Exploring the Effect of Vitamin C on Tooth Growth

Matthew Halbasch

February 25, 2019

Overview

In this report we analyse data from a treatment of guinea pigs with Vitamin C delivered in two forms - orange juice or ascorbic acid. We compare delivery methods and dosages to discern the effect of Vitamin C on tooth growth. We find that dosage is correlated with increased growth, but the effect of supplement type is not statistically significant.

Data Processing

The data for this report are available through the `ToothGrowth` dataset in R. We load this dataset in below, and examine its structure.

```
data(ToothGrowth)
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 ...
```

This dataset has 3 variables: the tooth length, the supplement type (OJ or VC for orange juice or ascorbic acid) and the dose in mg/day. We briefly overview this data with a plot of the length by supplement and dose.

```
library(ggplot2)
library(ggsci)

ggplot(data = ToothGrowth, aes(x=factor(dose), y=len,
                               fill=factor(dose))) +
  geom_violin(alpha=0.6) +
  facet_grid(.~supp) +
  scale_fill_uchicago() +
  xlab("Dose") +
  ylab("Tooth Length") +
  labs(title = "Violin Plot of Tooth Length by Dose and Supplement",
       fill = "Dose") +
  theme_minimal(10) +
  theme(plot.title=element_text(hjust=0.5))
```

Figure 1 shows that the tooth length appears to vary significantly with dose within each supplement class. The effect between supplement classes for fixed doses is less clear, and could well be due to noise. We examine these assessments quantitatively in the next section.

Results

For all the results in this section, we assume that each observation is drawn from the same population, so that e.g. the population standard deviations of each sub-group are equal.

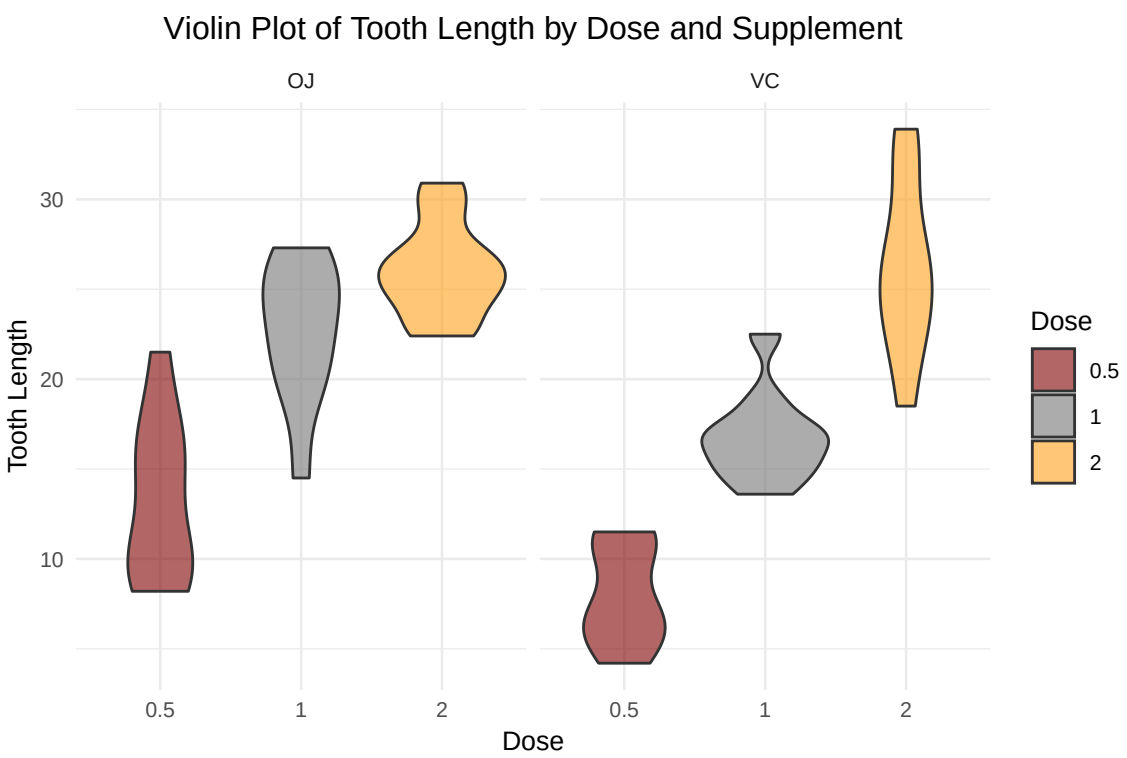

Figure 1: A violin plot of the tooth length as a function of dose and supplement. We see that within each supplement type, the dose appears to have a significant effect on the tooth length. The effect of the supplement itself is less clear from this plot.

We first consider the effect of the dosage of Vitamin C on tooth growth. In particular, we will consider the tooth growth of the guinea pigs who had 2 mg/day of Vitamin C against the growth of those with only 0.5 mg/day. For the purpose of this analysis we will ignore the 1 mg/day group, as we are only trying to determine if dosing has any effect, and Figure 1 shows that the largest effect will be between the 2 mg and 0.5 mg groups.

First, we create the two groups.

```
library(dplyr)
X <- ToothGrowth %>% filter(dose == 0.5)
length(X$len)
```

```
## [1] 20
```

```
Y <- ToothGrowth %>% filter(dose == 2)
length(X$len)
```

```
## [1] 20
```

Each of these groups has 20 members. We will now perform a 2-sided t-test on these data:

```
t.test(x=X$len, y=Y$len, paired=FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  X$len and Y$len
## t = -11.799, df = 38, p-value = 2.838e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15352 -12.83648
## sample estimates:
## mean of x mean of y
##    10.605    26.100
```

The t test shows that the 95% confidence interval for the means falls well outside of 0, and we find a p-value of $p < 1e-13$, so we reject the null hypothesis that dosage has no effect at any reasonable level of α.

So we find that a dosage of 2 mg/day is associated with higher levels of tooth growth than 0.5 mg/day.

Next, we consider the effect of the supplement type. We again construct our two subsets:

```
X2 <- ToothGrowth %>% filter(supp == "VC")
length(X2$len)
```

```
## [1] 30
```

```
Y2 <- ToothGrowth %>% filter(supp == "OJ")
length(Y2$len)
```

```
## [1] 30
```

Each group now has 30 observations. We perform a t-test on these two samples to compare the difference in their means to 0:

```
t.test(x=X2$len, y=Y2$len, paired = FALSE, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  X2$len and Y2$len
```

```
## t = -1.9153, df = 58, p-value = 0.06039
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -7.5670064  0.1670064
## sample estimates:
## mean of x mean of y
##  16.96333  20.66333
```

We see that the 95% confidence interval does contain 0 this time, with a p value of $p > .05$. So, with a standard value of $\alpha = .05$ we cannot reject the null hypothesis, and cannot conclude that either supplement is more effective at inducing tooth growth.

Conclusion

We find in our analysis that while the dosage of Vitamin C does have an effect on tooth growth (dosage of 2 mg/day is associated with higher tooth growth than 0.5 mg/day), we fail to reject the hypothesis that the supplement type has no effect on tooth growth.

Note that this analysis is under the assumption that all guinea pigs have the same underlying variance in tooth length, and that the tooth lengths are normally distributed. We also assume all observations are independent.

These findings are roughly in line with the intuition provided by Figure 1, which shows a strong effect in dosage but a weaker effect between supplements.

Further analysis could compare the effect within smaller subgroups (e.g. 2 mg/day OJ vs VC), but the sample sizes here are quite small for this sort of subgroup analysis, and we must be careful of the risks of multiple testing.