The Effect of Vitamin C on Tooth Growth in Guinea Pigs

Ivan Filho

Introduction

I have created a GitHub repository where I've put the files related to this homework. It can be accessed via the following url: https://github.com/ivanmantova/coursera-statsinference.

Exploratory Analysis

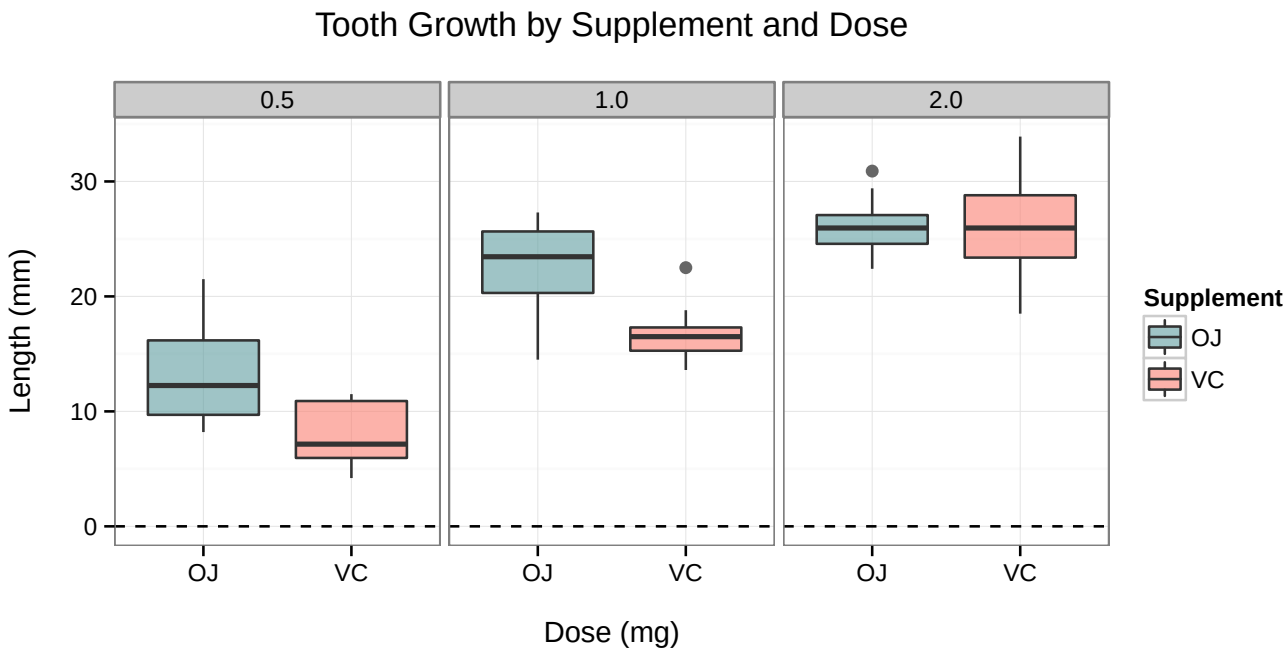
Let's load the **ToothGrowth** data and see some summary information from the data set.

The data set consists of measurements of length of guinea pigs teeth. It also states that there are two delivery methods of Vitamin C, orange juice (OJ) and ascorbic acid (VC). They were administered with three dose levels of 0.5, 1, and 2 mg.

Let's explore the data.

```
##     dose
## supp 0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

As the table above shows, there are just 10 observations for each group of delivery method (`supp`) and dose.



The plot shows that for doses of .5 and 1.0 mg, there seems to be a relatively large difference between the treatments. For doses of 2.0 mg, however, there doesn't seem to be much difference between the treatments. We are going to perform some analysis, do some hypothesis testing and calculate confidence intervals in order to quantify this uncertainty.

Confidence Intervals

We are going to produce t confidence intervals for tooth growth using the techniques presented in the class.

Given the small sample size, we could try to use Student's t-distribution for modeling the data, but **we can't assume equal variance for both groups**. We can, however, use an approximation of the t-distribution using the formula for Degrees of Freedom as shown in the lectures.

Here we compare the difference of the empirical mean between the two supplements for the various doses:

- t test between OJ and VC supplements for .5 mg:

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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:
##  1.719 8.781
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

- t test between OJ and VC supplements for 1 mg:

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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:
##  2.802 9.058
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

- t test between OJ and VC supplements for 2 mg:

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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:
##  -3.798  3.638
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

Hypothesis testing

For the first two doses, .5 and 1, the 95% Confidence Intervals we calculated do not include zero. We have enough evidence to reject the null hypothesis that the true difference between the two means is zero with *p-values* of `0.0064` and `0.0010`, respectively.

In the last t-test, for the 2 mg dose, however, we do not have enough evidence to reject the null hypothesis, as the 95% Confidence Interval does contain zero.

Conclusion

If we are willing to assume that these are random samples from a relavant population of Guinea Pigs receiving treatment, we can say that for small doses ($< 2mg$), the most effective treatment is with Orange Juice (OJ). For the 2 mg treatment we don't have enough data to make that conclusion and failed to reject the null hypothesis that there is no difference between the supplements.