Statistical Inference Project: Tooth Growth

Chris van Hasselt

Sunday, May 24, 2015

Overview

This report examines tooth growth data supplied in the R datasets library. As described in the help, the *ToothGrowth* dataset concerns 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)." *Note: The help file did not indicate units for the tooth length; this report will refer to the measurement as "units".*

Basic Exploratory Analysis

The data is a data.frame with 60 observations on 3 variables:

1. *len*: tooth length
2. *supp*: supplement type, VC (vitamin C) or orange juice (OJ)
3. *dose*: dose in milligrams

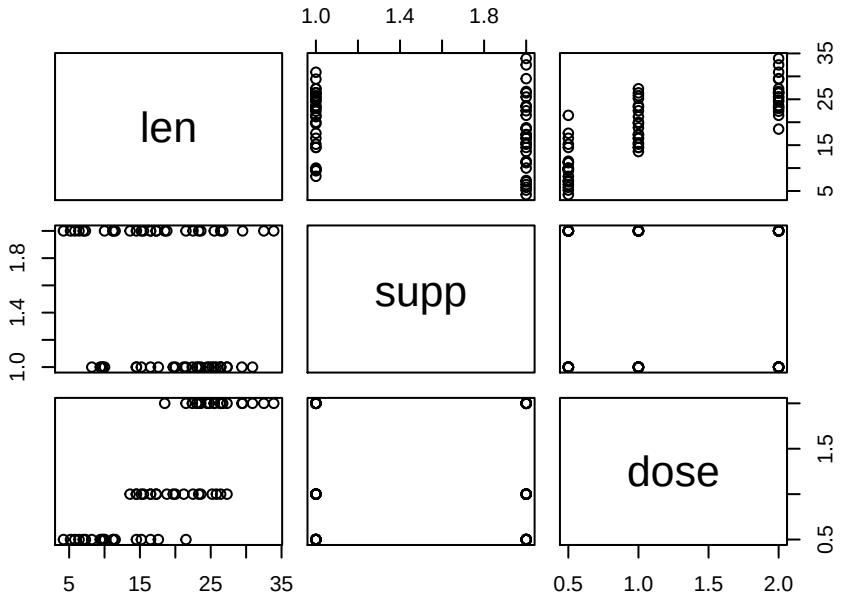
```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
```

A summary gives some indication of the overall variability of the data.

```
##       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
```

Plotting the data indicates some variations in the data to explore.

```
plot(ToothGrowth)
```



Looking at thie plots, it appears that there is an upward trend associating tooth length and the dosage of vitamin C, regardless of supplement choice. A question to consider is whether the different supplements promote a difference in growth, if one supplement promotes more growth than the other. Is the variation between growth associated with the supplement chosen, or are we simply seeing natural variation?

Tidying Data

For convenient analysis, I will group the guinea pig data into separate groups by supplement:

```
# group data by supplement factor
VCpigs <- ToothGrowth[ToothGrowth$supp=="VC",]
OJpigs <- ToothGrowth[ToothGrowth$supp=="OJ",]
```

Analysis

The mean of the two groups differs by about 4 units, without considering supplement dosage.

```
# calculate means and show the absolute difference in mean.
VCxbar <- mean(VCpigs$len)
OJxbar <- mean(OJpigs$len)

abs(VCxbar - OJxbar)
```

```
## [1] 3.7
```

The null hypothesis is that this difference is due to normal variation. Because there are a small number of observations (30) for each supplement, the t-test is the right computation to assess the validity of the

hypothesis. And because we are only interested at this time in assessing whether the difference is attributable to normal variation or the supplement, a two-sided t-test should be used. The data are not paired, and we

```
t.test(OJpigs$len - VCpigs$len, paired=FALSE,var.equal=FALSE)$conf
```

```
## [1] 1.408659 5.991341
## attr(,"conf.level")
## [1] 0.95
```

The confidence interval does not contain zero, so we can rule out the possibility that the difference in means is due to normal variation, rejecting the null hypotheis. Put another way, the difference in tooth length is attributable to the choice of supplement.

Conclusions

The analysis allows us to assess the efficacy of orange juice versus a vitamin C supplement in promoting tooth growth, with orange juice leading to greater tooth length. However, the differential between the two supplements may be dose dependent as well. Further analysis as demonstrated in this report could allow us to assess the effectiveness of each supplement at each dose.