Analyses of Tooth Growth

Author: Pieter Overdevest

Overview

In this short report we intend to investigate tooth growth for two supplements (OJ and VC) at 3 dosage levels (0.5, 1 and 2 mg).

Basic Exploratory Data Analyses

We load the data and perform some basic analyses to get a feel for the data. After loading the data, we take a look at the dimension and top part of the dataframe,

```
data(ToothGrowth)
dim(ToothGrowth)
```

```
## [1] 60  3
```

```
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
```

This shows the data consists of 60 rows and 3 columns. Now, we report the summary,

```
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
```

Tooth growth length is between 4.2 and 33.9, and dosage is between 0.5 and 2. There are 2 supplements, OJ and VC. Both occur 50% of the times. The ‘table’ function helps to understand how many data points we have per dose and supplement,

```
table(ToothGrowth[,2:3])
```

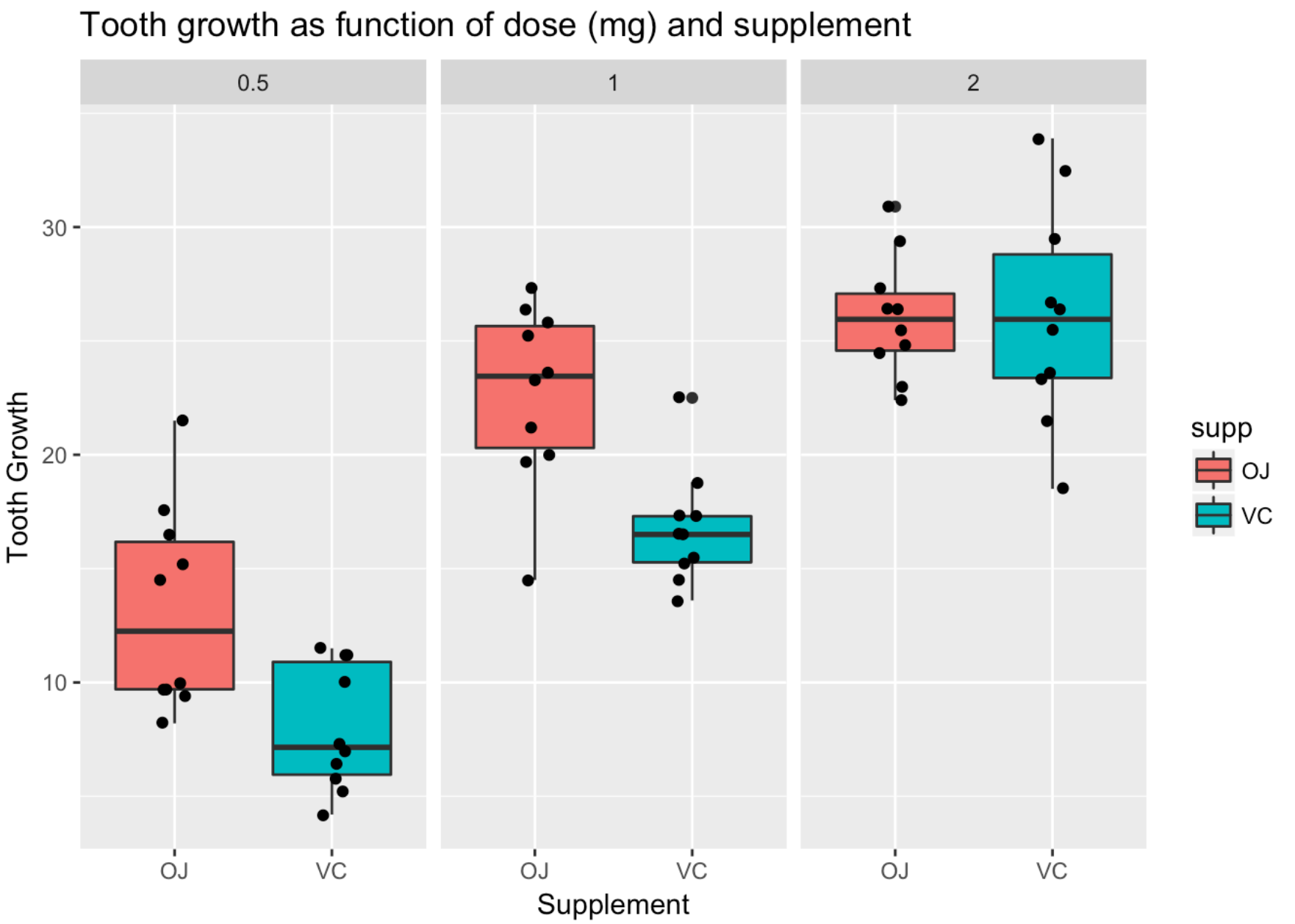
```
##     dose
## supp 0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

This shows that each permutation contains 10 observations.

After checking out the numbers, it is time to make a plot. Let's look at a boxplot where we split the data by dosage (0.5, 1, 2) and supplement (OJ, VC),

```
library(ggplot2)

g <-  ggplot(ToothGrowth, aes(supp, len)) +
        geom_boxplot(aes(fill = supp)) + geom_jitter(width = 0.1) +
        facet_wrap(~ dose, nrow = 1, ncol = 3) +
        labs(title = "Tooth growth as function of dose (mg) and supplement") +
        labs(x = "Supplement", y ="Tooth Growth")
g
```



The boxplots suggest that there may be a difference in tooth growth between supplements at dosage 0.5 and 1 mg. At a dosage of 2 mg, tooth growths appear to be the same for the two supplements.

Hypothesis Testing

Assumptions

When testing the difference between tooth growth, we consider the following,

- Tooth growth is normally distributed.
- Variables must be independent and identically distributed (i.i.d.).
- Variances of tooth growth depends on supplement and dosage, and can thus be different.

Test differences in tooth growths

From the boxplot we concluded that some combinations may be different. In order to test this with a t-test, we pull out the data for each of the combinations. This makes it easy to refer to the data in the t.test function, which we will do later. For example, GROWTHOJHALF is the tooth growth data for supplement OJ at dosage 0.5 mg,

```
GROWTHOJHALF <- ToothGrowth$len[(ToothGrowth$supp=="OJ" & ToothGrowth$dose==0.5)]
GROWTHVCHALF <- ToothGrowth$len[(ToothGrowth$supp=="VC" & ToothGrowth$dose==0.5)]
GROWTHOJONE <- ToothGrowth$len[(ToothGrowth$supp=="OJ" & ToothGrowth$dose==1)]
GROWTHVCONE <- ToothGrowth$len[(ToothGrowth$supp=="VC" & ToothGrowth$dose==1)]
GROWTHOJTWO <- ToothGrowth$len[(ToothGrowth$supp=="OJ" & ToothGrowth$dose==2)]
GROWTHVCTWO <- ToothGrowth$len[(ToothGrowth$supp=="VC" & ToothGrowth$dose==2)]
```

Let's compare supplements OJ and VC at dosage 0.5 mg,

```
t.test(GROWTHOJHALF, GROWTHVCHALF, alternative = "two.sided", mu=0, paired=FALSE,
var.equal=FALSE, conf.level=0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  GROWTHOJHALF and GROWTHVCHALF
## 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 of x mean of y
##     13.23      7.98
```

We observe a p-value of 0.0064. If we consider a significance level (alpha) of 0.05, we need to reject the null hypothesis. In other words, both supplements are indeed different at dosage level 0.5 mg.

Secondly, we compare supplements OJ and VC at dosage 1 mg,

```
t.test(GROWTHOJONE, GROWTHVCONE, alternative = "two.sided", mu=0, paired=FALSE, va
r.equal=FALSE, conf.level=0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  GROWTHOJONE and GROWTHVCONE
## 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 of x mean of y
##     22.70     16.77
```

We observe a p-value of 0.0010. If we consider a significance level (alpha) of 0.05, we need to reject the null hypothesis. In other words, both supplements are indeed different at dosage level 1 mg.

Finally, we compare supplements OJ and VC at dosage 2 mg,

```
t.test(GROWTHOJTWO, GROWTHVCTWO, alternative = "two.sided", mu=0, paired=FALSE, va
r.equal=FALSE, conf.level=0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  GROWTHOJTWO and GROWTHVCTWO
## 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
```

We observe a p-value of 0.96. If we consider a significance level (alpha) of 0.05, we cannot reject the null hypothesis. In other words, both supplements are the same at dosage level 2 mg.