Investigation on ToothGrowth Dataset

YeongWei

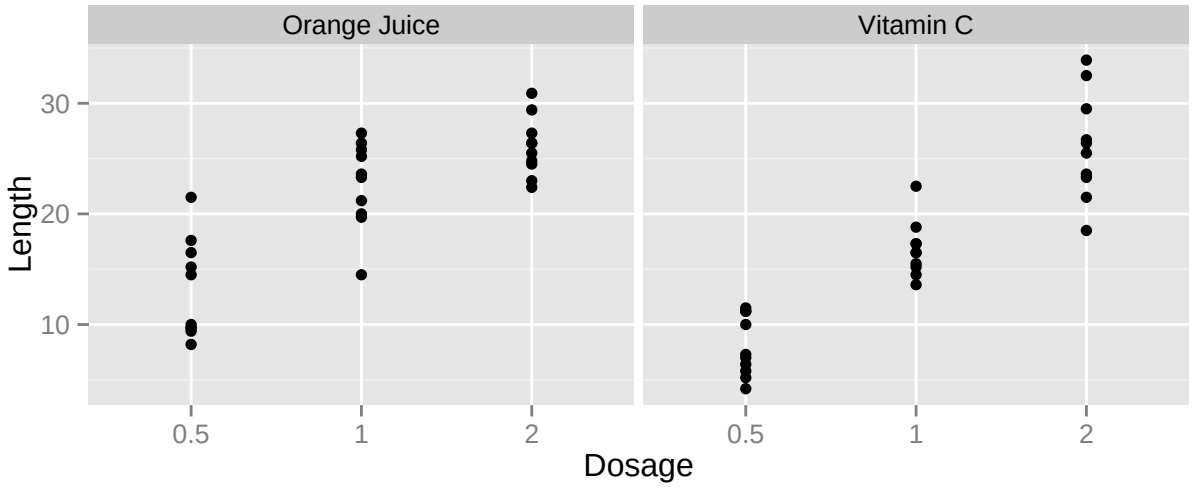
Sunday, November 22, 2015

Overview

This document aims to investigate the data associated to the ToothGrowth dataset.

Data Exploratory and Analysis

Scatter Plots



Box Plots

Observations and Discussions

1. As overview, the higher dosage of supplement, the higher the probability of getting higher length.
2. Vitamin C has the potential of getting higer length than Orange Juice.
3. Orange Juice has a tendency of logistic length growth and Vitamin C has a tendency of linear length growth.

Analysis on Tooth Growth by Supplement Type and Dosage

By Supplement Type

Equal Variance

```
stats <- t.test(ToothGrowth$Length[ToothGrowth$SupplementType == "Orange Juice"],
       ToothGrowth$Length[ToothGrowth$SupplementType == "Vitamin C"],
       paired = FALSE, var.equal = TRUE);
c(stats$p.value, stats$conf.int, stats$estimate);
```

```
##                                       mean of x   mean of y
##  0.06039337 -0.16700642  7.56700642 20.66333333 16.96333333
```

Unequal Variance

```
stats <- t.test(ToothGrowth$Length[ToothGrowth$SupplementType == "Orange Juice"],
       ToothGrowth$Length[ToothGrowth$SupplementType == "Vitamin C"],
       paired = FALSE, var.equal = FALSE);
c(stats$p.value, stats$conf.int, stats$estimate);
```

```
##                                       mean of x   mean of y
##  0.06063451 -0.17101562  7.57101562 20.66333333 16.96333333
```

By Dosage

Equal Variance

```
stats <- t.test(ToothGrowth$Length[ToothGrowth$Dosage == 0.5],
       ToothGrowth$Length[ToothGrowth$Dosage == 2], paired = FALSE, var.equal = TRUE);
c(stats$p.value, stats$conf.int, stats$estimate);
```

```
##                                           mean of x     mean of y
##  2.837553e-14 -1.815352e+01 -1.283648e+01  1.060500e+01  2.610000e+01
```

Unequal Variance

```
stats <- t.test(ToothGrowth$Length[ToothGrowth$Dosage == 0.5],
       ToothGrowth$Length[ToothGrowth$Dosage == 2], paired = FALSE, var.equal = FALSE);
c(stats$p.value, stats$conf.int, stats$estimate);
```

```
##                                                 mean of x     mean of y
##  4.397525e-14 -1.815617e+01 -1.283383e+01  1.060500e+01  2.610000e+01
```

Summary

1. All test consist of negative values for confidence interval implies a small sample size.
2. There is minimal impact from variance equality assumption.
3. Confidence Interval test by Supplement Type yields P-Value approximates to 0.06 and includes 0 in the Confidence Interval.
4. Confidence Interval test by Dosage yields P-Value approximates to 0 (multiplication with e-14) and does not include 0 in the Confidence Interval.

Conclusion

1. The Confidence Interval by Supplement Type includes 0 and P-Value large than 0.05 (standard cut-off point). Therefore the Supplement Type does not have apparent contibution to tooth length.
2. The Confidence Interval by Dosage does not include 0 and P-Value close to 0. Therefore the amount Dosage has apparent contribution to tooth length.
3. Regardless of the Supplement Type, the amount of Dosage has a better impact on tooth growth.