Vitamin C delivery impact on Odontoblast growth

Chris O'Brien

23 July 2015

Contents

Synopsis

In this report we aim to determine whether the dosage and delivery method of ascorbic acid (Vitamin C) has any impact on the length of odontoblasts (Teeth). This is an exploratory analysis and inference on the ToothGrowth dataset source from C.I. Bliss (1952) *The Statistics of Bioassay.* Academic Press.

Once we have loaded the data be perform a very basic transformation on the dosage column converting it to a factor as this represents discrete rather than continuous data. Our exploratory analysis helped us come to two hypotheses:

1. That the method of delivery regardless of dosage impacts odontoblast length
2. The dosage of ascorbic acid regardless of delivery method impacts growth

Applying a T test to our two hypotheses we find that both the method of delivery and the dosage between 0.5 and 2.0mg differences are statistically significant using a 95% confidence interval

Data Processing

Loading the data

The data for this assignment was sourced from the datasets library available in the Base R packages.

Processing the data

From figure 1 the only transformation we applied to the data was to convert the dosage level into a factor. As dosage represented 3 discrete values it made more sense to treat this as such rather than continuous values on a scale.

Exploratory analysis

From figure 2 we can see from our dataset consists of three data points, the length of the odontoblast, the two methods of delivery and the three levels of ascorbic acid dosage

Performing a complete break down of the two facets (dose and supp) from figure 3 it is hard to come to any real conclusions about the data other than possibly as dosage increases that the mean of the odontoblast lengths gets larger.

Exploring further into each factor from figure 4 and 5 further, we certainly get a better feeling that there is a potential difference in the method of delivery and the dosage level.

Confidence testing

Method of delivery

Our hypothesis is that the difference in average odontoblast length between those dosed with orange juice is greater from those given ascorbic acid

$$H_o : \mu_{oj} - \mu_{aa} = 0$$
$$H_a : \mu_{oj} - \mu_{aa} > 0$$

When we perform a two-sample one-sided t-test on the difference in length of odontoblasts given the two different delivery methods we cannot say to a 95% confidence interval of 0.4982 to Inf. In this instance, given a 95% confidence interval, we can successfully reject the null hypothesis and accept our alternative that orange juice has a greater effect on odontoblast length than ascorbic acid.

It should also be noted that in this analysis we have specified that the variances are different as our assumptions for this analysis are that each sample is iid normal and have different means and variances.

Dosage

Our hypothesis in this instance is that the average odontoblast length between those given 0.5mg of Vitamin C is less than those given 2.0mg

$$H_o : \mu_{0.5} - \mu_{2.0} = 0$$
$$H_a : \mu_{0.5} - \mu_{2.0} < 0$$

For our two-sample t-test on the variance in dosage we select the lowest and highest dosage levels, 0.5 and 2.0mg respectively. Similar to before we perform a two-sample one-sided t-test on the difference in odontoblast length between our dosage levels and we find that our 95% confidence intervals are from -Inf to -13.279.

Given our 95% confidence interval on the difference between two to means we can successfully reject our null hypothesis and accept the alternative that those given 0.5mg of Vitamin C have a smaller odontoblast length than those given 2.0mg

Similar to before this assumes that each sample is iid normal and that each has different variances, which is why in our t-test we set that the variances are different.

Further analysis

After reading the original study (Page 499) it was interesting to note a few points that aren't mentioned in the help information for this dataset.

- The was an equal split in the sex of the Guinea Pigs and that this could be neglected.
- The original aim of the experiment was to determine whether the antiscorbutic activity of orange juice was greater than that of the ascorbic acid alone.
- They concluded that there was a difference between the dosage of Orange Juice and Ascorbic Acid alone.
  - They ignored the 2.0mg dosage of orange juice as the data suggested there was a ceiling at 2.0mg of Vitamin C. See figure 8
  - *Note*: This is a reproduction of a graph from Page 501 of the original paper.
  - This same behaviour is demonstrated in figure 3 where the distributions for the 2.0mg dosages are very similar.
- While we havent added the data here, performing a two-sided t-test on the method of delivery acutally fails to reject the null hypothesis. As the exploratory analysis suggested that there wasn't just a difference but that one had a greater effect than the other we chose a one-sided test.

Appendix

Data processing

Figure 1 .

```
# As mentioned we mutate the dose into a factor to allow better comparisons
tg <- data.table(ToothGrowth)
tg <- tg %>%
    mutate(dose=factor(dose))
```

Exploratory analysis

Figure 2 .

```
summary(tg)
```

```
##       len        supp     dose
##  Min.   : 4.20   OJ:30   0.5:20
##  1st Qu.:13.07   VC:30   1  :20
##  Median :19.25           2  :20
##  Mean   :18.81
##  3rd Qu.:25.27
##  Max.   :33.90
```

Figure 3 .

Figure 4 .

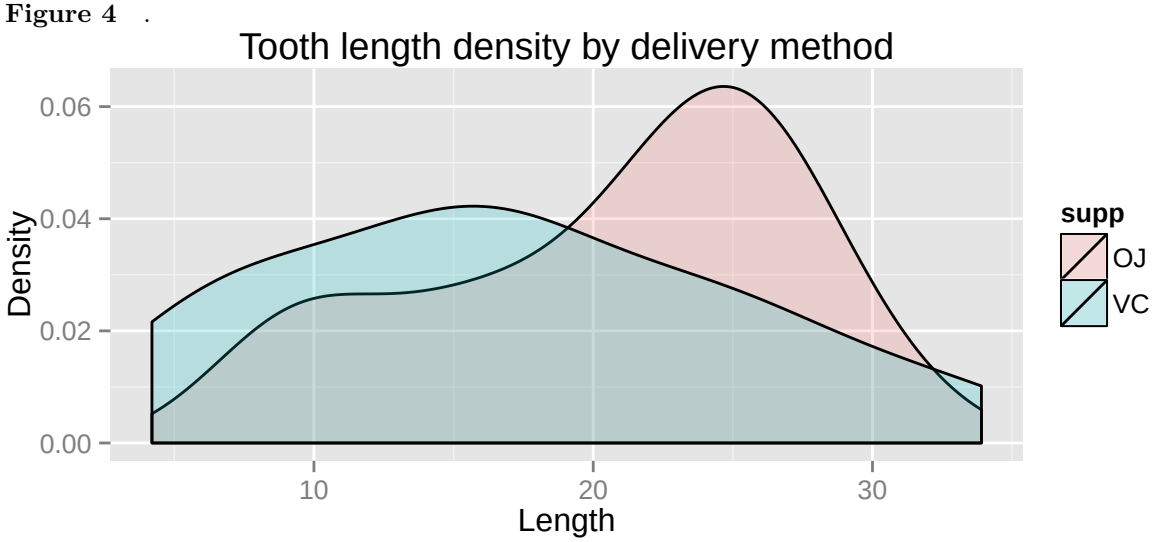
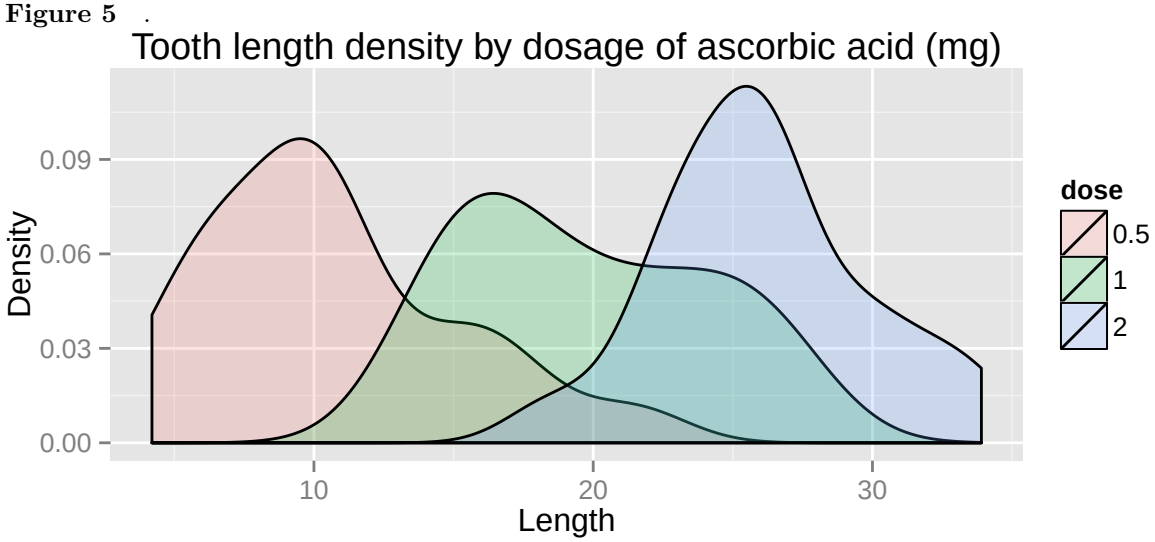


Figure 5 .



Confidence testing

Figure 6 .

```
t.test(len ~ supp, data=tg, paired=F, var.equal=F, alternative = 'g')
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 55.309, p-value = 0.03032
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4682687       Inf
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

Figure 7 .

```
tg1 <- tg %>% filter(dose %in% c(0.5,2.0))
t.test(len ~ dose, data=tg1, paired=F, var.equal=F, alternative='l')
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -11.799, df = 36.883, p-value = 2.199e-14
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##       -Inf -13.27926
## sample estimates:
## mean in group 0.5   mean in group 2
##            10.605            26.100
```

Further analysis

Figure 8 .