Statistical Inference Course Project Part 2

Sarah Fogel

Thursday, December 18, 2014

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

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

First load the data

```
library(datasets)
data(ToothGrowth)
```

Some basic exploratory analysis:

```
dim(ToothGrowth)
```

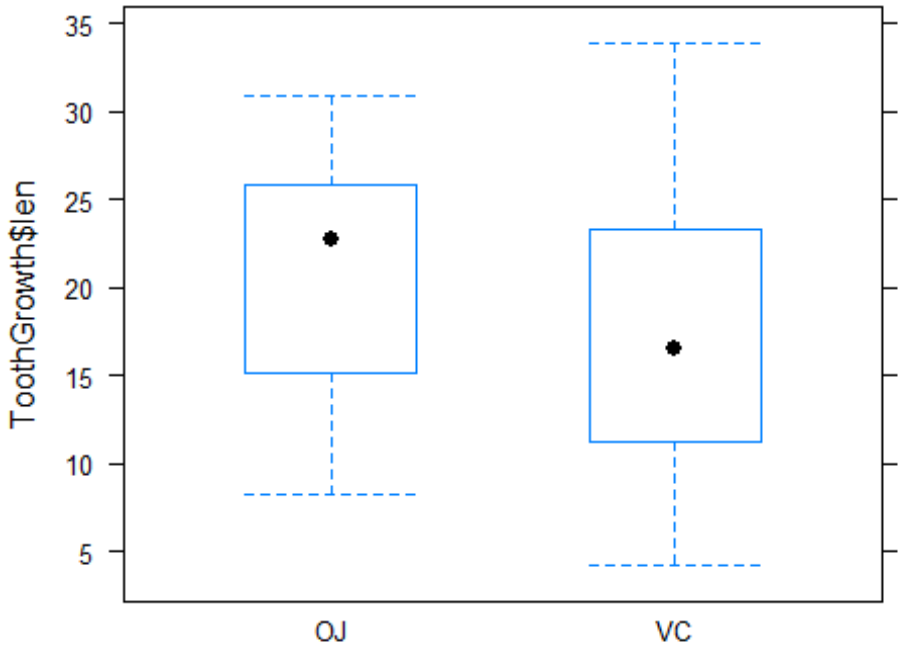
```
## [1] 60  3
```

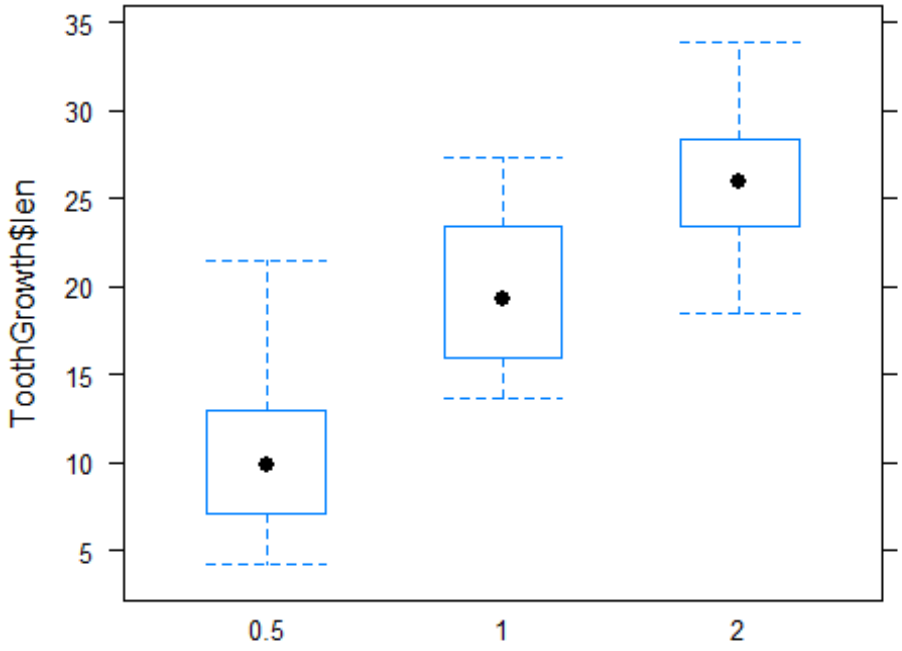
```
str(ToothGrowth)
```

```
## '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
```





Part 2. Provide a basic summary of the data.

This data set shows the length of teeth in each of 10 guinea pigs at each of three dose levels of Vitamin C with each of two delivery methods.

```
##       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
```

Part3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)

First prepare to compare the growth by supp and dose by splitting the length by those variables.

```
Tooth.split.supp<-split(ToothGrowth$len, f=ToothGrowth$supp)
Tooth.split.dose<-split(ToothGrowth$len, f=ToothGrowth$dose)
```

Now compare the growth by supp via a t-test

```
t.test(Tooth.split.supp$OJ, Tooth.split.supp$VC)
```

```
##
##  Welch Two Sample t-test
##
## data:  Tooth.split.supp$OJ and Tooth.split.supp$VC
## 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 of x mean of y
##  20.66333  16.96333
```

Next, compare the growth by dose via three t-tests

```
t.test(Tooth.split.dose$`0.5`, Tooth.split.dose$`1`)
```

```
##
##  Welch Two Sample t-test
##
## data:  Tooth.split.dose$`0.5` and Tooth.split.dose$`1`
## 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 of x mean of y
##    10.605    19.735
```

```
t.test(Tooth.split.dose$`1`, Tooth.split.dose$`2`)
```

```
##
##  Welch Two Sample t-test
##
## data:  Tooth.split.dose$`1` and Tooth.split.dose$`2`
## 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 of x mean of y
##    19.735    26.100
```

```
t.test(Tooth.split.dose$`0.5`, Tooth.split.dose$`2`)
```

```
##
##  Welch Two Sample t-test
##
## data:  Tooth.split.dose$`0.5` and Tooth.split.dose$`2`
## 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 of x mean of y
##    10.605    26.100
```

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

The assumptions requred include that t intervals assume that data are iid normal, though it works well whenever the distribution is roughly symmetric and mound shaped - which the length variable is

The first conclusion is that, since the 95% confidence interval for the difference in means between the supplement types includes 0, and the p-value is greater than .05, we fail to reject the null hypothesis that the means are the same. Therefore, we cannot say that the supplement had an effect on tooth growth.

For the dosages, each of the 95% confidence intervals did not contain 0, and each of the p-values was less than .05. Therefore, we reject the null hypothesis for each case. We can state that there was an increase in tooth length with a dose of 1 over a dose of .5, and with a dose of 2 over 1 or .5.