Statistical Inference Assignement
ToothGrowth Datas Analysis

by Bruno Berrehuel

November 14, 2016

1 Overview

The datas are about the length of odontoblast (cells responsible for tooth growth) for 60 pigs after an experimental threatment in vitamin C. Each animal received a dose of vitamin C, from 0.5 to 2 mg/day, by orange juice (OJ) or ascorbic acid (VC)[1]. Each experiment deals with 10 pigs.
I'll try to answer the two following questions :

1. Q1 : is the vitamin C effective for tooth growth ?
2. Q2 : what is the best delivery method between orange juice or ascorbic acid ?

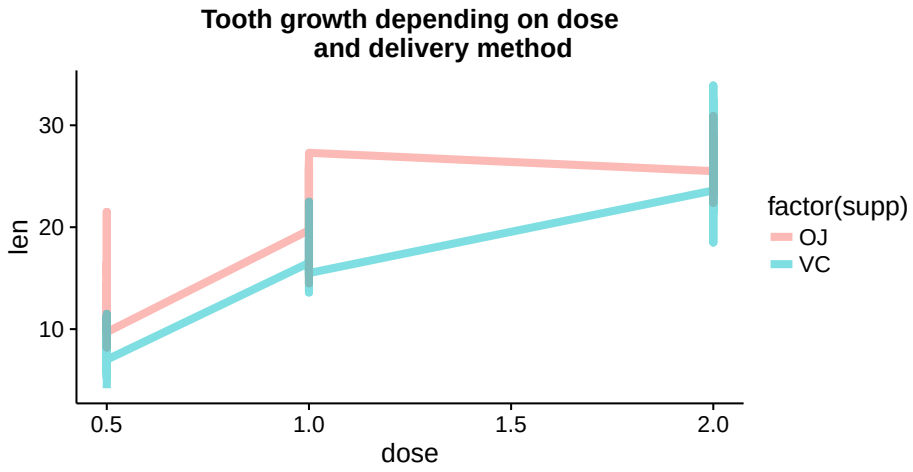
2 Data exploratory

First I'll take some informations about the means and the standard deviation of each dose, and plot the evolution of the tooth growth depending on the dose and the delivery method :

```
ToothGrowth %>% group_by(supp,dose) %>%
    summarise(mean(len), round(sd(len),3))

## Source: local data frame [6 x 4]
## Groups: supp [?]
##
##     supp  dose `mean(len)` `round(sd(len), 3)`
##   <fctr> <dbl>       <dbl>               <dbl>
## 1     OJ   0.5       13.23               4.460
## 2     OJ   1.0       22.70               3.911
## 3     OJ   2.0       26.06               2.655
## 4     VC   0.5        7.98               2.747
## 5     VC   1.0       16.77               2.515
## 6     VC   2.0       26.14               4.798
```

[1] Source : https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/ToothGrowth.html



It seems that :

- H1 : odontoblasts are taller with a bigger dose of vitamin C, no matter of the delivery method.
- H2 : orange juice (OJ) has better results than ascorbic acid (VC).

3 Probability tests

I define some variables form the ToothGrowth datasets, in order to compare with the R function **t.test** :

- H1 : the mean of odontoblasts length at 0.5 mg/day dose with the mean at 2 mg/day
- H2 : the means of odontoblasts length for the 2 delivery method ar each vitamin C dose.

3.1 Tests for H1 hypothesis

First filter the ToothGrowth dataset depending on dose of vitamin C, for both of the delivery method, then do t.test to confirm that the mean at 2 mg/day is more important than the mean at 0.5 mg/day.

```
vita05 <- ToothGrowth %>% filter(dose==0.5)
vita2 <- ToothGrowth %>% filter(dose==2)
t.test(vita2$len-vita05$len, alternative="greater")

##
##  One Sample t-test
##
## data:  vita2$len - vita05$len
## t = 11.291, df = 19, p-value = 3.595e-10
```

```
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  13.12216      Inf
## sample estimates:
## mean of x
##    15.495
```

3.2 Tests for H2 hypothesis

First filter the ToothGrowth dataset depending on dose of vitamin C and delivery method, then compare them with the t.test function.

```
indice <- c(0.5,1,2)
OJ <- lapply(indice, function(i) {ToothGrowth %>% filter(supp=="OJ") %>%
             filter(dose==i)})
VC <- lapply(indice, function(i) {ToothGrowth %>% filter(supp=="VC") %>%
             filter(dose==i)})
t.test(OJ[[1]]$len-VC[[1]]$len, alternative="greater")

##
##  One Sample t-test
##
## data:  OJ[[1]]$len - VC[[1]]$len
## t = 2.9791, df = 9, p-value = 0.007736
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  2.019552      Inf
## sample estimates:
## mean of x
##      5.25

t.test(OJ[[2]]$len-VC[[2]]$len, alternative="greater")

##
##  One Sample t-test
##
## data:  OJ[[2]]$len - VC[[2]]$len
## t = 3.3721, df = 9, p-value = 0.004115
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  2.706401      Inf
## sample estimates:
## mean of x
##      5.93

t.test(OJ[[3]]$len-VC[[3]]$len)
```

```
##
##  One Sample t-test
##
## data:  OJ[[3]]$len - VC[[3]]$len
## t = -0.042592, df = 9, p-value = 0.967
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -4.328976  4.168976
## sample estimates:
## mean of x
##     -0.08
```

We have the following results :

- we can reject $\mu_{OJ05} = \mu_{VC05}$ and accept $\mu_{OJ05} > \mu_{VC05}$
- we can reject $\mu_{OJ1} = \mu_{VC1}$ and accept $\mu_{OJ1} > \mu_{VC1}$
- we failed to reject $\mu_{OJ2} = \mu_{VC2}$

4 Conclusions

4.1 Is the vitamin C effective for tooth growth ?

As the t.test shows that the real mean of the 2 mg/day tooth growth is greater than the 0.5 mg/day tooth growth, the vitamin C threatment is effective for the pigs.

4.2 What is the best delivery method ?

The t.tests show that :

- the orange juice delivery method is more effective than the ascorbic acid method for small doses, as the real tooth growth means are greater for orange juice for 0.5 and 1 mg/day doses.
- the methods are equivalent for the dose 2 mg/day, as we failed to reject the null hypothesis that the means are equal.

I recommend the delivery of vitamin C, by orange juice, for the tooth growth of the guinea pigs.