Analysis of ToothGrowth data

Synopsis

We're going to analyze the ToothGrowth data in the R datasets package.

Basic Summary

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).

A data frame with 60 observations on 3 variables.

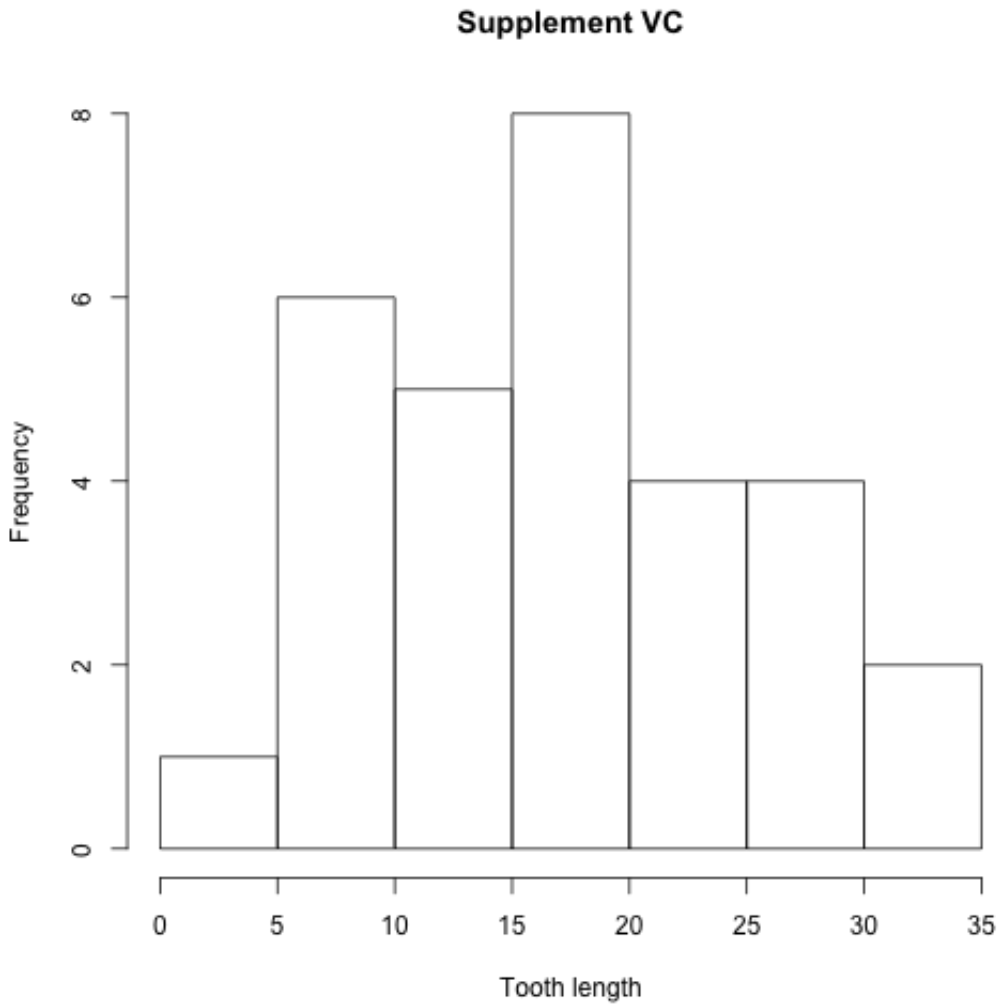
- **len** numeric Tooth length
- **supp** factor Supplement type (VC or OJ).
- **dose** numeric Dose in milligrams.

Data analysis

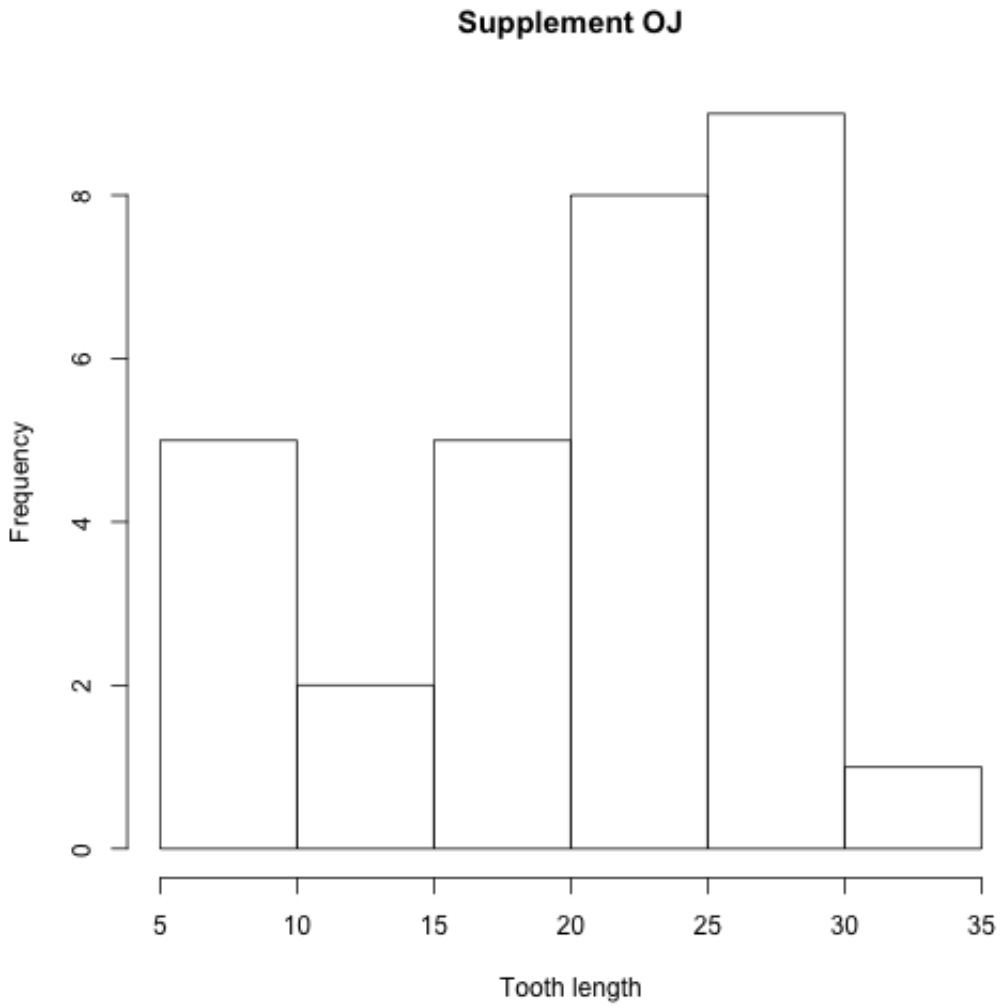
Tooth growth by supplement

There are two supplements, going by the names VC and OJ. Both of them seem to follow a normal distribution

```
ToothGrowth.vc <- ToothGrowth[ToothGrowth$supp=="VC",]
hist(ToothGrowth.vc$len,xlab="Tooth length", main=
"Supplement VC")
```



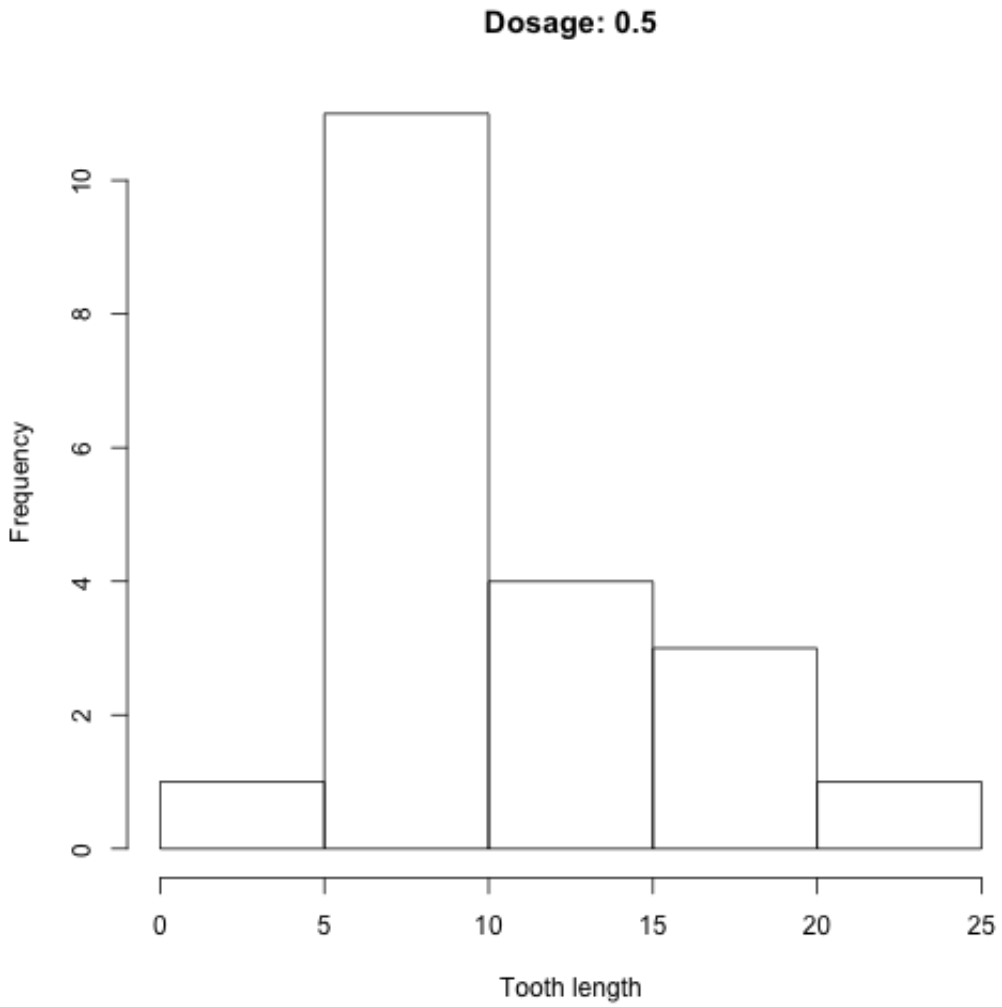

```
ToothGrowth.oj <- ToothGrowth[ToothGrowth$supp=="OJ",]
hist(ToothGrowth.oj$len, xlab="Tooth length", main=
"Supplement OJ")
```



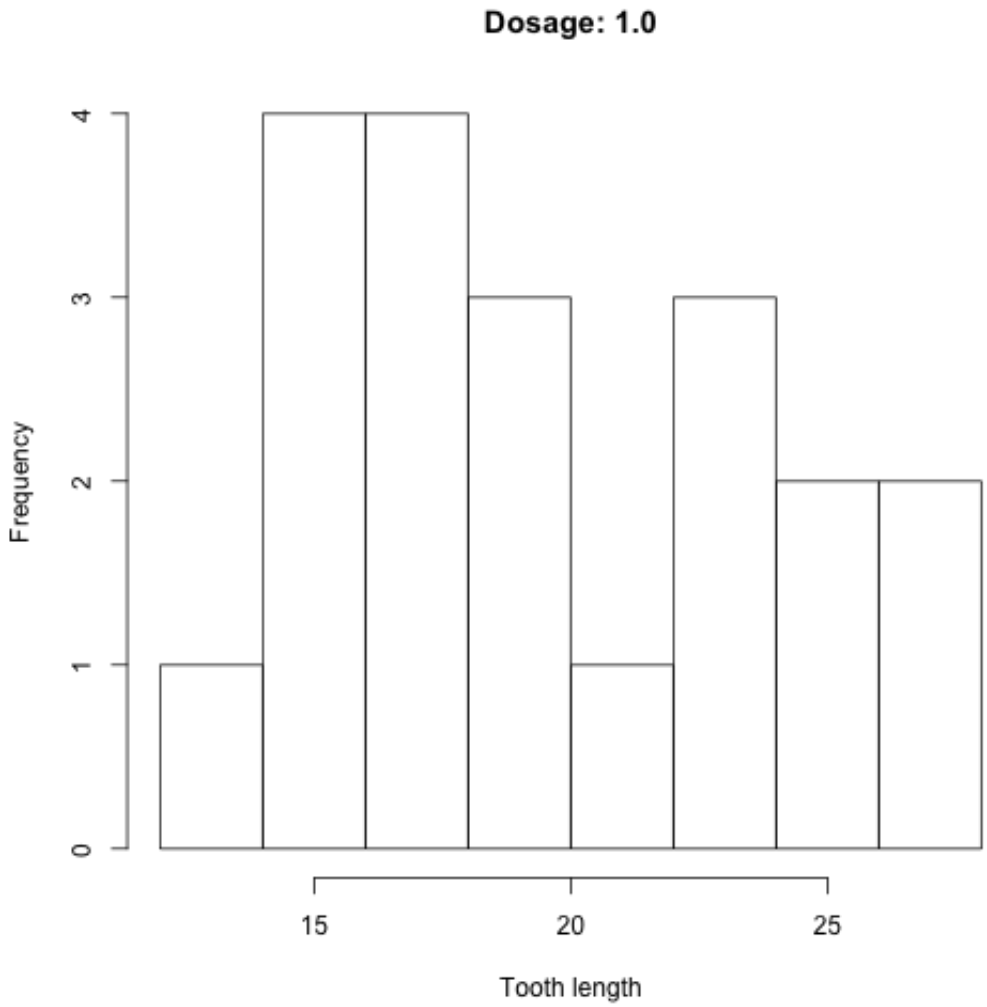
Toothgrowth by dosage

The same holds true for the dosages:

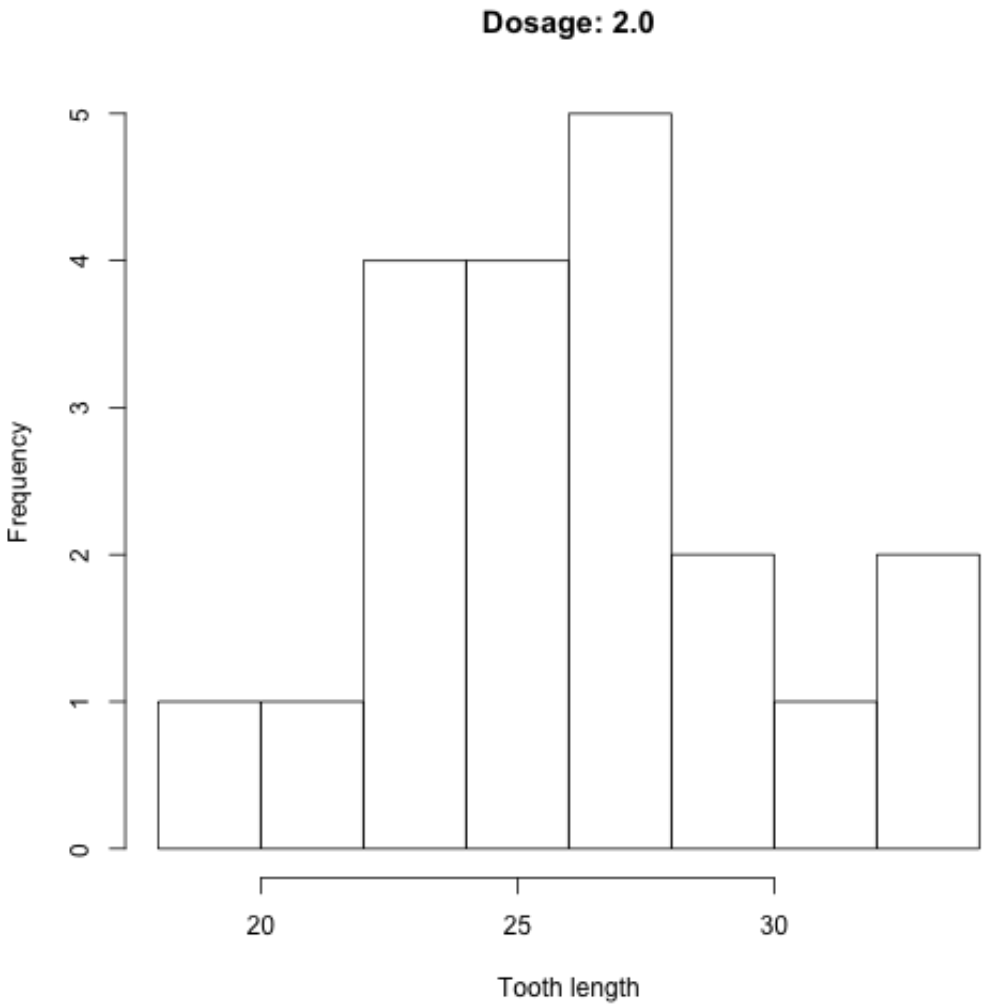
```
ToothGrowth.lowdose <- ToothGrowth[ToothGrowth$dose==0.5,]
hist(ToothGrowth.lowdose$len, xlab="Tooth length", main=
"Dosage: 0.5")
```




```
ToothGrowth.middose <- ToothGrowth[ToothGrowth$dose==1.0,]
hist(ToothGrowth.middose$len, xlab="Tooth length", main=
"Dosage: 1.0")
```




```
ToothGrowth.highdose <- ToothGrowth[ToothGrowth$dose==2.0,]
hist(ToothGrowth.highdose$len, xlab="Tooth length", main=
"Dosage: 2.0")
```



Hypothesis tests for supplement and dosage types

Supplement

Let us use the two sample T test to determine with conviction the fact that one supplement is more effective than the other.

Basically, let the null hypothesis be that both the dosage types have the same mean. Alternative hypothesise, hence will be that both the dosage types have unequal means.

We use R's Welch's T-test to determine this.

```
tTest <- t.test(ToothGrowth.oj$len, ToothGrowth.vc$len)
tTest
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth.oj$len and ToothGrowth.vc$len
## t = 1.915, df = 55.31, p-value = 0.06063
## alternative hypothesis: true difference in means is not
equal to 0
## 95 percent confidence interval:
##  -0.171  7.571
## sample estimates:
## mean of x mean of y
##     20.66     16.96
```

The p-value obtained is 0.0606 . Since it is greater than the 0.05, on a 95% confidence interval, we cannot reject the null hypothesis. It is however pretty close to that. Depending on the criticality, we can consider it so.

Dosage

Let's do this pairwise between, 0.5 and 1, and 1 and 2. If we reject the null hypothesis. We can assume transitivity, for all practical purposes.

Between 0.5mg and 1.0mg

```
tTest <- t.test(ToothGrowth.lowdose$len,
ToothGrowth.middose$len)
tTest
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth.lowdose$len and
ToothGrowth.middose$len
## t = -6.477, df = 37.99, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not
equal to 0
## 95 percent confidence interval:
##  -11.984  -6.276
## sample estimates:
## mean of x mean of y
##     10.61     19.73
```

The p-value obtained is 1.2683×10^{-7} . Since it is significantly less than the 0.05, on a 95% confidence interval, we can safely reject the null hypothesis.

This essentially means that 1.0 mg, with the tooth length mean as 19.735 has a definitively high effectiveness vis-a-vis 0.5 mg (mean: 10.605).

Between 1.0mg and 2.0mg

```
tTest <- t.test(ToothGrowth.middose$len,
ToothGrowth.highdose$len)
tTest
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth.middose$len and
ToothGrowth.highdose$len
## t = -4.901, df = 37.1, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not
equal to 0
## 95 percent confidence interval:
##  -8.996 -3.734
## sample estimates:
## mean of x mean of y
##     19.73     26.10
```

The p-value obtained is 1.9064×10^{-5} . Since it is significantly less than the 0.05, on a 95% confidence interval, we can safely reject the null hypothesis.

This essentially means that 1.0 mg, with the tooth length mean as 19.735 has a definitively lower effectiveness vis-a-vis 2.0 mg (mean: 26.1).

Conclusion

This effectively means that orange juice, with mean 20.6633 has a higher mean than ascorbic acid, but the p value test proves that it is not conclusive enough for 95% confidence interval. Meanwhile, it has been shown above that greater the dosage, greater is the effect on the tooth length.