Statistical Inference Course Project

Edilmo Palencia

August 22, 2015

Tooth Growth Analysis

The data present the next structure:

```
## '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 ...
```

We have 60 observations with tree variables: length is numeric, supplement is a factor and dose is numeric as well.

A quick summary of the data:

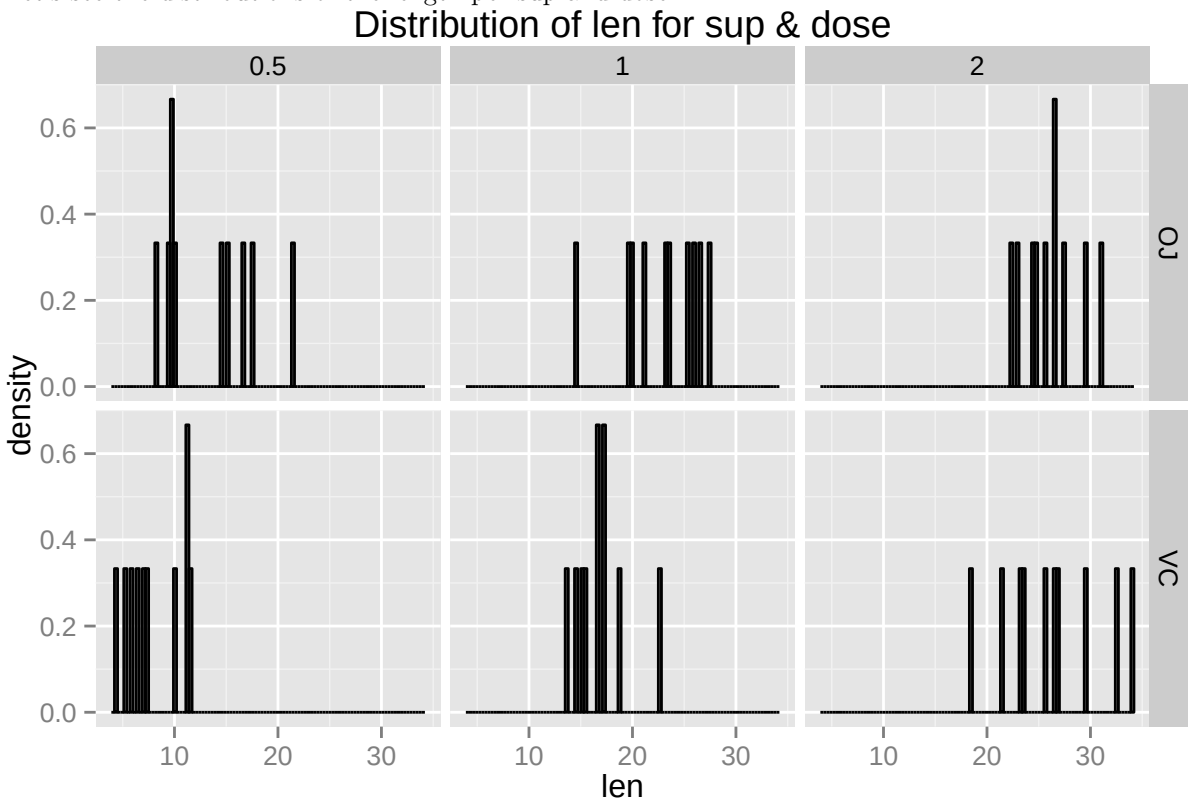
len	supp	dose
Min. : 4.20	OJ:30	Min. :0.500
1st Qu.:13.07	VC:30	1st Qu.:0.500
Median :19.25	NA	Median :1.000
Mean :18.81	NA	Mean :1.167
3rd Qu.:25.27	NA	3rd Qu.:2.000
Max. :33.90	NA	Max. :2.000

In the table we can see:
- There are 30 observation for each dosis.
- The minimum, maximun, mean, median, first and third quantile for the length and the doses.

Let's see how many observations we have per sup and types of doses:

```
##      dose 0.5  1  2
## supp
## OJ         10 10 10
## VC         10 10 10
```

Let's see the distributions of the length per sup and dose:



Let's compute the t-test over len for each group of sup and dose:

```
##          OJ.0.5       VC.0.5       OJ.1         VC.1         OJ.2
## conf.int 10.03972     6.015176     19.90227     14.97066     24.16069
## conf.int 16.42028     9.944824     25.49773     18.56934     27.95931
## estimate 13.23        7.98         22.7         16.77        26.06
## p.value  6.074068e-06 7.209903e-06 1.933448e-08 5.69872e-09 1.833351e-10
##          VC.2
## conf.int 22.70791
## conf.int 29.57209
## estimate 26.14
## p.value  3.367967e-08
```

In the above table we can se for each group of sup and dose: the sample mean, the t confidence interval those means, and the p values.

Asumptions

From the data presented above we precibe that the sup of value OJ is more efective than VC. To support this hipotesis let's compare the means.

```
##
##  Paired t-test
##
## data:  OJ.means and VC.means
## t = 1.9472, df = 2, p-value = 0.1909
```

```
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.475757 11.875757
## sample estimates:
## mean of the differences
##                     3.7
```

In the results presented we can see that the 95% confidence interval includes 0, with a p-value of .19 that is greater than .05

Conclusions

There are few samples to confirm the assumption stated above. This conclussion is derived from the fact that the confidence interval includes the 0 and it is quite wide.