Coursera Statistical Inference Course Project:

The Effect of Vitamin C on Tooth Growth in Guinea Pigs

November 2015

Anthony L.

Overview

This report is going to investigate the effect of different vitamin C delivery methods on tooth growth in guinea pigs. The response is the length of odontoblasts (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).

Section 1 - Loading Tooth Growth Dataset

The Tooth Growth dataset can be found in the datasets package in R. GGplot2 is also loaded to perform some of the analysis.

```
library(datasets)
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

```
tooth <- ToothGrowth
```

Section 2 - Summary of Tooth Growth Dataset

The Tooth Growth dataset has 60 observations of 3 variables. The 3 variables are len for length of teeth, supp for the delivery method of Vitamin C, and dose for the dose level of Vitamin C.

```
str(tooth)
```

```
## '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 ...
```

```
summary(tooth)
```

```
##       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
```

The graph below compares the tooth length when the delivery method is oranage to the tooth length when the method is ascorbic acid. Overall, it appears that an increase in Vitamin C increases tooth length regardless of delivery method. More importantly though, it appears that orange juice is more effective at lower dose levels but about the same at the 2.0 level.

```
p <- ggplot(tooth, aes(dose, len)) + geom_point()
p + facet_grid(. ~ supp) + geom_smooth(method=lm)
```
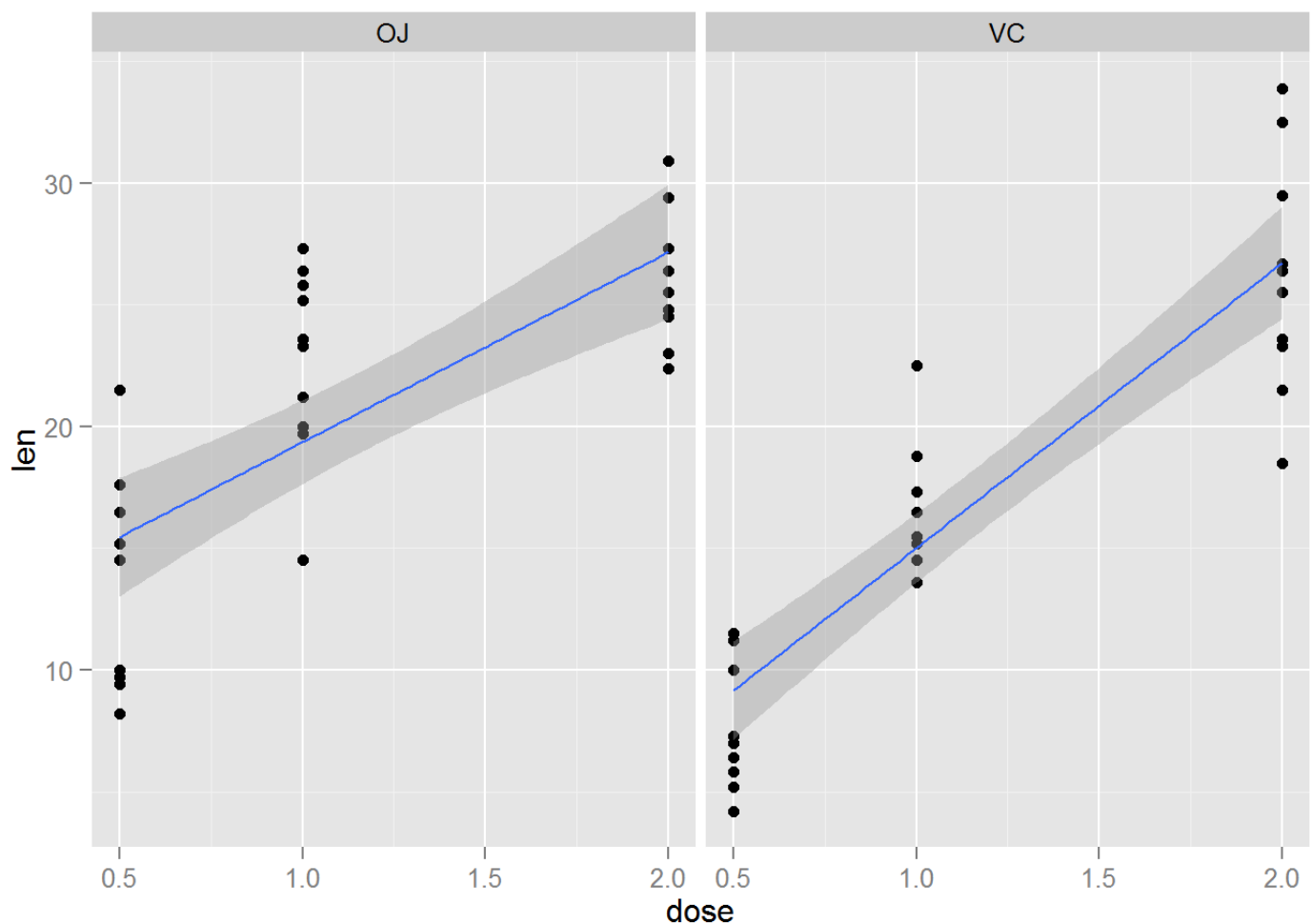

Section 3 - Comparison by Vitamin C Supplement and Dose Level

To take a look at the differences within each dose level, the data is subsetted into 3 datasets. A T-test is then performed for each of the 3 datsets to compare for differences in effectiveness between orange juice and ascorbic acid.

```
tooth0.5 <- subset(tooth, dose == 0.5)
tooth1.0 <- subset(tooth, dose == 1.0)
tooth2.0 <- subset(tooth, dose == 2.0)
```

The T-tests returned a t-value of 3.17 at the 0.5 dose level, 4.03 at the 1.0 dose level, and a -0.05 at the 2.0 level. These results indicate that there are differences at the 0.5 and 1.0 dose level but no difference at the 2.0 level.

```
t.test(tooth0.5$len ~ tooth0.5$supp, paired=F, var.equal=T)
```

```
##
##   Two Sample t-test
##
## data:  tooth0.5$len by tooth0.5$supp
## t = 3.1697, df = 18, p-value = 0.005304
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   1.770262 8.729738
## sample estimates:
## mean in group OJ mean in group VC
##             13.23             7.98
```

```
t.test(tooth1.0$len ~ tooth1.0$supp, paired=F, var.equal=T)
```

```
##
##   Two Sample t-test
##
## data:  tooth1.0$len by tooth1.0$supp
## t = 4.0328, df = 18, p-value = 0.0007807
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   2.840692 9.019308
## sample estimates:
## mean in group OJ mean in group VC
##             22.70            16.77
```

```
t.test(tooth2.0$len ~ tooth2.0$supp, paired=F, var.equal=T)
```

```
##
##   Two Sample t-test
##
## data:  tooth2.0$len by tooth2.0$supp
## t = -0.0461, df = 18, p-value = 0.9637
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.722999  3.562999
## sample estimates:
## mean in group OJ mean in group VC
##             26.06            26.14
```

Section 4 - Conclusion

The assumptions required for a two sample T-test to be valid is normal distribution in samples, the data is independently sampled, and that the variances are homogenous. The analysis of the data indicates that at the 0.5 and 1.0 mg dose level, orange juice is a more effective delivery method. However, at the 2.0 dose mg level, there is no difference between the two delivery methods.