal variances for t distribution confidence interval
3. Assume the Central Limit Theorm
4. n must be large enough to be statistically signifcant
5. Since n is small for each set of tests use Gossett's T test
6. Assuming a constant variance between groups of Guinea Pigs receiving difference amounts of treatment and different supplements