Statistical Inference - Course Project Part 2

Michael Celeste

Saturday, November 22, 2014

Load data and perform a basic exploratory analysis of the ToothGrowth data

```
library(ggplot2)

## Data frame structure
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 ...
```

```
## Data frame values
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
```

```
## Show first few rows of data
head(ToothGrowth)
```

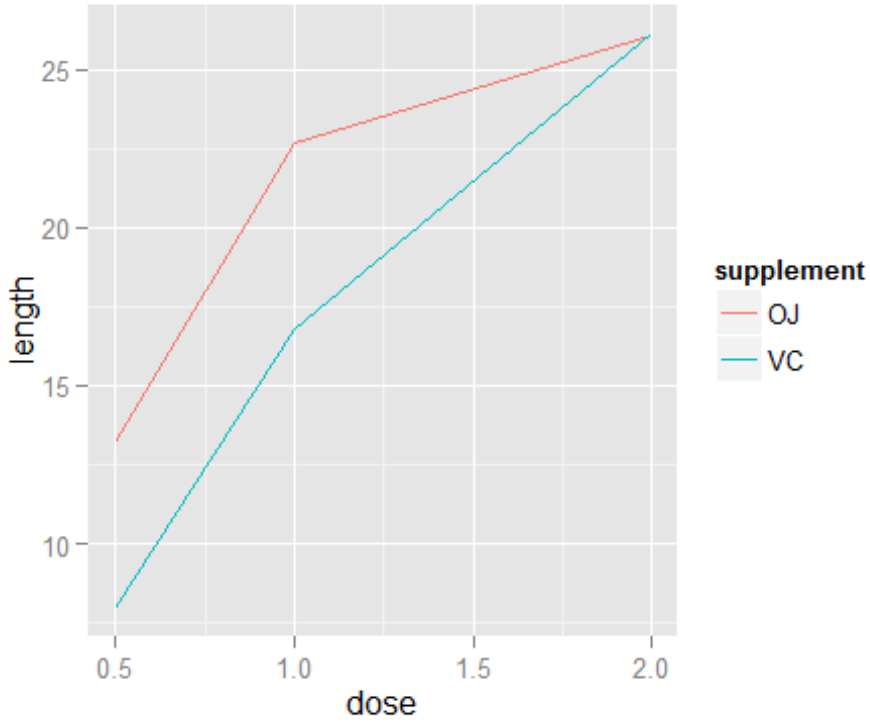
```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
## Aggregate the mean length grouped by supplement and dose
tg <- aggregate(ToothGrowth$len, list(ToothGrowth$supp, ToothGrowth$dose),
mean)

## Rename the columns in tg
names(tg)[1] <- "supplement"
```

```
names(tg)[2] <- "dose"
names(tg)[3] <- "length"

## Plot the data
ggplot(tg, aes(x=dose, y=length, colour=supplement)) + geom_line()
```



Thirty subjects took OJ and another 30 took Vitamin C. The means for tooth length differ for various doses except for the 2.0 dose.

A test comparing the means of each supplement will be performed to see if there is enough of a difference of one supplement to the other.

```
## Calculate the mean lenght for OJ
ojmean <- mean( ToothGrowth[ToothGrowth$supp == "OJ", "len"] )

ojsd <- sd( ToothGrowth[ToothGrowth$supp == "OJ", "len"] )

## Calculate the mean length for VC
vcmean <- mean( ToothGrowth[ToothGrowth$supp == "VC", "len"] )

vcsd <- sd( ToothGrowth[ToothGrowth$supp == "VC", "len"] )

## Calculate the confidence level (95%) of the mean OJ supplement and compare
it to the mean of VC supplement
ojmean + c(-1,1)*qnorm(0.95, lower.tail=T)*ojsd/sqrt(30)

## [1] 18.68 22.65
```

The VC supplement mean (16.9633) is below the confidence interval therefore it seems that the VC supplement does not increase the length as much as OJ. Because the size of the sample for each supplement is 30 it may be better to do T tests. A t-test will be done for all doses then one for each dose level.

```
## Create a t-test for the type of supplement and dose (each)
t.test(ToothGrowth[ToothGrowth$supp == "OJ", "len"] -
ToothGrowth[ToothGrowth$supp == "VC", "len"])

##
##  One Sample t-test
##
## data:  ToothGrowth[ToothGrowth$supp == "OJ", "len"] -
ToothGrowth[ToothGrowth$supp ==     "VC", "len"]
## t = 3.303, df = 29, p-value = 0.00255
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.409 5.991
## sample estimates:
## mean of x
##       3.7

t.test(ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose==0.5, "len"] -
ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose==0.5, "len"])

##
##  One Sample t-test
##
## data:  ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 0.5,
"len"] - ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose ==     0.5,
"len"]
## t = 2.979, df = 9, p-value = 0.01547
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.263 9.237
## sample estimates:
## mean of x
##      5.25

t.test(ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose==1, "len"] -
ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose==1, "len"])

##
##  One Sample t-test
##
## data:  ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 1,
"len"] - ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose ==     1,
"len"]
## t = 3.372, df = 9, p-value = 0.008229
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
```

```
##  1.952 9.908
## sample estimates:
## mean of x
##      5.93
```

```
t.test(ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose==2, "len"] -
ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose==2, "len"])
```

```
##
##  One Sample t-test
##
## data:  ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 2,
"len"] - ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose ==     2,
"len"]
## t = -0.0426, df = 9, p-value = 0.967
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -4.329  4.169
## sample estimates:
## mean of x
##     -0.08
```

Conclusions

From the results of the T-tests in each dose it can be inferred that the effectiveness of the VC supplement is not as effective as the OJ until dose number 2. Therefore, the minimum amount of VC supplement that should be taken if trying to supplement OJ is a dose of 2.