# The Effect on Vitamin C on Guinea Pigs' Tooth Growth

Giulio Mario Martena

6/4/2020

## Overview

The general goal of this analysis is to analyze the `ToothGrowth` dataset, part of the `datasets` library preincluded in R. From the documentation of the dataset, we learn that among 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).*

```
data("ToothGrowth")
ToothGrowth$dose <- factor(ToothGrowth$dose)
ToothGrowth<-as_tibble(ToothGrowth);head(ToothGrowth)
```

```
## # A tibble: 6 x 3
##     len supp  dose
##   <dbl> <fct> <fct>
## 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   VC    0.5
```

## Exploratory Analysis

From *figure 1* (cfr. Appendix session below) it appears that a clear separation between the two treatments is only visible with a dose of 1 mg/day; for 2 mg/day the separation is not clear, even though there it appears that ascorbic acid has a bigger variance.

## Hypothesis Testing

First of all, I wonder if mean from the tooth growth is the same **globally** and **divided by dosage**.

- *I will assume that the variance is the same and that they used 6 different sets of guinea pigs in order not to have results affected by previous treatments.*

### Hypothesis 1: no dosage separation

$$H_0 : \mu_{OJ} = \mu_{VC}$$

$$H_1 : \mu_{OJ} \neq \mu_{VC}$$

```
t.test(len ~ supp, ToothGrowth, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 58, p-value = 0.06039
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1670064  7.5670064
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

Since $p > 0.05$ and the confidence interval includes 0, so even though $\mu_{OJ} > \mu_{VC}$, $H_0$ *cannot* be rejected; hence, globally, we can assume that the means are in fact equal.

**Hypothesis 2: 0.5mg/day dosage**

$$H_0 : \mu_{OJ,.5} = \mu_{VC,.5}$$
$$H_1 : \mu_{OJ,.5} \neq \mu_{VC,.5}$$

```
t.test(len ~ supp, data = filter(ToothGrowth, dose == 0.5), paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 18, p-value = 0.005304
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.770262 8.729738
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

The p-value is smaller than .05, and the confidence interval does not include 0. In this case we can reject $H_0$ and accept that $\mu_{OJ,.5} > \mu_{VC,.5}$, which means that for a dosage of 0.5mg/day the orange juice causes more teeth growth than ascorbic acid.

**Hypothesis 3: 1mg/day dosage**

$$H_0 : \mu_{OJ,1} = \mu_{VC,1}$$
$$H_1 : \mu_{OJ,1} \neq \mu_{VC,1}$$

```
t.test(len ~ supp, data = filter(ToothGrowth, dose == 1), paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 18, p-value = 0.0007807
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.840692 9.019308
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

Just as in hypothesis 2, T-Test suggests that $p < 0.05$ and the confidence interval does not include 0. $H_0$ is rejected, hence for a dosage of 1mg/day it appears that orange juice grows teeth more than its rival.

**Hypothesis 4: 2mg/day dosage**

$$H_0 : \mu_{OJ,2} = \mu_{VC,2}$$
$$H_1 : \mu_{OJ,2} \neq \mu_{VC,2}$$

```
t.test(len ~ supp, data = filter(ToothGrowth, dose == 2), paired = FALSE, var.equal = TRUE)
```

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

```
## data:  len by supp
## t = -0.046136, df = 18, p-value = 0.9637
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.722999  3.562999
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

Since $p > 0.05$ and the confidence interval includes 0, so even though $\mu_{OJ,2} > \mu_{VC,2}$, $H_0$ *cannot* be rejected; hence, globally, we can assume that the both supplements bring the same growth for a 2mg/day dosage.

From the boxplot it would also appear a similar growth for 1mg/day dosage of orange juice and 2mg/day dosage of ascorbic acid. Is it true?

**Hypothesis 5: 1mg/day of orange juice vs 2mg/day of ascorbic acid**

$$H_0 : \mu_{OJ,1} = \mu_{VC,2}$$
$$H_1 : \mu_{OJ,1} \neq \mu_{VC,2}$$

```
h <- bind_rows(ToothGrowth %>%
                  filter((dose==1 & supp=='OJ') | (dose==2 & supp=='VC'))
                  )
t.test(len ~ supp, h, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = -1.7574, df = 18, p-value = 0.09584
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -7.552325  0.672325
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            26.14
```

The $p > 0.05$ and the C.I. includes 0. The $H_0$ cannot be rejected, which means that 2 mg/day of ascorbic acid seems to affect the teeth as much 1 mg/day of orange juice.

A similar pattern also appears in 1mg/day of orange juice and 2mg/day of the same supplement.

**Hypothesis 6: 1mg/day vs 2mg/day of orange juice**

$$H_0 : \mu_{OJ,2} = \mu_{OJ,1}$$
$$H_1 : \mu_{OJ,2} \neq \mu_{OJ,1}$$

```
h <- bind_rows(ToothGrowth %>%
                  filter((dose== c(1,2) & supp=='OJ'))
                  )
t.test(len ~ dose, h, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by dose
```

```
## t = -2.4241, df = 8, p-value = 0.04158
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.1221999  -0.2778001
## sample estimates:
## mean in group 1 mean in group 2
##           20.72           26.42
```

$p < 0.05$ and the confidence interval does not include 0 (which would be the difference in means if they were equal), hence $H_0$ is rejected. 2mg/day of orange juice appear to bring more tooth growth to the table than half that dosage.

## Conclusion

The teeth length seems to **grow with higher dosage of vitamin C**. Furthermore, orange juice appears to influence it more significantly on lower dosages, whereas on 2mg/day dosage the effect is pretty much similar (cfr. hypothesis 4).

## Appendix

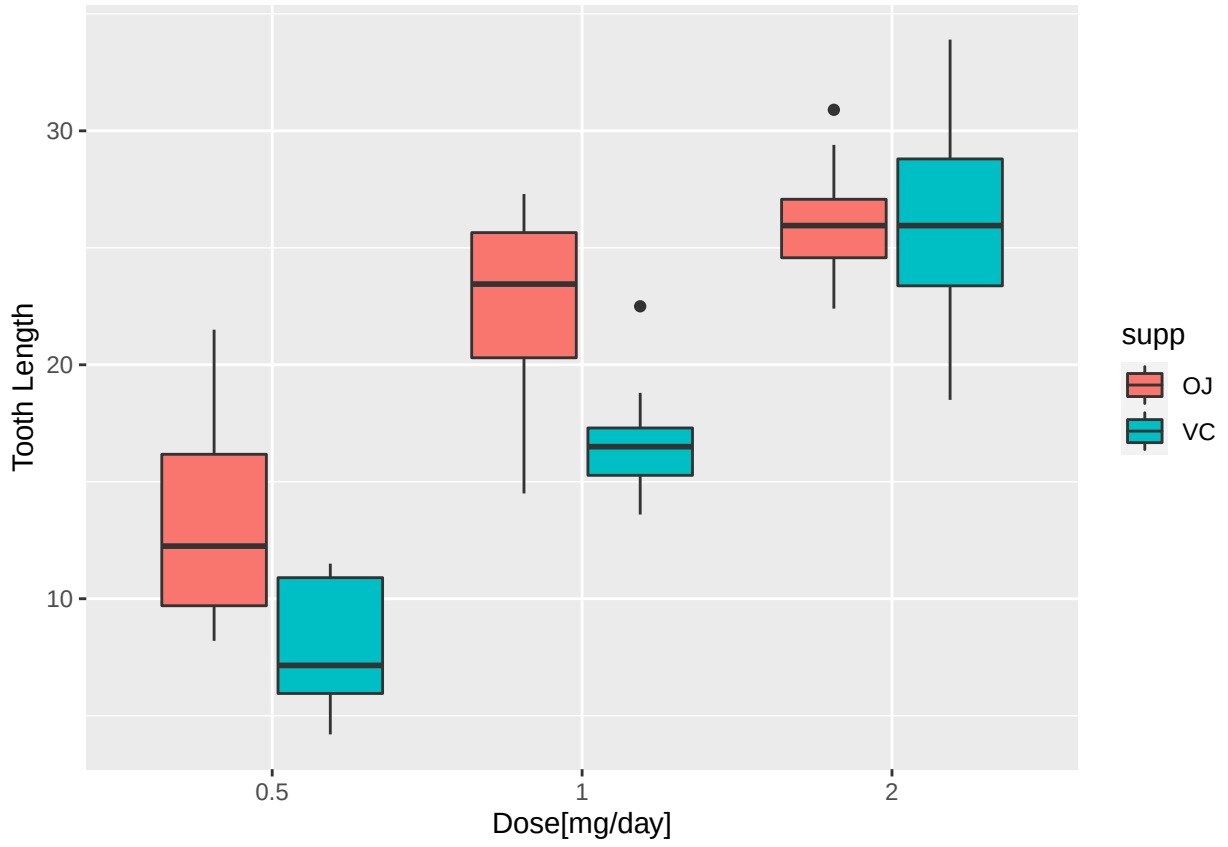


Figure 1: Tooth Length vs Supplement and Dose Levels