Statistical Inference - Final Project: Part 2
Basic Inferential Data Analysis with R

Anastasia Thöns

December 26, 2016

Introduction

In the project `ToothGrowth` data set is used - *The Effect of Vitamin C on Tooth Growth in Guinea Pigs* (Ref.: ETH Zürich). 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). The data `ToothGrowth` is a data frame with 60 observations on 3 variables:

[,1] len - numeric Tooth length

[,2] supp - factor Supplement type (VC or OJ).

[,3] dose - numeric Dose in milligrams/day

and it is a part of the library `datasets`.

Source

- C. I. Bliss (1952) The Statistics of Bioassay. Academic Press.

References

- McNeil, D. R. (1977) Interactive Data Analysis. New York: Wiley.

Implementation in R

The following R libraries are used:

```
library(datasets)
library(data.table)
library(ggplot2)
```

Load the `ToothGrowth` data as a `data.table` structure and give descriptive names.

```
data(ToothGrowth)
ToothGrowth <- data.table(ToothGrowth)
names(ToothGrowth) <- c("tooth.length", "supplement.type", "dose")
```

Odontoblast length in the data set varies in the range $4.2 - -33.9$ in microns with the mean value $\mu = 18.8$ and the variance $\sigma^2 = 58.5$.

```
mean(ToothGrowth$tooth.length)
```

```
## [1] 18.81333
```

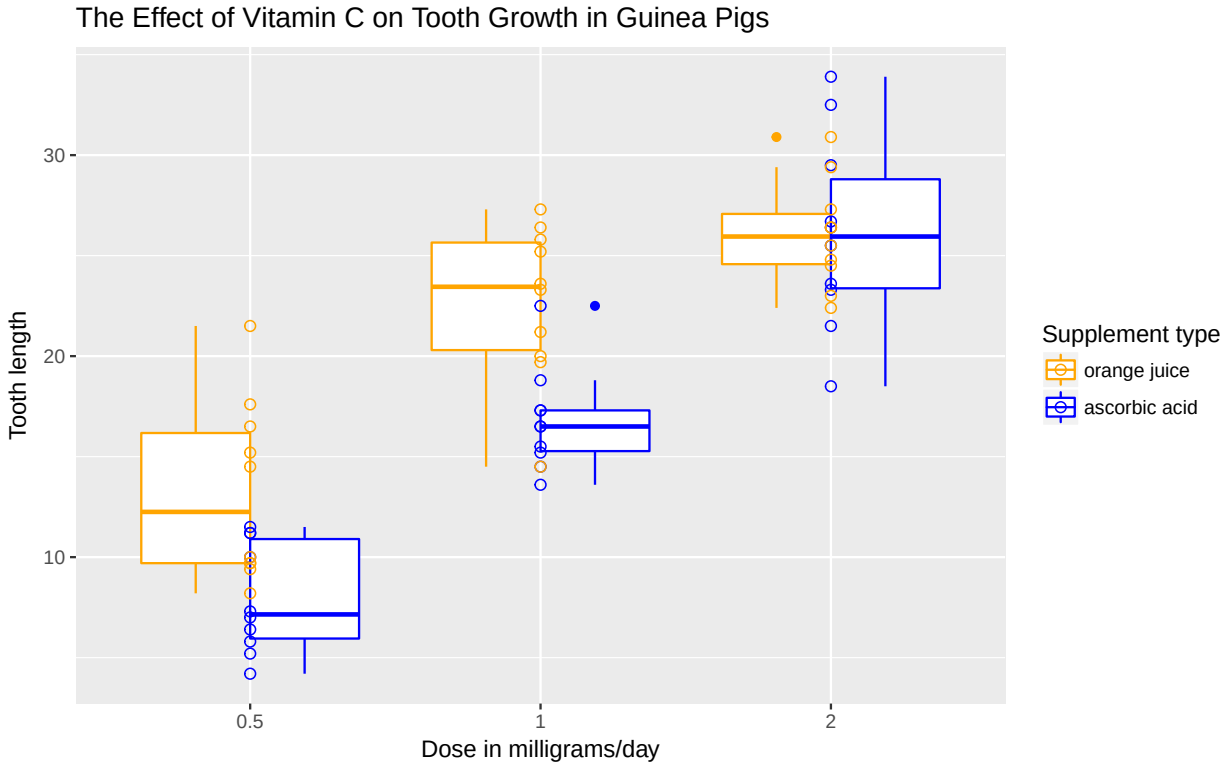
```
var(ToothGrowth$tooth.length)
```

```
## [1] 58.51202
```

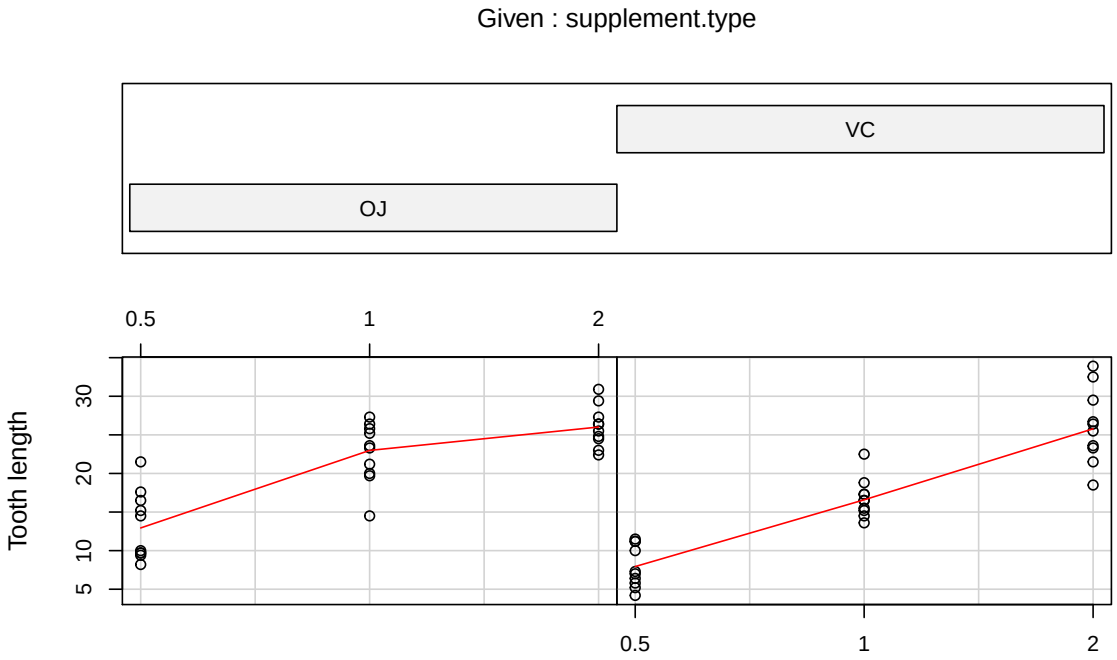
Basic summary of the data set can be found in the following table:

tooth.length	supplement.type	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

Basic scatterplot of the data set divided by the supplement type:



Conditioning plot of the data sets:



ToothGrowth data: length vs dose, given type of supplement

Looking at the graphics we can assume that the effect of vitamin C on tooth growth is proportional to the dose and also appears to be higher with orange juice than with ascorbic acid at least at lower dose (< 2 mg/day). In the following we will provide some statistical inference analysis to confirm of reject the statement.

Confidence Intervals

We now analyse the correlation between the tooth length and the supplement type by Welch's unpaired t-test:

```
t.test(tooth.length ~ supplement.type, paired = FALSE, ar.equal = TRUE,  data = ToothGrowth)
```

```
##
##  Welch Two Sample t-test
##
## data:  tooth.length by supplement.type
## 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
```

The 95 percent confidence interval $[-0.17, 7.57]$ is not fully positive.

Let us now perform the hypothesis test. The null hypothesis is that there is no correlation between the tooth grows and the delivery method, i.e. the means are the same. In the previous we can see already that the mean in case of orange juice supplement is higher then in case of ascorbic acid. The values Z and Z_α can be found as following:

```r
g1 <- ToothGrowth$tooth.length[ToothGrowth$supplement.type=="VC"]
g2 <- ToothGrowth$tooth.length[ToothGrowth$supplement.type=="OJ"]
difference <- g2 - g1
mn <- mean(difference)
s <- sd(difference)
n <- 30
Z <- sqrt(n)*(mn - 0)/s
Z
```

```
## [1] 3.302585
```

```r
Z.alpha <- qnorm(0.95)
Z.alpha
```

```
## [1] 1.644854
```

And so we can see that $Z = 3.3 > 1.64$ and we reject the null hypothesis.

Conclusions

From the performed statistical analysis on the data `ToothGrowth` we can conclude the following:

- The tooth length in tested guinea pigs is proportional to the dose supplemented of vitamin C based on exploratory data analyses and regression line.
- Based on hypothesis testing we conclude that the natural supplement type, i.e. orange juice, has a stronger affect on the tooth growth then the synthetic one, i.g. ascorbic acid.