Exploring the ToothGrowth Dataset

Marco Totolo

Overview

In this document we will analyze the ToothGrowth dataset and perform hypothesis testing in order to decide which of two delivery methods is more effective. The conclusion is that the first delivery method is better than the second for lower dose levels, but we cannot say if the first delivery method is better or worse than the second for the highest dose level.

Exploratory Analysis

First let's load the data and some usuful libraries.

```
library(datasets)
library(dplyr)
library(ggplot2)
data("ToothGrowth")
```

From the description of the dataset we know that

"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)."

In fact, we have 60 records of three variables:

```
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
```
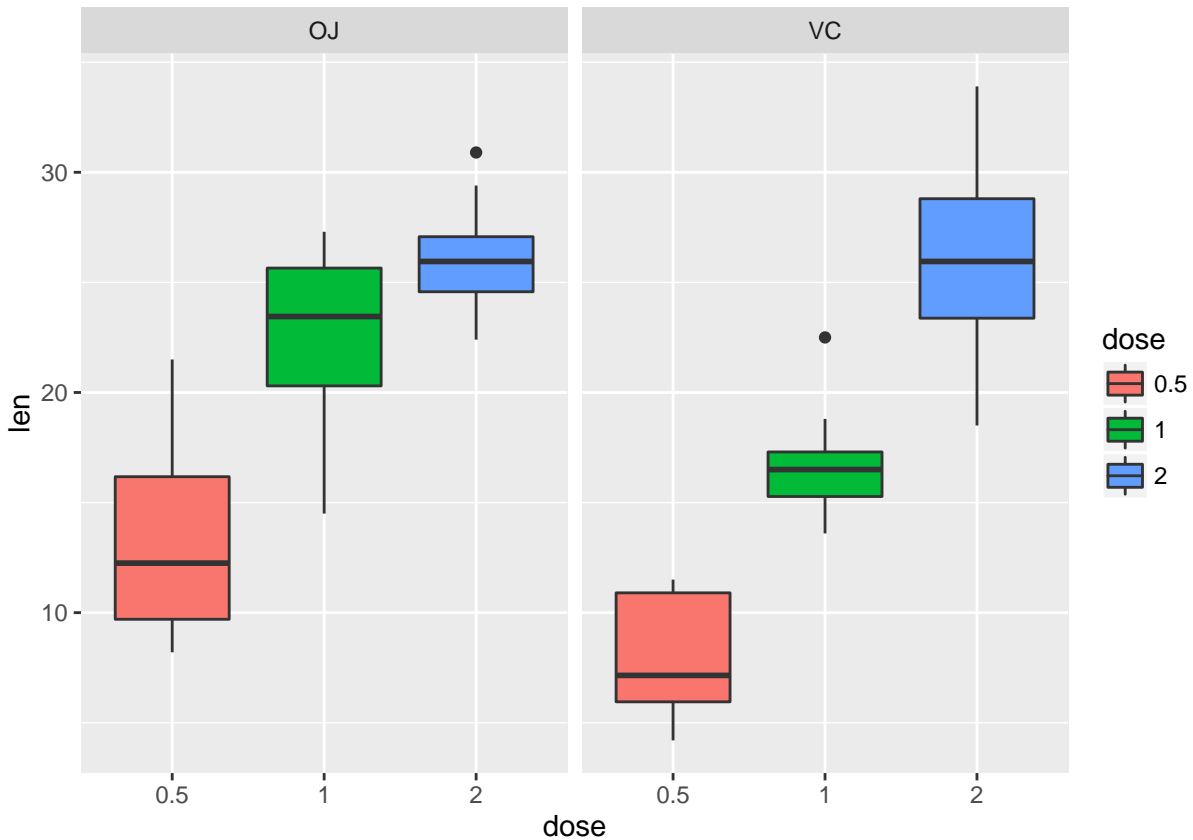
And we have 10 records for each supp/dose combination:

```
with(ToothGrowth,table(supp,dose))
```

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

Let's plot them and see if there's a recognizable pattern.

```
ToothGrowth$dose<-as.factor(ToothGrowth$dose)
g<-ggplot(ToothGrowth,aes(x=dose,y=len, fill=dose))
g+geom_boxplot()+facet_grid(.~supp)
```


It seems that for both delivery methods, increasing the dose also increses the tooth growth. Orange juice seems to be more effective, except for the higher dose level where they have comparable effects.

Hypothesis Testing

Let's try and test one of the above observations.

We want to test whether the two delivery methods have the same effect for dose=1. The null and alternative hypothesis are the following:

H_0: the means of the two samples are the same

H_a: the means of the two samples are different

We will therefore perform a two sided t-test, assuming unequal variances (it's the more general case) and not paired samples (the guinea pigs were different for the two trials).

```
g1<-ToothGrowth %>% filter(supp=="OJ" & dose==1) %>% select(len)
g2<-ToothGrowth %>% filter(supp=="VC" & dose==1) %>% select(len)
t.test(g1,g2,var.equal=F,paired=F)
```

```
##
##  Welch Two Sample t-test
##
## data:  g1 and g2
## 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
```

The 5% confidence interval lies above zero: I have to reject the null hypothesis and the alternative hypothesis is likely to be true.

Now let's perform multiple testing on all possible combinations of the 6 subgroups (dose/supp pairs). This means 15 t-tests. We will use the same assumptions as before.

```r
pValues<-NULL
TestName<-NULL
g<-NULL
## create the 6 subgroups (supp/dose pairs) and their names
for (i in unique(ToothGrowth$supp)) {
    for (j in unique(ToothGrowth$dose)) {
        g1<- ToothGrowth %>% filter(supp==i & dose==j) %>%
            select(len)
        g<-c(g,g1)
        TestName<-c(TestName,paste0(i,j))
    }
}

## create the 15 combinations to be tested and their names

g.pairs<-t(combn(g,2))
TestName.pairs<-apply(t(combn(TestName,2)),1,paste0,collapse="=")
#TestName.pairs<-c(TestName.pairs,
#                  apply(t(combn(TestName,2)),1,paste0,collapse="<"))

## perform 15 t-tests

for (i in 1:nrow(g.pairs)) {
    pValues<-c(pValues,t.test(g.pairs[i,][[2]],g.pairs[i,][[1]],
                              alt="two.sided")$p.value)
}
#for (i in 1:nrow(g.pairs)) {
#    pValues<-c(pValues,t.test(g.pairs[i,][[2]],g.pairs[i,][[1]],
#                              alt="less")$p.value)
#}
```

The list of 15 null hypothesis tested is the following:

```r
TestName.pairs
```

```
## [1] "VC0.5=VC1"   "VC0.5=VC2"   "VC0.5=OJ0.5" "VC0.5=OJ1"   "VC0.5=OJ2"
```

```
## [6] "VC1=VC2"       "VC1=OJ0.5"     "VC1=OJ1"       "VC1=OJ2"       "VC2=OJ0.5"
## [11] "VC2=OJ1"       "VC2=OJ2"       "OJ0.5=OJ1"     "OJ0.5=OJ2"     "OJ1=OJ2"
```

Of those 15, we have successfully rejected with a 5% level 13 of them:

```
sum(pValues<0.05)
```

```
## [1] 13
```

Actually, since we have performed multiple testing, we have to correct the p-values to ensure our 5% error rate. I will use the "BH" method.

```
pValues<-p.adjust(pValues, method = "BH")
```

The new number of significant results is 12:

```
sum(pValues<0.05)
```

```
## [1] 12
```

The list of *rejected* null hyopthesis is therefore the following:

```
TestName.pairs[(pValues<0.05)]
```

```
##  [1] "VC0.5=VC1"    "VC0.5=VC2"    "VC0.5=OJ0.5" "VC0.5=OJ1"    "VC0.5=OJ2"
##  [6] "VC1=VC2"      "VC1=OJ1"      "VC1=OJ2"      "VC2=OJ0.5"    "OJ0.5=OJ1"
## [11] "OJ0.5=OJ2"    "OJ1=OJ2"
```

And the hypothesis that we failed to reject are:

```
TestName.pairs[(pValues>0.05)]
```

```
## [1] "VC1=OJ0.5" "VC2=OJ1"    "VC2=OJ2"
```

Conclusions

We can therefore safely state (with 5% error rate) that our inital observations were correct:

- For both delivery methods, a higher dose produces a higher effect
- Doses being equal, orange juice is more effective, except for the higher dose level