

Statistical Inference Course Project Part 2

domesc

June 16, 2015

Assignment

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

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

We load first the data and we look at the variable types through str function:

```
data("ToothGrowth")
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 ...
```

We have 2 numerical columns “len” and “dose” and a factor variable “supp”. To have more information on the data set:

```
head(ToothGrowth)
```

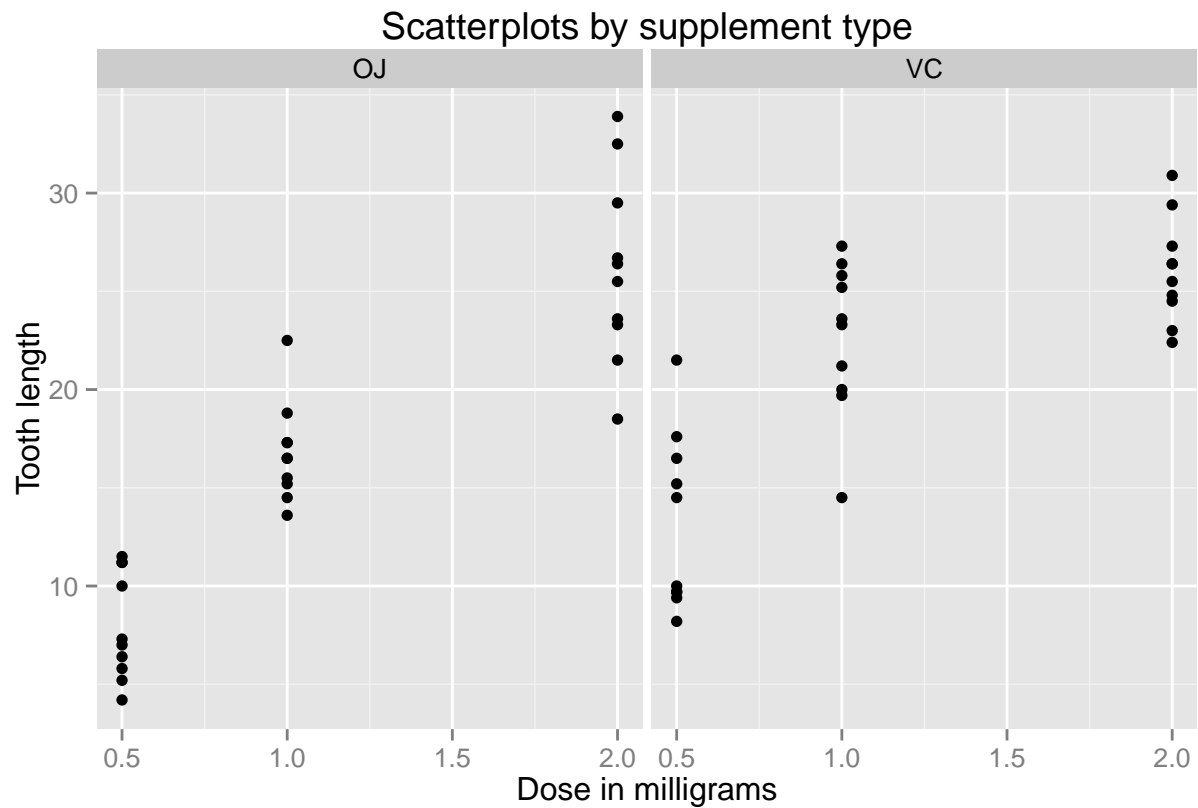
```
##      len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
help("ToothGrowth")
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
```

2. Provide a basic summary of the data

```
library(ggplot2)
plot <- ggplot(data = ToothGrowth, aes(x=ToothGrowth$dose, y = ToothGrowth$len))+
  ggtitle("Scatterplots by supplement type") +
  xlab("Dose in milligrams") +
  ylab("Tooth length") +
  geom_point(color = 'black', fill = 'white') +
  facet_wrap(~supp)
plot
```



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)

First of all we use the t.test to check if the difference of Vitamin C supplement type influence the tooth growth. We do this by looking at the p value:

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

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

In this case it is not < 0.05 , so we cannot reject the null hypothesis and we can conclude that there is no evidence for a correlation between supplement type and tooth growth.

We now check if the dose can influence tooth growth always by looking at the p value:

```
t.test(ToothGrowth$len[ToothGrowth$dose==2],
       ToothGrowth$len[ToothGrowth$dose==1],
       paired = FALSE,
       var.equal = TRUE)

##
## Two Sample t-test
##
## data: ToothGrowth$len[ToothGrowth$dose == 2] and ToothGrowth$len[ToothGrowth$dose == 1]
## t = 4.9005, df = 38, p-value = 1.811e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 3.735613 8.994387
## sample estimates:
## mean of x mean of y
## 26.100 19.735
```

this time since the p value is < 0.05 we have enough evidence to reject the null hypothesis. This means that there is a correlation between a dose change from 1 to 2 mg and the tooth growth.

```
t.test(ToothGrowth$len[ToothGrowth$dose==1],
       ToothGrowth$len[ToothGrowth$dose==0.5],
       paired = FALSE,
       var.equal = TRUE)

##
## Two Sample t-test
##
## data: ToothGrowth$len[ToothGrowth$dose == 1] and ToothGrowth$len[ToothGrowth$dose == 0.5]
## t = 6.4766, df = 38, p-value = 1.266e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 6.276252 11.983748
## sample estimates:
## mean of x mean of y
## 19.735 10.605
```

The confidence interval does not contain zero. There is no need to compare the dosage of 0.5 and 2 mg.

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

We can conclude saying that there is no correlation between supplement type and tooth growth. On the other hand while increasing the dosage of vitamin C we can see a growth in the guinea pig's teeth.