

ToothGrowth Data Analysis

Hai Bacti

August 26, 2020

Synopsis

In this report, we're going to analyze the ToothGrowth data in the R datasets package. After make some basic exploratory of the data, we will use confidence intervals and hypothesis tests to conclude the impact on tooth growth by different supplement types and doses.

Data Analysis

The `ToothGrowth` dataset contains 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).

```
library(datasets)
dim(ToothGrowth)
```

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

```
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
```

A data frame with 60 observations on 3 variables: tooth length **len** (in numeric), supplement type **supp** (in factor VC or OJ) and dose **dose** (numeric in milligrams/day).

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

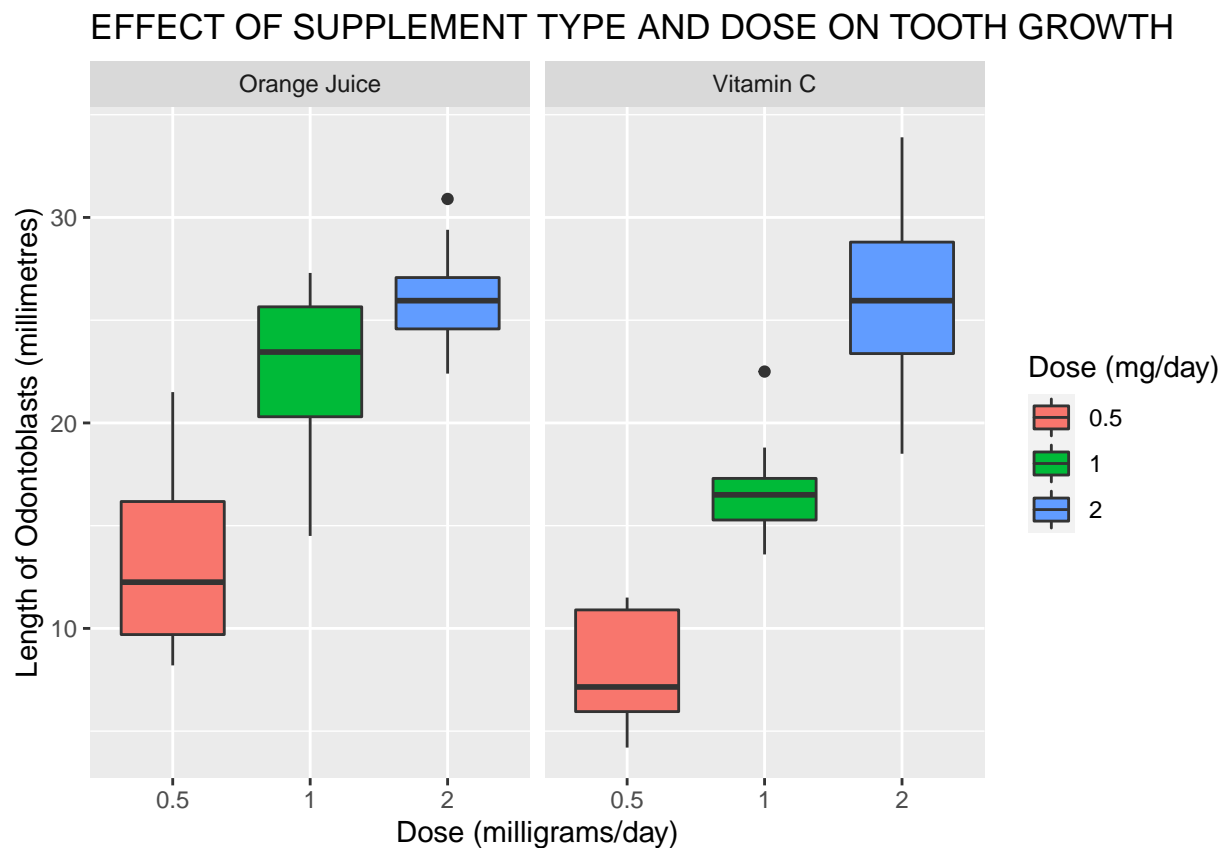
```
##      dose
## supp 0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

So there are 10 observations for each combination of supplement type and dose.

Hypothesis-Testing

```
library(dplyr)
library(ggplot2)
```

```
ggplot(ToothGrowth, aes(x = factor(dose), y = len, fill = factor(dose))) +
  geom_boxplot() +
  facet_grid(.~if_else(supp == 'OJ', 'Orange Juice', 'Vitamin C')) +
  xlab('Dose (milligrams/day)') +
  ylab('Length of Odontoblasts (millimetres)') +
  labs(fill = 'Dose (mg/day)') +
  ggtitle('EFFECT OF SUPPLEMENT TYPE AND DOSE ON TOOTH GROWTH')
```



The above plot indicates that the higher dose may cause larger teeth growth. Beside that, the impacts of Orange Juice are higher than Vitamin C at dosage from 0.5 to 1 mg/day while it's quite balance at dose 2 mg/day.

We define some utility functions to check the hypothesis.

```
GetLength <- function(s = '', d)
  ToothGrowth %>% filter(grepl(s, supp), dose == d) %>% select(len)

Testing <- function(x, y, alternative) {
  test <- t.test(x = x, y = y, alternative = alternative)
  c(p.value = test$p.value, conf.low = test$conf[1], conf.high = test$conf[2])
}
```

```
Testing(x = GetLength(d = 1), y = GetLength(d = 0.5), alternative = 'greater')
```

```
##      p.value      conf.low      conf.high
## 6.341504e-08 6.753323e+00      Inf
```

```
Testing(x = GetLength(d = 2), y = GetLength(d = 1), alternative = 'greater')
```

```
##      p.value      conf.low      conf.high
## 9.532148e-06 4.173870e+00      Inf
```

It failed to reject the alternative hypothesis H_a : **Higher dose may cause larger teeth growth.**

```
Testing(x = GetLength('OJ', 0.5), y = GetLength('VC', 0.5), alternative = 'greater')
```

```
##      p.value      conf.low      conf.high
## 0.003179303 2.346040347      Inf
```

It failed to reject the alternative hypothesis H_a : **Orange Juice cause larger impact than Vitamin C at dose 0.5 mg/day.**

```
Testing(x = GetLength('OJ', 1), y = GetLength('VC', 1), alternative = 'greater')
```

```
##      p.value      conf.low      conf.high
## 0.0005191879 3.3561576363      Inf
```

It failed to reject the alternative hypothesis H_a : **Orange Juice cause larger impact than Vitamin C at dose 1 mg/day.**

```
Testing(x = GetLength('OJ', 2), y = GetLength('VC', 2), alternative = 'two.sided')
```

```
##      p.value      conf.low      conf.high
## 0.9638516 -3.7980705 3.6380705
```

It rejected the alternative hypothesis H_a : **Orange Juice cause different impact from Vitamin C at dose 2 mg/day.**

Conclusion

The mainly results are:

- Higher dose may cause larger teeth growth no matter what supplement method.
- At dosage 0.5 and 1 mg/day, the impacts of Orange Juice are higher than Vitamin C.
- At dose 2 mg/day, both supply method show similar result.