Tooth Growth Data Analysis

Nicholas Nappy

September 2014

Introduction

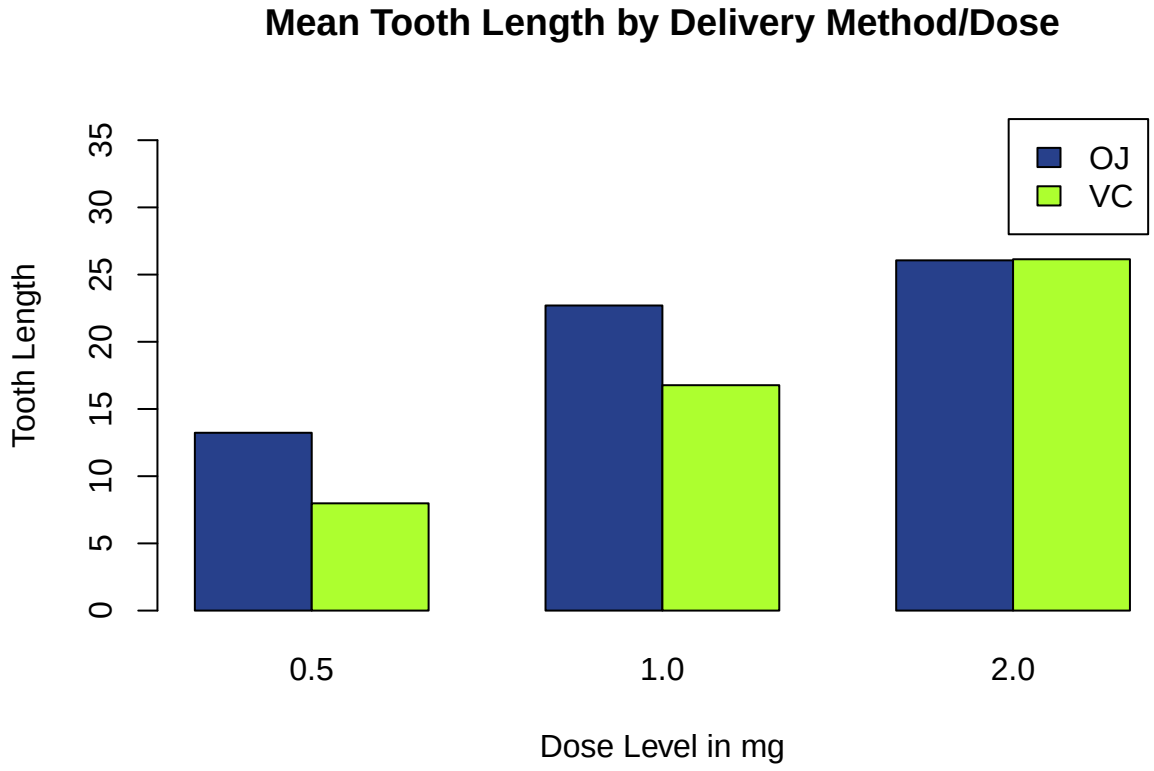
This is a basic statistical inferential data analysis completed for the Johns Hopkins University Statistical Inference course project on Coursera (https://www.coursera.org/course/statinference) in September 2014.

It utilizes the ToothGrowth data set from the datasets package. Information on the data can be found online at http://stat.ethz.ch/R-manual/R-devel/library/datasets/html/ToothGrowth.html

The R code for this analysis can be found at https://github.com/nnappy/JHU-Statistical-Inference/blob/master/ToothGrowth.R

I assume that during the observation periods during which this data was collected, all other variables were held constant. In another words, there were no changes to diet, lifestyle, etc that may have affected tooth growth or length. This assumption is required to draw any conclusions regarding supplement, dosage and tooth growth.

Please note that for the purposes of this analysis: "dose_one" = 0.5mg, "dose_two" = 1.0mg and "dose_three" = 2.0mg



Data Summary

Tooth Length by Subjects Supplementing with Orange Juice

```
##     dose_one       dose_two      dose_three
```

```
##  Min.   : 8.2   Min.   :14.5   Min.   :22.4
##  1st Qu.: 9.7   1st Qu.:20.3   1st Qu.:24.6
##  Median :12.2   Median :23.4   Median :25.9
##  Mean   :13.2   Mean   :22.7   Mean   :26.1
##  3rd Qu.:16.2   3rd Qu.:25.6   3rd Qu.:27.1
##  Max.   :21.5   Max.   :27.3   Max.   :30.9
```

Tooth Length by Subjects Supplementing with Ascorbic Acid

```
##     dose_one        dose_two      dose_three
##  Min.   : 4.20   Min.   :13.6   Min.   :18.5
##  1st Qu.: 5.95   1st Qu.:15.3   1st Qu.:23.4
##  Median : 7.15   Median :16.5   Median :25.9
##  Mean   : 7.98   Mean   :16.8   Mean   :26.1
##  3rd Qu.:10.90   3rd Qu.:17.3   3rd Qu.:28.8
##  Max.   :11.50   Max.   :22.5   Max.   :33.9
```

Statistical Tests

The data in this set measures supplementation in ten subjects with different doses and delivery methods. In order to measure differences in tooth growth between the two methods at various dosages, a paired t-test at each dosage level was performed.

```
##
##  Paired t-test
##
## data:  oj$dose_one and vc$dose_one
## t = 2.979, df = 9, p-value = 0.01547
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.263 9.237
## sample estimates:
## mean of the differences
##                    5.25
```

At a 0.5mg dose, we would reject the null hypothesis that there is no difference in tooth growth based on the type of supplementation. The mean of the differences between the two supplements is 5.25 which is well above 0.

Alternatively, the p value is .01547, which is well below a p value of .05, a standard benchmark of statistical significance. Therefore, assuming the null hypothesis is true, there would only be a .01547 probability of getting a mean value as extreme as 5.25 which would lead me to reject the null hypothesis.

```
##
##  Paired t-test
##
## data:  oj$dose_two and vc$dose_two
## t = 3.372, df = 9, p-value = 0.008229
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.952 9.908
## sample estimates:
## mean of the differences
##                    5.93
```

At a 1.0mg dose, we would reject the null hypothesis that there is no difference in tooth growth based on the type of supplementation. The mean of the differences between the two supplements is 5.93 which is well above 0.

Alternatively, the p value is .008229, which is well below a p value of .05, a standard benchmark of statistical significance. Therefore, assuming the null hypothesis is true, there would only be a .008229 probability of getting a mean value as extreme as 5.93 which would lead me to reject the null hypothesis.

```
##
##  Paired t-test
##
## data:  oj$dose_three and vc$dose_three
## t = -0.0426, df = 9, p-value = 0.967
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.329  4.169
## sample estimates:
## mean of the differences
##                   -0.08
```

At a 2.0mg dose, we would fail to reject the null hypothesis that there is no difference in tooth growth based on the type of supplementation. The mean of the differences between the two supplements is -.08 which is close to 0.

Alternatively, the p value is .967, which is above a p value of .05, a standard benchmark of statistical significance. Therefore, assuming the null hypothesis is true, there would be a .967 probability of getting a difference in mean value as extreme as -.08 which would lead me to fail to reject the null hypothesis.

Final Conclusion

There is a statistically significant improvement in tooth growth at 0.5mg and 1.0mg when supplementation is done with orange juice. At a 2.0/mg dose the type of supplementation does not appear to matter.

For questions regarding this analysis or to contact me please email nnappy@gmail.com