Tooth Growth Dataset Analysis

Efstathios Fotiadis

Overview

Our objective was to provide a summary of the experimental data set and conduct hypothesis testing between the different groups. Results show that dosage plays a role on teeth growth in guinea pigs while supplement type is important only on lower doses.

Dataset

The ToothGrowth dataset has been collected from an experiment with 60 guinea pigs. The study meant to identify if there supplement type and the dose affected the legnth of the teeth of the subjects. Two different supplements types were used: Orange Juice(OJ) and Vitamin C (VC) and three different doses 0.5,1.0 and 2.0 mg. There are thus 6 groups - OJ 0.5, VC 0.5, OJ 1, VC 1, OJ 2, VC 2 - each cointaining 10 pigs. The groups are not-paired.

Basic statistics

We present the 6 different groups and some basic statistics of them.

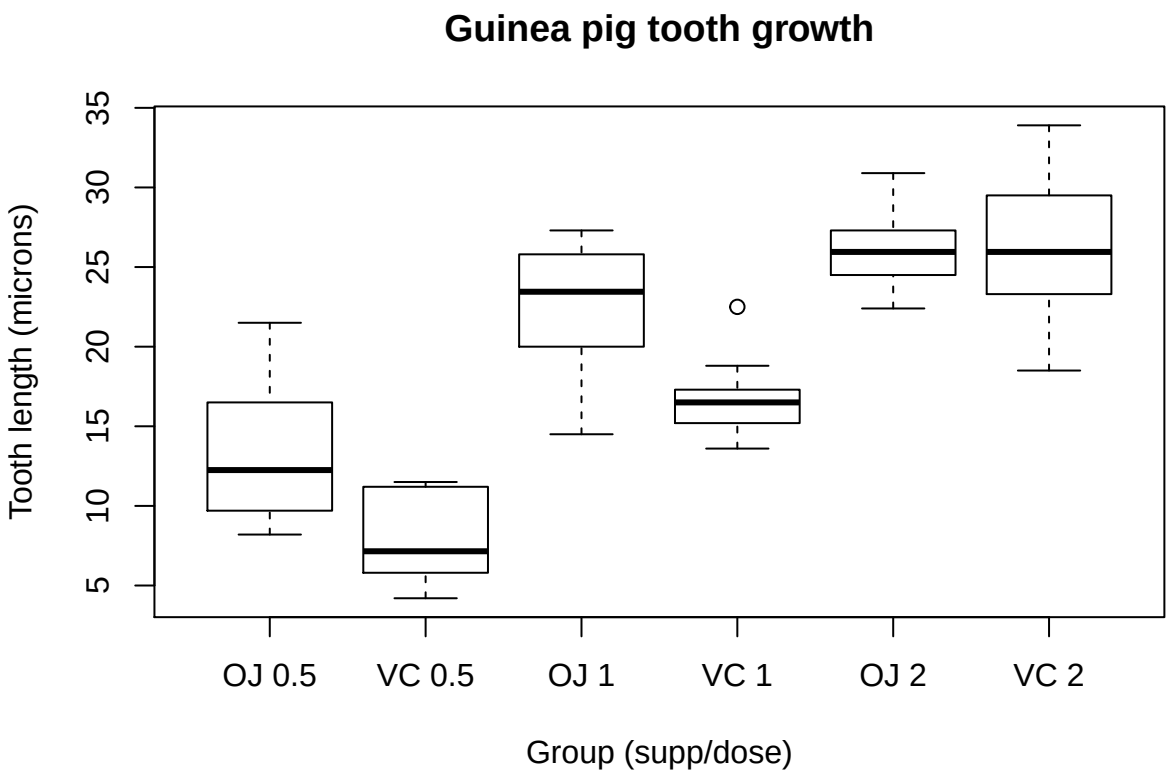
```
library(psych)
stats2 <- describeBy(len, interaction(supp, dose, sep=" "), mat = TRUE)
```

Group	Samples	Mean	Std. dev.	Median	Min	Max	Skew	Kurtosis	SE
OJ 0.5	10	13.23	4.459708	12.25	8.2	21.5	0.4381154	-1.3697072	1.4102837
VC 0.5	10	7.98	2.746634	7.15	4.2	11.5	0.1330745	-1.8068950	0.8685620
OJ 1	10	22.70	3.910953	23.45	14.5	27.3	-0.6804992	-0.6799774	1.2367520
VC 1	10	16.77	2.515309	16.50	13.6	22.5	0.9255236	0.0762364	0.7954104
OJ 2	10	26.06	2.655058	25.95	22.4	30.9	0.3685108	-1.0857354	0.8396031
VC 2	10	26.14	4.797731	25.95	18.5	33.9	0.1605264	-1.2320527	1.5171757

Visualising groups

When in comes to comparing groups, visulisations often can help. In this case we use boxplots that provide a comprehensive overview of the data.

```
boxplot(len ~ interaction(supp, dose, sep=" "),data=ToothGrowth,
        xlab="Group (supp/dose)",
        ylab="Tooth length (microns)",
        main='Guinea pig tooth growth')
```



Pair-wise t-tests

We proceed to make $\binom{6}{2} = 15$ tests between the 6 groups. We assume that each group is normally distributed and since the groups are independent we use non-paired tests. Because $n = 10$ we prefer to go with the t-test. To account for multiple testing we apply the Bonferroni correction to the p-values. The null hypothesis is that the groups come from the same distribution and we will reject the null hypothesis if the data reveal this to have a p-value < 0.05. We use a two-tailed test because we are interested in differences and we can't assume beforehand which combination is more effective.

```
library(stats)
t <-pairwise.t.test(len, interaction(supp, dose, sep=" "), paired=FALSE, p.adjust="bonferroni" )
kable(t$p.value)
```

	OJ 0.5	VC 0.5	OJ 1	VC 1	OJ 2
VC 0.5	0.0313870	NA	NA	NA	NA
OJ 1	0.0000048	0.00e+00	NA	NA	NA
VC 1	0.5047975	2.19e-05	0.0088455	NA	NA
OJ 2	0.0000000	0.00e+00	0.6502822	7.2e-06	NA
VC 2	0.0000000	0.00e+00	0.5816391	6.0e-06	1

In the table above the p-values for all the t-tests between the 6 groups can be found. For example the p-value for the t-test between the group `VC 0.5` and `OJ 0.5` is 0.0313870.

Results

From the results table we note that changing the dosage is important for both supplement types ($p<0.05$). Furthermore from the boxplot we see that increasing the dosage seems to have a positive effect on tooth

length.

Also we observe that the supplement is important ($p<0.05$) for low doses (0.05 & 1.0 mg). Judging by the box plot, Orange Juice seems to increase tooth growth more that Vitamin C. When the dose is 2.0 mg the supplement type does not affect tooth growth ($p = 1.0$).