Title

An Analysis Of The ToothGrowth Data

Overview

In this exercise, firstly I perform exploratory analysis on the ToothGrowth Data, showing basic charactaristics of the data. Then I use confidence intervals and hypothesis tests to compare tooth growth of Guinea pigs by supp and dose.

Summary of the Data

The ToothGrowth Data is a data frame with 60 observations on 3 variables, recording the lenght of odontoblasts (teeth) in each of 60 guinea pigs, 10 for each combination of dose level of Vitamin C (*0.5*, *1.0*, *2.0*) and delivory method (*orange juice* and *ascorbid acid*).

Here is a brief explation about the data's 3 variables: 1. len: numeric variable recording the length of odontoblasts of guinea pigs 2. supp: factor variable representing two supplement type (VC an OJ) 3. dose: numeric variable recording the volume of dose in milligrams

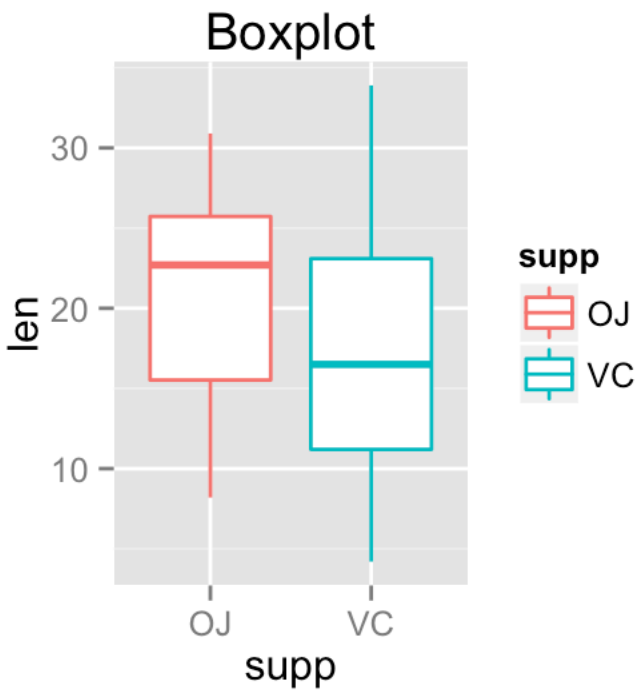
The data is part of the original study by Crampton, E.W. The academic essay pertaining the study, "The growth of the odontoblasts of the incisor tooth as a criterion of the vitamin C intake of the guinea pig.", was published in The Journal of Nutrition, vol. 33, issue 5, May 1947, pp. 491-504.

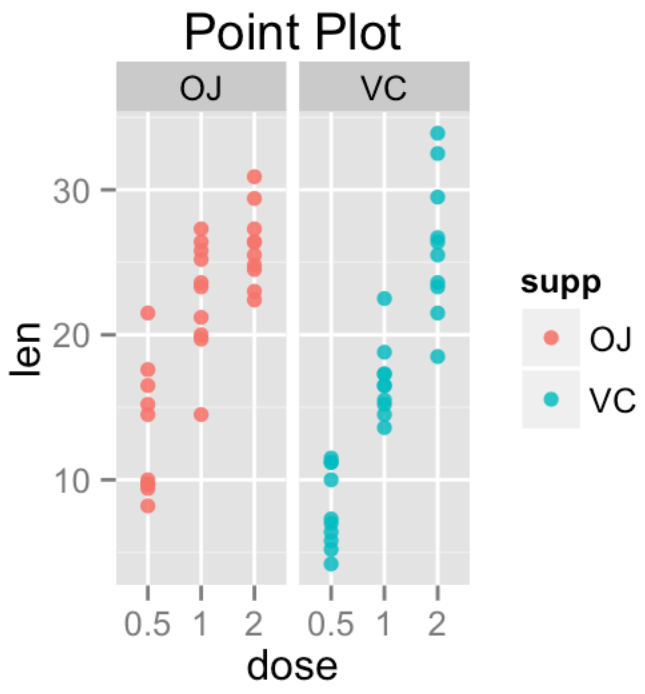
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
##       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
```

Description of the Data

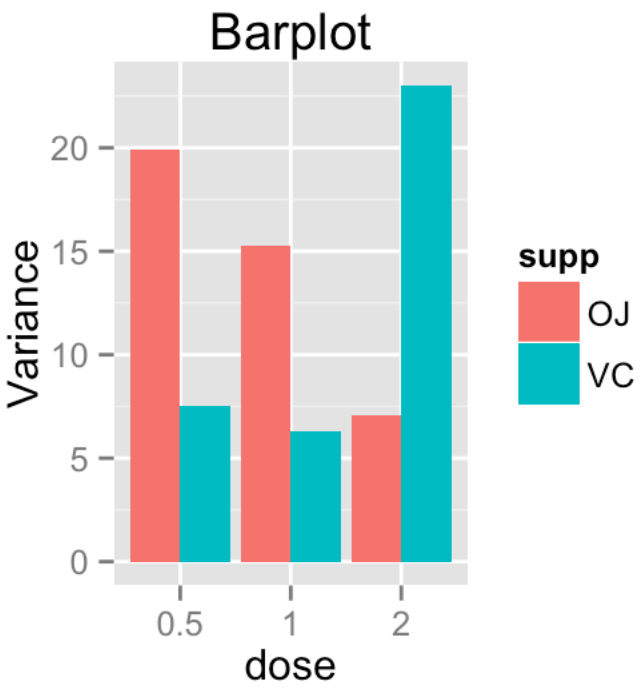
I draw boxplot, point plot, barplot and histogram to explore the data and the differences between the teeth lengths of guineas grouped in different combinations of doses of Vitamin C and delivery methods.

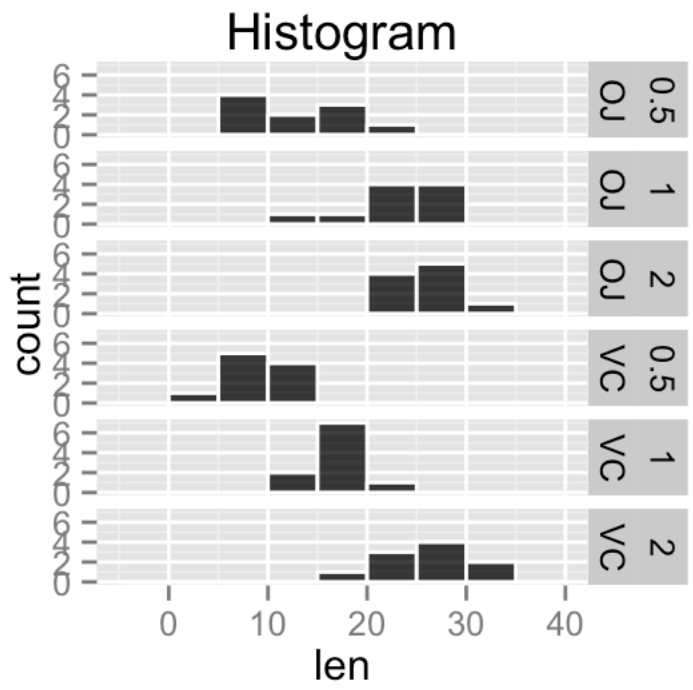




The boxplot shows the median of teeth length of guinea pigs which intook orange jucie is larger than that of those which intook ascorbid acid.

The point plot shows the distribution of teeth length of each of 60 guinea pigs. The points in the plot are grouped by different combinations of doses of Vitamin C and delivery methods.





Grouping by dose and supp variables, the 60 guinea pigs in the data are divided into 6 groups. I calculate the variance value for each of the groups, and draw a barplot to show the comparison of the variances between each pair of groups which share the same dose volume but are different in delivery methods.

From the barplot, we can see within each pair of groups, the variance of one group is significantly different from the other. Therefore we should treat the 6 groups as independent groups with unequal variances when we conduct t test on them.

The histogram shows the distribution of teeth lengths of each group. We can see in the histogram, the distribution of teech length of groups with the same delivery method moves to the right, or say, the a guinea pig grows longer at large as the dose volume increases.

Comparing Tooth Growth

To see the different growth effect a combination of dose volume and delivery method can take on guinea pigs, we compare each pair of groups which have the same delivery method but different dose volume. To be more specific, we perform comparison between (OJ, 0.5) and (VC, 0.5), (OJ, 1) and (VC, 1), (OJ, 2) and (VC, 2) by taking a t test on each of the 3 groups.

Each observation corresponds to a distinct guinea pig, and all observations are independent and nearly normally distributed. Therefore, we evaluate the mean difference via t test on each of the 3 groups under the assumption that the distribution of length of each group follows T distribution, and that each of groups are not paired and have unequal variances.

The t test on the group (OJ, 0.5) and the group (VC, 0.5):

Confidence Interval of the t test

```
## [1] 1.719057 8.780943
## attr(,"conf.level")
## [1] 0.95
```

The p-value of the t test

```
## [1] 0.006358607
```

According the test result above, for 95 percent of times the difference in means of the two groups falls within the confidence interval [1.72, 8.78]. The p-value of the test is around 0.006, far less than 0.05, which means we can reject the null hypothesis that the difference in means of the two groups is equal to 0.

The t test on the group (OJ, 1) and the group (VC, 1):

Confidence Interval of the t test

```
## [1] 2.802148 9.057852
## attr(,"conf.level")
## [1] 0.95
```

The p-value of the t test

```
## [1] 0.001038376
```

According the test result above, for 95 percent of times the difference in means of the two groups falls within the confidence interval [2.8, 9.06]. The p-value of the test is around 0.001, far less than 0.05, which means we can reject the null hypothesis that the difference in means of the two groups is equal to 0.

The t test on the group (OJ, 2) and the group (VC, 2):

Confidence Interval of the t test

```
## [1] -3.79807  3.63807
## attr(,"conf.level")
## [1] 0.95
```

The p-value of the t test

```
## [1] 0.9638516
```

According the test result above, for 95 percent of times the difference in means of the two groups falls within the confidence interval [-3.8, 3.64]. Because the interval contains 0, we can not confirm whether the mean difference is significant. Furthermore, the p-value of the test is around 0.964, far more than 0.05, which means we fail to reject the null hypothesis that the difference in means of the two groups is equal to 0.

Conclusion of the Comparison

Under the assumption that observations in the ToothGrowth Data follows the T distribution, and that each of the 6 groups can not be paired with one another and has a distinct value of variance, we conducted a t test on each pair of groups that intook the same dose volume of Vitamin C with different delivery methods (OJ or VC). In the test on the group (OJ, 0.5) and the group (VC, 0.5), we rejected the null hypothesis that the difference in means between the two groups is equal to 0. Similarly, we also reject the null hypothesis in the test on the group (OJ, 1) and the group (VC, 1). However, we fail to reject the null hypothesis in the test on the group (OJ, 2) and the group (VC, 2).

Based on the test results, we conclude that teeth grow longer with statistical significance when guinea pigs intake 0.5 or 1.0 miligrams of Vitamin C via orange juice than they do when their owners intake the same volume of Vitamin C via ascorbid acid. When guinea pigs intake 2.0 miligrams of Vitamin C, however, the delivery methods have no statistical significant effect on the teeth growth of guinea pigs. We can infer from this fact that 2.0 miligrams of Vitamin C is a volume above a guinea pig's basic demand for Vitamin C on average, which can explain effects the delivery methods take on teeth growths fades as the intake of Vitamin C increases.