Effects of Vitamin C on Tooth Growth

Lucas Qualmann

9/23/2019

Overview

This report utilized the dataset ToothGrowth in R to look at the effects of Vitamin C on toothgrowth for guineau pigs. The analysis shows the type of supplement does not have a statistically significant impact on tooth length, but the dosage rate does have a statistically significant impact on tooth length.

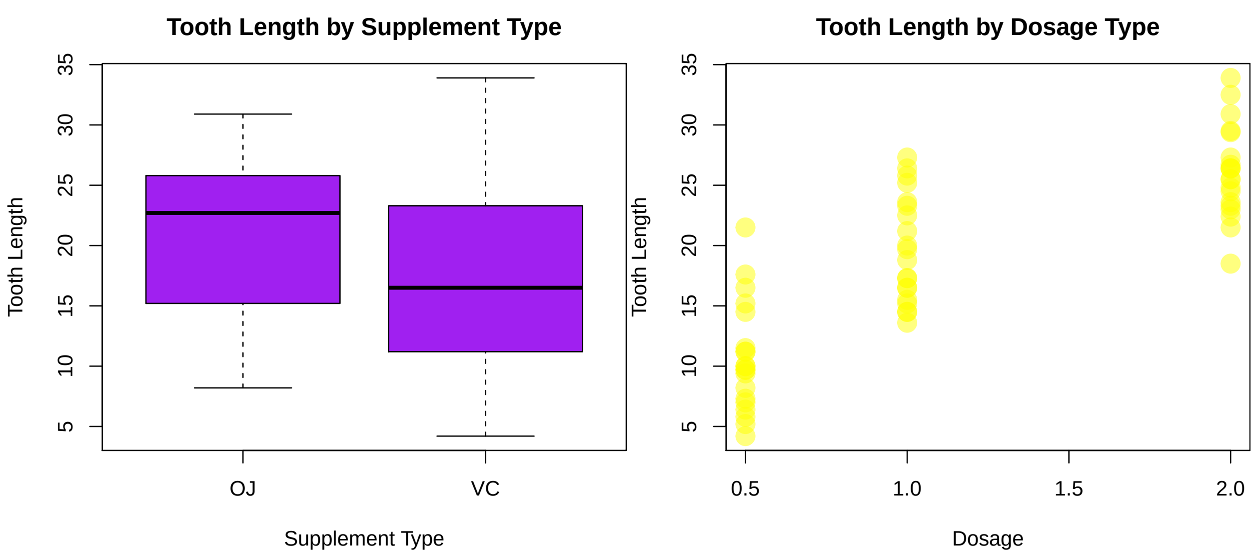
Summary

From the R help page, the data contains 60 observations on 3 variables: len, supp, and dose. Len is the tooth length of the guineau pig, supp is the supplement type, and dose is the dose in milligrams/day. The supplement type is either OJ (orange juice) or VC (acorbic acid). The dose of the supplement is either 0.5, 1.0, or 2.0 milligrams/day.

```
##   Statistic Tooth_Length
## 1      Mean    18.813333
## 2    Median    19.250000
## 3       Min     4.200000
## 4       Max    33.900000
## 5        SD     7.649315
```

The table above shows the basic summary statistics on the length of the teeth measured in the dataset, not broken out by any of the other variables. The mean of 18.81 is slightly below the mediam of 19.25. The range of lengths is 4.2 to 33.9. Finally, the standard deviation is 7.65.

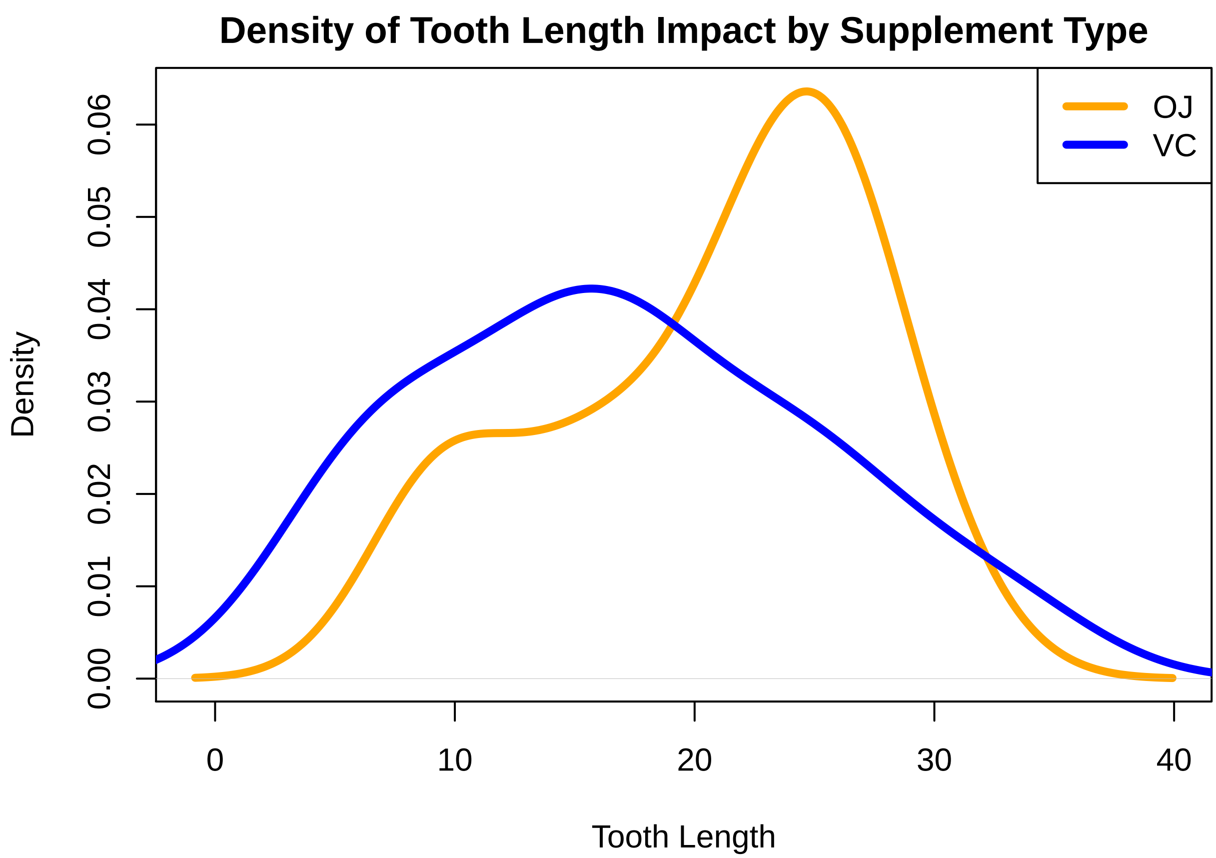
Now let's break down the data in terms of supplement type and dosage type. To do this, we'll make two plots showing the values of tooth length by supplement type and dosage type.



The first plot shows OJ having a higher average tooth length than VC, but the large range of VC values do call into question whether or not supplement type is a major factor for tooth growth. The dosage plot, on the other hand, does seem to show a higher dosage resulting in more tooth growth.

Delivery Method Effects

First we want to look at whether or not the delivery method has an impact on tooth length. Our first analysis will be to look at the density plots of OJ and VC supplement type to get an idea as to whether or not it looks like there is an impact.



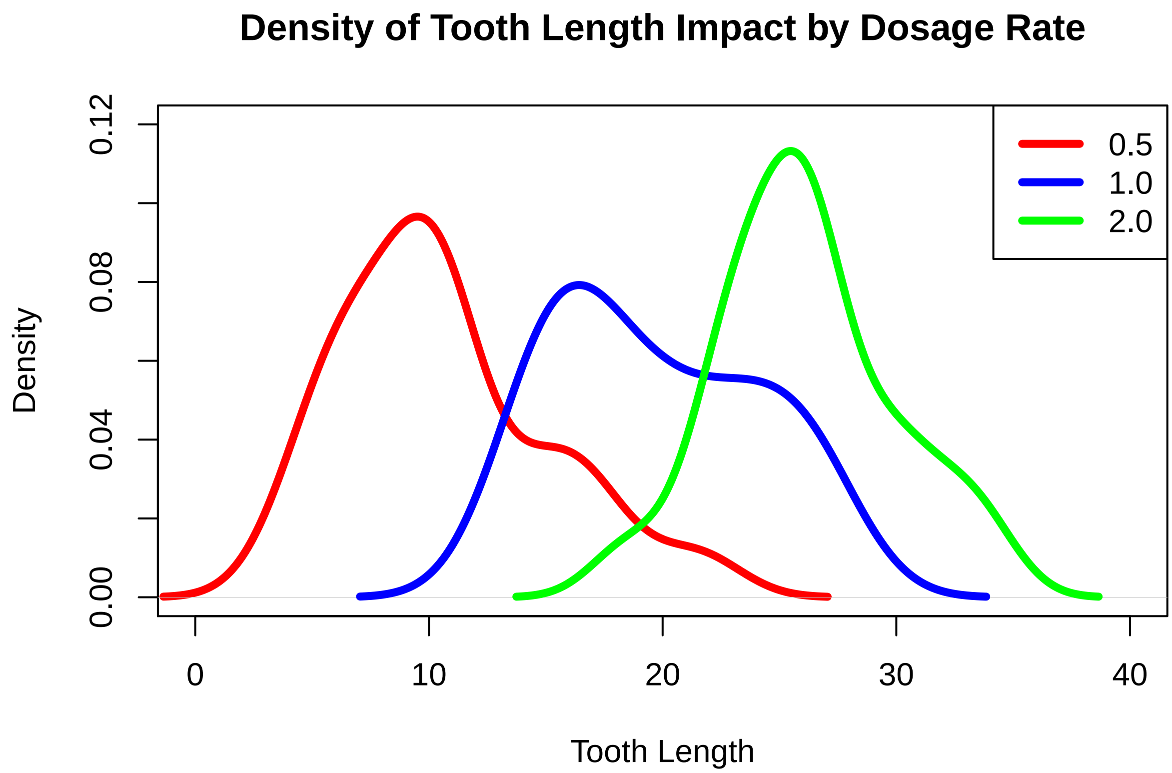
The above chart shows that the two supplement distributions overlap quite a bit. This chart would support the null hypothesis that there isn't much of an impact caused by the supplement type. To make sure this is correct, we will run a t.test.

```
##    Statistic   Low_Value High_Value
## 1    P-Value  0.06039337         NA
## 2 Conf.Level -0.16700642   7.567006
```

The main items we are looking for on the t.test is the p-value and the 95% confidence interval. The p-value tells us the probability of the data occuring assuming the null hypothesis (no effect) is correct. The p-value needs to be less than 5% in order for us to reject the null hypothesis. In addition, the 95% confidence interval also lets us know if we should accept or reject the null hypothesis. If the 95% confidence interval contains 0, that is a sign that we should accept the null hypothesis. If the interval doesn't contain 0, we should reject it. Since the t.test shows the p-value is above 5% and the 95% confidence interval contains 0, the null hypothesis of the type of supplement not making an impact is accepted.

Dose Effects

Next, we want to look at whether or not the dose of vitamin C has an impact on tooth length. Our first analysis will be to look at the density plots of the three rates to get an idea as to whether or not it looks like there is an impact.



The plot does suggest a relationship between dosage rate and tooth length. As the dosage rate increases, tooth length increases as well. In addition, there is not a lot of overlap between the densities of the 3 rates. This is especially obvious when looking at the difference between 0.5 and 2.0 milligrams per day. Based on the chart, it looks like we should reject the null hypothesis of dosage rate not affecting the tooth length. Let's run some t.tests to verify. We'll need to run a t.test 3 times for every combination of rates since we can only compare 2 rates at a time.

```
##            Statistic     Low_Value High_Value
## 1     P-Value (.5-1)  1.266297e-07         NA
## 2 Conf.Level (.5-1) -1.198375e+01  -6.276252
## 3     P-Value (.5-2)  2.837553e-14         NA
## 4 Conf.Level (.5-2) -1.815352e+01 -12.836481
## 5      P-Value (1-2)  1.810829e-05         NA
## 6  Conf.Level (1-2) -8.994387e+00  -3.735613
```

The t.tests confirm our hypothesis of dosage rate making an impact. All of the p-values are well below our 5% threshold, and none of the 95% confidence intervals contain 0. This means we can reject the null hypothesis and accept the alternative hypothesis that dosage rate affects tooth length.

Conclusions

The conclusions that we are able to draw from our analysis is the supplement type of orange juice or ascorbic acid does not have a statistically significant impact on tooth length. However, we have shown that dosage rate of vitamin C does have a statistically significant impact. The impact is the higher the dosage rate, the more tooth length there is. These conclusions are based on the assumption of the observations being independent of each other.

Appendix

R Code

```
#summary
stats <- data.frame("Statistic" = c("Mean", "Median", "Min", "Max", "SD"),
                    "Tooth_Length" = c(mean(ToothGrowth$len),
                                       median(ToothGrowth$len),
                                       min(ToothGrowth$len),
                                       max(ToothGrowth$len),
                                       sd(ToothGrowth$len)))
print(stats)

par(mfrow = c(1, 2), mar = c(4, 4, 3, 0))
plot(ToothGrowth$supp,ToothGrowth$len, col = "purple", xlab = "Supplement Type",
     ylab = "Tooth Length", main = "Tooth Length by Supplement Type")
plot(ToothGrowth$dose,ToothGrowth$len,
     col = rgb(red = 1, green = 1, blue = 0, alpha = 0.5), xlab = "Dosage",
     ylab = "Tooth Length", main = "Tooth Length by Dosage Type", pch = 19,
     cex = 2)

par(mfrow = c(1, 1), mar = c(4,4,2,1))
library(dplyr)
oj <- ToothGrowth %>% select(len, supp) %>% filter(supp == "OJ") %>% pull(len)
vc <- ToothGrowth %>% select(len, supp) %>% filter(supp == "VC") %>% pull(len)
plot(density(oj), col = "orange", lwd = 4, xlab = "Tooth Length",
     main = "Density of Tooth Length Impact by Supplement Type")
lines(density(vc), col = "blue", lwd = 4)
legend("topright", legend = c("OJ", "VC"),
       col = c("orange", "blue"), lwd = 4)

ptype <- t.test(oj, vc, paired = FALSE, var.equal = TRUE,
                data = ToothGrowth)$p.value
ctype <- t.test(oj, vc, paired = FALSE, var.equal = TRUE,
                data = ToothGrowth)$conf
typestats <- data.frame("Statistic" = c("P-Value", "Conf.Level"),
                        "Low_Value" = c(ptype, ctype[1]),
                        "High_Value" = c(NA, ctype[2]))
print(typestats)

#dose effects
d5 <- ToothGrowth %>% select(len, dose) %>% filter(dose == 0.5) %>% pull(len)
d1 <- ToothGrowth %>% select(len, dose) %>% filter(dose == 1) %>% pull(len)
d2 <- ToothGrowth %>% select(len, dose) %>% filter(dose == 2) %>% pull(len)
plot(density(d5), col = "red", lwd = 4, xlab = "Tooth Length", ylim = c(0, .12),
     main = "Density of Tooth Length Impact by Dosage Rate", xlim = c(0, 40))
lines(density(d1), col = "blue", lwd = 4)
lines(density(d2), col = "green", lwd = 4)
legend("topright", legend = c("0.5", "1.0", "2.0"),
       col = c("red", "blue", "green"), lwd = 4)

prate1 <- t.test(d5, d1, paired = FALSE, var.equal = TRUE,
                data = ToothGrowth)$p.value
crate1 <- t.test(d5, d1, paired = FALSE, var.equal = TRUE,
```

```
                data = ToothGrowth)$conf
prate2 <- t.test(d5, d2, paired = FALSE, var.equal = TRUE,
                 data = ToothGrowth)$p.value
crate2 <- t.test(d5, d2, paired = FALSE, var.equal = TRUE,
                 data = ToothGrowth)$conf
prate3 <- t.test(d1, d2, paired = FALSE, var.equal = TRUE,
                 data = ToothGrowth)$p.value
crate3 <- t.test(d1, d2, paired = FALSE, var.equal = TRUE,
                 data = ToothGrowth)$conf

ratestats <- data.frame("Statistic" = c("P-Value (.5-1)", "Conf.Level (.5-1)",
                                        "P-Value (.5-2)", "Conf.Level (.5-2)",
                                        "P-Value (1-2)", "Conf.Level (1-2)"),
                        "Low_Value" = c(prate1, crate1[1], prate2, crate2[1],
                                        prate3, crate3[1]),
                        "High_Value" = c(NA, crate1[2], NA, crate2[2],
                                         NA, crate3[2]))
print(ratestats)
```

R Session Info

```
print(sessionInfo())
```

```
## R version 3.6.1 (2019-07-05)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 18362)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] dplyr_0.8.3
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.2        crayon_1.3.4     digest_0.6.20    assertthat_0.2.1
##  [5] R6_2.4.0          magrittr_1.5     evaluate_0.14    pillar_1.4.2
##  [9] rlang_0.4.0       stringi_1.4.3    rmarkdown_1.15   tools_3.6.1
## [13] stringr_1.4.0     glue_1.3.1       purrr_0.3.2      xfun_0.9
## [17] yaml_2.2.0        compiler_3.6.1   pkgconfig_2.0.2  htmltools_0.3.6
## [21] tidyselect_0.2.5 knitr_1.24       tibble_2.1.3
```