The Effect of Vitamin C on teeth growth

DanH

August 14, 2016

Overview

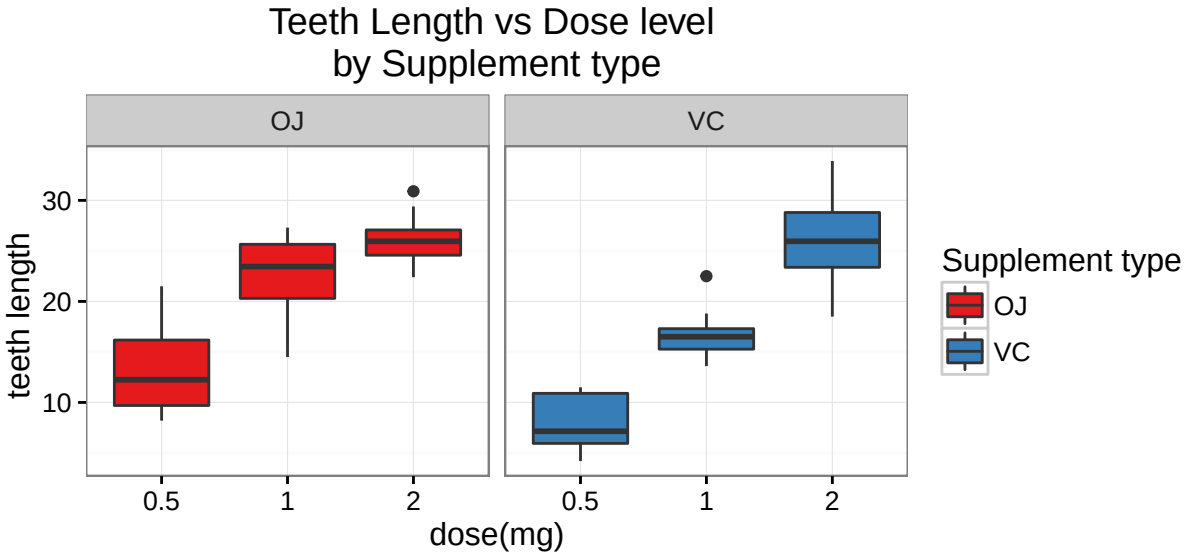
Here we want to analyze the effectiveness of vitamin c on teeth growth in guinea pigs. We'll do this by using t-tests, comparing teeth length by supplement type and dose level.

Exploratory data analysis

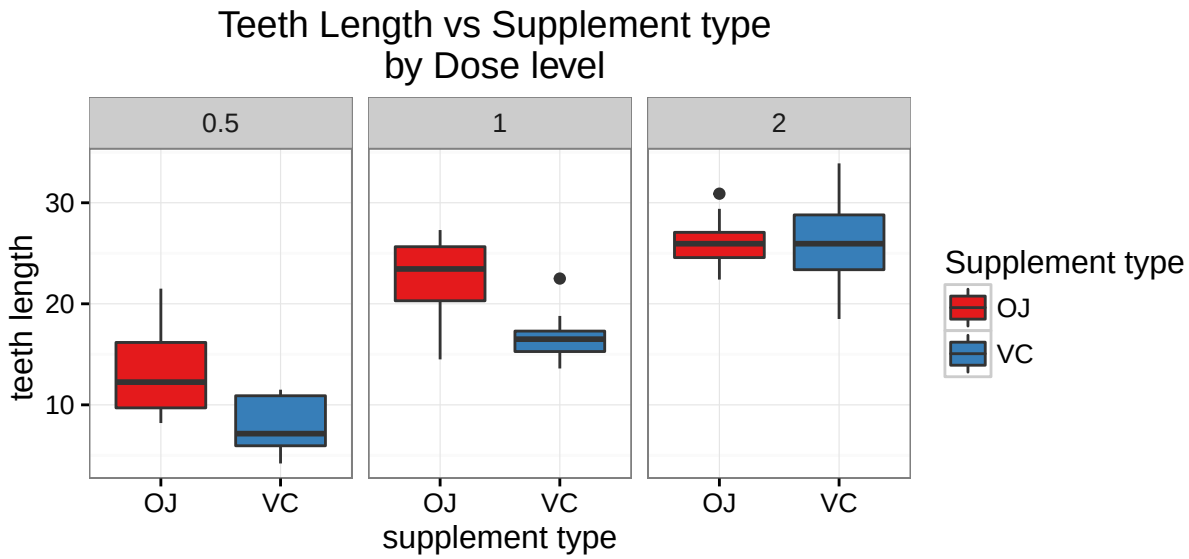
We have a dataset of 60 observations of 3 variables:

- **len**: teeth length, numeric variable
- **supp**: supplement type (VC: vitamin C or OJ: orange juice), factor variable
- **dose**: dose(in milligrams), numeric variable

The numeric variable *dose* contains only 3 unique values: 0.5, 1, 2. We can conveniently convert it to a factor variable with three levels



This multipanel plot emphasizes the relationship between teeth length and dose level, for each supplement type.



This second plot shows the relationship between teeth length and supplement type emphasizing direct comparison between supplement types.

Hypothesis Test

We will run a two-sample t-test for factor *supp* and one two-sample t-test for each possible pair of the 3 levels in the factor *dose*, that is, we will run a total of 4 t-tests.

Test A, dose = 0.5 and dose = 1

```
##
##  Welch Two Sample t-test
##
## data:  len_a by dose_a
## 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 in group 0.5   mean in group 1
##            10.605            19.735
```

Test B, dose = 0.5 and dose = 2

```
##
##  Welch Two Sample t-test
##
## data:  len_b by dose_b
## 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 in group 0.5   mean in group 2
##            10.605            26.100
```

Test C, dose = 1 and dose = 2

```
##
##  Welch Two Sample t-test
##
## data:  len_c by dose_c
## 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 in group 1 mean in group 2
##          19.735          26.100
```

We went through all possible combinations of levels from the factor variable dose and in all cases the p-value is lower than the default signficance level 0.05. Thus, we reject Ho. In other words there appears to be a positive relationship between dose level and teeth length

Testing by Supplement

```
##
##  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
```

We can see that the p-value of the test is 0.06. Since the p-value is greater than 0.05 and the confidence interval of the test contains zero, we can reject the null hypothesis and say that supplement types don't seem to have any impact on teeth growth. In other words, there's no significant statistical difference between them

Conclusions

Before using t-tests we should always make sure that the following conditions are met:

- Independence: there must be random sampling/assignment
- Normality: observations must come from a normal or nearly-normal distribution

Assuming all the previous conditions are met we can conclude that:

It appears that there is a statistically significant difference between teeth length and dose levels across both delivery methods, in other words, as the dose increases so does teeth length.

On the other hand, there doesn't seem to be a statistically significant difference between delivery methods, with Orange juice apparently more effective at dose levels 0.5 and 1, and VC slightly more effective at dose level 2

APPENDIX

Load dataset, convert, summarize and inspect it

```
#Load required packages
library(dplyr, warn.conflicts = F)
library(ggplot2)
library(ggthemes)

#Load data and convert to tbl format
ToothGrowth <- tbl_df(ToothGrowth)

#Structure of the dataframe
ToothGrowth %>% str()

#Summary
ToothGrowth %>% summary()

#Unique values in the dose vector
ToothGrowth %>% select(dose) %>% unique()

#Convert dose to factor
ToothGrowth <- ToothGrowth %>% mutate(dose = as.factor(dose))
```

Plot 1

```
ToothGrowth %>%
ggplot(aes(x=dose, y=len, fill = supp)) +
geom_boxplot() +
facet_grid(. ~ supp) +
scale_fill_brewer(palette = "Set1") +
theme_bw() +
ggtitle("Teeth Length vs Dose level \nby Supplement type") +
labs(x="dose(mg)", y= "teeth length ") +
guides(fill=guide_legend(title="Supplement type"))
```

Plot 2

```
ToothGrowth %>%
ggplot(aes(x = supp, y = len)) +
geom_boxplot(aes(fill = supp)) +
facet_wrap(~ dose) +
scale_fill_brewer(palette = "Set1") +
theme_bw() +
ggtitle("Teeth Length vs Supplement type \nby Dose level ") +
labs(x="supplement type", y= "teeth length ") +
guides(fill=guide_legend(title="Supplement type"))
```

Difference in avg. *len* by *supp* type, at dose level 2

```
ToothGrowth  %>% filter(dose == 2)  %>% group_by(supp)   %>%  summarise(avg.length = mean(len))
```

Hypothesis tests

Test A, dose = 0.5 and dose = 1

```
#Exract the len and dose vectors from the df ToothGrowth
len_a <- ToothGrowth %>% filter(dose %in% c(0.5,1)) %>% select(len) %>% unlist()
dose_a <- ToothGrowth %>% filter(dose %in% c(0.5,1)) %>% select(dose) %>% unlist()
#Test
(Test.a <- t.test(len_a~dose_a, paired = FALSE))
```

Test B, dose = 0.5 and dose = 2

```
#Exract the len and dose vectors from the df ToothGrowth
len_b <- ToothGrowth %>% filter(dose %in% c(0.5,2)) %>% select(len) %>% unlist()
dose_b <- ToothGrowth %>% filter(dose %in% c(0.5, 2)) %>% select(dose) %>% unlist()
#Test
(Test.b <- t.test(len_b~dose_b, paired = FALSE))
```

Test C, dose = 1 and dose = 2

```
#Exract the len and dose vectors from the df ToothGrowth
len_c <- ToothGrowth %>% filter(dose %in% c(1,2)) %>% select(len) %>% unlist()
dose_c <- ToothGrowth %>% filter(dose %in% c(1,2)) %>% select(dose) %>% unlist()
#Test c
(Test.c <- t.test(len_c~dose_c, paired = FALSE))
```

Testing by Supplement type

```
#Exract the len and supp vectors from the df ToothGrowth
len <- ToothGrowth %>% select(len) %>% unlist()
supp <- ToothGrowth %>% select(supp) %>% unlist()
#Test
t.test(len~supp, paired=F)
```