Analysis and testing of tooth growth in Guinea Pigs

Rohil Narula

20 July 2018

Simulation exercise

Overview

This paper will attempt to document the analyses and hypothesis tests carried out in relation with supply and dosage of Vitamin C and its effects on the growth of teeth in Guinea pigs. Student's t-tests will be carried out and subsequent confidence intervals will be calculated.

Processing

In this section, the initial variables are set, and the dataset is loaded. Then, the dataset is coerced into a more plyable and ordered form and printed.

Variable initialization

Loading initial packages.

```
packages <- c('ggplot2', 'dplyr', 'tidyr', 'datasets')
sapply(packages, require, character.only = TRUE, quietly = TRUE)
```

```
##  ggplot2    dplyr    tidyr datasets
##     TRUE     TRUE     TRUE     TRUE
```

The dataset is then loaded.

```
data("ToothGrowth")
ToothGrowth <- tbl_df(ToothGrowth)
ToothGrowth <- mutate(ToothGrowth, dose = as.factor(dose))
ToothGrowth <- group_by(ToothGrowth, supp, dose)
ToothGrowth
```

```
## # A tibble: 60 x 3
## # Groups:   supp, dose [6]
##      len supp  dose
##    <dbl> <fct> <fct>
##  1   4.2 VC    0.5
##  2  11.5 VC    0.5
##  3   7.3 VC    0.5
##  4   5.8 VC    0.5
##  5   6.4 VC    0.5
##  6  10   VC    0.5
##  7  11.2 VC    0.5
##  8  11.2 VC    0.5
##  9   5.2 VC    0.5
## 10   7   VC    0.5
## # ... with 50 more rows
```

Exploratory analysis

Quantitative analysis

From the description of the dataset, The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, orange juice or ascorbic acid (a form of vitamin C and coded as VC).

The following lines of code and plots will attempt to provide a quantitative summary of the data.

```r
summary(ToothGrowth)
```

```
##       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
```

```r
summarise(ToothGrowth, mean(len))
```

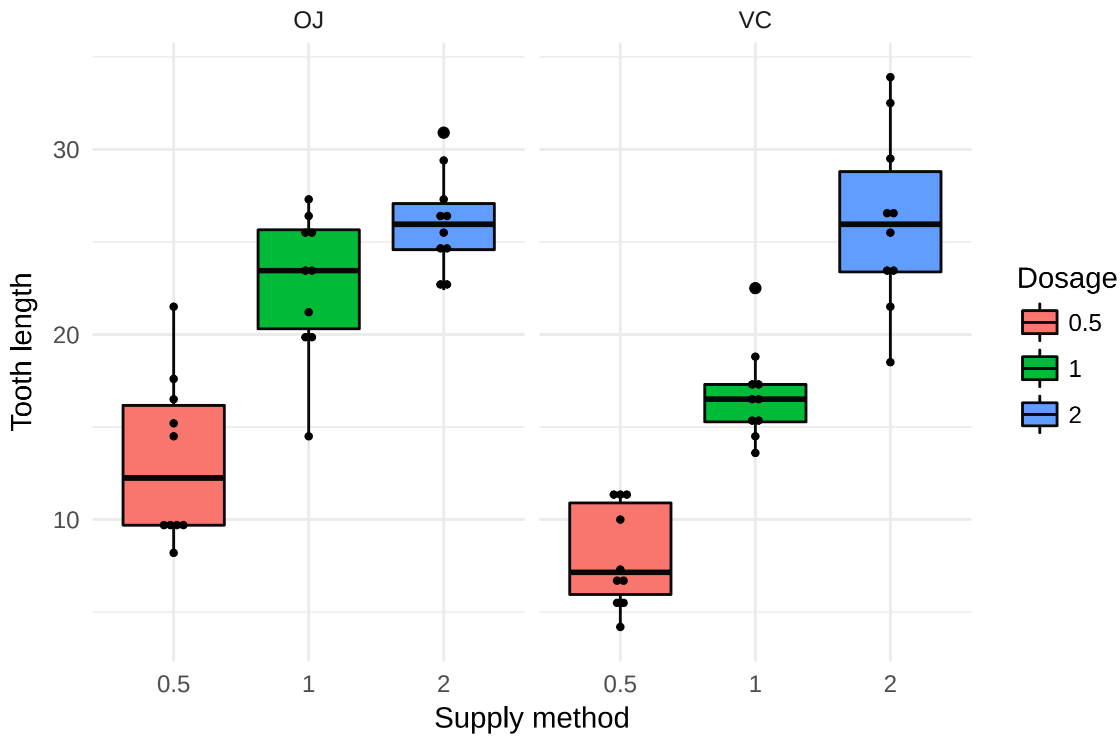
```
## # A tibble: 6 x 3
## # Groups:   supp [?]
##   supp  dose  `mean(len)`
##   <fct> <fct>       <dbl>
## 1 OJ    0.5         13.2
## 2 OJ    1           22.7
## 3 OJ    2           26.1
## 4 VC    0.5          7.98
## 5 VC    1           16.8
## 6 VC    2           26.1
```

```r
ToothGrowth[which(ToothGrowth$len == max(ToothGrowth$len)), ]
```

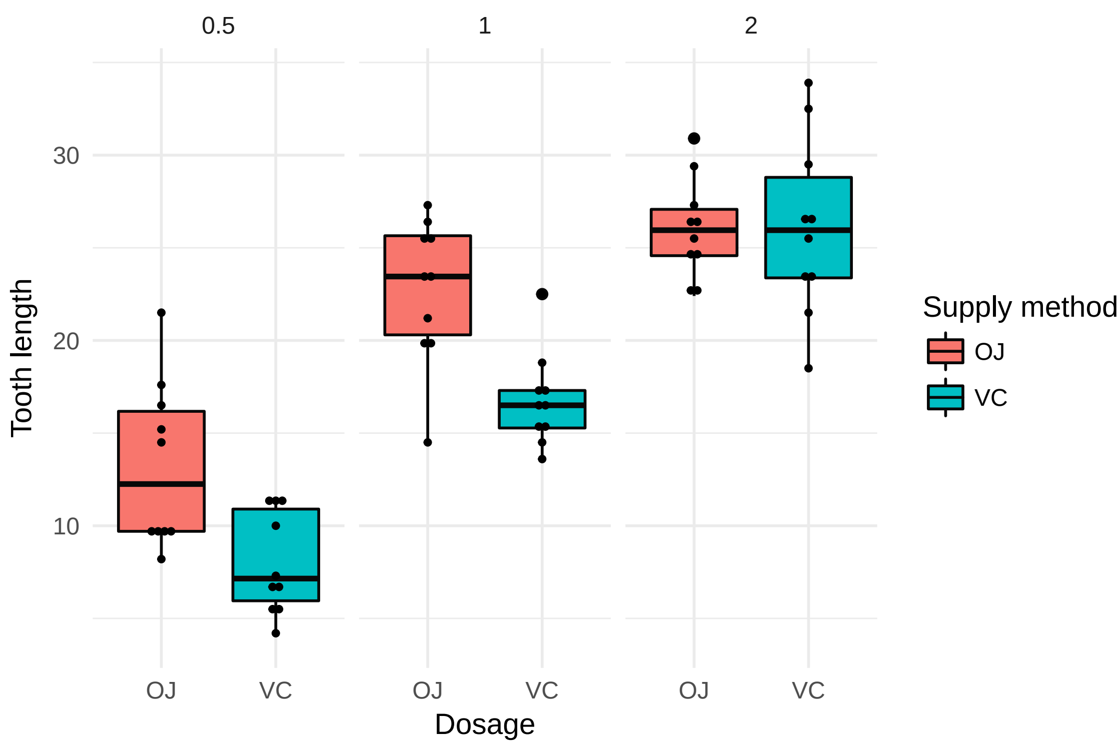
```
## # A tibble: 1 x 3
## # Groups:   supp, dose [1]
##     len supp  dose
##   <dbl> <fct> <fct>
## 1  33.9 VC    2
```

Plotting the data

A few plots to understand the data distribution.



Now plotting the same data, but factoring the dosage.



Summary

Initial inferences from exploratory work

Look at the quantitative exploratory analyses. If the supply method of orange juice (OJ) is examined, it can be noted that the tooth length steadily increases with dose. The same is the case for the ascorbic acid (VC) supply method. Therefore, it seems initially, that increase in dosage has a positive effect on tooth length.

Supply method comparison

Factoring out the dosage, it seems that in the lower dosages (0.5 and 1), the OJ supply method leads to larger tooth length. However, at the dosage of 2, they seem to be about the same, but VC has more variability to it. It can also be noted that the maximum tooth length was obtained by a VC dose of 2 (33.9).

More detailed analysis is needed to be carried out to draw informed conclusions.

Student's t-test and confidence intervals

First, the two supply methods are seperated and tested assuming that the variances are different. Hence, pooled variance analysis is carried out.

```
test <- t.test(ToothGrowth$len[ToothGrowth$supp == "OJ"],
               ToothGrowth$len[ToothGrowth$supp == "VC"],
               paired = FALSE, var.equal = FALSE)
test$p.value
```

```
## [1] 0.06063451
```

```
test$conf.int
```

```
## [1] -0.1710156  7.5710156
## attr(,"conf.level")
## [1] 0.95
```

Assuming $\alpha = 0.05$ and the p-value of the test being 0.06063, we can conclude that there is not enough evidence to reject the null hypothesis. Therefore, statistically, Student's t-test suggests that there is no statistical difference in the choice of supply methods (the value 0 lies in the 95% confidence interval).

Now, the dosage will be tested in a similar way, however the variance of the two dosages are assumed to be the same.

```
test <- t.test(ToothGrowth$len[ToothGrowth$dose == 0.5],
               ToothGrowth$len[ToothGrowth$dose == 1],
               paired = FALSE, var.equal = TRUE)
test$p.value
```

```
## [1] 1.266297e-07
```

```
test$conf.int
```

```
## [1] -11.983748  -6.276252
## attr(,"conf.level")
## [1] 0.95
```

The extremely low p-value in this case is suggestive of the correlation between a higher dosage and a larger tooth length. The t-test seems to suggest with a lot of certainty that the null-hypothesis should be rejected. Let the next dosage interval also be considered in the following t-test.

```r
test <- t.test(ToothGrowth$len[ToothGrowth$dose == 1],
               ToothGrowth$len[ToothGrowth$dose == 2],
               paired = FALSE, var.equal = TRUE)
test$p.value
```

```
## [1] 1.810829e-05
```

```r
test$conf.int
```

```
## [1] -8.994387 -3.735613
## attr(,"conf.level")
## [1] 0.95
```

Again, a statistically significant p-value is obtained that rejects the null-hypothesis, similar to the previous t-test. Therefore, it can be concluded that an increase in dosage of the source of Vitamin C does in fact have a statistically significant increasing effect on the tooth length in guinea pigs.

Conclusion

From the initial exploratory plots, it seemed like the OJ supply method would be method would be better for smaller doses, however, t-tests seem to show that this difference is statistically insignifcant, and in reality, the supply method does not have a meaningful effect on tooth length (assuming that the variances are different in each distribution).

Again, from the plots, it seemed that the increase in dosage positively affects the tooth length in guinea pigs. This was deemed to be a statistically significant effect as confirmed by the t-tests performed on each interval (assuming that the variances of each dose distribution is the same).

In conclusion, the choice of supply method is not statistically significant, but the increase in dosage of either supply method has a positive impact on guinea pig tooth length.