# Basic Inferential Data Analysis Exercise: The Effect of Vitamin C on Tooth Growth in Guinea Pigs

Andrés Hidalgo Vargas

14/7/2020

## Exploratory Analysis

```
library(ggplot2)
library(dplyr)

data(ToothGrowth)
```

### Data description

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).

```
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 ...
```

There are 2 levels of the supp variable. "OJ" and "VC". These stand for *Orange Juice* and *Vitamin C* respectively.

```
unique(ToothGrowth$dose)
```

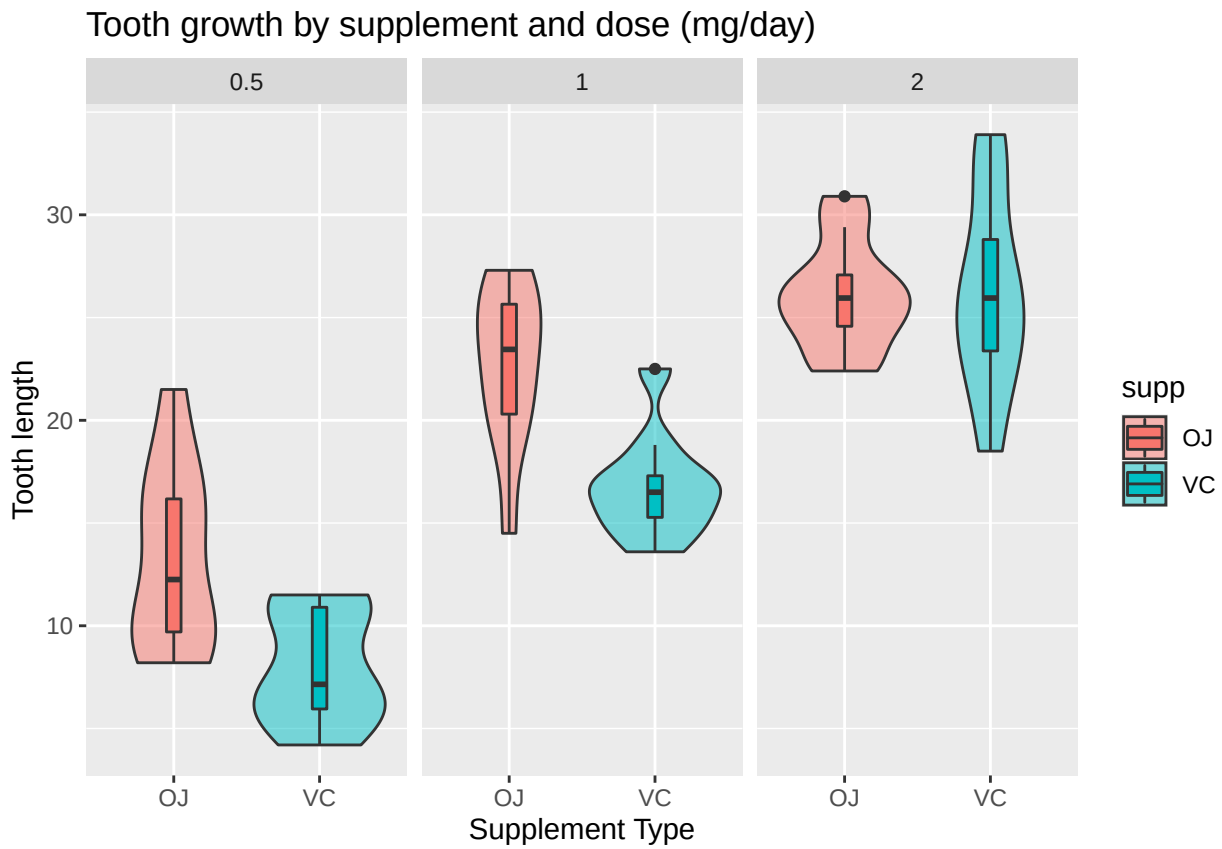
```
## [1] 0.5 1.0 2.0
```

There were 3 doses used. 0.5, 1.0 and 2.0.

```
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
```

```
ggplot(ToothGrowth, aes(x = supp, y = len, fill = supp)) +
  geom_violin(alpha = 1/2) +
  geom_boxplot(width = 0.1) +
  facet_wrap(~ dose)+
  xlab("Supplement Type")+
  ylab("Tooth length")+
  labs(title = "Tooth growth by supplement and dose (mg/day)")
```




Analyzing the violin plots, we will assume unequal variances in our hypothesis testing.

# Comparisons

## Supplement and lenght relationship

**H0:** mean odontoblast length of vitamin C subjects = mean odontoblas length of orange juice subjects

```
t.test(len ~ supp, data = ToothGrowth)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##         20.66333         16.96333
```

Given the result of the t test the p-value calculated 0.061 is greater tan our alpha 0.05. Therefore there is not sufficient statistical evidence to to reject the null hypothesis of equal means.

## Dose and length relationship

Due to the fact that in the analysis 3 different doses where used, we need to perform a `t.test` for each combination.

```
dose0.5 <- ToothGrowth %>%
  filter(dose == 0.5) %>%
  select(len)

dose1 <- ToothGrowth %>%
  filter(dose == 1) %>%
  select(len)

dose2 <- ToothGrowth %>%
  filter(dose == 2) %>%
  select(len)
```

### 0.5 with 1 mg/day

**H0:** mean odontoblast length of 0.5 mg/day dose = mean odontoblast length of 1 mg/day dose

```
t.test(dose0.5,dose1)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose0.5 and dose1
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean of x mean of y
##    10.605    19.735
```

p-value 1.268e-07 $< 0.05$, therefore there is enough statistical evidence to reject the null hypothesis.

### 0.5 with 2 mg/day

**H0:** mean odontoblast length of 0.5 mg/day dose = mean odontoblast length of 2 mg/day dose

```r
t.test(dose0.5,dose2)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose0.5 and dose2
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean of x mean of y
##    10.605    26.100
```

p-value 4.898e-14 $< 0.05$, therefore there is enough statistical evidence to reject the null hypothesis.

#### 1 with 2 mg/day

**H0:** mean odontoblast length of 1 mg/day dose = mean odontoblast length of 2 mg/day dose

```r
t.test(dose1,dose2)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose1 and dose2
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean of x mean of y
##    19.735    26.100
```

p-value 1.906e-05 $< 0.05$, therefore there is enough statistical evidence to reject the null hypothesis.

#### Dose and lenght relationship conclusion

The data suggests that there is a relationship between the length of odontoblasts and the dosage administered to the subjects.

## Supplement dosage level comparisons

Given that dosage seems like a variable that affects the odontoblasts length, we need to analyze if there is a relationship between the length and the combination of supplements and dosage levels.

```r
VC0.5 <- ToothGrowth %>%
  filter(supp == "VC" & dose == 0.5) %>%
  select(len)
```

```r
VC1 <- ToothGrowth %>%
  filter(supp == "VC" & dose == 1) %>%
  select(len)

VC2 <- ToothGrowth %>%
  filter(supp == "VC" & dose == 2) %>%
  select(len)

OJ0.5 <- ToothGrowth %>%
  filter(supp == "OJ" & dose == 0.5) %>%
  select(len)

OJ1 <- ToothGrowth %>%
  filter(supp == "OJ" & dose == 1) %>%
  select(len)

OJ2 <- ToothGrowth %>%
  filter(supp == "OJ" & dose == 2) %>%
  select(len)
```

### 0.5 mg/day Vitamin C with 0.5 mg/day Orange Juice

**H0:** mean odontoblast length of 0.5 mg/day dose of Vitamin C = mean odontoblast length of 0.5 mg/day dose of Orange Juice.

```r
t.test(VC0.5,OJ0.5)
```

```
##
##  Welch Two Sample t-test
##
## data:  VC0.5 and OJ0.5
## 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:
##  -8.780943 -1.719057
## sample estimates:
## mean of x mean of y
##      7.98     13.23
```

p-value $0.006 < 0.05$. There is enough statistical evidence to reject the null hypothesis.

### 1 mg/day Vitamin C with 1 mg/day Orange Juice

**H0:** mean odontoblast length of 1 mg/day dose of Vitamin C = mean odontoblast length of 1 mg/day dose of Orange Juice.

```r
t.test(VC1,OJ1)
```

```
##
##  Welch Two Sample t-test
```

```
##
## data:  VC1 and OJ1
## 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:
##  -9.057852 -2.802148
## sample estimates:
## mean of x mean of y
##     16.77     22.70
```

p-value $0.001 < 0.05$. There is enough statistical evidence to reject the null hypothesis.

### 2 mg/day Vitamin C with 2 mg/day Orange Juice

**H0:** mean odontoblast length of 2 mg/day dose of Vitamin C = mean odontoblast length of 2 mg/day dose of Orange Juice.

```
t.test(VC2,OJ2)
```

```
##
##  Welch Two Sample t-test
##
## data:  VC2 and OJ2
## 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:
##  -3.63807  3.79807
## sample estimates:
## mean of x mean of y
##     26.14     26.06
```

p-value $0.964 > 0.05$. The confidence interval includes 0.

Therefore there is no statistical significant difference between the means Vitamin C and Orange Juice whit a dose level of 2 mg/day.

## Conclusions

Dosage levels of the supplements given affect tooth growth significantly. While, supplements by themselves do not seem to affect the length of odontoblasts.

Orange Juice when given in 0.5 and 1 mg/day doses, promotes bigger tooth growth compared to Vitamin C. Thou, when administered at 2 mg/day, both supplements seem to promote a similar tooth growth effect.