The effect of vitamin C on tooth growth in Guinea Pigs

Ilya Semenov

March 29, 2016

Overview

Original dataset 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). Goal of analysis is to identify whether the effect of vitamin C on teeth growth is positive and whether the method of recieving vitamin C matters.

Initial data processing

Following procedure were used for loading data.

```
data(ToothGrowth)
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: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
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
```

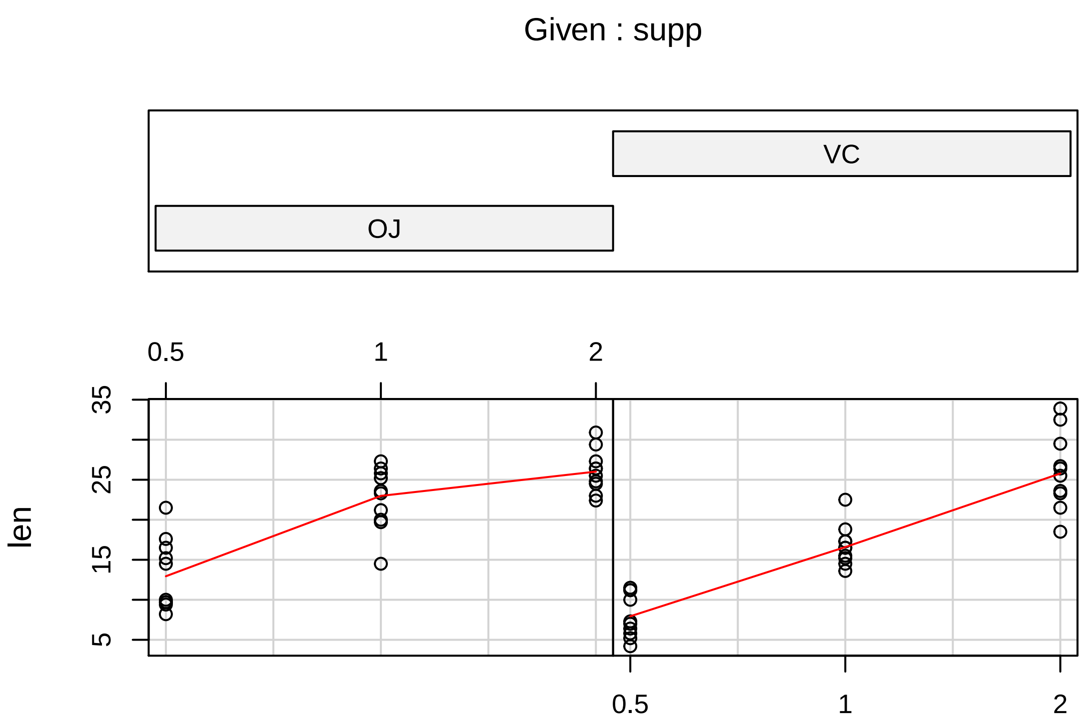
Since we are going to reveal the influence of a vitamin C dose and observations include only three dose types, let factorize field 'dose'.

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

Exploratory data analysis

First let plot the data separately for vitamin C supplement types.

```
coplot(len ~ dose | supp, data = ToothGrowth, panel = panel.smooth,
       xlab = "ToothGrowth data: length vs dose, given type of supplement")
```



TooothGrowth data: length vs dose, given type of supplement

Plot suggests hypothesis of positive vitamin C effect with larger doses. Let examine delivery methods more closely in order to clarify null hypothesis.

```
library(dplyr, quietly = TRUE, warn.conflicts = FALSE)
ToothGrowth %>% group_by(supp, dose) %>% summarise(mean(len))
```

```
## Source: local data frame [6 x 3]
## Groups: supp [?]
##
##     supp   dose mean(len)
##   (fctr) (fctr)     (dbl)
## 1     OJ    0.5     13.23
## 2     OJ      1     22.70
## 3     OJ      2     26.06
## 4     VC    0.5      7.98
## 5     VC      1     16.77
## 6     VC      2     26.14
```

It seems that orange juice is more effective than ascorbic acid, at least with small or medium doses. This consideration causes separate hypothesis check for different dose levels.

Positive vitamin C effect test using Student's t-Test

First let check if supplement method of vitamin C matters. We'll test null hypothesis that orange juice supplement has same mean as ascorbic acid. Let split tooth lengthes into two vectors according to supplement method.

```r
splitsupp <- split(ToothGrowth$len, ToothGrowth$supp)
```

According to dataset description we can assume experiments weren't conducted pairwise. Let check if variances for different methods are close enough to assume them equal.

```r
sapply(splitsupp, var)
```

```
##       OJ       VC
## 43.63344 68.32723
```

Since two variances seem pretty different, we'll set var.equal parameter in t-Test to FALSE.

```r
t.test(splitsupp$OJ, splitsupp$VC, paired = FALSE, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  splitsupp$OJ and splitsupp$VC
## 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 of x mean of y
##  20.66333  16.96333
```

According to t-Test we fail to reject the null hypothesis that different supplement methods have same mean since test confidence interval (-0.17, 7.57) contains zero.

Let consider now if enough evidence are presented for positive effect of vitamin C dose on tooth growth.

Let's try to reject null hypothesis that pigs recieved vitamin dose of 0.5 and 1.0 have same mean of tooth length. We are going to perform two-sided test, our pigs are not paired and we can assume estimation of variance with pooled variance.

```r
t.test(ToothGrowth$len[ToothGrowth$dose == .5],
       ToothGrowth$len[ToothGrowth$dose == 1],
       paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$dose == 0.5] and ToothGrowth$len[ToothGrowth$dose == 1]
## t = -6.4766, df = 38, p-value = 1.266e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983748  -6.276252
## sample estimates:
## mean of x mean of y
##    10.605    19.735
```

P-value is below 5% significance level, there is no zero inside confidence interval of (-11.98, -6.28), so we can conclude positive effect of 1.0 mg vitamin dose over 0.5 mg dose.

Let's do the same in order to compare 1.0 mg dose with 2.0 mg dose.

```
t.test(ToothGrowth$len[ToothGrowth$dose == 1],
       ToothGrowth$len[ToothGrowth$dose == 2],
       paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$dose == 1] and ToothGrowth$len[ToothGrowth$dose == 2]
## t = -4.9005, df = 38, p-value = 1.811e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.994387 -3.735613
## sample estimates:
## mean of x mean of y
##    19.735    26.100
```

Once again p-value is below 5% significance level, there is no zero inside (-8.99, -3.74) confidence interval so we can confirm positive effect of 2.0 mg vitamin C dose over 1.0 mg dose.

Multiple comparisons

Let check our results for family wise error rate (FWER) and false discovery rate (FDR). We need to create a vector of four p-values for all the tests made.

```
pVector <- c(
  t.test(splitsupp$OJ, splitsupp$VC,
         paired = FALSE, var.equal = FALSE)$p.value,
  t.test(ToothGrowth$len[ToothGrowth$dose == .5],
       ToothGrowth$len[ToothGrowth$dose == 1],
       paired = FALSE, var.equal = TRUE)$p.value,
  t.test(ToothGrowth$len[ToothGrowth$dose == 1],
       ToothGrowth$len[ToothGrowth$dose == 2],
       paired = FALSE, var.equal = TRUE)$p.value
)
```

FWER control sequence:

```
sum(p.adjust(pVector, method = 'bonferroni') < 0.05)
```

```
## [1] 2
```

```
round(p.adjust(pVector, method = 'bonferroni'), 4)
```

```
## [1] 0.1819 0.0000 0.0001
```

FDR control sequence:

```r
sum(p.adjust(pVector, method = 'BH') < 0.05)
```

```
## [1] 2
```

```r
round(p.adjust(pVector, method = 'BH'), 4)
```

```
## [1] 0.0606 0.0000 0.0000
```

Only one p-value corresponding to supplement method test can be considered significant enough.

Summary

According to provided data and our study we are able to state the following:

- positive effect of vitamin C on Guinea pigs tooth growth confirmed. Higher vitamin dose making this effect more noticable.
- not enough data were provided to claim that vitamin C supplement method has any influence on pigs tooth growth.