Tooth Growth.R

Harold Trammel

August 20, 2015

Purpose for the Study

According to patched R Reference Manual entry for the Tooth Growth dataset, the citation for the original study from which this data was obtained is Crampton, E. W. (1947) The growth of the odontoblast of the incisor teeth as a criterion of vitamin C intake of the guinea pig. *The Journal of Nutrition* 33(5): 491–504. [link] (http://jn.nutrition.org/content/33/5/491.full.pdf).

In this paper, the authors state that the purpose of the study was to evaluate the use of odontoblast growth as a bioassay for vitamin C levels. The procedures used in the study preclude paired comparisons of the results. The data in this dataset compares the bioassay for vitamin C in orange juice and in an aqueous solution of pure vitamin C. In reviewing the paper, I have inferred a secondary purpose, which is to prove the bioassay using odontoblasts was superior to the bioassay using weight gain. Much of the statistical evidence presented in the paper compared the precision of the two bioassays. Little statistical attention was given to the accuracy and precison within the odontoblast bioassay in regard to amounts determined.

Based on the stated purpose, I would expect that the bioassay would accurately, consistently, and precisely indicate the amount of vitamin C in a given source. The study data contains odontoblast lengths from six groups of ten guinea pigs. Three groups received an aqueous solution of vitamin C in three different doses. The other three groups received orange juice in three diffent dosages. Since the authors of the study state that they had seen a dose-related response in the bioassay and that there was a range for which it was satisfactory, it seemed important to compare the results for the full range as well as for the individual doses.

Data Exploration

The dataset provided was poorly documented. The description that came with the dataset is confusing as it can be interpreted to say that each guinea received each of the three doses from both vitamin C sources. From the paper, I determined that there were a total of sixty guinea pigs with each set of ten subject receiving only one dosing regimen. Paired comparisons are not appopriate. While the dataset is called **Tooth Growth**, the actual measures where changes in ondontobast length which are inside teeth.

Figure 1 shows a box plot of the Tooth Growth data comparing changes in odontoblast length to the two different supplements received. This graph indicates similar changes in length from both supplements but the vitamin C group had a wider interquartile range.

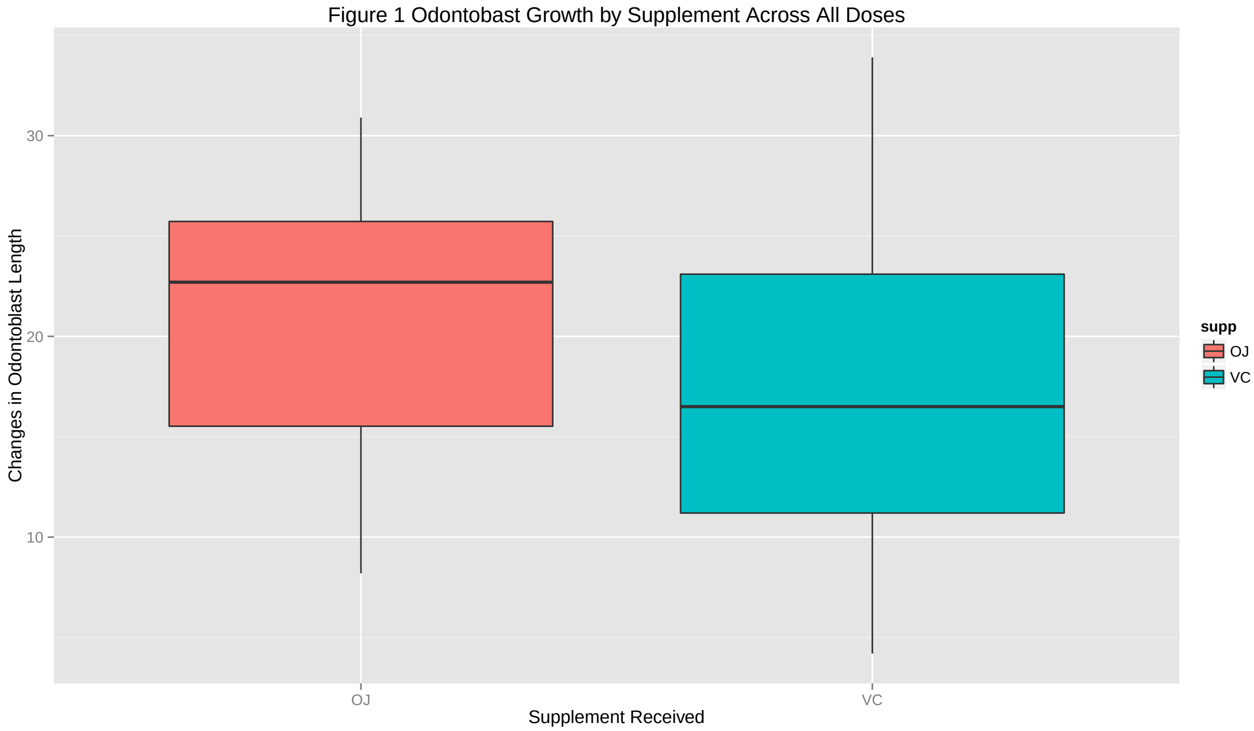

Figure 2 shows a box plot of the Tooth Growth data but graph compares the changes in length with both dose and supplement. The results from the 2.0 mg dose are similar between the supplements. This is not the case for the lower doses as the orange juice groups had more change in length. This suggests that the dose-related response previously seen by the authors may be supported.

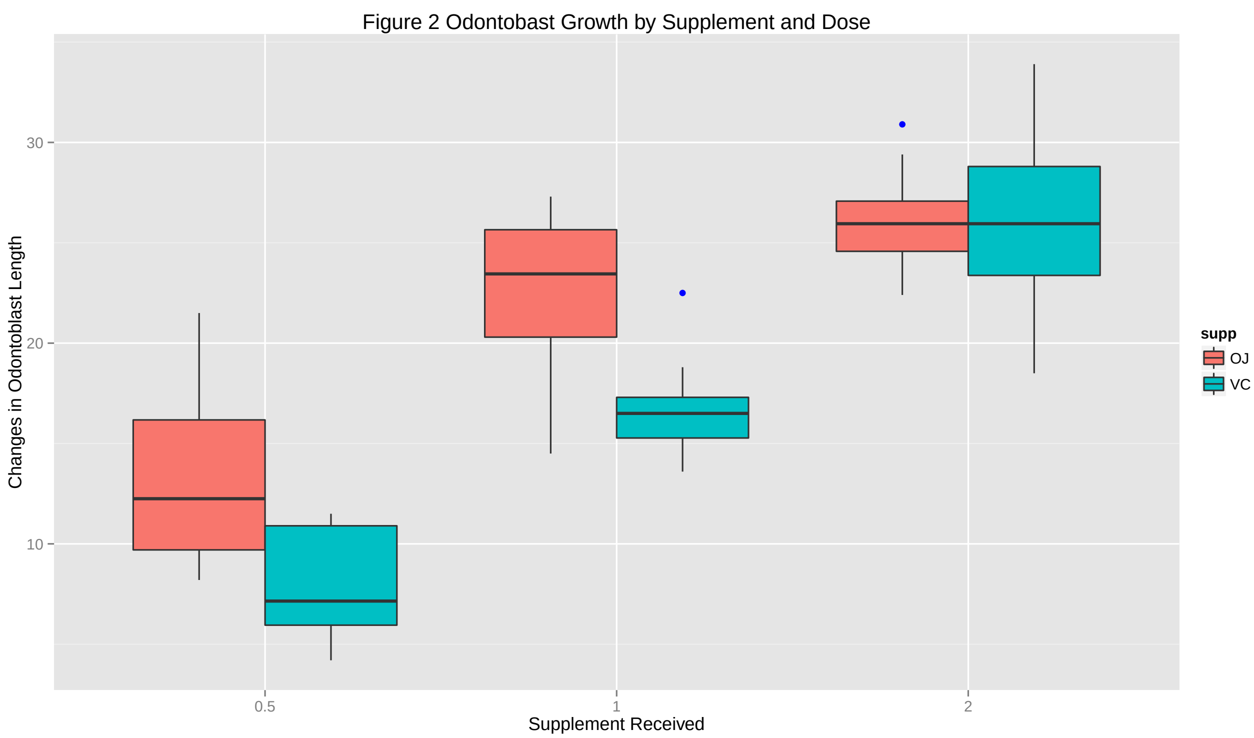

Table 1 shows the basis statistics for the dose specific data. As might be expected from the graphs, the means at the 2.0 dose are

Supplement	Dose	Mean	Standard Deviation	Standard Error
Vitamin C	0.5	7.98	2.746634	0.8685620
Orange Juice	0.5	13.23	4.459708	1.4102837
Vitamin C	1.0	16.77	2.515309	0.7954104
Orange Juice	1.0	22.70	3.910953	1.2367520
Vitamin C	2.0	26.14	4.797731	1.5171757
Orange Juice	2.0	26.06	2.655058	0.8396031

Table 1 Summary of Statistics

Data Analysis

I chose to have two null hypotheses since there are two different appoaches to using this reseach.

My first null hypothesis (H_O) is that there is no difference between the bioassay of orange juice and an aqueous solution of vitamin C across a 0.5 mg to 2.0 mg dosing range.

My first alternative hypothesis (H_A) is that there is a difference between the bioassay of orange juice and an aqueous solution of vitamin C across a 0.5 mg to 2.0 mg dosing range.

My second null hypothesis (H_{Oa}) is that there is no difference between the bioassay of orange juice and an aqueous solution of vitamin C at each of the three dosing regimens.

My second alternative hypothesis (H_{Aa}) is that there is a difference between the bioassay of orange juice and an aqueous solution of vitamin C at each of the three dosing regimens.

H_O versus H_A

Applying the Student's T-test to the odontoblast length versus the supplement given resulted in the following:

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len by ToothGrowth$supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

The H_0 was accepted for the full dosage range.

H_{0a} versus H_{Aa}

Applying the Student's T-test to the odontoblast length versus the dose and supplement given resulted in the following:

```
##
##  One Sample t-test
##
```

```
## data:  d05_diff
## t = 2.9791, df = 9, p-value = 0.01547
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.263458 9.236542
## sample estimates:
## mean of x
##      5.25
```

The H_{0a} hypothesis is rejected at the 0.5 mg dose.

```
##
##  One Sample t-test
##
## data:  d1_diff
## t = 3.3721, df = 9, p-value = 0.008229
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.951911 9.908089
## sample estimates:
## mean of x
##      5.93
```

The H_{0a} hypothesis is rejected at the 1.0 mg dose.

```
##
##  One Sample t-test
##
## data:  d2_diff
## t = -0.042592, df = 9, p-value = 0.967
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -4.328976  4.168976
## sample estimates:
## mean of x
##     -0.08
```

The H_{0a} hypothesis is accepted at the 2.0 mg dose.

Appendix

Setup

```r
data("ToothGrowth")
library(dplyr)
library(ggplot2)
library(knitr)

se <- function(x) sqrt(var(x)/length(x))
fig <- 0
```

```r
tbl <- 0

tg <- ToothGrowth %>%
    group_by(supp, dose) %>%
    summarise_each(funs(mean, sd, se)) %>%
    rename(tg_mean = mean, tg_sd = sd, tg_se = se)

tg_stats <- rbind(tg[4,], tg[1,], tg[5,], tg[2,], tg[6,], tg[3,])
tg_stats <- rbind(tg[4,], tg[1,], tg[5,], tg[2,], tg[6,], tg[3,])
names(tg_stats) <- c("Supplement","Dose","Mean","Standard Deviation","Standard Error")
tg_stats$Supplement <- as.character(tg_stats$Supplement)

for (i in c(1,3,5)) tg_stats[i,1] <-"Vitamin C"
for (i in c(2,4,6)) tg_stats[i,1] <-"Orange Juice"

vc05 <- ToothGrowth$len[ 1 : 10]
vc1  <- ToothGrowth$len[11 : 20]
vc2  <- ToothGrowth$len[21 : 30]
oj05 <- ToothGrowth$len[31 : 40]
oj1  <- ToothGrowth$len[41 : 50]
oj2  <- ToothGrowth$len[51 : 60]
```

Exploratory Data Analysis

```r
fig <- fig + 1
pl0 <- ggplot(ToothGrowth, aes(factor(supp), y=len, fill = supp)) +
    geom_boxplot(stat="boxplot", outlier.colour = "blue") +
    ggtitle(paste("Figure", fig, "Odontobast Growth by Supplement Across All Doses")) +
    xlab("Supplement Received") +
    ylab("Changes in Odontoblast Length")
print(pl0)

fig <- fig + 1
pl1 <- ggplot(ToothGrowth, aes(factor(dose), y=len, fill = supp)) +
    geom_boxplot(stat="boxplot", outlier.colour = "blue", notch = FALSE) +
    ggtitle(paste("Figure", fig, "Odontobast Growth by Supplement and Dose")) +
    xlab("Supplement Received") +
    ylab("Changes in Odontoblast Length")
```

Analysis

```r
t_all <- t.test(ToothGrowth$len~ToothGrowth$supp, conf=0.95)

print(t_all)

d05_diff <- oj05 - vc05
t05 <-  t.test(d05_diff)
print(t05)

d1_diff <- oj1 - vc1
```

```r
t1 <-  t.test(d1_diff)
print(t1)

d2_diff <- oj2 - vc2
t2 <-  t.test(d2_diff)
print(t2)
```