Tooth Growth Analysis

Michael Houck

August 23, 2015

Summary

The ToothGrowth dataset contains the length of teeth in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid). This analysis will look at those results and determine the statistical differences between the methods.

Data and Library Import

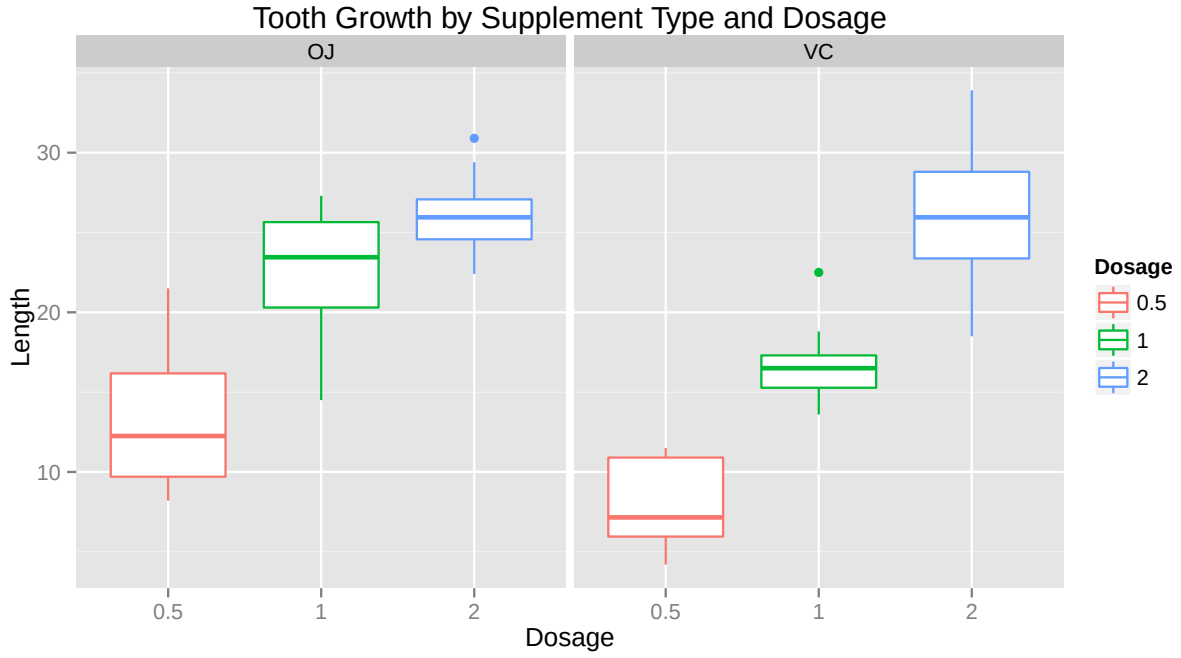
We will import the ToothGrowth dataset as well as the ggplot2 library to display some exploratory graphs.

```
library(ggplot2)
library(datasets)
data("ToothGrowth")
```

Exploratory Analysis

The ToothGrowth dataset contains the supplement type, the dosage and the resulting tooth growth amount. Here is an overall summary of the data.

```
qplot(as.factor(dose), len,
      data = ToothGrowth,
      geom="boxplot",
      facets=~ supp,
      color = as.factor(dose)) +
  labs(x = "Dosage", y = "Length", color = "Dosage",
       title = "Tooth Growth by Supplement Type and Dosage")
```



A visual examination seems to show OJ .5mg and 1mg performing better than VC, but VC and OJ are very close for the 2mg dose.

Statistical Analysis

Dose by Dose

.5 mg Dosage

```
dose.m <- .5

g1 <- ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == dose.m,"len"]
g2 <- ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose == dose.m,"len"]

t.test(g1,g2)
```

```
##
##  Welch Two Sample t-test
##
## data:  g1 and g2
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean of x mean of y
##     13.23      7.98
```

The 95% confidence interval being entirely positive shows that OJ's mean is greater than VC for dosage .5mg.

1 mg Dosage

```
dose.m <- 1

g1 <- ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == dose.m,"len"]
g2 <- ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose == dose.m,"len"]

t.test(g1,g2)
```

```
##
##  Welch Two Sample t-test
##
## data:  g1 and g2
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean of x mean of y
##     22.70     16.77
```

The 95% confidence interval being entirely positive shows that OJ's mean is greater than VC for dosage 1mg.

2 mg Dosage

```
dose.m <- 2

g1 <- ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == dose.m,"len"]
g2 <- ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose == dose.m,"len"]

t.test(g1,g2)
```

```
##
##  Welch Two Sample t-test
##
## data:  g1 and g2
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean of x mean of y
##     26.06     26.14
```

The 95% confidence interval spans 0, which shows that we cannot say that there is a significant difference in the means of OJ and VC for dosage 2mg.

Overall Supplement Comparison

We can also compare the supplements overall across all dosages.

```
g1 <- ToothGrowth[ToothGrowth$supp == "OJ","len"]
g2 <- ToothGrowth[ToothGrowth$supp == "VC","len"]

t.test(g1,g2)
```

```
##
##  Welch Two Sample t-test
##
## data:  g1 and g2
## 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 of x mean of y
##  20.66333  16.96333
```

The 95% confidence interval is not entirely above 0, so we cannot say that the mean of OJ is always above that of VC.

Assumptions

The statistical analysis did not assume paired samples as the description of the dataset did not mention if the samples were paired. We will also assume that the tooth growth amounts are normally distributed. We know the sample sizes were equal in all tests. We will also assume that the trials were performed independently from one another, so none of the samples were dependent.