Tooth Growth Analysis

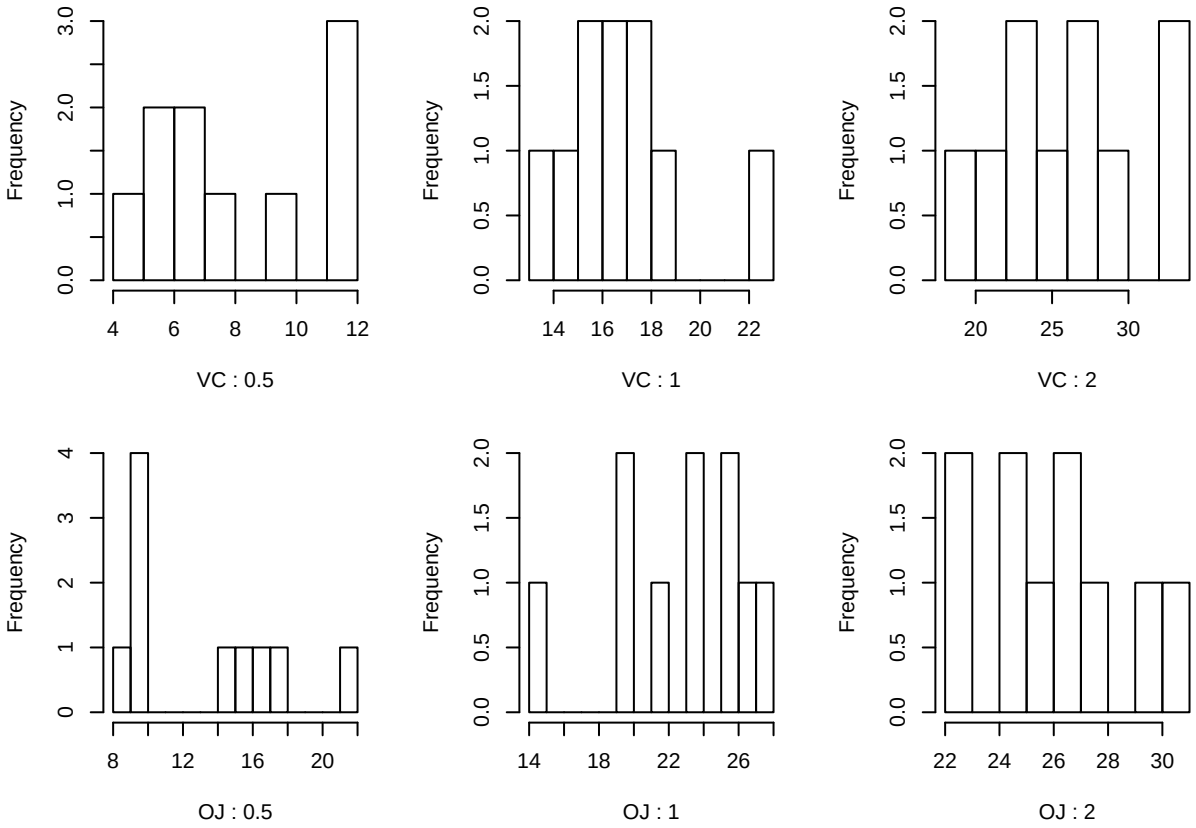
Emploration of the R ToothGrowth data set, which tracks the effect of vitamin C on tooth growth in guinea pigs. The response is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods: orange juice (**OJ**), or ascorbic acid (**VC**).

1. Basic exploratory data analysis

```
data("ToothGrowth")
head(ToothGrowth)
```

```
##    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
## 6 10.0   VC  0.5
```

We have 10 samples per dose for each of the two supplements. Below is a histogram that gives an idea of the spread in tooth length for each of the 6 supplement:dose combinations. As we can see the number of observations is small, and it is not really clear whether they have a normal distribution. **Assumption:** the data has a normal distribution.



2. Basic summary of the data.

```
tg_oj <- subset(ToothGrowth, supp == "OJ")
tg_vc <- subset(ToothGrowth, supp == "VC")
tg_vc$difference = tg_vc$len - tg_oj$len
tg_oj$difference = tg_oj$len - tg_vc$len

n <- 10
tg_oj <- summarise(group_by(tg_oj, Supplement = supp, Dose = dose),
                   MeanToothLength = mean(len), MeanDiff = mean(difference),
                   SDDiff = sd(difference), N=length(len))
tg_vc <- summarise(group_by(tg_vc, Supplement = supp, Dose = dose),
                   MeanToothLength = mean(len), MeanDiff = mean(difference),
                   SDDiff = sd(difference), N=length(len))
tg_oj
```

```
## Source: local data frame [3 x 6]
## Groups: Supplement
##
##   Supplement Dose MeanToothLength MeanDiff   SDDiff  N
## 1         OJ  0.5           13.23     5.25 5.572801 10
## 2         OJ  1.0           22.70     5.93 5.560985 10
## 3         OJ  2.0           26.06    -0.08 5.939660 10
```

```
tg_vc
```

```
## Source: local data frame [3 x 6]
## Groups: Supplement
##
##   Supplement Dose MeanToothLength MeanDiff   SDDiff  N
## 1         VC  0.5            7.98    -5.25 5.572801 10
## 2         VC  1.0           16.77    -5.93 5.560985 10
## 3         VC  2.0           26.14     0.08 5.939660 10
```

Here we see the two groups, one that received vitamin C through asorbic acid (supp = "VC"), and a group that received vitamin C in orange juice (supp = "OJ"). We also see that the average toot length increases with the dose, and finally that average growth with asorbic acid only eceeds orange juice at the 2 mg dose.

3. Comparison of average tooth growth by Supplement and Dose

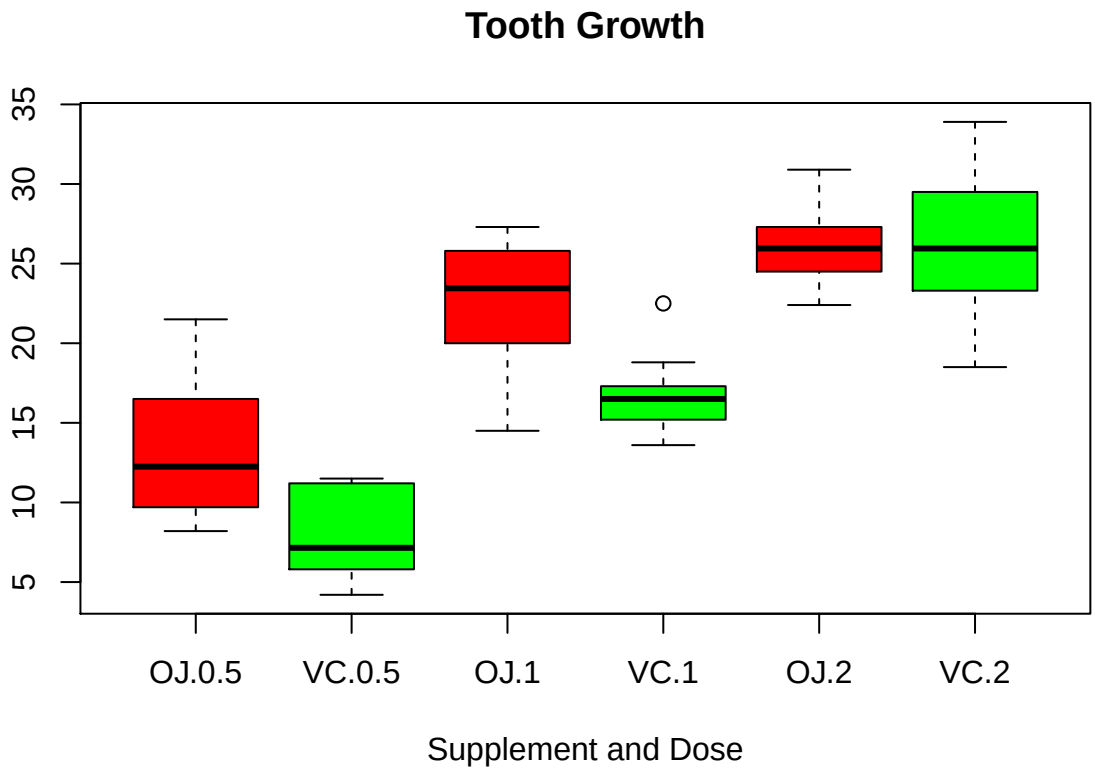
```
tga <- summarise(group_by(ToothGrowth, Supplement = supp, Dose = dose),
                 Mean = mean(len), N = length(len), SD = sd(len))

tgt <- summarise(group_by(tga, Supplement, Dose, Mean),
                  Interval = qnorm(0.95)*SD/sqrt(N))
summarise(group_by(tgt, Supplement, Dose, Mean),
          LowerConf = round(Mean - Interval, 2) ,
          UpperConf = round(Mean + Interval, 2))
```

Using mean tooth length as the comparison

```
## Source: local data frame [6 x 5]
## Groups: Supplement, Dose
##
##   Supplement Dose  Mean LowerConf UpperConf
## 1         OJ  0.5 13.23     10.91     15.55
## 2         OJ  1.0 22.70     20.67     24.73
## 3         OJ  2.0 26.06     24.68     27.44
## 4         VC  0.5  7.98      6.55      9.41
## 5         VC  1.0 16.77     15.46     18.08
## 6         VC  2.0 26.14     23.64     28.64
```

The above summay shows average tooth growth grouped by supplement and dose. It also includes the lower and upper boundaries for the 95% confidence interval of that mean. This boxplot below illustrates this data:




```
tg_oj_s <- summarise(group_by(tg_oj, Supplement, Dose, MeanDiff),
                     Interval = qnorm(0.95)*SDDiff/sqrt(N))
summarise(group_by(tg_oj_s, Supplement, Dose, MeanDiff),
          LowerConf = round(MeanDiff - Interval, 2) ,
          UpperConf = round(MeanDiff + Interval, 2))
```

Using mean growth difference as the comparison

```
## Source: local data frame [3 x 5]
## Groups: Supplement, Dose
##
##   Supplement Dose MeanDiff LowerConf UpperConf
## 1         OJ  0.5     5.25      2.35      8.15
## 2         OJ  1.0     5.93      3.04      8.82
## 3         OJ  2.0    -0.08     -3.17      3.01
```

4. Conclusions.

The basic rule of thumb is that $\frac{1}{\sqrt{n}}$ gives us a margin of error. Each supplement only have 10 data points for each dose, which gives a vey high margin of error: **31.6%**. I am basing my conclusion on the assumption that the data has a **normal distribution**. That said it seems that OJ is the most effective supplement for promoting tooth growth, and since the mean at 2 mg for OJ is included in the upper confidence level for 1.0 mg of OJ it may be that **the optimal dose is 1mg of orange** juice, **assuming** we want to maximize the effect and minimize the dose.