Nicholas Stanford
Statistical Inference
Data Science Specialization
Johns Hopkins University
Coursera
July 11, 2017

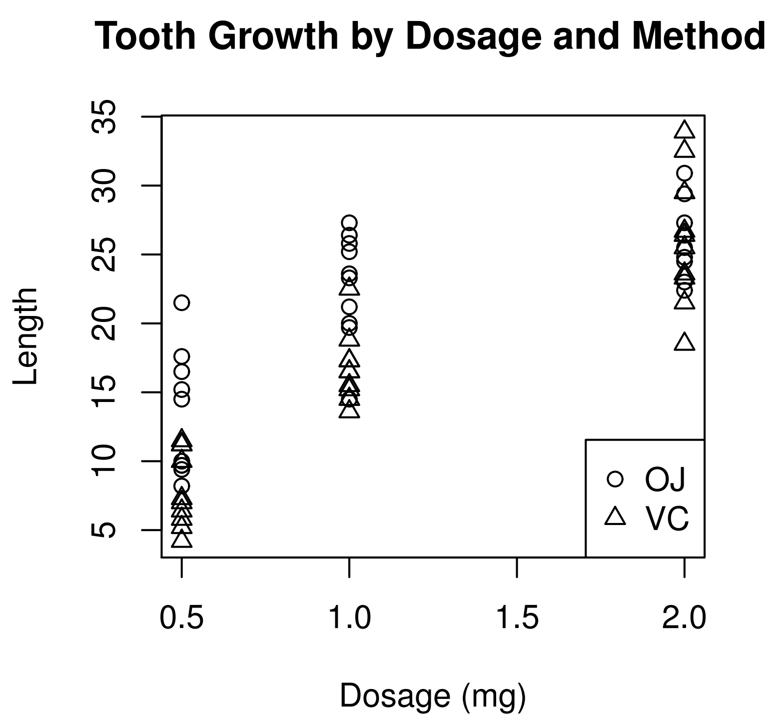
Course Project Part 2: Basic Inferential Data Analysis

1. Overview

The ToothGrowth data set gives information regarding the effect that vitamin C has on the growth of odontoblasts in guinea pigs. There are 60 observations total, which are evenly divided both between receiving the supplement either through orange juice (OJ) or ascorbic acid (VC) and the size of the daily supplement (0.5, 1.0, or 2.0 mg). The goal of the following is to determine which vitamin C delivery method is most efficient in promoting tooth growth by using confidence intervals and p-values.

2. Basic Data Summary

First we present a scatter plot from the experiment in order to give an initial visual sense of the data. Dosage is the independent variable and tooth growth the dependent variable. Each data point is further identified as coming from one delivery method or the other using differing symbols to represent the point.



Below is a table that summarizes the data of the experiment in terms of the averages for each test class. It seems apparent that, for the dosage levels studied, a higher dose on average

	0.5	1.0	2.0	Marginal
OJ	13.23	22.70	26.04	20.66
VC	7.98	16.77	26.14	16.96
Marginal	10.60	19.74	26.10	18.81

TABLE 1. Tooth Growth Averages

resulted in more tooth growth. Delivery of vitamin C via OJ was far more effective for daily dosages of 0.5mg and 1.0mg, yet the methods appear to be equally effective at the level of 2.0mg/day. These initial observations will be further analyzed in the following section. For now, we give a table of standard deviations to give at least an initial sense of reliability of the data.

	0.5	1.0	2.0	Marginal
OJ	4.46	3.91	2.66	6.61
VC	2.75	2.52	4.80	8.27
Marginal	4.50	4.42	3.77	7.65

TABLE 2. Tooth Growth Standard Deviations

3. Confidence Intervals and Hypothesis Tests

There is no reason that we should be able to assume that the distributions underlying any of the six test groups (or their marginals) are normal, so it will be best to form T confidence intervals. The T distribution has wider tails than normal distributions, which results in larger intervals for fixed confidence percentages. In other words, they form a more conservative estimate. The T interval also tends to be more accurate for smaller data sets. Its assumption that the data are iid normal, or roughly symmetric and mound shape, is good enough. In proceeding with this technique we also assume that this sample is representative of the population.

Below is a table of data giving the 95% confidence intervals for each of the 6 types of treatment, as well as the marginal confidence intervals. A sample of the R code used to obtain these statistics is as follows:

```
> t.test(ToothGrowth[ToothGrowth$supp=='VC' & ToothGrowth$dose==0.5,]$len)
```

	0.5	1.0	2.0	Marginal
OJ	(10.04, 16.42)	(19.90, 25.50)	(24.16, 27.96)	(18.20, 23.13)
VC	(6.02, 9.94)	(14.97, 18.57)	(22.71, 29.57)	(13.88, 20.05)
Marginal	(8.50, 12.71)	(17.67, 21.80)	(24.33, 27.87)	(16.84, 20.79)

TABLE 3. Tooth Growth 95% Confidence Intervals

4. Conclusions

It is worth commenting that it was likely illogical to compute marginal statistics. The primary assumption that such a statistic could violate is that the data we have is representative of the entire population. For instance, it would not be unreasonable to assume that there are other methods of delivering vitamin C besides orange juice and ascorbic acid, and it is also possible that different dosages could have been given. In summary, such marginal statistics may not come from any sort of physically meaningful distribution. The following analysis will therefore be restricted to comparing cross-sectional data.

Given that the confidence intervals of OJ given at the 0.5 and 1.0 levels do not overlap, it seems highly likely that the latter comes from a distribution with a higher mean. Indeed, the p-value given when a T test is performed on these two data sets is extremely small, on the order of 10^{-4}. Similar conclusions can be reached for all dosage levels comparisons of a VC supplement.

On the other hand, the confidence intervals of OJ given at 1.0mg and 2.0mg doses do overlap. However, their p-value assuming unequal variances is $0.039 < 0.05$. This number is even lower, 0.037, assuming equal variance. Therefore, we can still conclude with over 95% confidence that giving 2.0mg/day doses of OJ is more effective in promoting tooth growth than 1.0mg/day doses of OJ. This discrepancy is due to the fact that the small overlap occurs at the extremes of each interval, so both means being such an extreme value is unlikely.

It can thus be concluded that for the dosage levels studied, a higher dose is always more effective. It would be wrong though to try to extrapolate this data to make predictions about even larger dosages.

At both the 0.5mg and 1.0mg levels, it is clear that an OJ supplement is more effective than VC. However, the 2.0mg level is far more complicated. The mean tooth growth with a 2.0mg dosage is 26.04 and 26.14 for OJ and VC respectively, two extremely close numbers. Further, the confidence interval for OJ is actually a subset of VC's confidence interval. With a p-value of 0.964, no conclusion can be reached regarding the superior method. As another way of putting it, the 95% confidence interval of the difference between OJ and VC lengths is $(-3.80, 3.64)$. Since this interval contains 0, the null hypothesis that the two distributions have an equal mean cannot be rejected.

In conclusion, at lower dosage levels using an orange juice supplement will result in more ondotoblast growth than an ascorbic acid supplement. However, if the goal is to maximize tooth growth one should always use a 2.0mg/day dosage. In this case it does not matter which type of supplement is used, as one cannot conclude that one method is superior to the other. However, it may be safer to use a 2.0mg OJ supplement since it has a smaller variance. More accurate conclusions could be drawn with a larger sample size, as comparing two sets of ten observations is not particularly reliable. It would also be of interest to study higher dosage levels.