The Effect of Vitamin C on Tooth Growth in Guinea Pigs

TK

9 septembre 2016

Statistical Inference Course Project Part 2

Basic Inferential Data Analysis Instructions

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

1. We 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.
4. State our conclusions and the assumptions needed for your conclusions.

Description data:

The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs.

Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, (orange juice or ascorbic acid (a form of vitamin C and coded as VC).

Usage:

ToothGrowth

Format:

A data frame with 60 observations on 3 variables.

[,1] len numeric Tooth length [,2] supp factor Supplement type (VC or OJ). [,3] dose numeric Dose in milligrams/day

Load the ToothGrowth data

```
library("ggplot2")
library("datasets")
data <- ToothGrowth
str(data)
```

```
## '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 ...
```

```r
names(data)
```

```
## [1] "len"  "supp" "dose"
```

```r
head(data)
```

```
##    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
```

Basic summary of the data

```r
summary(data)
```

```
##       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
```

```r
unique(data$len)
```

```
##  [1]  4.2 11.5  7.3  5.8  6.4 10.0 11.2  5.2  7.0 16.5 15.2 17.3 22.5 13.6
## [15] 14.5 18.8 15.5 23.6 18.5 33.9 25.5 26.4 32.5 26.7 21.5 23.3 29.5 17.6
## [29]  9.7  8.2  9.4 19.7 20.0 25.2 25.8 21.2 27.3 22.4 24.5 24.8 30.9 29.4
## [43] 23.0
```

```r
unique(data$supp)
```

```
## [1] VC OJ
## Levels: OJ VC
```

```r
unique(data$dose)
```

```
## [1] 0.5 1.0 2.0
```

```r
table(data$dose, data$supp)
```

```
##
##       OJ VC
##   0.5 10 10
##   1   10 10
##   2   10 10
```

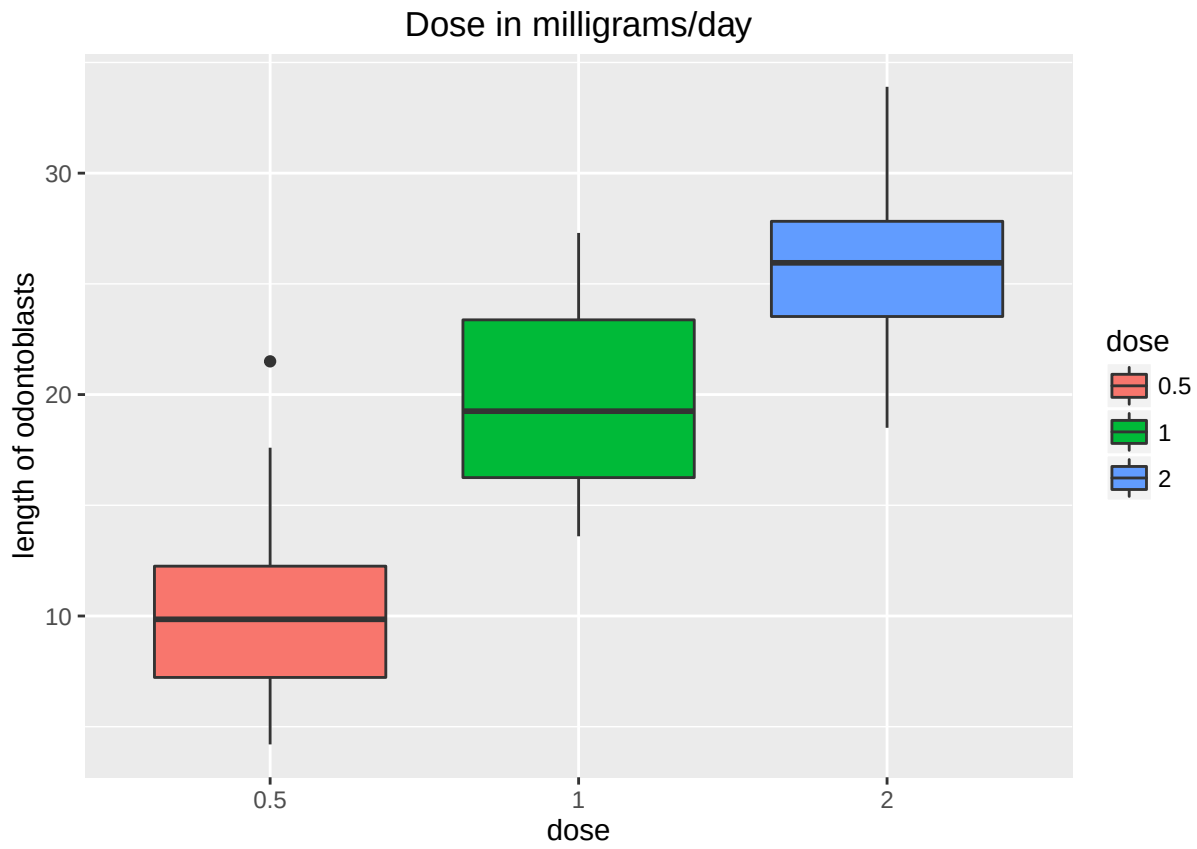
```
# We convert variable "dose" to a factor

data$dose <- as.factor(data$dose)
```

Exploration of data and construction of plot

1 plot: dose of vitamin C ~ length of odontoblasts

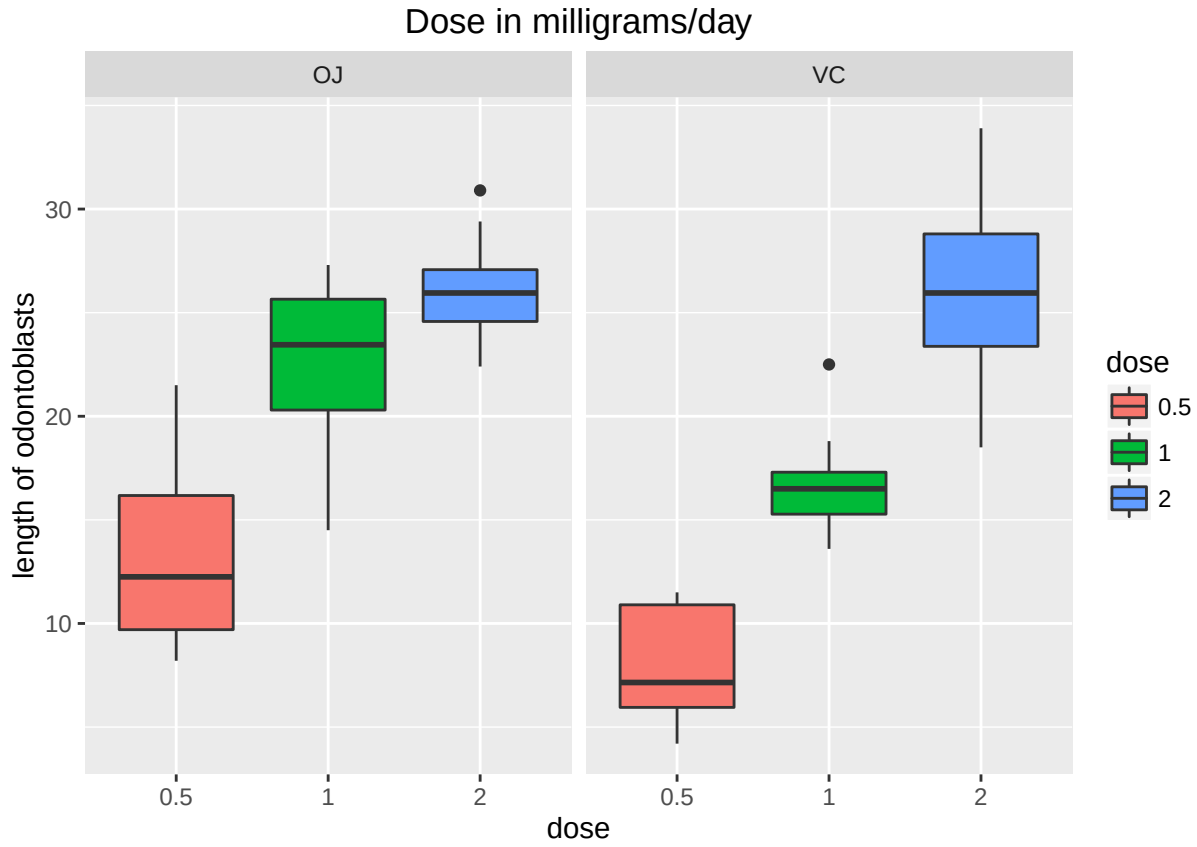
```
g <- ggplot(aes(x=dose, y=len), data= data) + geom_boxplot(aes(fill=dose))+
        labs(title="Dose in milligrams/day", x= "dose", y= "length of odontoblasts")
g
```



Length of odontoblasts increase if dose increase.

2 plot: how it depends on the type of supplement

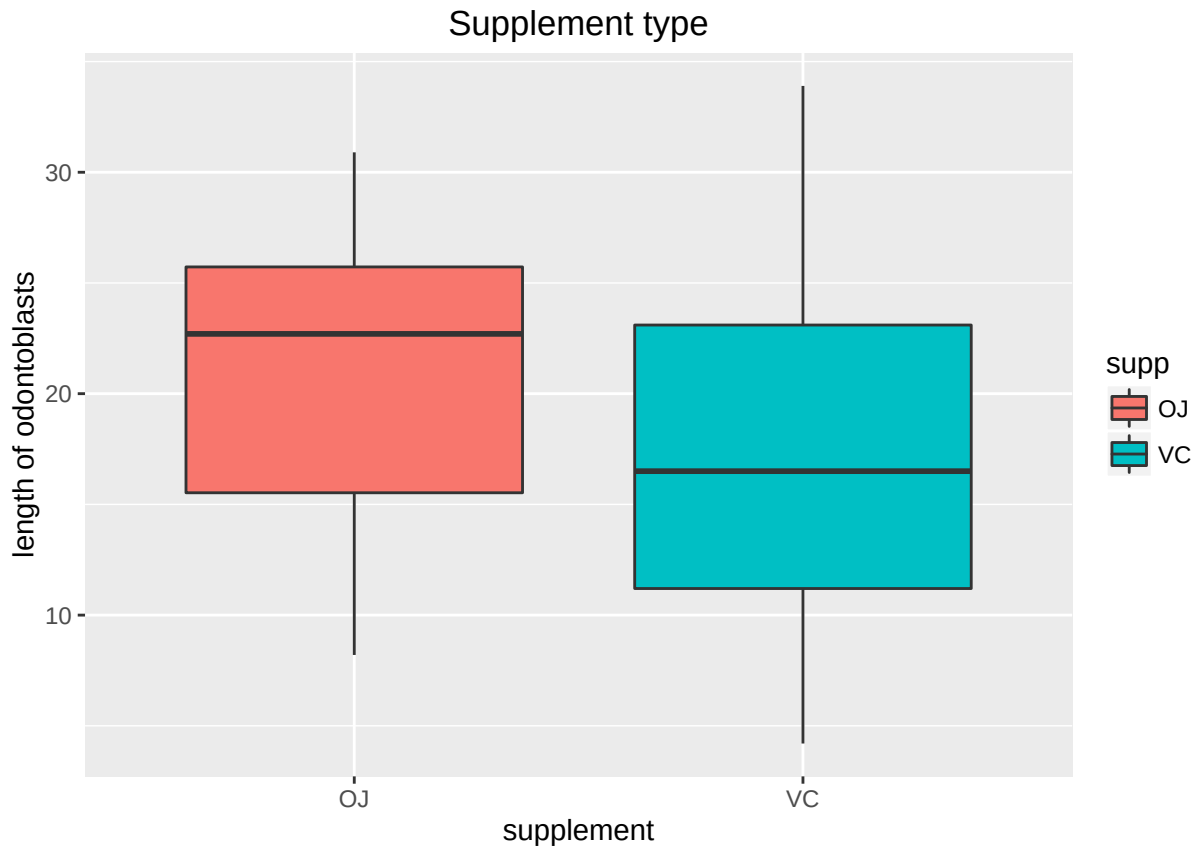
```
x <- ggplot(aes(x=dose, y=len), data= data) + geom_boxplot(aes(fill=dose))+
        facet_grid(.~supp)+
        labs(title="Dose in milligrams/day", x= "dose", y= "length of odontoblasts")
x
```



With dose = 2 mg the difference between average of the two groups decreases.

3 plot: type of supplement ~ length of odontoblasts

```
p <- ggplot(aes(x=supp, y=len), data= data) + geom_boxplot(aes(fill=supp))+
        labs(title="Supplement type", x= "supplement", y= "length of odontoblasts")
p
```



Length of odontoblasts increase if vitamin C delivery with orange juice.

Analyse tooth growth by supplement and dose of vitamin C

Now we start analyse:

We want to estimate the difference in tooth growth with the administration of vitamin C with orange juice and how it depends on the dose of vitamin C.

```
t.test(len ~ supp, data=data)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##         20.66333         16.96333
```

```r
t.test(len ~ supp, data=data[data$dose == 0.5,])
```

```
##
##  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
```

```r
t.test(len ~ supp, data=data[data$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
```

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

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

This test shows us that the group with the given vitamin C with orange juice has an average value of length of odontoblasts more important than the VC group.

But if the dose of vitamin C increases this difference is less important.