Investigating the Effects of Vitamin C Supplements on Tooth Growth in Guinea Pigs

Connor Claypool

19 June 2018

Overview

In this analysis we will use t-tests to compare the mean tooth length of groups of guinea pigs which have been given different doses of different types of vitamin C supplements, in order to draw conclusions about the relative effect of different dose levels and supplement types on tooth growth. We will assume that tooth length is independent between subjects, and that confounding variables have been properly addressed through random allocation of subjects to groups. Additionally, we assume that our sample sizes are large enough given the underlying distribution such that the various differences between sample means follow the t-distribution with the relevant degrees of freedom.

Loading the Data and Exploratory Analysis

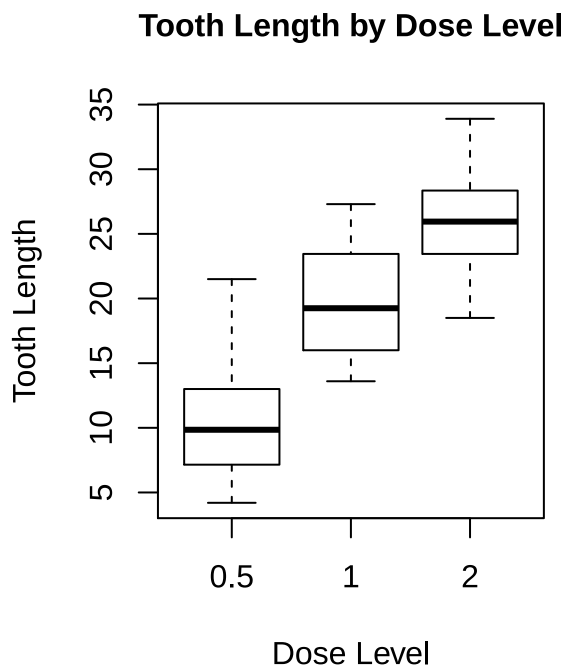
The first step is to load the data and determine its structure.

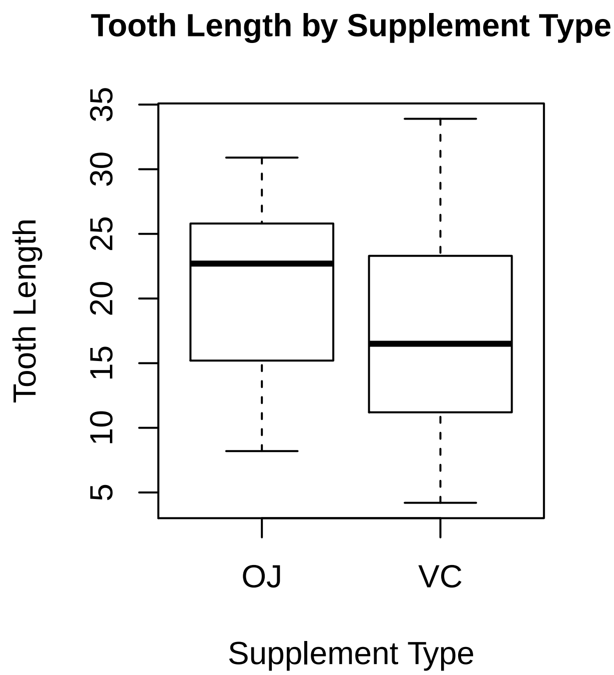
```r
library(datasets)
data(ToothGrowth)
str(ToothGrowth)
```

```
## '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 ...
```

We can see that the dataset contains 60 observations of 3 variables, which are tooth length, supplement type and dose. Next, we create some exploratory box plots to give a visual representation of the distributions of tooth lengths based on dose and supplement type.

```r
par(mfrow=c(1, 2))
boxplot(len ~ dose, ToothGrowth)
title(main = list("Tooth Length by Dose Level", cex=1),
      xlab = "Dose Level", ylab = "Tooth Length")
boxplot(len ~ supp, ToothGrowth)
title(main = list("Tooth Length by Supplement Type", cex = 1),
      xlab = "Supplement Type", ylab = "Tooth Length")
```





We can see from these plots that larger doses seem to be clearly associated with greater tooth length. However, the relationship between tooth length and supplement type appears less clear. It is also important to note that there is no baseline condition to which no supplements were administered, meaning there is no information on absolute increases in tooth growth from normal. Therefore, we will compare mean tooth length by dose and by supplement type and conclude from these comparisons which treatments cause more tooth growth than others.

Comparing the Effects of Dose Levels

First, we will test the hypothesis that a dose level of 1 is associated with a greater mean tooth length than that of dose level 0.5, and obtain a 95% confidence interval for the difference in the population means for these doses. We perform a one-sided t-test with no assumption of equal variance.

```
len_dose_0.5 <- subset(ToothGrowth, dose == 0.5)$len
len_dose_1 <- subset(ToothGrowth, dose == 1)$len

test_dose_1 <- t.test(len_dose_1, len_dose_0.5, paired = FALSE,
                      var.equal = FALSE, alternative = "greater")

print(paste("p-value:", test_dose_1$p.value))
```

```
## [1] "p-value: 6.34150360086923e-08"
```

```
print(paste("95% confidence interval:", paste(test_dose_1$conf.int[1],
                                              test_dose_1$conf.int[2])))
```

```
## [1] "95% confidence interval: 6.75332267058178 Inf"
```

With a significance level of 0.05, we reject the null hypothesis, and see that with 95% confidence the true difference in mean tooth length between these treatment levels is greater than 6.75. We interpret this to mean that dose level 1 causes significantly more tooth growth than dose level 0.5 across supplement types.

Similarly, we will test the hypothesis that a dose level of 2 is associated with a greater mean tooth length than that of dose level 1, and obtain a 95% confidence interval for this difference. As previously, we perform a one-sided t-test without assuming equal variance.

```
len_dose_2 <- subset(ToothGrowth, dose == 2)$len

test_dose_2 <- t.test(len_dose_2, len_dose_1, paired = FALSE,
                        var.equal = FALSE, alternative = "greater")

print(paste("p-value:", test_dose_2$p.value))
```

```
## [1] "p-value: 9.53214756835902e-06"
```

```
print(paste("95% confidence interval:", paste(test_dose_2$conf.int[1],
                                               test_dose_2$conf.int[2])))
```

```
## [1] "95% confidence interval: 4.1738696921444 Inf"
```

We reject the null hypothesis at a significance level of 0.05 and see that with 95% confidence the true difference in means is at least 4.17, concluding that dose level 2 causes significantly more tooth growth than dose level 1 across supplement types.

Comparing the Effects of Supplement Types

Finally, we will test the extent of any significant difference between the mean tooth lengths for each supplement type. Given that there appears no clear relationship between tooth length and supplement type, we will use a two-sided t-test, again assuming unequal variance.

```
len_oj <- subset(ToothGrowth, supp == "OJ")$len
len_vc <- subset(ToothGrowth, supp == "VC")$len

test_supp <- t.test(len_oj, len_vc, paired = FALSE,
                     var.equal = FALSE, alternative = "two.sided")

print(paste("p-value:", test_supp$p.value))
```

```
## [1] "p-value: 0.0606345078809341"
```

```
print(paste("95% confidence interval:", paste(test_supp$conf.int[1],
                                               test_supp$conf.int[2])))
```

```
## [1] "95% confidence interval: -0.171015618367165 7.57101561836716"
```

At a significance level of 0.05, we fail to reject the null hypothesis, and see that with 95% confidence the true difference in mean tooth length by supplement type across dose levels is between -0.17 and 7.57. It appears that orange juice (OJ) may be associated with a slightly larger mean tooth length than ascorbic acid (VC), but this relationship is not quite statistically significant.