

Does Orange Juice help Tooth Growth?

Yangjing Long

19.09.2014

Introduction

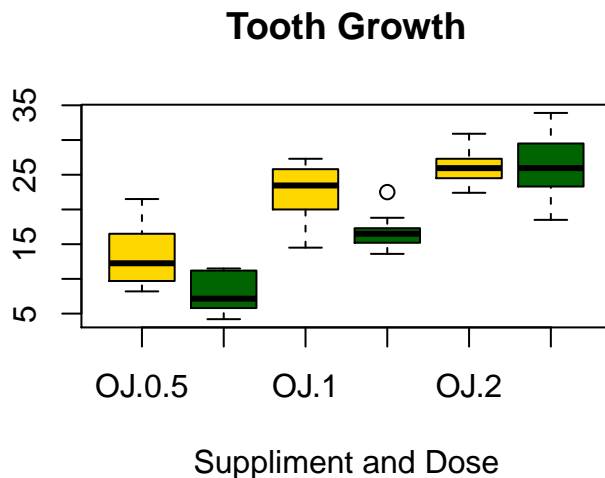
Does Orange Juice helps Tooth Growth? In this project, by analyzing the ToothGrowth data in the R data sets package, using statistical inference, we found that Vitamin C and Orange Juice do help Tooth Growth. Under the constraint of data range in our data set, the more orange Juice you drink, or the more Vitamin you consume, your tooth will grow more.

Pre-processing and exploratory analysis.

At first, we load the ToothGrowth data and perform some basic exploratory data analyses.

We can see that the data set contains 3 columns, there are no Null or NA values. “len” stands for Tooth length, which are numeric values. “supp” stands for supplements, there are two values: OJ for Orange Juice and VC for Vitamin C. “dose” is a numeric variable representing the amount of supplement in milligrams.

I used boxplots to show the how length is related with three dosages and two supplement types. This exercise gives me a couple of useful hypotheses to work with for the T tests- namely whether the dosage has a significant impact and whether the dosage type is significant.



Provide a basic summary of the data.

The output of the summary of the data is:

```
##      len      supp      dose
##  Min.   : 4.2    OJ:30   Min.   :0.50
##  1st Qu.:13.1   VC:30   1st Qu.:0.50
##  Median :19.2                Median :1.00
##  Mean   :18.8                Mean   :1.17
##  3rd Qu.:25.3                3rd Qu.:2.00
##  Max.   :33.9                Max.   :2.00
```

The mean of Tooth length is 18.81 and its standard deviation is 7.649

The Mean of ToothGrowth with supp OJ is 20.66 and the Mean of ToothGrowth with supp VC is 16.96. The sd of ToothGrowth with supp OJ is 6.606, and the standard derivation of ToothGrowth with supp VC is 8.266.

The Mean of ToothGrowth with dosage level 0.5mg is 10.61, 1mg is 19.73, 2 mg is 26.1. The standard derivation of the ToothGrowth with the corresponding dosage levels are 4.5, 4.415, 3.774.

Confidence Intervals and Hypothesis Tests

We use confidence intervals and hypothesis tests to compare tooth growth by supp and dose. We did 3 difference kinds of test as follows. The 95% T confidence intervals of the difference are given for each tests. For convenience, all the null hypothesis we made is that the two groups are the same, even though base on the boxplot, for the case of dose 2, we should assume that OJ and VC for the tooth growth are similar.

1. Tooth Growth by supp (3 2-sample tests)

First, we assume that dose is fixed and compare the impact of ToothGrowth length with difference of two supplement groups OJ and VS in dosage levels of 0.5, 1, 2 (results show in Appendix 1).

We use two sample unpaired T-test, because we do not know that the data is from the same sample. For each dosage level, we computer the 95% t confidence Interval for the length difference of OJ and VC. The corresponding intervals are (3.169 7.331), (3.853 8.007), (-2.298, 2.138). And the p values are 0.006, 0.001 and 0.96. so for the dosage level of 0.5 and 1, we should reject null hypothesis. However, in the case of dose level 2, the impact of OJ and VS are similar. The p value for two sample t-test is 0.96, so we accept null hypothesis, OJ and VC's impact to toothgrowth is not significant.

2. Tooth Growth by dose (4 2-sample tests)

We assume that supply is fixed and compare the impact of ToothGrowth for difference dose. (detailed results show in Appendix 2).

two sample unpaired T test	95% CI	p value
OJ (0.5mg) vs OJ (1.0mg)	5.524 13.416	8.785e-05
OJ (1.0mg) vs OJ (2.0mg)	0.1886 6.5314	0.0392
VC (0.5mg) vs VC (1.0mg)	11.52 17.92	3.655e-08
VC (1.0mg) vs VC (2.0mg)	5.686 13.054	9.156e-05

Based on the table, we reject all null hypethesis for these cases.

3. Extra: test when dose change from 1 to 2, whether OJ has better impact for the teethgrowth than VC.

OJ (1mg to 2mg) vs VC (1mg to 2mg), this result is in Appendix 3. The 95% CI is (-11.1823, -0.8377), and p value is 0.0252. We reject null hypethsis, that means that OJ and VC increasing from 1 mg to 2mg has significantly different impact to the Toothgrowth.

Conclusions

For all the tests, I assumption unpaired. unequal variance.

The impact of OJ and VC for the toothgrowth is significantly different for dose levels 0.5 and 1. In these cases, p values are 0.006 and 0.001 (two sample t-test), we should reject null hypothesis. However, in the case of dose level 2, the impact of OJ and VS are similar. the p value for two sample t-test is 0.96.

And both the increase of the dose for OJ and VC has significantly impact to the toothgrowth. And for the test of the toothgrowth length increase for OJ and VC, from dose level increasing of 1 to 2, the p value is 0.0252, it means that it is significantly difference, when dose level increasing from 1 to 2, VC has higher impact to the teethgrowth than OJ.

Appendix:

This Appendix includes the code which we could not include in the main report.

Appendix 1:

t-test for the difference of the supply for three groups of dose level: dose = 0.5, dose =1, dose =2.

```
## t-test for dose = 0.5

## [1] 3.169 7.331

##
## Welch Two Sample t-test
##
## data: g2$len and g1$len
## t = -3.17, df = 14.97, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -8.781 -1.719
## sample estimates:
## mean of x mean of y
##      7.98      13.23

## t-test for dose = 1

## [1] 3.853 8.007

##
## Welch Two Sample t-test
##
## data: g2$len and g1$len
## t = -4.033, df = 15.36, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -9.058 -2.802
## sample estimates:
## mean of x mean of y
##     16.77     22.70

## t-test for dose = 2

## [1] -2.298 2.138

##
## Welch Two Sample t-test
##
## data: g2$len and g1$len
## t = 0.0461, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
##  -3.638  3.798
## sample estimates:
## mean of x mean of y
##      26.14      26.06
```

Appendix 2:

T-test for fixed supply and difference dose. The difference from dose =0.5 to dose =1, for OJ.

```
## [1] -14.615 -4.325

##
## Welch Two Sample t-test
##
## data: g2$len and g1$len
## t = 5.049, df = 17.7, p-value = 8.785e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  5.524 13.416
## sample estimates:
## mean of x mean of y
##      22.70      13.23
```

The difference from dose =1 to dose =2, for OJ.

```
## [1] -7.2709  0.5509

##
## Welch Two Sample t-test
##
## data: g2$len and g1$len
## t = 2.248, df = 15.84, p-value = 0.0392
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1886 6.5314
## sample estimates:
## mean of x mean of y
##      26.06      22.70
```

The difference from dose =0.5 to dose =1, for VC.

```
## [1] -17.44 -12.00

##
## Welch Two Sample t-test
##
## data: g2$len and g1$len
## t = 9.74, df = 16.14, p-value = 3.655e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```

```
## 11.52 17.92
## sample estimates:
## mean of x mean of y
##      22.70      7.98
```

The difference from dose =1 to dose =2, for VC.

```
## [1] -13.335 -5.405

##
## Welch Two Sample t-test
##
## data: g2$len and g1$len
## t = 5.47, df = 13.6, p-value = 9.156e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  5.686 13.054
## sample estimates:
## mean of x mean of y
##      26.14      16.77
```

Appendix 3:

T-test for the change difference of VC and OJ when dose increasing from 1 to 2.

```
## [1] -12.8834  0.8634

##
## Welch Two Sample t-test
##
## data: difference_VC and difference_OJ
## t = -2.441, df = 18, p-value = 0.0252
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -11.1823 -0.8377
## sample estimates:
## mean of x mean of y
##      -9.37      -3.36
```