Exploration of Tooth Growth Data

jmtaysom

05/23/2015

Executive Summary

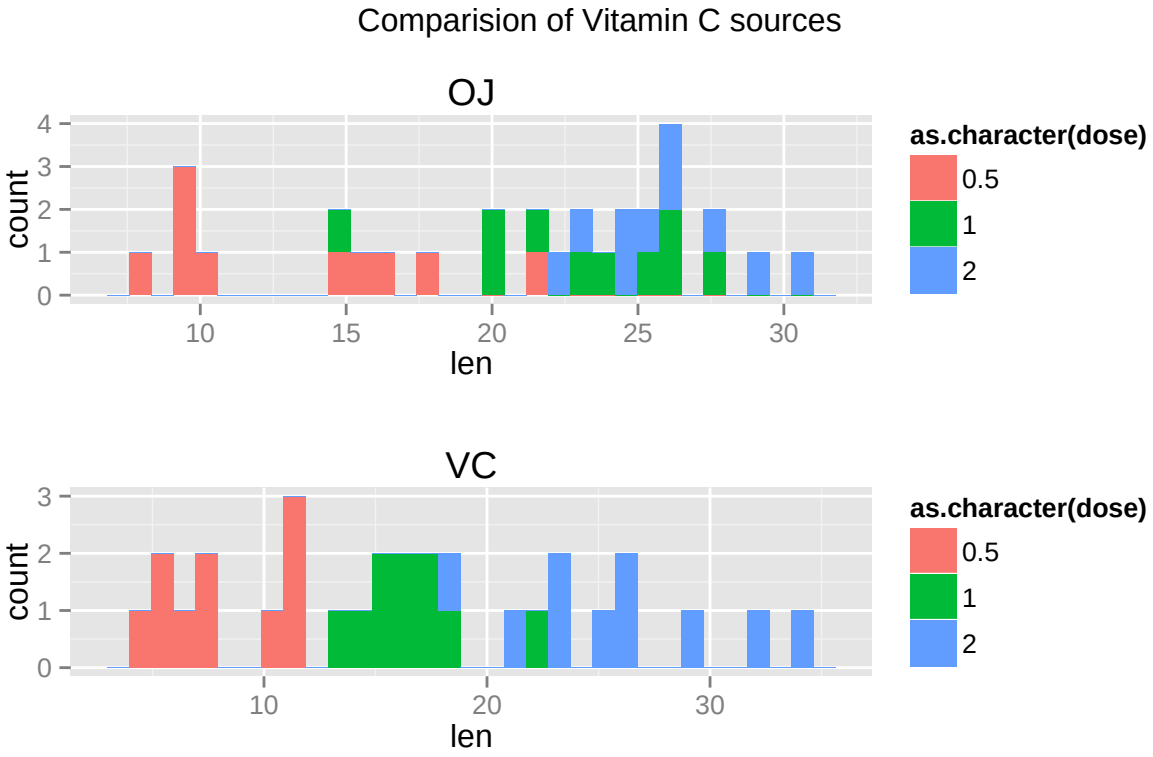
Based on the results of the Tooth Growth dataset it can be said that as the subjects recieved higher doses of vitamin c the length of their teeth increased. At the dosage levels of 0.5 and 1.0 mg orange juice induced longer tooth length than ascorbic acid. At the 2.0 dosage the difference between the two was statistically insignificant. For both orange juice and ascorbic acid as the dose increased so did the length of the tooth.

Introduction

The tooth growth dataset is a standard dataset that comes in the datasets library of R. It is the results of the "The Effect of Vitamin C on Tooth Growth in Guinea Pigs" and is descibed as "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)."

Summary of the dataset

The data shows that as the the dosage of vitamin c increases so dose the mean length of the teeth (see table 1). At the lower two dosage levels the length of teeth had more variation for orange juice than ascorbic acid but the reverse is true at the 2.0 mg dosage. Both of these are visible in the histogram below.



Interpreting the T-tests

1. When comparing orange juice to ascorbic acid at the 0.5 mg dosage we find that orange juice is correlated with longer tooth length than ascorbic acid. The test resulted in a p-value of 0.0053 and a 95% confidence interval of 1.77 - 8.73 longer teeth. This is statistically significant.
2. When comparing orange juice to ascorbic acid at the 1.0 mg dosage we find that orange juice is correlated with longer tooth length than ascorbic acid. The test resulted in a p-value of 0.0008 and a 95% confidence interval of 2.84 - 9.02 longer teeth. This is statistically significant.
3. When comparing orange juice to ascorbic acid at the 2.0 mg dosage we find that orange juice is correlated with longer tooth length than ascorbic acid. The test resulted in a p-value of 0.96 and a 95% confidence interval of -3.72 - 3.56 longer teeth. This is not statistically significant.
4. When comparing orange juice to ascorbic acid at all dosage levels we find that orange juice is correlated with longer tooth length than ascorbic acid. The test resulted in a p-value of 0.06 and a 95% confidence interval of -0.17 - 7.57 longer teeth. This is not statistically significant at the 95% level but would be at the 90% level. The statistical insignicance of the difference of the 2.0 mg dosage is masking the significance at the other levels.
5. When comparing orange juice between the 1.0 mg dosage and 0.5 mg dosage we find that the 1.0 mg is correlated with longer tooth length than 0.5 mg dosage. The test resulted in a p-value of 0.00008 and a 95% confidence interval of 5.53 - 13.41 longer teeth. This is statistically significant.
6. When comparing orange juice between the 2.0 mg dosage and 1.0 mg dosage we find that the 2.0 mg is correlated with longer tooth length than 1.0 mg dosage. The test resulted in a p-value of 0.037 and a 95% confidence interval of 6.32 - 11.26 longer teeth. This is statistically significant.
7. When comparing ascorbic acid between the 1.0 mg dosage and 0.5 mg dosage we find that the 1.0 mg is correlated with longer tooth length than 0.5 mg dosage. The test resulted in a p-value of 0.0000006 and a 95% confidence interval of 6.32 - 11.26 longer teeth. This is statistically significant.
8. When comparing ascorbic acid between the 2.0 mg dosage and 1.0 mg dosage we find that the 2.0 mg is correlated with longer tooth length than 1.0 mg dosage. The test resulted in a p-value of 0.00003 and a 95% confidence interval of 5.77 - 12.97 longer teeth. This is statistically significant.

Appendix

Required libraries

```
library(ggplot2)
library(grid)
library(gridExtra)
library(datasets)
```

Load the data

```
tg <- ToothGrowth

#Subsetting the data
OJ <- subset(tg, supp=='OJ')
VC <- subset(tg, supp=='VC')
tg.05 <-subset(tg, dose==0.5)
tg.10 <-subset(tg, dose==1.0)
tg.20 <-subset(tg, dose==2.0)
```

Summary of the data

Table 1

```
# Summaries of the means of the data by each unique group
aggregate(tg$len,by=list(tg$sup, factor(tg$dose)), mean)
```

```
##   Group.1 Group.2     x
## 1      OJ     0.5 13.23
## 2      VC     0.5  7.98
## 3      OJ       1 22.70
## 4      VC       1 16.77
## 5      OJ       2 26.06
## 6      VC       2 26.14
```

Table 2

```
# Summaries of the standard deviations of the data by each unique group
aggregate(tg$len,by=list(tg$sup, factor(tg$dose)), sd)
```

```
##   Group.1 Group.2        x
## 1      OJ     0.5 4.459709
## 2      VC     0.5 2.746634
## 3      OJ       1 3.910953
## 4      VC       1 2.515309
## 5      OJ       2 2.655058
## 6      VC       2 4.797731
```

Plot the data

```
# Plot the six groups
OJplot <- ggplot(OJ, aes(x=len, fill=as.character(dose))) + geom_histogram() + ggtitle('OJ')
VCplot <- ggplot(VC, aes(x=len, fill=as.character(dose))) + geom_histogram() + ggtitle('VC')

grid.arrange(OJplot, VCplot, nrow=2, main='Comparision of Vitamin C sources')
```

T-tests

```
# 1. Comparing Orange Juice to Ascorbic Acid at 0.5 mg dose
t.test(tg.05$len[tg.05$supp=="OJ"], tg.05$len[tg.05$supp=="VC"], var.equal=T, paired=F)
```

```
##
##  Two Sample t-test
##
## data:  tg.05$len[tg.05$supp == "OJ"] and tg.05$len[tg.05$supp == "VC"]
## 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 of x mean of y
##     13.23      7.98
```

```
# 2. Comparing Orange Juice to Ascorbic Acid at 1.0 mg dose
t.test(tg.10$len[tg.10$supp=="OJ"], tg.10$len[tg.10$supp=="VC"], var.equal=T, paired=F)
```

```
##
##  Two Sample t-test
##
## data:  tg.10$len[tg.10$supp == "OJ"] and tg.10$len[tg.10$supp == "VC"]
## 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 of x mean of y
##     22.70     16.77
```

```
# 3. Comparing Orange Juice to Ascorbic Acid at 2.0 mg dose
t.test(tg.20$len[tg.20$supp=="OJ"], tg.20$len[tg.20$supp=="VC"], var.equal=T, paired=F)
```

```
##
##  Two Sample t-test
##
## data:  tg.20$len[tg.20$supp == "OJ"] and tg.20$len[tg.20$supp == "VC"]
## t = -0.046136, 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 of x mean of y
##     26.06     26.14
```

```
# 4. Comparing Orange Juice to Ascorbic Acid for all doses
t.test(OJ$len, VC$len, var.equal=T, paired=F)
```

```
##
##  Two Sample t-test
##
## data:  OJ$len and VC$len
## t = 1.9153, df = 58, p-value = 0.06039
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1670064  7.5670064
## sample estimates:
## mean of x mean of y
##  20.66333  16.96333
```

```r
# 5. Comparing 1.0 mg dose to 0.5 mg dose for orange juice
t.test(OJ$len[OJ$dose==1.0], OJ$len[OJ$dose==0.5],var.equal=T, paired=F)
```

```
##
##  Two Sample t-test
##
## data:  OJ$len[OJ$dose == 1] and OJ$len[OJ$dose == 0.5]
## t = 5.0486, df = 18, p-value = 8.358e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   5.529186 13.410814
## sample estimates:
## mean of x mean of y
##     22.70     13.23
```

```r
# 6. Comparing 2.0 mg dose to 1.0 mg dose for orange juice
t.test(OJ$len[OJ$dose==2.0], OJ$len[OJ$dose==1.0],var.equal=T, paired=F)
```

```
##
##  Two Sample t-test
##
## data:  OJ$len[OJ$dose == 2] and OJ$len[OJ$dose == 1]
## t = 2.2478, df = 18, p-value = 0.03736
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2194983 6.5005017
## sample estimates:
## mean of x mean of y
##     26.06     22.70
```

```r
# 7. Comparing 1.0 mg dose to 0.5 mg dose for asorbic acid
t.test(VC$len[VC$dose==1.0], VC$len[VC$dose==0.5],var.equal=T, paired=F)
```

```
##
##  Two Sample t-test
##
## data:  VC$len[VC$dose == 1] and VC$len[VC$dose == 0.5]
## t = 7.4634, df = 18, p-value = 6.492e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   6.315654 11.264346
## sample estimates:
## mean of x mean of y
##     16.77      7.98
```

```r
# 8. Comparing 2.0 mg dose to 1.0 mg dose for ascorbic acid
t.test(VC$len[VC$dose==2.0], VC$len[VC$dose==1.0],var.equal=T, paired=F)
```

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

```
## data:  VC$len[VC$dose == 2] and VC$len[VC$dose == 1]
## t = 5.4698, df = 18, p-value = 3.398e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   5.77104 12.96896
## sample estimates:
## mean of x mean of y
##     26.14     16.77
```