Statistical Inference Course Project Part 2

Semin (Sammie) Bae
University of Waterloo

Overview:

This report will explore the ToothGrowth data in the R dataset package. The given dataset will be analyzed with a brief summary, confidence intervals, as well as hypothesis tests. These analysis methods will be used to help briefly summarize the information about the dataset.

Simulations

Code:

```
#Load the datasets package                    #-comment/explanation
library(datasets)
#Load the ToothGrowth data
data(ToothGrowth)
#Print the ToothGrowth data out
(ToothGrowth)
```

The above code prints the sample ToothGrowth dataset. It has 3 columns and 60 entries

Result:

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 7 11.2 VC 0.5 8 11.2 VC 0.5 9 5.2 VC 0.5 10 7.0 VC 0.5 11 16.5 VC 1.0 12 16.5 VC 1.0 13 15.2 VC 1.0 14 17.3 VC 1.0 15 22.5 VC 1.0	16 17.3 VC 1.0 17 13.6 VC 1.0 18 14.5 VC 1.0 19 18.8 VC 1.0 20 15.5 VC 1.0 21 23.6 VC 2.0 22 18.5 VC 2.0 23 33.9 VC 2.0 24 25.5 VC 2.0 25 26.4 VC 2.0 26 32.5 VC 2.0 27 26.7 VC 2.0 28 21.5 VC 2.0 29 23.3 VC 2.0 30 29.5 VC 2.0	31 15.2 OJ 0.5 32 21.5 OJ 0.5 33 17.6 OJ 0.5 34 9.7 OJ 0.5 35 14.5 OJ 0.5 36 10.0 OJ 0.5 37 8.2 OJ 0.5 38 9.4 OJ 0.5 39 16.5 OJ 0.5 40 9.7 OJ 0.5 41 19.7 OJ 1.0 42 23.3 OJ 1.0 43 23.6 OJ 1.0 44 26.4 OJ 1.0 45 20.0 OJ 1.0	46 25.2 OJ 1.0 47 25.8 OJ 1.0 48 21.2 OJ 1.0 49 14.5 OJ 1.0 50 27.3 OJ 1.0 51 25.5 OJ 2.0 52 26.4 OJ 2.0 53 22.4 OJ 2.0 54 24.5 OJ 2.0 55 24.8 OJ 2.0 56 30.9 OJ 2.0 57 26.4 OJ 2.0 58 27.3 OJ 2.0 59 29.4 OJ 2.0 60 23.0 OJ 2.0

Data:

- len- Length of tooth (mm)
- supp: Type of supplement VC(Vitamin C) vs OJ (Orange Juice)
- dose: dose of supplement (mg)

Basic Summary

```
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
```

For the length of the tooth, lowest was 4.2 mm, median was 19.25 mm, mean was 18.81 mm and max was 33.90 mm.There were 30 Orange Juice data sets and 30 Vitamin C data sets. For the dosage, the minimum was 0.5 mg, median was 1.0 mg, mean was 1.167 mg, and maximum was 2 mg.

Graphing with GGPlot2

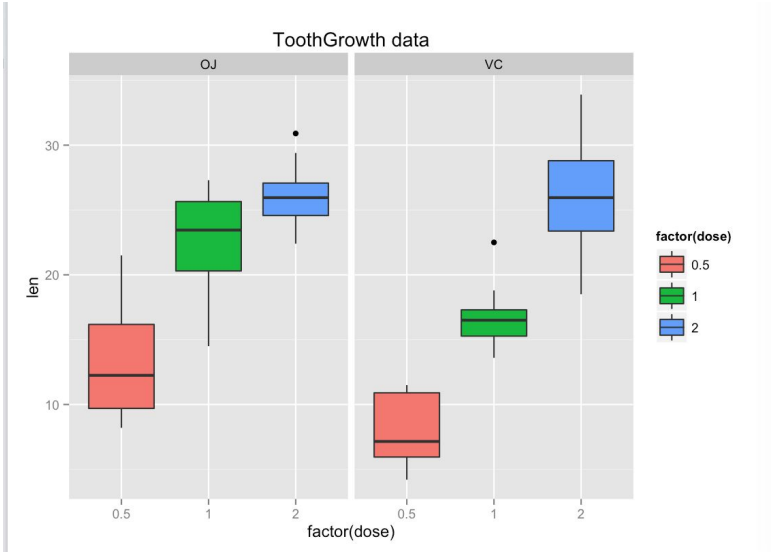

Figure 1, ToothGrowth Graph

Confidence Intervals Test:

Confidence interval is a type interval estimation of parameter of the population. t.test function in R provides variety of t-tests.

```
> 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
```

Now this analysis will be compared to three data sets provided by this code:

```
test1<-subset(ToothGrowth, dose==0.5)

test2<-subset(ToothGrowth, dose==1.0)

test3<-subset(ToothGrowth, dose==2.0)
```

Above code split the data set into only three possible values for dosage: 0.5, 1.0 and 2.0

```
> t.test(len~supp, data=test1, paired=FALSE)
...
mean in group OJ mean in group VC                #dataset 1
         13.23            7.98

> t.test(len~supp, data=test2, paired=FALSE)   #dataset 2
...
mean in group OJ mean in group VC
         22.70           16.77

> t.test(len~supp, data=test3, paired=FALSE)   #dataset 3
...
mean in group OJ mean in group VC
         26.06           26.14
```

Conclusion

In the first set, means were 13.23 for OJ and 7.98 for VC while the last set had means of 26.06 and 26.14.This shows that longer length of the tooth had more dosages of Orange Juice/Vitamin C. Furthermore, Orange Juice shows better effect on tooth growth in first two data sets but has similar amount in dataset 3, indicating its early efficiency diminishing to Vitamin C's level. These conclusions can be seen in Figure 1, where g