Tooth Growth Analysis

Mark Culp

March 22, 2017

Overview:

This analysis examines the effects of vitamin C on tooth growth in guinea pigs. Each of the 60 guinea pigs studied received one of three doses of vitamin C per day: 0.5, 1.0, or 2.0 mg/day. There were two delivery methods: orange juice (OJ) and ascorbic acid (VC). The response measures the growth in odontoblasts (cells responsible for tooth growth).

```
# Load dataset
library(datasets)
data("ToothGrowth")

# No scientific notation
options(scipen = 999)
```

First we conduct an exploratory analysis of the "ToothGrowth" dataset. We need to understand the state of our data.

Exploratory analysis

```
# List length of rows / columns
dim(ToothGrowth)
```

```
## [1] 60  3
```

```
# Determine column data types
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 ...
```

```
# Range of tooth growth
range(ToothGrowth$len)
```

```
## [1]  4.2 33.9
```

```
# Levels and counts of delivery methods/supplements
table(ToothGrowth$supp)
```

```
##
## OJ VC
## 30 30
```

```
# Levels and counts of dosages
table(ToothGrowth$dose, ToothGrowth$supp)
```

```
##
##       OJ VC
##   0.5 10 10
```

```
##    1   10 10
##    2   10 10
```

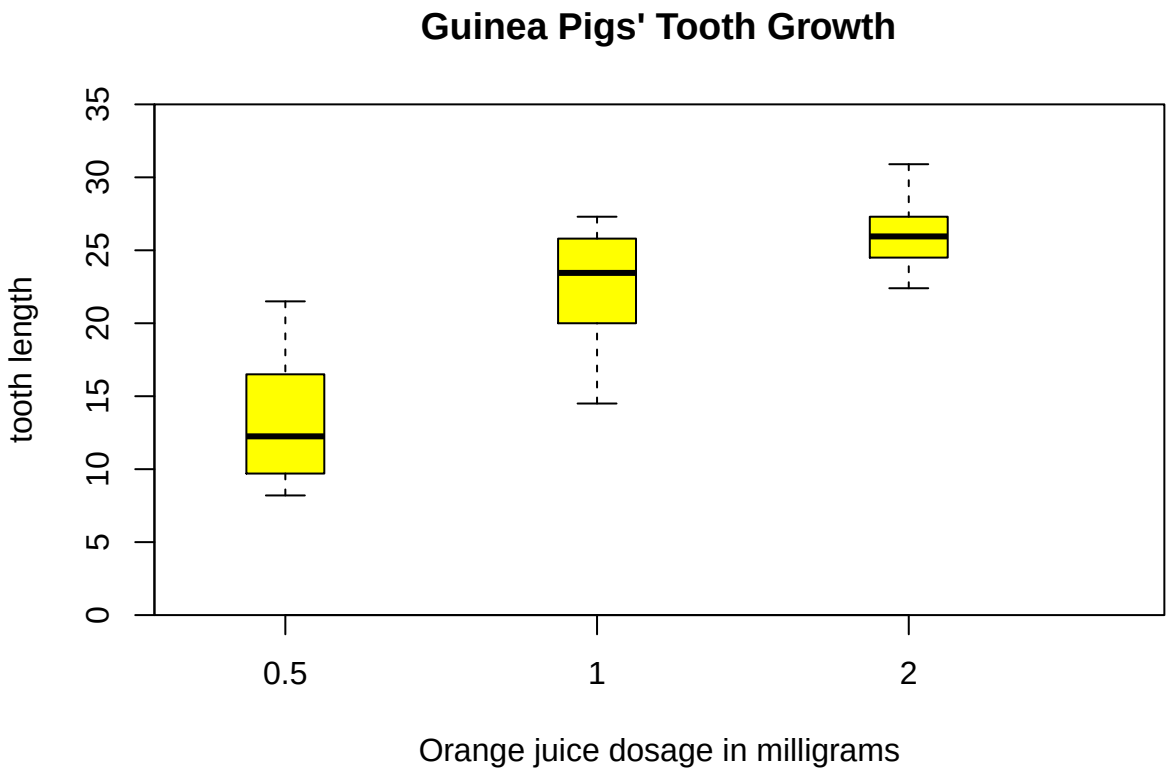
So this appears to be a very tidy data set: 30 guinea pigs were give orange juice and 30 were given ascorbic acid. 10 from each set of guinea pigs were given 0.5, 1.0, or 2.0 milligrams/day of one of the two supplements.

Data summary

So lets take a look at the characteristics of the data related to the 30 guinea pigs given orange juice.

```
# Subset the guinea pigs given orange juice,
# overall and by dosages
oj <- ToothGrowth[ToothGrowth$supp == "OJ",]
oj05 <- oj[oj$dose == 0.5,]
oj10 <- oj[oj$dose == 1.0,]
oj20 <- oj[oj$dose == 2.0,]

# Plot tooth odontoblast length in relation
# to increasing dosages of orange juice.
boxplot(len ~ dose, data = oj,
        boxwex = 0.25, at = 1:3 - 0.2,
        col = "yellow",
        main = "Guinea Pigs' Tooth Growth",
        xlab = "Orange juice dosage in milligrams",
        ylab = "tooth length",
        xlim = c(0.5, 3.5), ylim = c(0, 35), yaxs = "i")
```

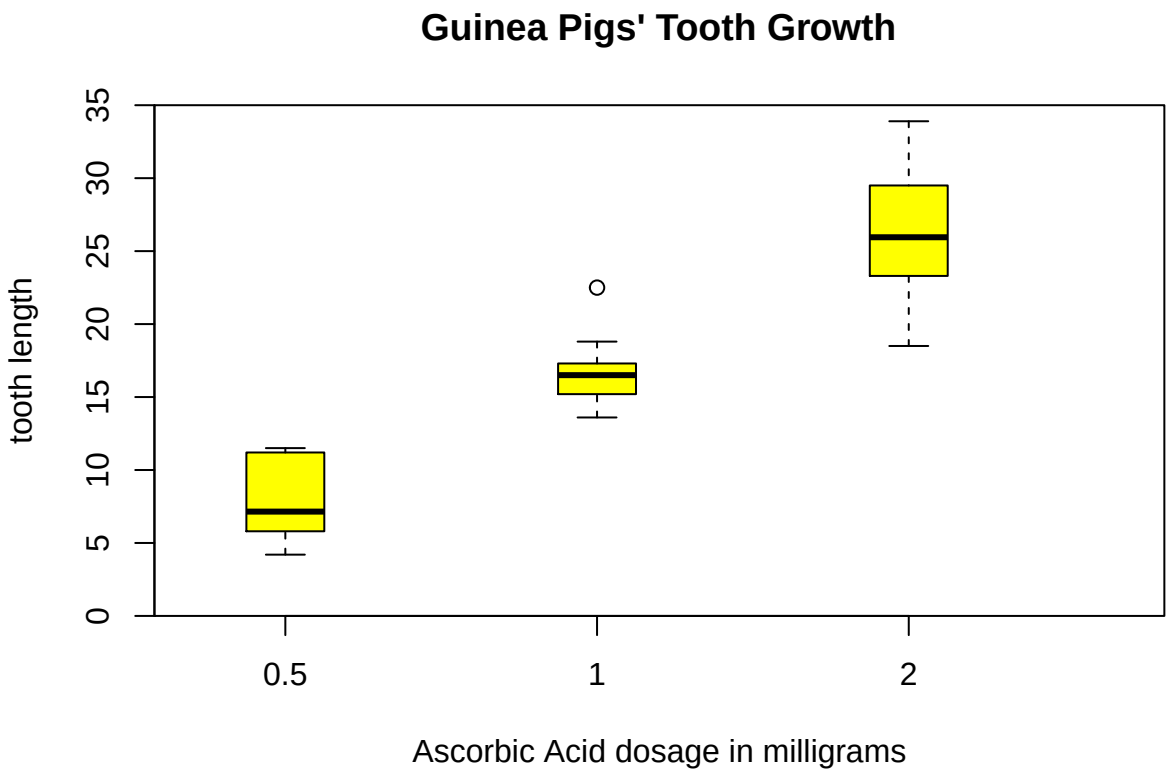


The histogram of tooth odontoblast length for the guinea pigs given orange juice appears to have a positive slope, possibly due to increasing dosages.

Now we look at the characteristics of the data related to the 30 guinea pigs given ascorbic acid.

```
# Subset the guinea pigs given ascorbic acid,
# overall and by dosages
vc <- ToothGrowth[ToothGrowth$supp == "VC",]
vc05 <- vc[vc$dose == 0.5,]
vc10 <- vc[vc$dose == 1.0,]
vc20 <- vc[vc$dose == 2.0,]

# Plot tooth odontoblast length in relation
# to increasing dosages of ascorbic acid.
boxplot(len ~ dose, data = vc,
        boxwex = 0.25, at = 1:3 - 0.2,
        col = "yellow",
        main = "Guinea Pigs' Tooth Growth",
        xlab = "Ascorbic Acid dosage in milligrams",
        ylab = "tooth length",
        xlim = c(0.5, 3.5), ylim = c(0, 35), yaxs = "i")
```



So now we compare tooth growth by supplement and dosage levels.

Comparison of tooth growth by supp and dose

```
# Load libraries
library(knitr)

# Means and standard deviation for the orange
# juice at the three dosage levels
mnsOJ <- rbind(
   c("0.5",mean(oj05$len),round(sd(oj05$len),3),round(var(oj05$len),3)),
   c("1.0",mean(oj10$len),round(sd(oj10$len),3),round(var(oj10$len),3)),
   c("2.0",mean(oj20$len),round(sd(oj20$len),3),round(var(oj20$len),3))
)

# Means and standard error for the ascorbic
# acid samples
mnsVC <- rbind(
  c("0.5",mean(vc05$len),round(sd(vc05$len),3),round(var(vc05$len),3)),
  c("1.0",mean(vc10$len),round(sd(vc10$len),3),round(var(vc10$len),3)),
  c("2.0",mean(vc20$len),round(sd(vc20$len),3),round(var(vc20$len),3))
)

# Display tables of calculated means, sd, variance
kable(mnsOJ, caption = "Orange Juice Means and Variance", col.names = c("Dossage","Mean","Std Dev","Vari
```

Table 1: Orange Juice Means and Variance

Dossage	Mean	Std Dev	Variance
0.5	13.23	4.46	19.889
1.0	22.7	3.911	15.296
2.0	26.06	2.655	7.049

```
kable(mnsVC, caption = "Ascorbic Acid Means and Variance", col.names = c("Dossage","Mean","Std Dev","Va
```

Table 2: Ascorbic Acid Means and Variance

Dossage	Mean	Std Dev	Variance
0.5	7.98	2.747	7.544
1.0	16.77	2.515	6.327
2.0	26.14	4.798	23.018

Tooth growth appears to increase with dosage levels under both supplements but the difference between the two supplements is much more significant at the 0.5 and 1.0 mg/day levels than it is at the 2.0 mg/day level. One concern here would be the increase in the variance of the ascorbic acid sample at the 2.0 level. One would expect it to decrease with an increased dossage level as the orange juice sample did.

Conclusions and assumptions

```
# So we conduct a t-test on the orange
# juice sample at the 0.5 mg/day dosage level.
```

```r
# This shows we can reject the null hypothesis
# that the orange juice had no impact on the
# guinea pigs' tooth growth.
t.test(oj05$len)
```

```
##
##  One Sample t-test
##
## data:  oj05$len
## t = 9.3811, df = 9, p-value = 0.000006074
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  10.03972 16.42028
## sample estimates:
## mean of x
##     13.23
```

```r
# So we conduct a t-test on the ascorbic
# acid sample at the 0.5 mg/day dosage level.
# This shows we can reject the null hypothesis
# that the ascorbic acid had no impact on the
# guinea pigs' tooth growth.
t.test(vc05$len)
```

```
##
##  One Sample t-test
##
## data:  vc05$len
## t = 9.1876, df = 9, p-value = 0.00000721
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  6.015176 9.944824
## sample estimates:
## mean of x
##      7.98
```

```r
# --------------------------------

# We now compare the guinea pigs' tooth growth
# under both supplements:
# ------------------------
# At the 0.5 mg/daydosage level
tg05 <- ToothGrowth[ToothGrowth$dose == 0.5,]
t.test(len ~ supp, paired = F, var.equal = F, data = tg05)
```

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

```
##            13.23             7.98
```

```
# At the 1.0 mg/daydosage level
tg10 <- ToothGrowth[ToothGrowth$dose == 1.0,]
t.test(len ~ supp, paired = F, var.equal = F, data = tg10)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##            22.70            16.77
```

```
# At the 2.0 mg/daydosage level
tg20 <- ToothGrowth[ToothGrowth$dose == 2.0,]
t.test(len ~ supp, paired = F, var.equal = F, data = tg20)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##            26.06            26.14
```

```
# The power of the orange juice samples at the
# 0.5 mg/day level are:
# power.t.test(n=10,delta = 13.23,sd= 4.459709, type="one.sample", alt = "one.sided")$power
```