Analysis of the Effects of Vitamin C to Guinea's Pigs Tooth Growth

Lou Marvin Caraig

Overview

The purpose of this analyis is to evaluate the effects of vitamin C to tooth growth of Guinea's pig with different combinations of supplement type and dosage.

Data exploration and assumptions

The `ToothGrowth` dataset can be loaded by simply loading the `datasets` library, while for the analysis we'll also use `dplyr` and `lsr`:

```
library(datasets)
library(dplyr)
library(lsr)
```

The dataset has 60 rows and 3 columns. The variable `len` represent the length of odontoblasts (teeth) in micron, `supp` represent the supplement type of the vitamin C and `dose` represent the amount of vitamin C dosed in milligrams. The variable `supp` is categorical and can be equal to `VC` (vitamin C dosed through ascorbic acid) or `OJ` (vitamin C dosed through orange juice). Here's a summary of the dataset:

	len	supp	dose
1	Min. : 4.20	OJ:30	Min. :0.500
2	1st Qu.:13.07	VC:30	1st Qu.:0.500
3	Median :19.25		Median :1.000
4	Mean :18.81		Mean :1.167
5	3rd Qu.:25.27		3rd Qu.:2.000
6	Max. :33.90		Max. :2.000

Table 1: Dataset summary.

In the following analysis we will perform t-tests and multiple t-tests. For these tests to be valid we will assume that the observations of the tooth length are iid drawn from a quasi-normal distribution (roughly symmetric and mound-shaped), and that the samples are representative for the entire population. This assumption is made also when the observations are grouped by supplement type and dose amount.

Comparison of tooth growth by supplement type, given dose amount

Before defining any hypothesis we first plot the tooth growth by supplement type. As we can see in Figure 1 (generated with Code 1) dosing vitamin C through orange juice seems more effective than ascorbic acid. Hence we test the hypothesis that the difference of means between the tooth length of those treated with orange juice and those treated with ascorbic acid is greater than 0. The test is performed at a 0.05 significance level.

	t-value	degrees of freedom	p-value	confidence-interval
t	1.92	55.31	0.03	(0.4682687, Inf)

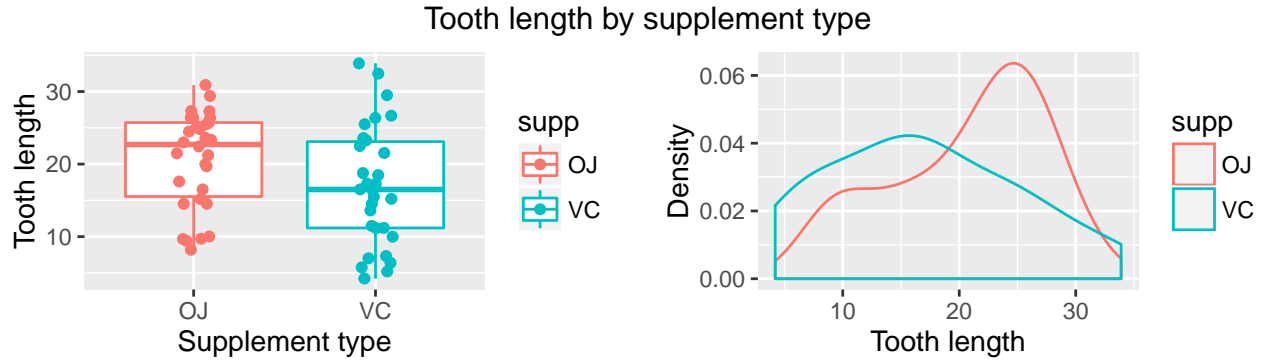Table 2: Result of t-test of the comparison of tooth growth by supplement type.

Figure 1: On the left the two boxplots. On the right the two smoothed density functions.

As we can see from Table 2 (calculated with Code 5) given that the p-value is equal to 0.0303173 we can reject the null hypothesis, meaning that orange juice affects tooth growth more then ascorbic acid. We can also calculate the effect size using Cohen's D with `cohensD(oj, vc)`, and it results 0.4945201 which can be interpreted as a medium effect.

Comparison of tooth growth by dose amount, given type of supplement

Now we plot the tooth growth by dose amount. As we can see in Figure 2 (generated with Code 2) bigger doses of vitamin C seem more effective than smaller ones. Hence we test the hypothesis that the difference of means between the tooth length of those treated with bigger doses and those treated with smaller ones is greater than 0. The test is performed at a 0.05 significance level. Given that we have three groups (three different possible values of `dose`) we perform a multiple t-test using Bonferroni correction.
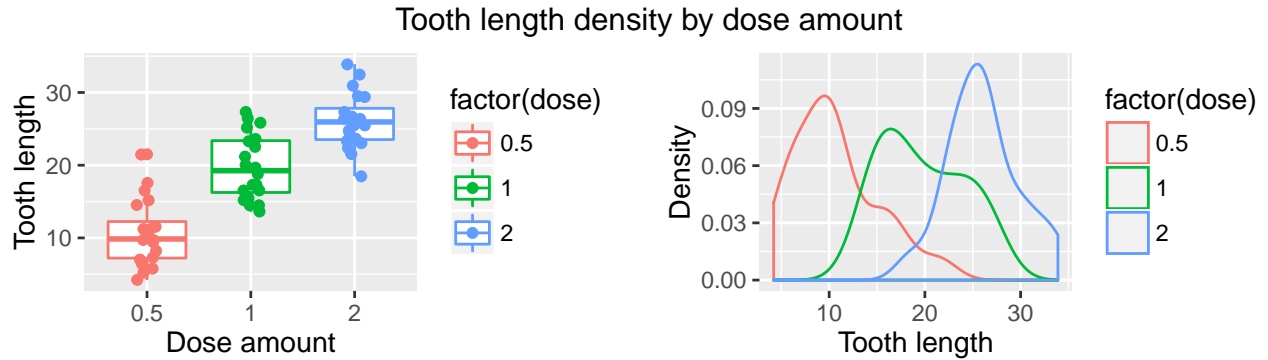

Figure 2: On the left the three boxplots. On the right the three smoothed density functions.

	dose 0.5	dose 1
dose 1	1.00E-08	
dose 2	2.20E-16	2.16E-05

Table 3: Result of multiple t-test of the comparison of tooth growth by dose amount.

As we can see from Table 3 (calculated with Code 6) all the comparisons (dose 1 vs. dose 0.5, dose 2 vs. dose 0.5, dose 2 vs dose 1) are statistically significant at a significance level of 0.05 even using such a conservative p-value adjusment method. This means that bigger doses affects tooth growth more then smaller ones. In Table 4 (calculated with Code 7) we can see that the effect size is pretty high in each case especially in the first case if we consider that the dosage difference is only 0.5 despite the third case where the dosage difference is 1.0.

dose 0.5 vs. dose 1.0	dose 0.5 vs. dose 2.0	dose 1.0 vs. dose 2.0
2.05	3.73	1.55

Table 4: Effect size of the difference of tooth growth between different dose amount.

Comparision of tooth growth by supplement type and dose amount

We can finally consider the tooth growth both by supplement type and dose amount. As we can see in Figure 3 (generatedy with Code 3) it seems that in some cases the tooth length achieved using ascorbic acid with a big dose is the same when using orange juice with a smaller dose. Hence it's interesting to perform a bidirectional hypothesis test on the difference of means considering both supplement type and dose amount. The test is performed at a 0.05 significance level. As before we have more than two groups so we perform a multiple t-test using Bonferroni correction. By looking at Table 5 (calculated with Code 8) we can confirm our supposition. Infact we can't reject the null hypothesis when doing the following comparisons: VC-1 vs. OJ-0.5, VC-2 vs. OJ-1, VC-2 OJ-2.
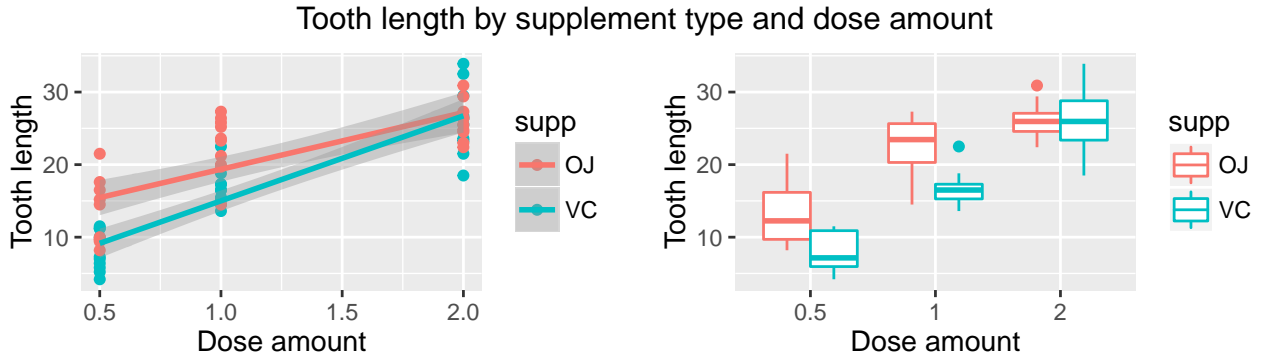

Figure 3: On the left the scatterplot mapping dose amount to tooth length with different colors for different supplement type with regression line. On the right the couples of boxplots, one for each supplement type, for each dose amount.

	OJ-0.5	OJ-1	OJ-2	VC-0.5	VC-1
OJ-1	4.76E-06				
OJ-2	2.14E-09	6.50E-01			
VC-0.5	3.14E-02	2.95E-11	2.01E-14		
VC-1	5.05E-01	8.85E-03	7.15E-06	2.19E-05	
VC-2	1.79E-09	5.82E-01	1.00E+00	1.70E-14	5.97E-06

Table 5: Result of multiple t-test of the comparison of tooth growth by supplement type and dose amount.

Conclusions

After these analysis we can conclude that dosing vitamin C through orange juice is more effective than ascorbic acid and also that bigger doses affect tooth growth more than smaller ones. In some cases the effect provided by a dose of ascorbic acid can be achieved by a smaller dose of orange juice.

Appendix

This appendix contains the code that has been used to generate the above presented plots and tables. Here's the required libraries loaded:

```
library(xtable)
options(xtable.comment=FALSE)
library(ggplot2)
library(gridExtra)
```

Code 1:

```
p1 <- ggplot(data=ToothGrowth, mapping=aes(supp, len, colour=supp)) +
    geom_boxplot() +
    geom_jitter(width=0.25) +
    labs(x='Supplement type', y='Tooth length')
p2 <- ggplot(data=ToothGrowth, mapping=aes(x=len, y=..density.., colour=supp)) +
    geom_density() +
    labs(x='Tooth length', y='Density')
grid.arrange(p1, p2, ncol=2, top='Tooth length by supplement type')
```

Code 2:

```
p1 <- ggplot(data=ToothGrowth, mapping=aes(factor(dose), len, colour=factor(dose))) +
    geom_boxplot() +
    geom_jitter(width=0.25) +
    labs(x='Dose amount', y='Tooth length')
p2 <- ggplot(data=ToothGrowth, mapping=aes(x=len, y=..density.., colour=factor(dose))) +
    geom_density() +
    labs(x='Tooth length', y='Density')
grid.arrange(p1, p2, ncol=2, top='Tooth length density by dose amount')
```

Code 3:

```
p1 <- ggplot(data=ToothGrowth, mapping=aes(dose, len, colour=supp)) +
    geom_point() +
    geom_smooth(method='lm') +
    labs(x='Dose amount', y='Tooth length')
p2 <- ggplot(data=ToothGrowth, mapping=aes(factor(dose), len, colour=supp)) +
    geom_boxplot() +
    labs(x='Dose amount', y='Tooth length')
grid.arrange(p1, p2, ncol=2, top='Tooth length by supplement type and dose amount')
```

Code 4:

```
print(xtable(summary(ToothGrowth), caption='Dataset summary.'))
```

Code 5:

```
oj <- (ToothGrowth %>% filter(supp=='OJ') %>% select(len))$len
vc <- (ToothGrowth %>% filter(supp=='VC') %>% select(len))$len

out <- t.test(oj, vc, alternative='greater')
outDt <- with(out, data.frame(t=statistic, dof=parameter, p=p.value,
                              ci=paste('(', format(conf.int[[1]]), ', ',
                                       format(conf.int[[2]]), ')', sep='')))
colnames(outDt) <- c('t-value', 'degrees of freedom', 'p-value', 'confidence-interval')
print(xtable(outDt, caption=paste('Result of t-test of the comparison of tooth growth',
                                  'by supplement type.')))
```

Code 6:

```
out <- with(ToothGrowth,
     pairwise.t.test(x=len, g=dose, p.adjust.method='bonferroni', alternative='greater'))
outPValue <- out$p.value
rownames(outPValue) <- paste('dose', rownames(outPValue))
colnames(outPValue) <- paste('dose', colnames(outPValue))
print(xtable(outPValue, display=rep('E', ncol(outPValue) + 1),
            caption=paste('Result of multiple t-test of the comparison of',
                          'tooth growth by dose amount.')))
```

Code 7:

```
dose0_5 <- (ToothGrowth %>% filter(dose==0.5) %>% select(len))$len
dose1_0 <- (ToothGrowth %>% filter(dose==1.0) %>% select(len))$len
dose2_0 <- (ToothGrowth %>% filter(dose==2.0) %>% select(len))$len

dt <- data.frame(d05_d10=cohensD(dose0_5, dose1_0),
                 d05_d20=cohensD(dose0_5, dose2_0),
                 d10_d20=cohensD(dose1_0, dose2_0))
colnames(dt) <- c(
    'dose 0.5 vs. dose 1.0', 'dose 0.5 vs. dose 2.0', 'dose 1.0 vs. dose 2.0')
rownames(dt) <- ''
print(xtable(dt, caption=paste('Effect size of the difference of tooth growth',
                               'between different dose amount.')))
```

Code 8:

```
dt <- ToothGrowth %>%
    mutate(group=paste(supp,dose,sep='-')) %>%
    select(len, group)
out <- with(dt, pairwise.t.test(x=len, g=group, p.adjust.method='bonferroni'))
outPValue <- out$p.value
print(xtable(outPValue, display=rep('E', ncol(outPValue) + 1),
            caption=paste('Result of multiple t-test of the comparison of',
                          'tooth growth by supplement type and dose amount.')))
```