Class06-PartII

Maria Koh

Saturday, September 20, 2014

This is the project Part 2 for the statistical inference class. 1.Load the ToothGrowth data and perform some basic exploratory data analyses

```
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 3.1.1
```

```
library(ggplot2)
library(plyr)
data(ToothGrowth)
```

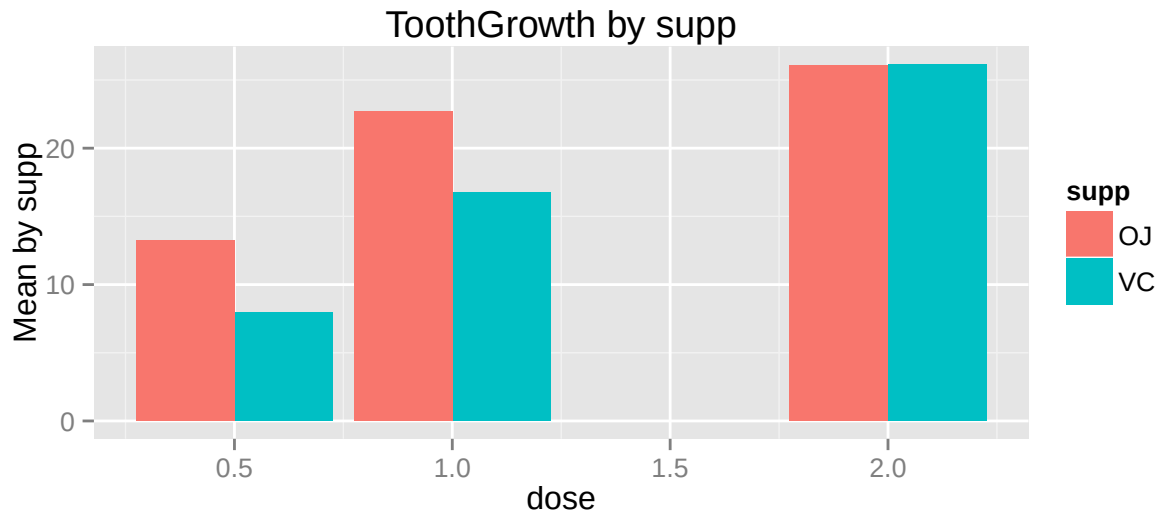
perform some basic exploratory data analyses using means for length by supp and dose

```
ToothGrowth.OJ <- subset(ToothGrowth, ToothGrowth$supp == "OJ")
ToothGrowth.VC <- subset(ToothGrowth, ToothGrowth$supp == "VC")

res.OJ <- data.frame( res.dose=unique(ToothGrowth.OJ$dose), res.mean=tapply(ToothGrowth.OJ$len, ToothGr
res.VC <- data.frame( res.dose=unique(ToothGrowth.VC$dose), res.mean=tapply(ToothGrowth.VC$len, ToothGr
res <- merge( res.OJ , res.VC, all = TRUE)
```

I embed plots for ToothGrowth.

```
ggplot( data=res) +
geom_bar(aes(x=as.numeric(res$res.dose), y=res$res.mean, fill=res$res.supp), stat="identity", position=
labs(fill="supp") + ggtitle("ToothGrowth by supp") + ylab("Mean by supp") + xlab("dose")
```

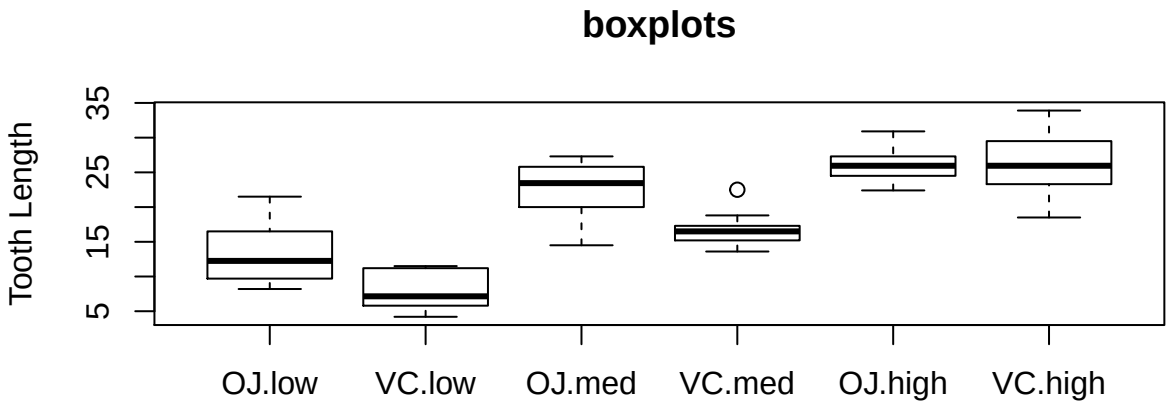


2.Provide a basic summary of the data.

```
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
```

```
ToothGrowth$dose = factor(ToothGrowth$dose, levels=c( 0.5, 1.0, 2.0), labels = c("low", "med", "high"))
boxplot(len ~supp*dose, data=ToothGrowth, ylab="Tooth Length", main = "boxplots")
```



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

3-1) Compare Orange Juice vs. Ascorbic Acid ignored dose type.

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

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "OJ"] and ToothGrowth$len[ToothGrowth$supp == "VC"]
## 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
```

H0 - There aren't significant differences in the impact of the Orange Juice and Ascorbic Acid.So, The differences of two group's means are 0. As the result the t test, p-value(0.06063) is larger than the 0.05, Ho accepted, There aren't the significant differences in the impact of the Orange Juice and Ascorbic Acid

3-2) Compare Orange Juice vs. Ascorbic Acid for only low(0.5mg) dose type.

```
t.test(ToothGrowth$len[ToothGrowth$supp=="OJ"&ToothGrowth$dose=="low"],ToothGrowth$len[ToothGrowth$supp
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose ==  and ToothGrowth$len[ToothGrow
## t = 3.17, df = 14.97, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719 8.781
## sample estimates:
## mean of x mean of y
##     13.23      7.98
```

H0 - There aren't significant differences in the impact of the Orange Juice and Ascorbic Acid for low(0.5 mg) dose.So, The differences of two group's means are 0. As the result the t test, p-value(0.006359) is less than the 0.05, Ho rejected, There are the significant differences in the impact of the Orange Juice and Ascorbic Acid for low dose.

3-3) Compare Orange Juice vs. Ascorbic Acid for only med(1mg) dose type.

```
t.test(ToothGrowth$len[ToothGrowth$supp=="OJ"&ToothGrowth$dose=="med"],ToothGrowth$len[ToothGrowth$supp
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose ==  and ToothGrowth$len[ToothGrow
## t = 4.033, df = 15.36, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802 9.058
## sample estimates:
## mean of x mean of y
##     22.70     16.77
```

H0 - There aren't significant differences in the impact of the Orange Juice and Ascorbic Acid for medium(1 mg) dose.So, The differences of two group's means are 0. As the result the t test, p-value(0.001038) is less than the 0.05, Ho rejected, There are the significant differences in the impact of the Orange Juice and Ascorbic Acid for low dose.

3-4) Compare Orange Juice vs. Ascorbic Acid for only high(2mg) dose type.

```
t.test(ToothGrowth$len[ToothGrowth$supp=="OJ"&ToothGrowth$dose=="high"],ToothGrowth$len[ToothGrowth$supp
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose ==  and ToothGrowth$len[ToothGrow
## t = -0.0461, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```

```
##  -3.798  3.638
## sample estimates:
## mean of x mean of y
##     26.06     26.14
```

H0 - There aren't significant differences in the impact of the Orange Juice and Ascorbic Acid for High(2 mg) dose.So, The differences of two group's means are 0. As the result the t test, p-value(0.9639) is larger than the 0.05, Ho accepted, There aren't the significant differences in the impact of the Orange Juice and Ascorbic Acid for high dose.

3-5) For Orange Juice, Compare low(0.5mg) dose type vs. med(1mg) dose type.

```
t.test(ToothGrowth$len[ToothGrowth$supp=="OJ"&ToothGrowth$dose=="low"],ToothGrowth$len[ToothGrowth$supp
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose ==  and ToothGrowth$len[ToothGrow
## t = -5.049, df = 17.7, p-value = 8.785e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -13.416  -5.524
## sample estimates:
## mean of x mean of y
##     13.23     22.70
```

H0 - There aren't significant differences in the impact for the Low(0.5mg) dose and Medium(1 mg) dose in the Orange Juice. So, The differences of two group's means are 0. As the result the t test, p-value(8.785e-05) is less than the 0.05, Ho rejected, There are the significant differences in the impact of the dose.

3-6) For Orange Juice, Compare low(0.5mg) dose type vs. high(2mg) dose type.

```
t.test(ToothGrowth$len[ToothGrowth$supp=="OJ"&ToothGrowth$dose=="low"],ToothGrowth$len[ToothGrowth$supp
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose ==  and ToothGrowth$len[ToothGrow
## t = -7.817, df = 14.67, p-value = 1.324e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -16.335  -9.325
## sample estimates:
## mean of x mean of y
##     13.23     26.06
```

H0 - There aren't significant differences in the impact for the Low(0.5mg) dose and Medium(1 mg) dose in the Orange Juice. So, The differences of two group's means are 0. As the result the t test, p-value(1.324e-06) is less than the 0.05, Ho rejected, There are the significant differences in the impact of the dose.

3-7) For Orange Juice, Compare med(1mg) dose type vs. high(2mg) dose type.

```
t.test(ToothGrowth$len[ToothGrowth$supp=="OJ"&ToothGrowth$dose=="med"],ToothGrowth$len[ToothGrowth$supp
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose ==  and ToothGrowth$len[ToothGrow
## t = -2.248, df = 15.84, p-value = 0.0392
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -6.5314 -0.1886
## sample estimates:
## mean of x mean of y
##     22.70     26.06
```

H0 - There aren't significant differences in the impact for the Low(0.5mg) dose and Medium(1 mg) dose in the Orange Juice. So, The differences of two group's means are 0. As the result the t test, p-value(0.0392) is less than the 0.05, Ho rejected, There are the significant differences in the impact of the dose.

3-8) For Ascorbic Acid, Compare low(0.5mg) dose type vs. med(1mg) dose type.

```
t.test(ToothGrowth$len[ToothGrowth$supp=="VC"&ToothGrowth$dose=="low"],ToothGrowth$len[ToothGrowth$supp
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "VC" & ToothGrowth$dose ==  and ToothGrowth$len[ToothGrow
## t = -7.463, df = 17.86, p-value = 6.811e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.266  -6.314
## sample estimates:
## mean of x mean of y
##      7.98     16.77
```

H0 - There aren't significant differences in the impact for the Low(0.5mg) dose and Medium(1 mg) dose in the Ascorbic Acid. So, The differences of two group's means are 0. As the result the t test, p-value(6.811e-07) is less than the 0.05, Ho rejected, There are the significant differences in the impact of the dose.

3-9) For Ascorbic Acid, Compare low(0.5mg) dose type vs. high(2mg) dose type.

```
t.test(ToothGrowth$len[ToothGrowth$supp=="VC"&ToothGrowth$dose=="low"],ToothGrowth$len[ToothGrowth$supp
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "VC" & ToothGrowth$dose ==  and ToothGrowth$len[ToothGrow
## t = -10.39, df = 14.33, p-value = 4.682e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -21.90 -14.42
## sample estimates:
## mean of x mean of y
##      7.98     26.14
```

H0 - There aren't significant differences in the impact for the Low(0.5mg) dose and high(2 mg) dose in the Ascorbic Acid. So, The differences of two group's means are 0. As the result the t test, p-value(4.682e-08) is less than the 0.05, Ho rejected, There are the significant differences in the impact of the dose.

3-10) For Ascorbic Acid, Compare med(1mg) dose type vs. high(2mg) dose type.

```
t.test(ToothGrowth$len[ToothGrowth$supp=="VC"&ToothGrowth$dose=="med"],ToothGrowth$len[ToothGrowth$supp
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$supp == "VC" & ToothGrowth$dose ==  and ToothGrowth$len[ToothGrow
## t = -5.47, df = 13.6, p-value = 9.156e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -13.054  -5.686
## sample estimates:
## mean of x mean of y
##     16.77     26.14
```

H0 - There aren't significant differences in the impact for the Medium(1mg) dose and High(2 mg) dose in the Ascorbic Acid. So, The differences of two group's means are 0. As the result the t test, p-value(9.156e-05) is less than the 0.05, Ho rejected, There are the significant differences in the impact of the dose.

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

There aren't the significant differences in the impact of the Orange Juice and Ascorbic Acid But, There are the significant differences in the impact for their dose.