

Analysis of Tooth Growth Data

Statistical Inference Course Project - Part 2

Brian Gulbis

October 03, 2016

Overview

In this report, the TooothGrowth data from the R datasets package will be analyzed to determine if there is a difference in the amount of tooth growth between subjects receiving different types of Vitamin C supplements and doses. The data will be evaluated using the t-test, and the results will show whether or not there is a significant difference in the mean odontoblast length between groups.

Exploratory Data Analysis

An exploration of the data reveals two data elements which could be used to group the data: supplement type and supplement dose. In an exploratory plot of the data, it appears there may be a difference in the mean odontoblast length based on the supplement dose and by the supplement type for at least some of the doses (see figure 1).

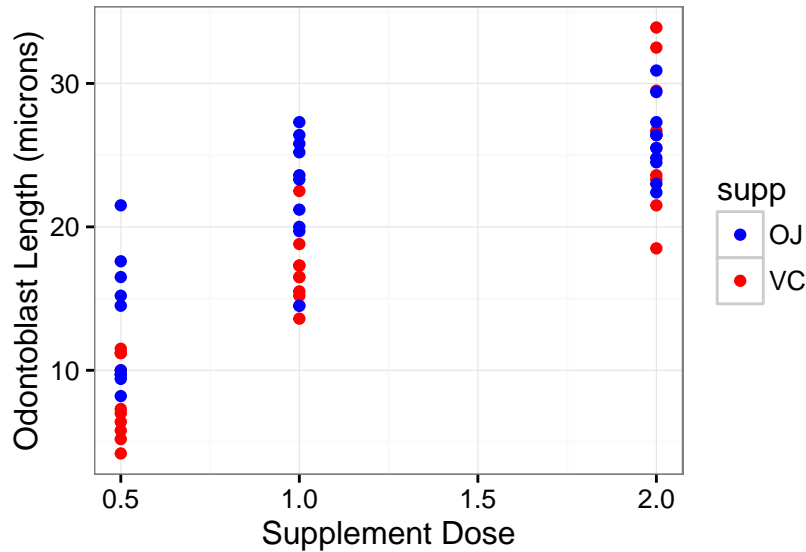


Figure 1: Odontoblast Length by Supplement Type and Dose

Through a plot of the distribution of odontoblast length for each group (by supplement type and dose), it becomes more apparent that the mean odontoblast length is greater in the subjects who received orange juice supplementation at doses of 0.5 and 1 mg/day, but not at 2 mg/day, when compared with those who received the chemically pure Vitamin C supplement at the equiavalent dosages (see figure 2). The dashed blue line indicates the mean in each group of subjects. Additionally, the data appear to be approximately normally distributed within each group.

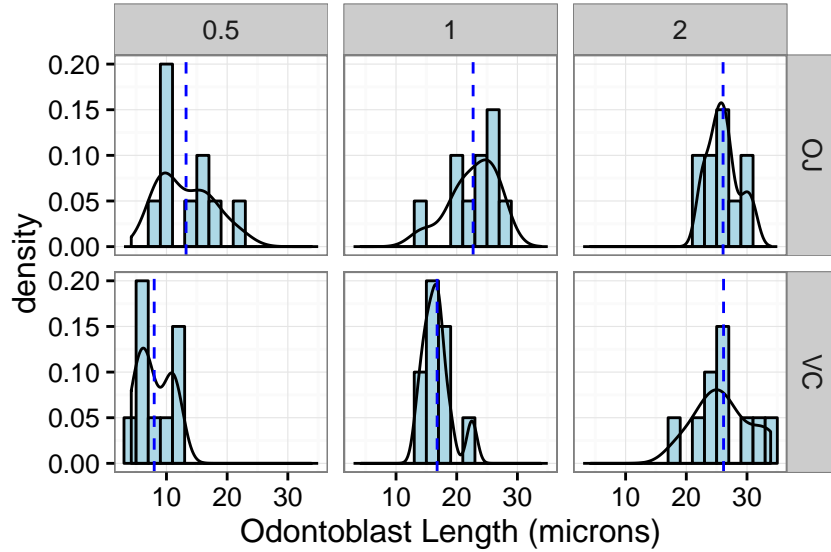


Figure 2: Comparison of Odontoblast Length by Supplement Type and Dosage

Hypothesis Testing

Since it appears that there may be a difference in the mean odontoblast length between supplement types and dosages, these differences will now be tested to determine if there truly is a significant difference. The data will be subsetting into six groups based on supplement type and supplement dose. The results of each supplement type will be compared at equivalent supplement dosages using the t-test. The null hypothesis for each comparison is that there is no difference in the mean odontoblast length for subjects receiving the different supplement types or dosages.

- $H_0 : \mu = \mu_0$
- $H_a : \mu \neq \mu_0$

The results of the comparisons are displayed in table 1 below.

Table 1: Comparison of Mean Odontoblast Length By Supplement Type within 3 Different Supplement Dosages

Dosage	Means (OJ vs. VC)	t-Statistic	95% Confidence Interval	p-Value
0.5	13.23, 7.98	3.17	1.719, 8.781	0.006
1.0	22.7, 16.77	4.033	2.802, 9.058	0.001
2.0	26.06, 26.14	-0.046	-3.798, 3.638	0.964

This reveals that there is a significant difference between the two supplement types at doses of 0.5 and 1 mg/day, because the 95% confidence interval does not cross 0, meaning that the means cannot be the same. Additionally, the p-value is < 0.05 , which also indicates that the means are significantly different. For these two groups, the null hypothesis is rejected.

However, for the 2 mg/day group, the 95% confidence interval does cross 0, which means that the means could be equivalent. Likewise, the p-value is ≥ 0.05 . Therefore, for the 2 mg/day group, the null hypothesis is accepted.

Conclusions

In this the report, the Tooth Growth data from the R datasets package was evaluated to determine if there was a difference in the amount of tooth growth, as measured by the odontoblast length in microns, between subjects receiving different types of Vitamin C supplementation and dosages.

The data in this analysis were assumed to be independent and identically distributed random variables. Additionally, the data were assumed to follow a normal distribution.

Since the data were assumed to be normally distributed, the t-test was used to perform comparisons. For the t-test calculations, it was assumed that the variances were not equal between the groups in each comparison.

The results of the hypothesis testing show that there is a difference in the mean odontoblast length in the groups receiving different supplement types at dosages of 0.5 and 1 mg/day, favoring larger tooth growth in the groups receiving orange juice supplementation. No difference was found in the groups receiving supplementation at a dose of 2 mg/day.

Appendix

The following is the R code used to produce the figures and statistical results in this report.

Calculating the mean in each group

```
tg.mean <- group_by(tg, supp, dose) %>%  
  summarize(mean.len = mean(len))
```

Figure 1

```
ggplot(tg, aes(x=dose, y=len, color=supp)) +  
  geom_point() +  
  scale_color_manual(values=c("OJ"="blue", "VC"="red")) +  
  xlab("Supplement Dose") +  
  ylab("Odontoblast Length (microns)") +  
  theme_bw()
```

Figure 2

```
ggplot(tg, aes(x=len)) +  
  geom_histogram(binwidth=2, color="black", fill="light blue", aes(y=..density..)) +  
  geom_density() +  
  geom_vline(data=tc.mean, aes(xintercept=mean.len), color="blue", linetype="dashed") +  
  facet_grid(supp ~ dose) +  
  xlab("Odontoblast Length (microns)") +  
  theme_bw()
```

Comparison of groups using the t-test

```
t1 <- t.test(len ~ supp, data=tc, subset=(dose == 0.5))  
t2 <- t.test(len ~ supp, data=tc, subset=(dose == 1))  
t3 <- t.test(len ~ supp, data=tc, subset=(dose == 2))
```