# Brief Analysis of Tooth Growth Data

Riccardo Finotello

19 June 2020

## Overview

In this brief analysis we consider data on the growth of odontoblasts cells in the teeth of guinea pigs under vitaminic treatment. We provide a short exploratory data analysis and compare tooth growth by **supp** and **dose** using basic statistical inference.

## Data

We first load the dataset from the `datasets` package in `R`.

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.6.3
```

```
data(ToothGrowth)

dat <- data.table(ToothGrowth, stringsAsFactors = FALSE)
names(dat) <- c("Length", "Supplement", "Dose")
```

The dataset describes the effects of the dosage of vitamin C or orange juice (the `Dose` column, numeric) in guinea pigs with different supplements type (`Supplement`, factor) on the length of teeth growth (the `Length` column, numeric). We provide a brief summary as a first reference:

```
## Classes 'data.table' and 'data.frame':   60 obs. of  3 variables:
##  $ Length    : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ Supplement: 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 ...
##  - attr(*, ".internal.selfref")=<externalptr>
```

We will be interested in a twofold analysis: we will investigate the outcome of administering the vitamins through *orange juice* (identified by `OJ`) or *vitamin C* (`VC`) on the length of the tooth, then we will try to infer something more on the dosage.

The data is made of both numeric and categorical variables. In particular the column `Supplement` can be used as a level to distinguish different treatments:

```
dat.VC <- dat[Supplement == "VC", .(Length, Dose)] # data for vitamin C
dat.OJ <- dat[Supplement == "OJ", .(Length, Dose)] # data for orange juice
```

At a visual level we can already appreciate the difference between different treatments on the length of the odontoblasts.

## Effect of Supplement Type on Tooth Growth

First of all we shall consider the impact of supplement type on the growth. We first plot the range of variation of the length per supplement type in Figure 2.

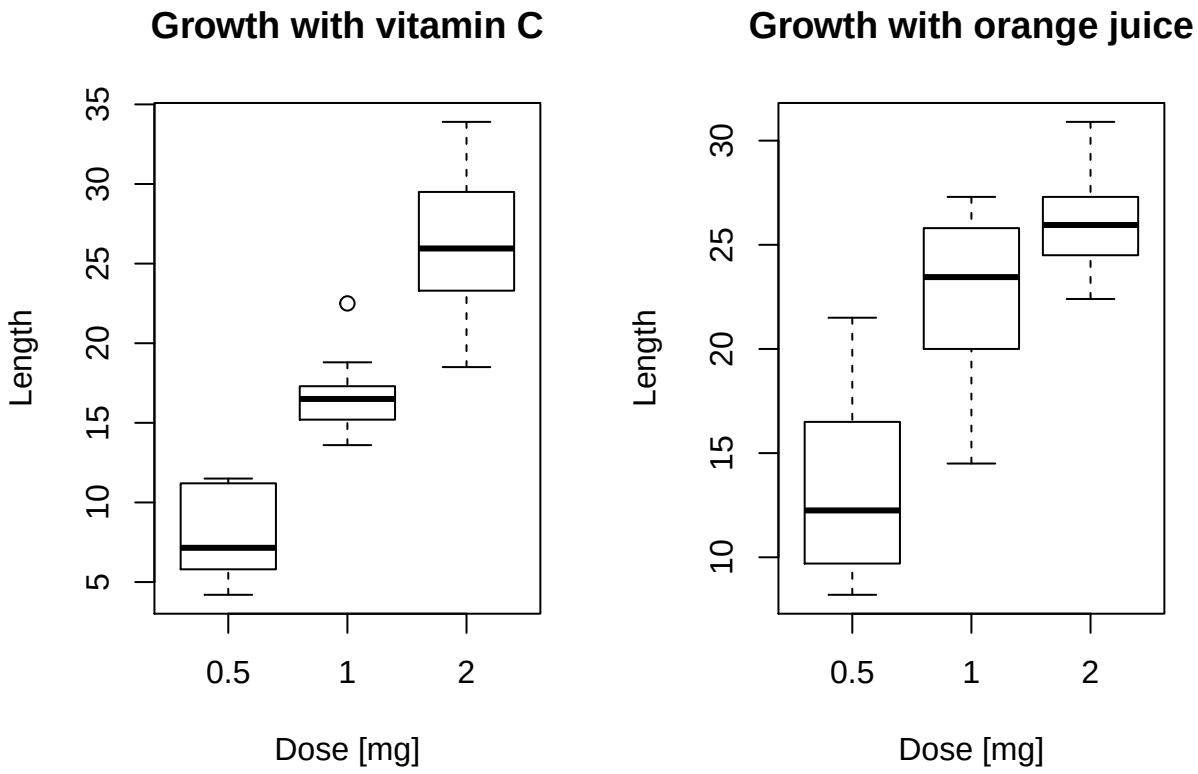


Figure 1: Growth length of odontoblasts for different treatments.

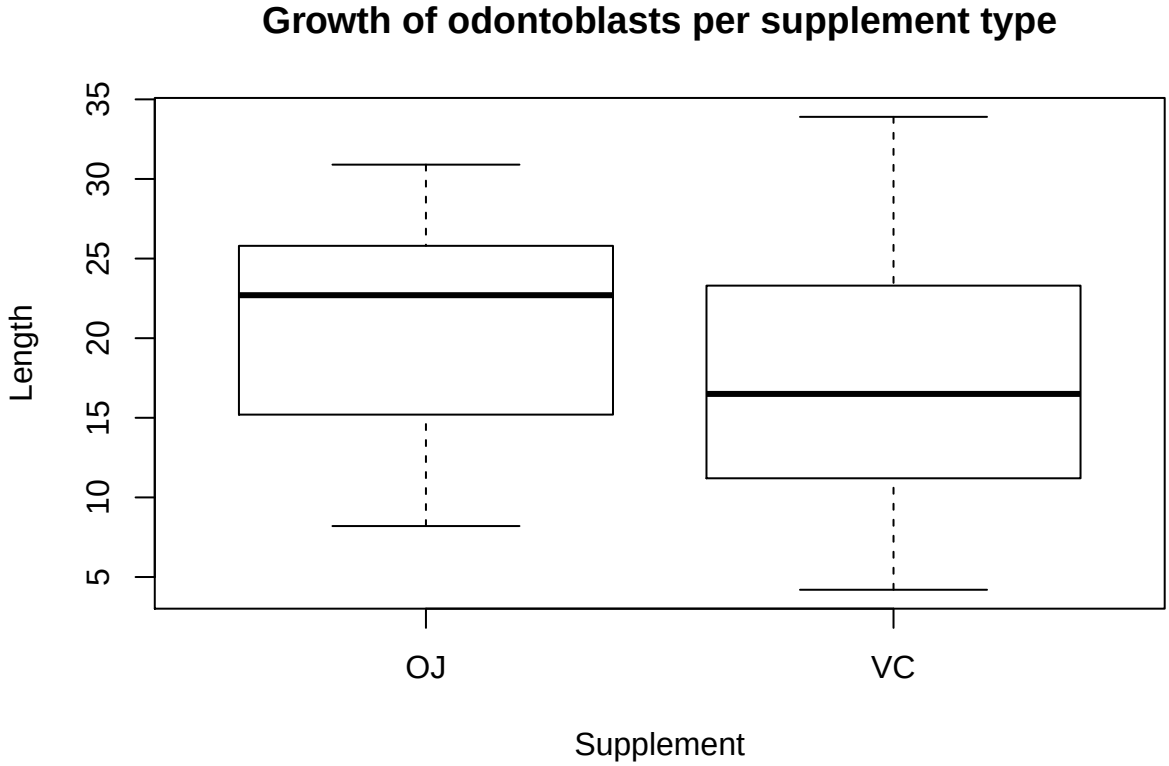


Figure 2: Growth of odontoblasts grouped by supplement type.

It seems that the largest effect on growth is given by giving supplements of orange juice to the guinea pigs. We can in fact see that the mean are relevantly higher in this case:

```
##
##  Welch Two Sample t-test
##
## data:  Length by Supplement
## t = 1.9153, df = 55.309, p-value = 0.03032
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4682687       Inf
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

The *t test* of unpaired, one-sided populations leads to a p-value less than 0.05 which leads to the rejection of the null hypothesis $H_0$ (i.e. there is no average difference), supported also by the confidence interval of the difference of the means which is strictly positive: the averange growth using orange juice seems to be larger than using vitamin C.

## Effect of Dosage of Orange Juice on Tooth Growth

We then explore the impact of the dosage of orange juice on tooth growth:

```
## Warning: package 'knitr' was built under R version 3.6.3
```

| Length | Dose |
|---|---|
| Min. : 8.20 | Min. :0.500 |
| 1st Qu.:15.53 | 1st Qu.:0.500 |
| Median :22.70 | Median :1.000 |
| Mean :20.66 | Mean :1.167 |
| 3rd Qu.:25.73 | 3rd Qu.:2.000 |
| Max. :30.90 | Max. :2.000 |

From the plot in Figure 1 we have already seen the difference in the outcome when different doses are administered. We then consider the average growth in length when using the orange juice:

| Dose | Average.Length |
|---|---|
| 0.5 | 13.23 |
| 1.0 | 22.70 |
| 2.0 | 26.06 |

The question that we ask is therefore whether a higher dosage of orange juice can lead to larger growths (assuming the null-hypothesis $H_0$ stating the opposite):

```
# the test is ordered such that higher doses are the RHS,
# thus we need the "less" alternative
t.1 <- t.test(Length ~ Dose,
              data = dat.OJ[Dose == 0.5 | Dose == 1.0],
              paired = FALSE,
              alternative = "less")
t.2 <- t.test(Length ~ Dose,
              data = dat.OJ[Dose == 0.5 | Dose == 2.0],
```

```
              paired = FALSE,
              alternative = "less")
t.3 <- t.test(Length ~ Dose,
              data = dat.OJ[Dose == 1.0 | Dose == 2.0],
              paired = FALSE,
              alternative = "less")

print(paste("P-value between 0.5 mg and 1.0 mg:", t.1$p.value))
```

```
## [1] "P-value between 0.5 mg and 1.0 mg: 4.39245952758075e-05"
```

```
print(paste("P-value between 0.5 mg and 2.0 mg:", t.2$p.value))
```

```
## [1] "P-value between 0.5 mg and 2.0 mg: 6.61891938848617e-07"
```

```
print(paste("P-value between 1.0 mg and 2.0 mg:", t.3$p.value))
```

```
## [1] "P-value between 1.0 mg and 2.0 mg: 0.0195975710231221"
```

It seems therefore that higher doses of orange juice may lead to benefit for the growth in length of the odontoblasts. We can also look at the confidence intervals to support the conclusion. In particular between the 95% CI between 0.5 mg and 1.0 mg is $-\infty$, -6.214316 and between 1.0 mg and 2.0 mg is $-\infty$, -0.7486236, which show that they do not include 0.

## Conclusions

From the simple data analysis we can therefore conclude that the growth of the odontoblasts is heavily influenced by the type of supplement and its dosage. In particular we showed that the best treatment for a larger growth is the administration of orange juice to the guinea pigs in larger quantities.