Statistical Inference Course Project Part 2

Luis Fernando Agottani

24/08/2019

For the second part, the ToothGrowth data in the R dataset package will be analyzed.

Part 2: Basic Inferential Data Analysis Instructions

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

```
library(datasets)

data(ToothGrowth)
```

2. Provide a basic summary of the data.

```
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
```

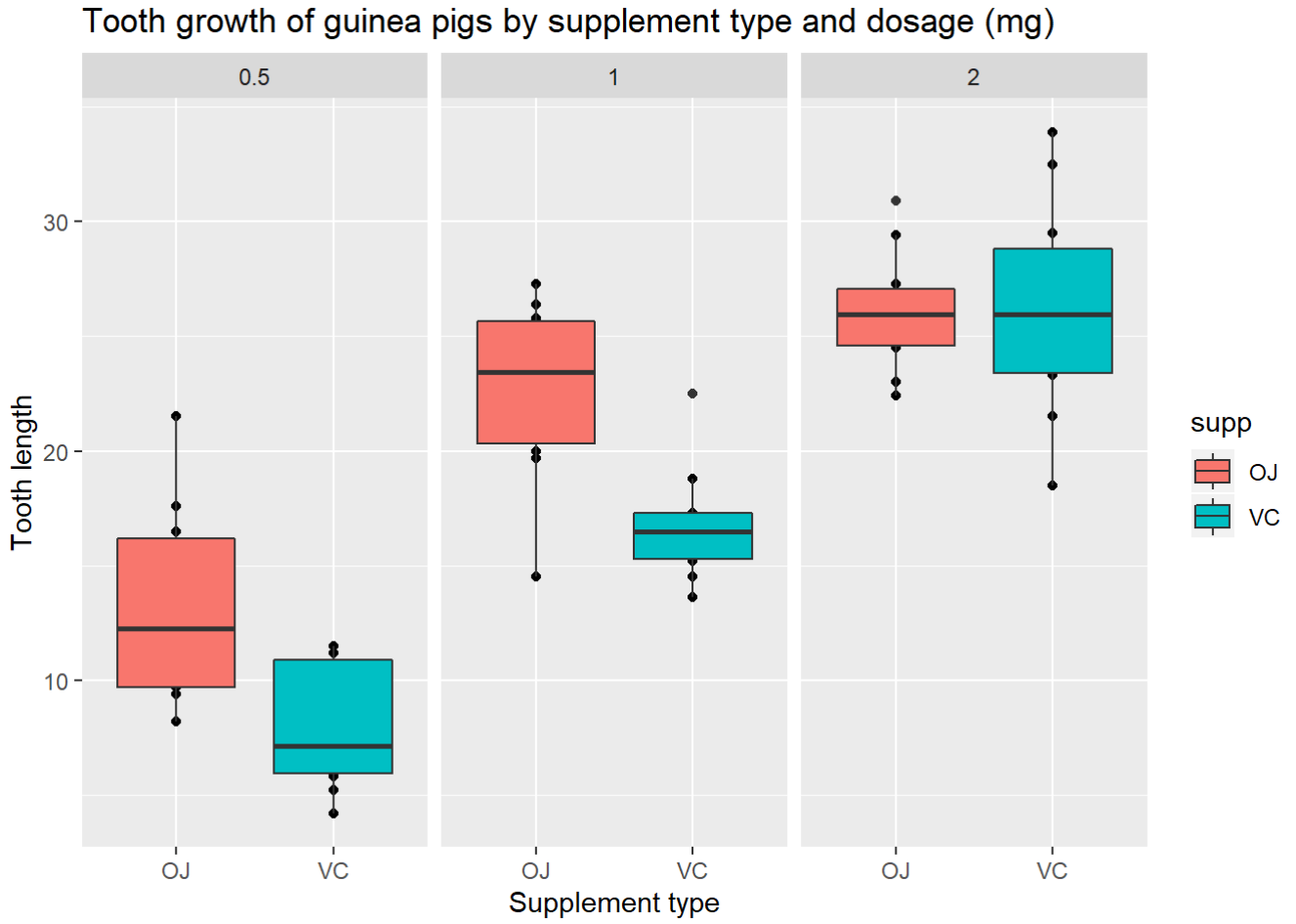
```
summary(ToothGrowth)
```

```
##       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
```

The data ToothGrowth give to us results from testing vitamin C in Guinea Pigs to measure the tooth growth in differents dosages (0,5; 1,0; 2,0 mg) and in differents types of supplements (VC and OJ).

```
library(ggplot2)

qplot(supp,len,data=ToothGrowth, facets=~dose, main="Tooth growth of guinea pigs by supplemen
t type and dosage (mg)",xlab="Supplement type", ylab="Tooth length") + geom_boxplot(aes(fill
 = supp))
```



The result of the exploratory analises from the data “ToothGrowth”, show that the tooth length increases when the dosage of vitamin C bigger, and for dosages of 0,5mg and 1mg, the supplement OJ has better results when comparing to VC. Lets confirm these with hypothesis.

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)

Hypothesis to the supplement.

Lets work with Hypothesis. Using our null hypothesis when using supplement OJ and VC.

So, Null Ho = lenth OJ = lenght VC, Alternative Ha = lenth OJ > lenght VC.

```
OJ = ToothGrowth$len[ToothGrowth$supp == 'OJ']
VC = ToothGrowth$len[ToothGrowth$supp == 'VC']

t.test(OJ, VC, alternative = "greater", paired = FALSE, var.equal = FALSE, conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  OJ and VC
## t = 1.9153, df = 55.309, p-value = 0.03032
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4682687       Inf
## sample estimates:
## mean of x mean of y
##  20.66333  16.96333
```

With the graphic, we can see that for dosage equal 2mg, there is not a significant difference. Lets check.

```
OJ2mg = ToothGrowth$len[ToothGrowth$supp == 'OJ' & ToothGrowth$dose == 2]
VC2mg = ToothGrowth$len[ToothGrowth$supp == 'VC' & ToothGrowth$dose == 2]

t.test(OJ2mg, VC2mg, alternative = "two.sided", paired = FALSE, var.equal = FALSE, conf.level
= 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  OJ2mg and VC2mg
## t = -0.046136, 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 of x mean of y
##     26.06     26.14
```

Hypothesis to dosage.

The null hypothesis now is that there is not difference in tooth growth between dosage and our alternative hypotesis is that when bigger the dosage greater the growth tooth effect.

```
dose05 = ToothGrowth$len[ToothGrowth$dose == '0.5']
dose10 = ToothGrowth$len[ToothGrowth$dose == '1']
dose20 = ToothGrowth$len[ToothGrowth$dose == '2']

t.test(dose10, dose05, alternative = "greater", paired = FALSE, var.equal = FALSE, conf.level
= 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose10 and dose05
## t = 6.4766, df = 37.986, p-value = 6.342e-08
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.753323      Inf
## sample estimates:
## mean of x mean of y
##    19.735    10.605
```

```
t.test(dose20, dose10, alternative = "greater", paired = FALSE, var.equal = FALSE, conf.level
= 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose20 and dose10
## t = 4.9005, df = 37.101, p-value = 9.532e-06
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  4.17387     Inf
## sample estimates:
## mean of x mean of y
##    26.100    19.735
```

State your conclusions and the assumptions needed for your conclusions.

In our t.test, p-values that are lower than 5% are reject the null hypothesis. we can conclude that more dosages of vitamin C will result in bigger tooth growth (p-value<0.05) and for dosages 0,5 mg and 1,0 mg the supplement OJ has more effect than VC (p-value<0.05. For dosage 2,0 mg results shows that is not difference between supplements for tooth growth (p-value>0.05.