Study of The Effect of Vitamin C on Tooth Growth in Guinea Pigs

Luis A Sandino

Created on 2016-03-20 00:41:07 using R version 3.2.3 (2015-12-10)

Synopsis

This project performs a basic inferential analysis of the ToothGrowth data in the R datasets package. According to the help file: *The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, (orange juice or ascorbic acid (a form of vitamin C and coded as VC).*

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

Load the dataset and convert variable dose to factor:

```
#
library(datasets)
data(ToothGrowth)
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
```

Inspect the structure of the data:

```
str(ToothGrowth)
```

```
## '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: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 1 ...
```

2. Provide a basic summary of the data

Obtain a summary of the data:

```
data <- ToothGrowth %>% group_by(supp, dose) %>% summarise(q25th = quantile(len,
    0.25), Median = quantile(len, 0.5), q75th = quantile(len, 0.75), Mean = mean(len),
    SD = sd(len))
data
```

1

```
## Source: local data frame [6 x 7]
## Groups: supp [?]
##
##     supp   dose  q25th Median  q75th  Mean        SD
##   (fctr) (fctr)  (dbl)  (dbl)  (dbl) (dbl)     (dbl)
## 1     OJ    0.5  9.700  12.25 16.175 13.23 4.459709
## 2     OJ      1 20.300  23.45 25.650 22.70 3.910953
## 3     OJ      2 24.575  25.95 27.075 26.06 2.655058
## 4     VC    0.5  5.950   7.15 10.900  7.98 2.746634
## 5     VC      1 15.275  16.50 17.300 16.77 2.515309
## 6     VC      2 23.375  25.95 28.800 26.14 4.797731
```

Visualize the data with box plots:

```
ggplot(data=ToothGrowth, aes(x=dose, y=len, fill=supp)) +
        geom_boxplot() +
        facet_grid(. ~ supp) +
        xlab("Dose [mg/day]") +
        ylab("Cell length") +
        guides(fill=guide_legend(title="Supplement type"))
```
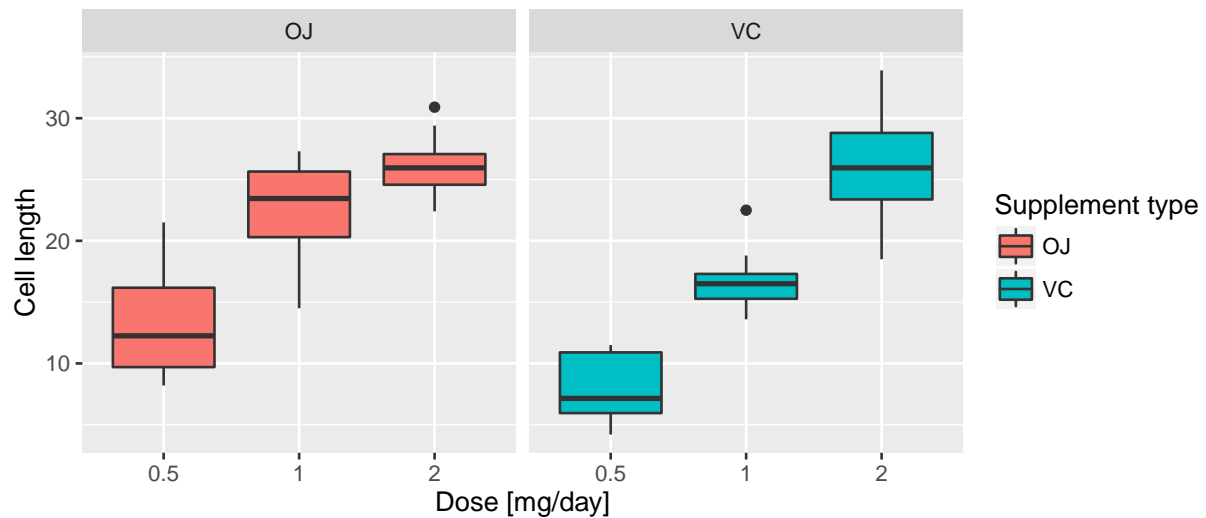

The figure shows that the odontoblasts lenght presents a positive correlation regarding the dosage for both supplement types (VC and OJ)

3. Use hypothesis tests to compare tooth growth by supplement type and dose

Table 1: Analysis of dose pairs for supplement type OJ

	P-Value	Conf. Int. low	Conf. Int. high
.5 vs 1	0.0000878	-13.42	-5.52
.5 vs 2	0.0000013	-16.34	-9.32
1 vs 2	0.0391951	-6.53	-0.19

As can be seen, all p-values are less than 0.05, therefore in all cases the null hypothesis is rejected. There is strong evidence in favour of the alternative hypothesis: *true difference in means is not equal to 0*

Table 2: Analysis of dose pairs for supplement type VC

	P-Value	Conf. Int. low	Conf. Int. high
.5 vs 1	0.0000007	-11.27	-6.31
.5 vs 2	0.0000000	-21.90	-14.42
1 vs 2	0.0000916	-13.05	-5.69

As can be seen, all p-values are less than 0.05, therefore in all cases the null hypothesis is rejected. There is strong evidence in favour of the alternative hypothesis: *true difference in means is not equal to 0*

Table 3: Analysis of supplement types OJ vs VC

	P-Value	Conf. Int. low	Conf. Int. high
.5	0.0063586	1.72	8.78
1	0.0010384	2.80	9.06
2	0.9638516	-3.80	3.64

As can be seen, p-values are less than 0.05 for doses .5 and 1. Therefore, in this two cases the null hypothesis is rejected. There is strong evidence in favour of the alternative hypothesis: *true difference in means is not equal to 0.* On the contrary, p-value is grater than 0.05 for dose 2. Thus, we fail to reject the null hypotesis.

4. State conclusions and the assumptions needed

Conclusions of the study:

1. There is strong evidence that increasing dose levels of Vitamin C leads to increased length of odontoblasts and therefore increased tooth growth in Guinea Pigs.
2. For doses of .5 and 1 mg/day, there is evidence that the administration via Oranje Juice is in favor of tooth growth
3. For doses of 2 mg/day, no difference can be satated between the two supplement types with the given data

Main assumptions:

1. The sample size of 60 Guinea Pigs is representative of the entire population of Guinea Pigs.
2. For the t-tests carried on, the variances are assumed to be different for the groups being compared.
3. For the t-tests carried on, the data is assumed to be unpaired.
4. The confidence level used is 95%

The code used for t test is:

```
t.test(x ~ y, data = data_subset,
       var.equal = FALSE, paired = FALSE,
       mu = 0,alternative = "two.sided",conf.level = 0.95)
```

where **x = len** and **y = dose** for *table 1* and *2*; and **y = supp** for *table 3*