ToothGrowth Data Analysis

Christopher Jones

November 21, 2017

In this report we examine ToothGrowth data, and draw some conclusions about it.

Data Loading

We load and look at the general shape of the data here:

```
# Data load code
library(datasets)
data(ToothGrowth)

str(ToothGrowth)
```

```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
head(ToothGrowth, 5)
```

```
##    len supp dose
## 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
```

Basic Summary & Data Analysis Of ToothGrowth Data

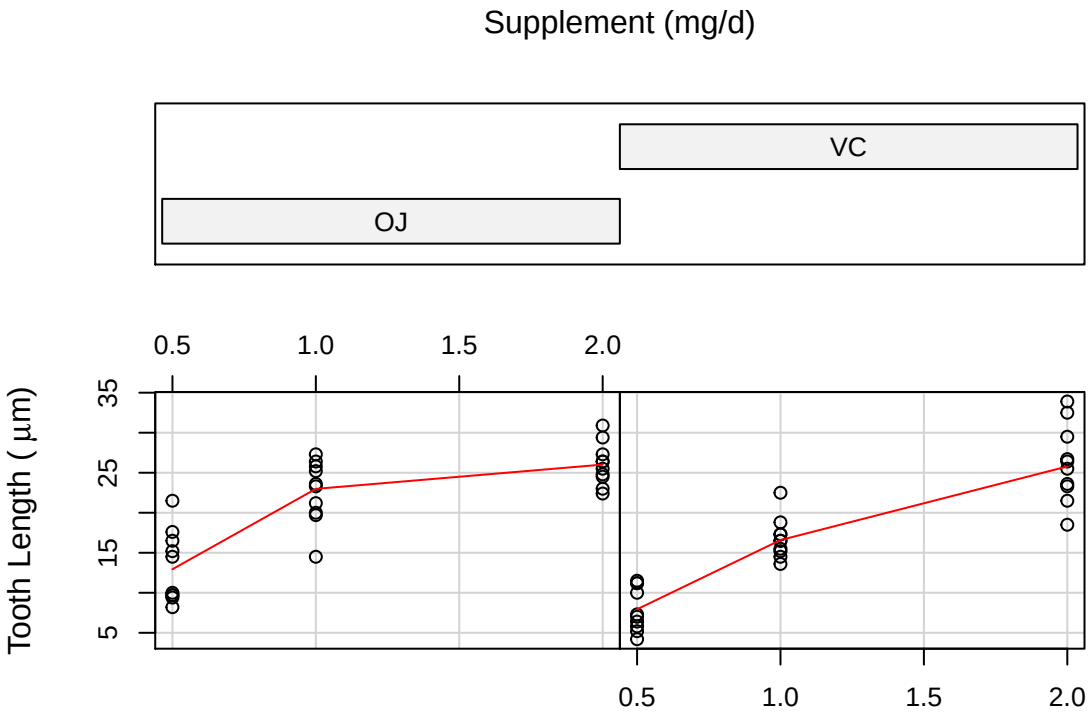
Basic summary statistics for the ToothGrowth dataset:

```
# Summary code
summary(ToothGrowth)
```

```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

And a visualization of how each supplement performs with dose/length:

```
# exploratory data analysis code
coplot(len ~ dose | supp, data = ToothGrowth, panel = panel.smooth,
       xlab = c("ToothGrowth data: length vs dose by supplement", "Supplement (mg/d)")
       , ylab = expression(paste("Tooth Length ( ", mu, "m)", sep="")))
```



ToothGrowth data: length vs dose by supplement

Tooth Growth Comparison By Supplement And Dose

We first use the t-test to compare the impact of supplement type on tooth growth:

```
# Compare supplements
t.test(len ~ supp, data = ToothGrowth)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##         20.66333         16.96333
```

Next we examine the impact of dosage on tooth growth

```
# 2 vs .5
t.test(ToothGrowth$len[ToothGrowth$dose == 2],  ToothGrowth$len[ToothGrowth$dose == 0.5])
```

```
##
##  Welch Two Sample t-test
##
```

```
## data:  ToothGrowth$len[ToothGrowth$dose == 2] and ToothGrowth$len[ToothGrowth$dose == 0.5]
## t = 11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  12.83383 18.15617
## sample estimates:
## mean of x mean of y
##    26.100    10.605
```

```
# 2 vs 1
t.test(ToothGrowth$len[ToothGrowth$dose == 2], ToothGrowth$len[ToothGrowth$dose == 1])
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$dose == 2] and ToothGrowth$len[ToothGrowth$dose == 1]
## t = 4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  3.733519 8.996481
## sample estimates:
## mean of x mean of y
##    26.100    19.735
```

```
# 1 vs dose .5
t.test(ToothGrowth$len[ToothGrowth$dose == 1], ToothGrowth$len[ToothGrowth$dose == 0.5])
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$dose == 1] and ToothGrowth$len[ToothGrowth$dose == 0.5]
## t = 6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   6.276219 11.983781
## sample estimates:
## mean of x mean of y
##    19.735    10.605
```

Conclusions

Conclusions:

1. From the supplement test, we reject the hypothesis that the mean tooth growth differs according to the supplement. Supplement type by itself does not affect tooth length.
2. From the dosage tests we reject the hypothesis that mean tooth growth is by chance, and conclude that tooth length is affected (positively) by dosage level.

Assumptions:

1. The guinea pig populations are random and independent.
2. Sampled guinea pigs are representive of guinea pigs generally.
3. The variances between the compared populations are different.