

Inferences on the Effect of Vitamin C on Tooth Growth

emmanuelq2

28 février, 2016

```
Sys.setlocale("LC_TIME", "English")
```

This analysis relates to “[The Effect of Vitamin C on Tooth Growth in Guinea Pigs](#)”. 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 - coded as OJ - or ascorbic acid (a form of vitamin C and coded as VC)). To compare differences of two means, we must **check three conditions**: - first, the guinea pigs were independent from each other - secondly, the sheep in each group (by level of dosage or by type of delivery method) were also independent of each other - randomized treatment should have been performed

1. LOADING DATA AND BASIC EXPLORATORY DATA ANALYSIS

```
data("ToothGrowth")
# head(ToothGrowth)
# str(ToothGrowth)
knitr::kable(head(ToothGrowth))
```

len	supp	dose
4.2	VC	0.5
11.5	VC	0.5
7.3	VC	0.5
5.8	VC	0.5
6.4	VC	0.5
10.0	VC	0.5

Since the dataset includes only three dose levels of Vitamin C (0.5, 1, and 2 mg), it seems relevant for the purposes of this analysis to convert the dose variable into a factor.

```
ToothGrowth$dose <- factor(ToothGrowth$dose, labels=c("0.5mg", "1mg", "2mg"))
```

2. BASIC SUMMARY

```
# basic summary statistics
summary(ToothGrowth)
```

```
##      len      supp      dose
## Min.   : 4.20    OJ:30    0.5mg:20
## 1st Qu.:13.07    VC:30     1mg  :20
## Median :19.25                2mg  :20
## Mean   :18.81
```

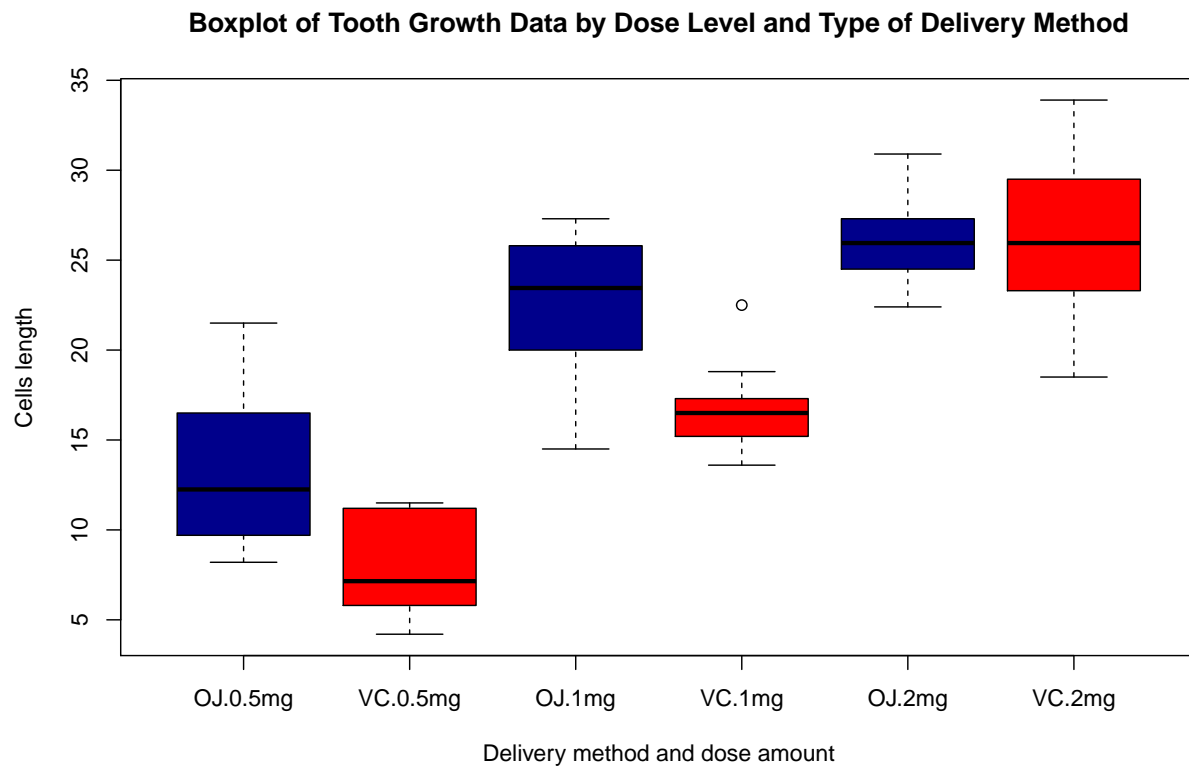
```
## 3rd Qu.:25.27
## Max.    :33.90
```

```
# knitr::kable(summary(ToothGrowth))

# summary statistics within each combination of dose method and delivery level
by(ToothGrowth$len, INDICES = list(ToothGrowth$supp, ToothGrowth$dose), summary)
```

```
## : OJ
## : 0.5mg
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   8.20   9.70   12.25   13.23   16.18   21.50
## -----
## : VC
## : 0.5mg
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   4.20   5.95   7.15    7.98   10.90   11.50
## -----
## : OJ
## : 1mg
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  14.50   20.30   23.45   22.70   25.65   27.30
## -----
## : VC
## : 1mg
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  13.60   15.27   16.50   16.77   17.30   22.50
## -----
## : OJ
## : 2mg
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  22.40   24.58   25.95   26.06   27.08   30.90
## -----
## : VC
## : 2mg
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  18.50   23.38   25.95   26.14   28.80   33.90
```

```
boxplot(len ~ supp * dose, data=ToothGrowth, xlab = 'Delivery method and dose amount', ylab= 'Cells length')
```



Box plot figure clearly shows that on average the length of odontoblasts (cells responsible for tooth growth) increases as the dose level grows. Therefore, we can observe a positive correlation between these two factors. On the other hand, it is not clear whether there is any correlation between the cells length and the type of delivery method (orange juice-OJ vs supplement-VC). Next we will use confidence intervals and hypothesis tests to compare tooth growth (or rather, cells length) by delivery method and dosage level.

3. COMPARING TOOTH GROWTH BY SUPP AND DOSE WITH CONFIDENCE INTERVALS AND/OR HYPOTESIS TESTS

- Checking for correlation between the delivery method and change in tooth growth (length of odontoblasts)

```
library(dplyr)
```

```
t.test(len ~ supp, paired = F, var.equal = F, data = ToothGrowth)
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.1710156 7.5710156
## sample estimates:
## mean in group OJ mean in group VC
## 20.66333 16.96333
```

A confidence interval of $[-0.171, 7.571]$ does not allow us to reject the null hypothesis since the interval includes 0. However, the p-value is slightly higher than .05.

- Checking the data for correlation between dose level and change in tooth growth within each dose level group

```
dose1 <- filter(ToothGrowth, dose == "0.5mg")
dose2 <- filter(ToothGrowth, dose == "1mg")
dose3 <- filter(ToothGrowth, dose == "2mg")
```

```
t.test(len ~ supp, paired = F, var.equal = F, data = dose1)
```

```
##
## 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
```

```
t.test(len ~ supp, paired = F, var.equal = F, data = dose2)
```

```
##
## 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
```

```
t.test(len ~ supp, paired = F, var.equal = F, data = dose3)
```

```
##
## 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
```

The confidence intervals for dose levels 0.5mg ([1.72, 8.78] and 1mg [2.80, 9.06]) allow us to reject the null hypothesis. Furthermore, the p-value for these two categories is very low. However, the confidence interval for dose level 2.0 ([-3.80, 3.64]) would not permit us to reject the null hypothesis. Therefore, regarding this criteria, we can not reject the null hypothesis and shall consider that there are no correlations between dose level and change in length of odontoblasts.

- Checking for correlation between the dose level and change in tooth growth

Since there are three levels of dosage, we need to perform three tests (one vs another: comparing data for dose 0.5mg vs dose 1mg, for dose 0.5mg vs dose 2mg and lastly for dose 1mg vs dose 2mg).

```
complevel1 <- filter(ToothGrowth, dose=="0.5mg"|dose=="1mg")
complevel2 <- filter(ToothGrowth, dose=="0.5mg"|dose=="2mg")
complevel3 <- filter(ToothGrowth, dose=="1mg"|dose=="2mg")
tapply(ToothGrowth$len, ToothGrowth$dose, mean)
```

```
## 0.5mg    1mg    2mg
## 10.605 19.735 26.100
```

```
# comparing data for dose 0.5mg vs dose 1mg
t.test(len ~ dose, paired = F, var.equal = F, data = complevel1)
```

```
##
## Welch Two Sample t-test
##
## data: len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -11.983781 -6.276219
## sample estimates:
## mean in group 0.5mg    mean in group 1mg
##          10.605          19.735
```

```
diff1 <- 10.605-19.735
```

The difference between the two means for the two groups equates to -9.13. The confidence interval for this difference ranges from [-11.983: -6.276]. So, there is a clear difference between the two groups and we are likely to reject the null hypothesis. Furthermore, the p-value is very low and less than .05.

```
# comparing data for dose 0.5mg vs dose 2mg
t.test(len ~ dose, paired = F, var.equal = F, data = complevel2)
```

```
##
## Welch Two Sample t-test
##
## data: len by dose
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -18.15617 -12.83383
## sample estimates:
## mean in group 0.5mg    mean in group 2mg
##          10.605          26.100
```

```
diff2 <- 10.605-26.1
```

The difference between the two means for the two groups equates to -15.495. The confidence interval for this difference ranges from [-18.156: -12.833]. So, there is a clear difference between the two groups and we are likely to reject the null hypothesis. Furthermore, the p-value is very low and much less than .05.

```
# comparing data for dose 1mg vs dose 2mg
t.test(len ~ dose, paired = F, var.equal = F, data = complevel3)

##
## Welch Two Sample t-test
##
## data: len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -8.996481 -3.733519
## sample estimates:
## mean in group 1mg mean in group 2mg
##          19.735          26.100
```

```
diff3 <- 19.735-26.1
```

The difference between the two means for the two groups equates to -6.365. The confidence interval for this difference ranges from [-8.996: -3.733]. So, there is a clear difference between the two groups and we are likely to reject the null hypothesis. Furthermore, the p-value is very low and much less than .05.

4. CONCLUSION

There is a **significant correlation between the dose level and change in tooth growth** (or, more exactly, on length of odontoblasts). A higher dose level consistently yields longer teeth.

When checking the data for correlation between dose level and change in tooth growth within each dose level group, we can observe some correlations for dose levels 0.5mg and 1mg, but not 2mg. It means that increasing in dosage for both delivery methods has a positive correlation with tooth growth but to a lesser extent for 2mg.