Part B, Analysis - Tooth growth in guinea pigs

A Lopez

20 October 2015

1. Load the ToothGrowth data and perform some basic exploratory data analyses

```
#Loading required data
library("datasets")
data(ToothGrowth)
toothGrowth <- ToothGrowth
#Taking 3 samples from the top and bottom of the dataset:
head(toothGrowth,n=3); tail(toothGrowth,n=3);
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
```

```
##     len supp dose
## 58 27.3   OJ    2
## 59 29.4   OJ    2
## 60 23.0   OJ    2
```

```
#Brief analysis of the dataset
describe(toothGrowth)
```

```
## toothGrowth
##
##  3  Variables      60  Observations
## --------------------------------------------------------------------------
## len
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##      60       0      43       1   18.81    6.37    8.11   13.07   19.25
##     .75     .90     .95
##   25.27   27.30   29.57
##
## lowest :  4.2  5.2  5.8  6.4  7.0, highest: 29.4 29.5 30.9 32.5 33.9
## --------------------------------------------------------------------------
## supp
##       n missing  unique
##      60       0       2
##
## OJ (30, 50%), VC (30, 50%)
## --------------------------------------------------------------------------
## dose
##       n missing  unique    Info    Mean
```

```
##      60       0       3    0.89   1.167
##
## 0.5 (20, 33%), 1 (20, 33%), 2 (20, 33%)
## ---------------------------------------------------------------------------
```

The dataset contains a data frame with 60 observations associated to 3 variables: length, dose and method of supplying. The basic exploratory analyses shows the response 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), resulting in different length of teeth in each one (min val is 4.2 and higher 33.9). To reveal their relationship of such variables, a combination of plots as depicted in plot 1 (see apendix at the end of this)

The relationship of "supply vs length" and "dose vs length"" are plotted independently in plot 2. Further testing based on Analysis of Variance (ANOVA) is also executed.

The analysis yields the following results:

1- The 2 mg dose has larger impact on tooth growth than 1mg and 0.5mg. The 1mg dose has more impact than 0.5mg dose.

2- The relation of the amount of doses is proportional to the growth. For instance at application of larger doses the growth of teeth results bigger.

3- The Supplement type has a clear influence too, although the OJ has a greater teeth growth average in combination with dosages 0.5 and 1. The VC supplement and the OJ in combination with dosage 2, present a small difference, that is the mean and the confidence interval are approximate in value.

ANOVA yields the following results:

1 - F ratio (1,54) = 12.317; Signifcance, Pr $<$ 0.01: A notable interaction between the length and dose.

2 - F ratio, (1,54) = 133.415, Pr $<$ 0.01: A clear effect of the supplement on length.

3 - F ratio (1,54) = 5.33, Pr $>$ 0.01: There is a minor interaction between the combination of dosage + supplement compared to the length.

2. Provide a basic summary of the data.

```
# Summary of Dataset and basic stats
summary(toothGrowth);stat.desc(toothGrowth)
```

```
##       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
```

```
##                      len supp        dose
## nbr.val        60.0000000   NA 60.00000000
## nbr.null        0.0000000   NA  0.00000000
## nbr.na          0.0000000   NA  0.00000000
## min             4.2000000   NA  0.50000000
## max            33.9000000   NA  2.00000000
## range          29.7000000   NA  1.50000000
```

```
## sum          1128.8000000   NA 70.00000000
## median         19.2500000   NA  1.00000000
## mean           18.8133333   NA  1.16666667
## SE.mean         0.9875223   NA  0.08118705
## CI.mean.0.95    1.9760276   NA  0.16245491
## var            58.5120226   NA  0.39548023
## std.dev         7.6493152   NA  0.62887219
## coef.var        0.4065901   NA  0.53903330
```

This simple exploratory analysis reveals the supp "OJ" has the better growth rates overall. However, if we look at dose size "2.0" we see that "VC" reaches a larger values than "OJ". The results show there is a notable interaction between the length (len) and dosage (dose), being the highest growth of 3.9 at does=2 under the VC supplement. Refer to plot 0.

3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose.

The following t-Test compares the tooth growth by supp and dose and verifies that for certain dosage levels there is a significant difference in tooth growth between the two supplement types. For each dose level the test verifies where there is a significant difference in tooth growth between VC and OJ supplement types. Please refer to appendix for the data computing.

Analysis:

a - With a confidence interval of [1.71, 8,78] for mean(OJ)-mean(VC) at dose level 0.5 mg, the null hypothesis can be rejected as there is a significant difference in tooth length between the two supplement types at this dose level.

b - With a confidence interval of [2.80, 9.05] for mean(OJ)-mean(VC) at dose level 1.0 mg, the null hypothesis can be rejected as there is a significant difference in tooth length between the two supplement types at this dose level.

c - With a confidence interval of [-3.79, 3.63] for mean(OJ)-mean(VC) at dose level 2.0 mg, the null hypothesis can not be rejected as there is not a significant difference in tooth length between the two supplement types at this dose level.

Conclusion

The analysis of the results yield the following conclusions: For doses of 0.5 and 1.0, OJ has a greater effect on Tooth Growth than VC, We know this by the p.value indicators being less than 5% and the confidence intervals of the test do not contain 0.

For the test at dose 2.0 Ho can't be rejected, as the p.value is greater than 5% and the confidence test contains 0.

For the test OJ vs VC we cannot reject the Ho either, as the p.value is 6% (greater than the 5% threshold used for statistical analysis) and the confidence interval contains 0

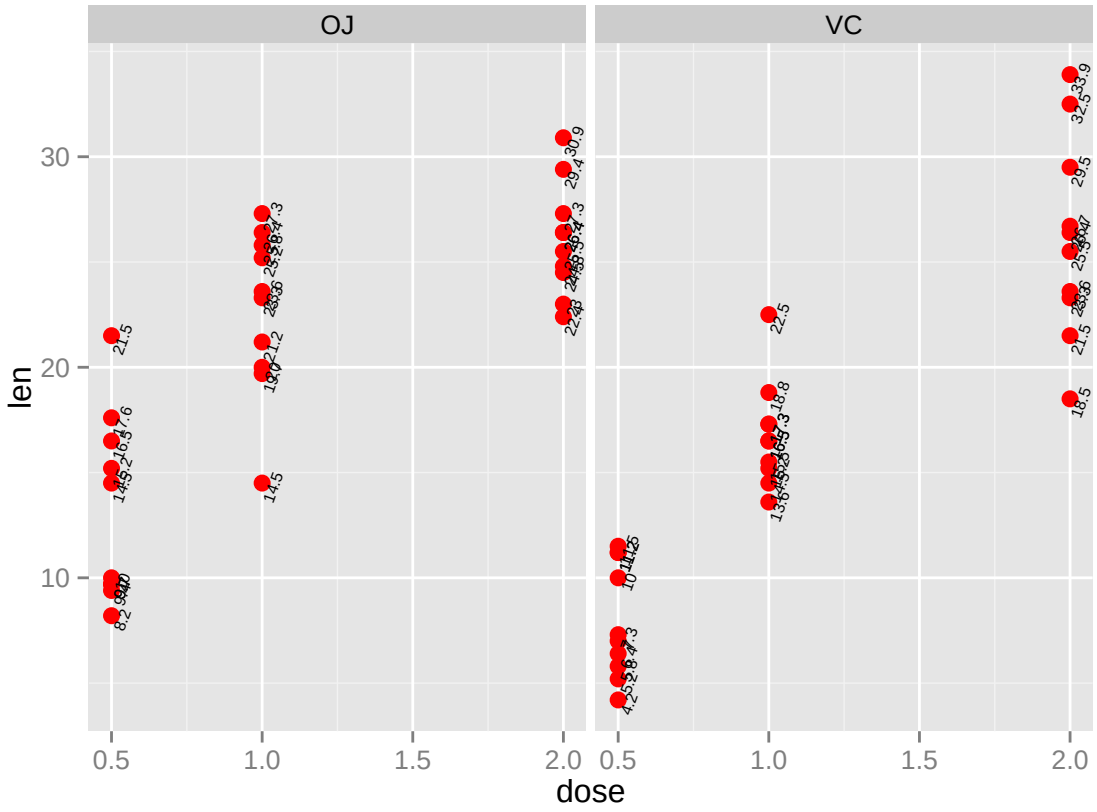
We can then conclude that to get greater tooth growth with low levels of dosage (0.5 & 1.0) the usage of OJ is recommended as opposed to VC. At greater levels of dosage (e.g 2), it's possible attain more growth using VC.

Assumptions: The populations are independent. e.g at least 60 guinea pigs are considered. We assume statistical principles are followed, e.g random populations of guinea pigs, the guinea pigs are a uniform population. The researchers were objective, took accurate measurements, and were not biased by any means.

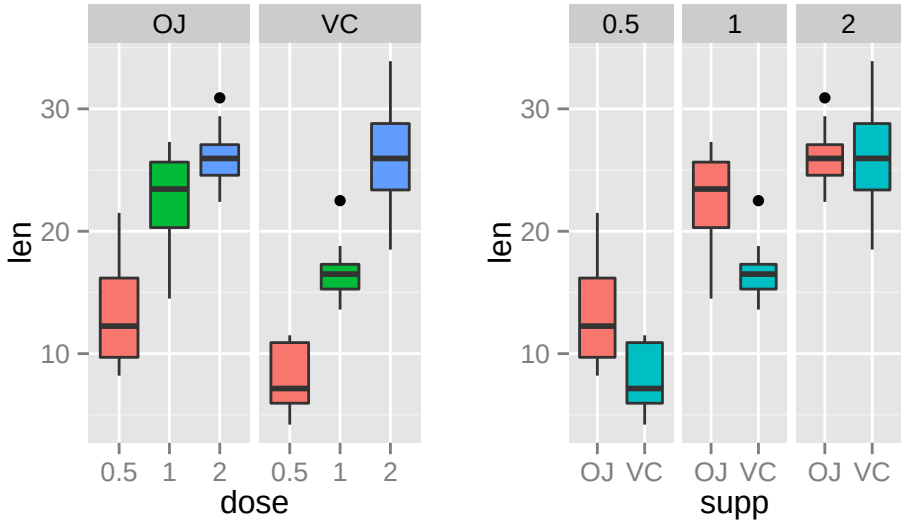
Appendix

Section Plots

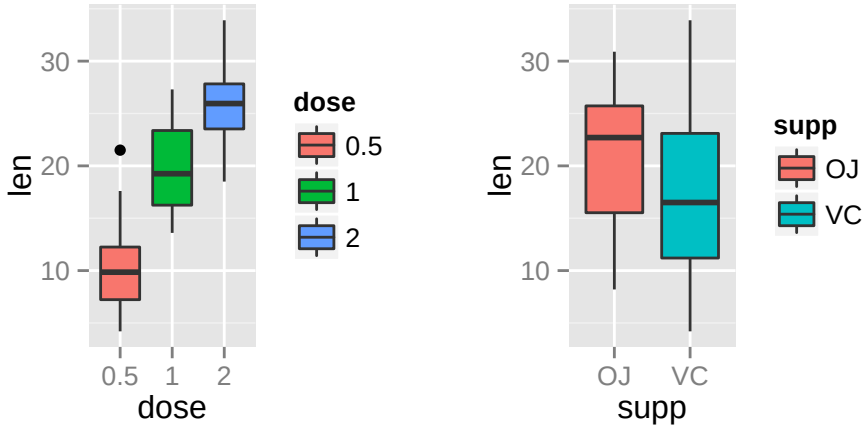
Plot 0. Scatter plot of len, dose and tooth growth



PLOT 1 . Relation of len, dose and supplement type



PLOT 2 . Relation of “len vs supplement” type and “dose vs len”"



Section Computations

1 -T test computing

```
##
##  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
```

```
##
##  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
```

```
##
##  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
```

2 -ANOVA analysis

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## supp         1  205.4   205.4  15.572 0.000231 ***
## dose         2 2426.4  1213.2  92.000  < 2e-16 ***
## supp:dose    2  108.3    54.2   4.107 0.021860 *
## Residuals   54  712.1    13.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```