Statistical Inference Project 2

Thursday, November 13, 2014

Requirements

Now in the second portion of the class, we're going to analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses

2. Provide a basic summary of the data.

3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only

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

Some criteria that you will be evaluated on

*Did you perform an exploratory data analysis of at least a single plot or table highlighting basi

*Did the student perform some relevant confidence intervals and/or tests?

*Were the results of the tests and/or intervals interpreted in the context of the problem correctly

*Did the student describe the assumptions needed for their conclusions?

```
#Set globals, load libraries, and load and attach the data
# Turn off scientific notations for numbers
options(scipen = 999)

if(!require("doBy")){
        Install.packages("doBy")
}
```

```
## Loading required package: doBy
## Loading required package: survival
## Loading required package: splines
## Loading required package: MASS
```

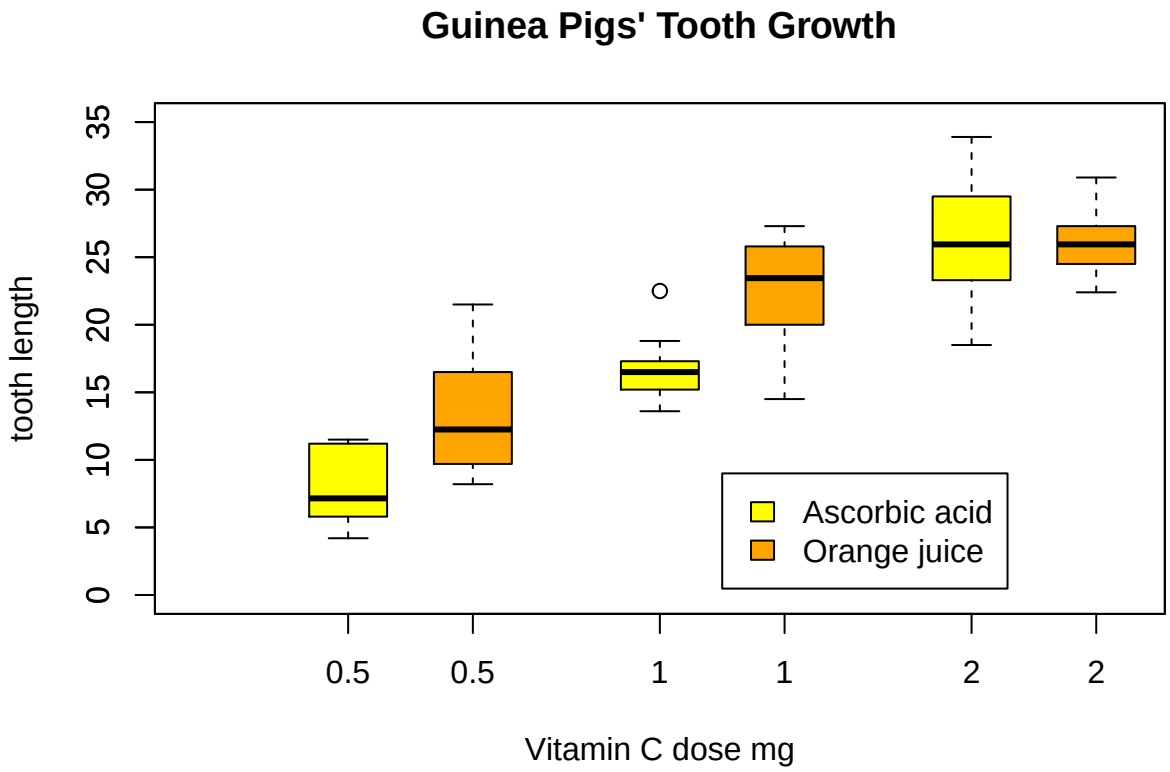
```
##Attach the Data
attach(ToothGrowth)
```

Exploratory Data Analysis

The ToothGrowth dataset consists of 60 observations on 3 variables:

```
len      numeric    Tooth length
supp     factor         Supplement type (VC or OJ).
dose     numeric    Dose in milligrams.
```

The ToothGrowth dataset explains the relation between 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) with each of two supplement types (orange juice and ascorbic



acid).

The boxplot above indicates that, on average, the length of the tooth increases as the dose also increases. Supplement type appears to make a difference at the 0.5 and 1 mg levels, but not at the 2 mg level.

Basic Summary of the Data

Below, I have calculated the mean and standard deviation for the length of the tooth by dose and supplement type.

```
summaryBy(len~dose+supp, data=ToothGrowth, FUN=c(mean,sd,length))
```

```
##   dose supp len.mean   len.sd len.length
## 1  0.5   OJ    13.23 4.459709         10
## 2  0.5   VC     7.98 2.746634         10
## 3  1.0   OJ    22.70 3.910953         10
## 4  1.0   VC    16.77 2.515309         10
## 5  2.0   OJ    26.06 2.655058         10
## 6  2.0   VC    26.14 4.797731         10
```

The summary data above indicates that, on average, the length of the tooth increases as the dose also increases. Supplement type appears to make a difference at the 0.5 and 1 mg levels, but not at the 2 mg level.

Hypothesis Tests and Confidence Intervals

The first hypothesis test was that there was no difference in tooth length between orange juice and ascorbic acid. The alternative hypothesis was that orange juice increased tooth length more than ascorbic acid did.

```
results <-t.test(len[supp=="OJ"],len[supp=="VC"],paired=TRUE, alternative = c("greater"))
results
```

```
##
##  Paired t-test
##
## data:  len[supp == "OJ"] and len[supp == "VC"]
## t = 3.3026, df = 29, p-value = 0.001275
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  1.796409      Inf
## sample estimates:
## mean of the differences
##                     3.7
```

The t-value of 3.303 with a p-value of 0.0013 indicated that the null hypothesis should be rejected. There is evidence to suggest that orange juice increased tooth length more than ascorbic acid did.

The null hypothesis for the second test was that there was no difference in tooth length between 2 mg and 1 mg dose levels. The alternative hypothesis was that the 2 mg dose increased tooth length more than 1 mg dose did.

```
results <-t.test(len[dose==2],len[dose==1],paired=TRUE, alternative = c("greater"))
results
```

```
##
##  Paired t-test
##
## data:  len[dose == 2] and len[dose == 1]
## t = 4.6046, df = 19, p-value = 0.00009671
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  3.974821      Inf
## sample estimates:
## mean of the differences
##                   6.365
```

The t-value of 4.605 with a p-value of 0.0001 indicated that the null hypothesis should be rejected. There is evidence to suggest that the 2 mg dose increased tooth length more than 1 mg dose did.

The null hypothesis for the final test was that there was no difference in tooth length between the 1 mg and .5 mg dose levels. The alternative hypothesis was that the 1 mg dose increased tooth length more than .5 mg dose did.

```
results <-t.test(len[dose==1],len[dose==.5],paired=TRUE, alternative = c("greater"))
results
```

```
##
```

```
##  Paired t-test
##
## data:  len[dose == 1] and len[dose == 0.5]
## t = 6.9669, df = 19, p-value = 0.0000006127
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.863996      Inf
## sample estimates:
## mean of the differences
##                    9.13
```

The t-value of 6.967 with a p-value of 0 indicated that the null hypothesis should be rejected. There is evidence to suggest that the 1 mg dose increased tooth length more than .5 mg dose did.

Conclusions and the Assumptions

With the values obtained, it can be assumed that there is a different in the growth of the tooth while the doses are larger. There is also evidence to suggest that supplement type has an effect on tooth growth.