Statistical Inference Course Project 2

Andres Mauricio Castro

Tooth Growth Data Analysis

Assignment Description

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 use the techniques from class, even if there's other approaches worth considering)
4. State your conclusions and the assumptions needed for your conclusions.

Load Data

```
library(datasets)
```

Exploratory Data Analysis

We see the documentation, to get a rough idea of what the data represents.

```
?ToothGrowth
```

```
## starting httpd help server ... done
```

Summary Of The Dataset

```
str(ToothGrowth)
```

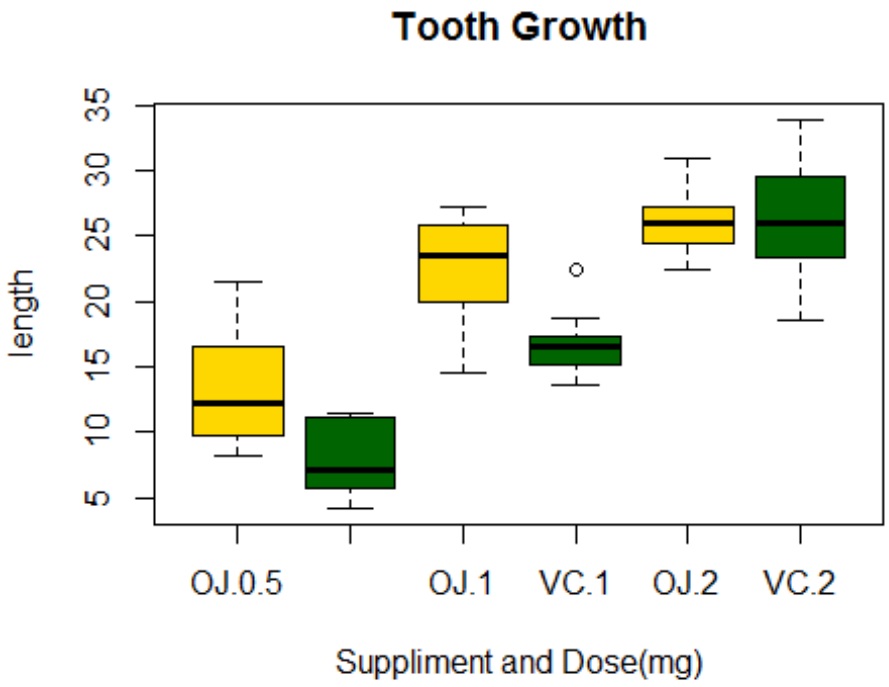
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
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
```

Create a box plot, to have an idea of how the dose is related to the tooth grows.

```
boxplot(len~supp*dose, data=ToothGrowth,
col=(c("gold","darkgreen")),main="Tooth Growth", xlab="Suppliment and
Dose(mg)", ylab="length")
```



As per the plot, the higher the dose, the longer the tooth grows. It seems Orance Juice (OJ) has a better effect on teeth growth, than the other supplement type (VC).

Compare Tooth Growth By Supp And Dose

We carry out T-tests at each dosage level for the two supplement type, in order to verify the support of our hyphotesis, from the previous plot.

The T-test at 0.5 mg:

```
t.test(len ~ supp, ToothGrowth[ToothGrowth$dose == .5, ])
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

The T-test at 1 mg:

```
t.test(len ~ supp, ToothGrowth[ToothGrowth$dose == 1, ])
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

The T-test at 2 mg:

```
t.test(len ~ supp, ToothGrowth[ToothGrowth$dose == 2, ])
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

Conclusions

The conclusions are based on the following assumptions:

```
- The guinea pigs are repesentative of the whole population.
- The distribution of means is supposed to be normal.
- The sample is supposed to be random.
```

From the exploratory data analysis, the higher the dose, the longer the tooth grows. It seems Orance Juice (OJ) has a better effect on teeth growth, than the other supplement type (VC).

From the T-test analysis, according to the p-values that for dosages of 0.5 mg and 1 mg, orange juice is a more effective delivery method of vitamin C, to promote the tooth grow, so, we can validate the hypothesis for dosages at 0.5 and 1 mg. From the p-value for the 2 mg, we cannot conclude that orange juice is a better delivery method of vitamin C, to promote tooth growth, than the other delivery method (VC).