Statistical Inference Course Project 2

Analysis of the ToothGrowth of Guinea pigs

Humbert Costas

22 de agosto de 2015

Overview

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

Basic exploratory data analyses

Data shows the results of a study on the effects that Vitamic C has on the growth (length) of teeth of guinea pigs. Vitamin C was provided either in the form of Orange Juice (OJ) or ascorbic acid (VC), at three different dose levels to 10 animals, giving a total of 60 different sample values.

```
data(ToothGrowth)
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
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: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 1 ...
```

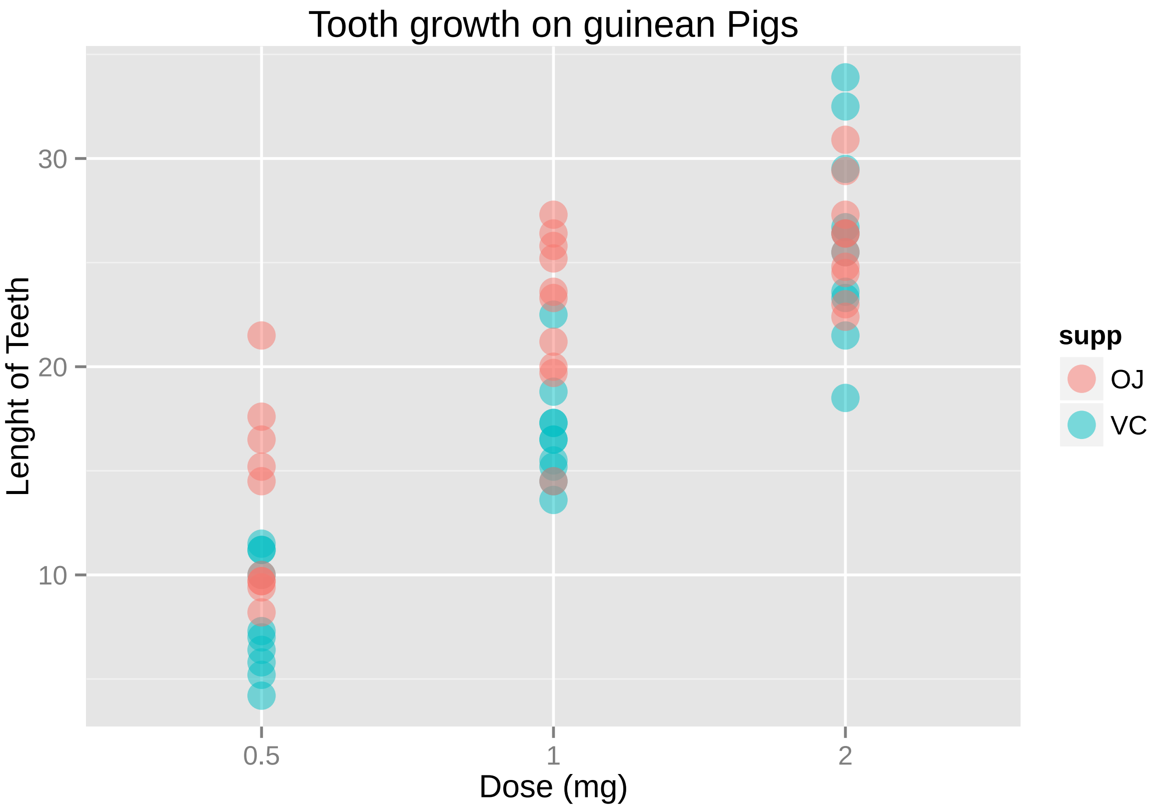
```
summary(ToothGrowth)
```

```
##       len        supp     dose
##  Min.   : 4.20   OJ:30   0.5:20
##  1st Qu.:13.07   VC:30   1  :20
##  Median :19.25           2  :20
##  Mean   :18.81
##  3rd Qu.:25.27
##  Max.   :33.90
```

Summary of the data.

Let's take a look of the data distribution with a scattered points graph

```
library(ggplot2)
g <- ggplot(ToothGrowth, aes(x = dose,y = len, color = supp))
g <- g + geom_point(alpha = 0.5, size = 5)
g <- g + ylab("Lenght of Teeth")
g <- g + xlab("Dose (mg)")
g <- g + ggtitle("Tooth growth on guinean Pigs")
g
```



By looking at the graph, we could say that at doses =0,5 and =1.0 OJ seems to have a clear effect on the length of teeth, compared to VC. On the other hand, when the dose =2.0 the results are less clear.

Comparison tooth growth by supp and dose. (using confidence intervals and hypothesis tests)

To compare the means of length of tooth when given OJ and VC at each dose level, we will let μ_{oj} be the mean of tooth growth with OJ and μ_{vs} the mean of tooth growth with VC. We will assume unequal variance (we can't assure that guinea pigs are from the same population) and will perform a t test, because each test was ran in groups of 10 guinea pigs.

Test if OJ and VC has equality for each dosage

$H_0 : \mu_{oj} - \mu_{vc} = 0$
$H_a : \mu_{oj} - \mu_{vc} \neq 0$

```
t_0.5 <- subset(ToothGrowth, dose == "0.5")
t <- t.test(len ~ supp, data = t_0.5, paired = FALSE, var.equal = FALSE)
t
```

Dose of 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
```

The null hypothesis is rejected because p-value (*0.0063586*) is less than 0.05. So, we can say that the tooth growth from orange juice (OJ) is different from ascorbic acid (VC)

```r
t_1 <- subset(ToothGrowth, dose == "1")
t <- t.test(len ~ supp, data = t_1, paired = FALSE, var.equal = FALSE)
t
```

Dose of 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 null hypothesis is rejected because p-value (*0.0010384*) is less than 0.05. So, we can say that the tooth growth from orange juice (OJ) is different from ascorbic acid (VC)

```r
t_2 <- subset(ToothGrowth, dose == "2")
t <- t.test(len ~ supp, data = t_2, paired = FALSE, var.equal = FALSE)
t
```

Dose of 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
```

The null hypothesis is accepted because p-value (*0.9638516*) is greater than 0.05. So, we can say that the tooth growth from orange juice (OJ) has no differences from ascorbic acid (VC)

Test if OJ affects more on tooth growth than VC

$H_0 : \mu_{oj} - \mu_{vc} \leq 0$
$H_a : \mu_{oj} - \mu_{vc} > 0$

```r
t <- t.test(len ~ supp, data = t_0.5, alternative = "greater", paired = FALSE, var.equal = FALSE)
t
```

Dose of 0.5

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.003179
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  2.34604     Inf
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

The null hypothesis is rejected because p-value(*0.0031793*) is smaller than 0.05. So, we can say that orange juice (OJ) affects more on tooth growth than ascorbic acid (VC).

```r
t <- t.test(len ~ supp, data = t_1, alternative = "greater", paired = FALSE, var.equal = FALSE)
t
```

Dose of 1

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.0005192
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
```

```
## 3.356158      Inf
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

The null hypothesis is rejected because p-value(*0.00052*) is smaller than 0.05. So, we can say that orange juice (*OJ*) affects more on tooth growth than ascorbic acid (*VC*).

```
t <- t.test(len ~ supp, data = t_2, alternative = "greater", paired = FALSE, var.equal = FALSE)
t
```

Dose of 2

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.0461, df = 14.04, p-value = 0.5181
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -3.1335     Inf
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

The null hypothesis is accepted because p-value(*0.5180742*) is greater than 0.05. So, we can't reach a conclusion if orange juice (*OJ*) affects more or less on tooth growth than ascorbic acid (*VC*).

Conclusions

- **On doses of less than 2mg** we can say that **Orange Juice has more effect** of tooth growth **than Ascorbic Acid**
- On doses of 2mg OJ and VC seem to have equal effect on tooth growth