ToothGrowth Data Exploratory Analysis

Alexander Tuzhikov

September 14, 2015

1 Synopsis

We are now moving to the part 2 of the task. Below we will explore the `ToothGrowth` data set from R `datasets` package. The headers below will correspond to the tasks. The data comes from the study "The Effect of Vitamin C on Tooth Growth in Guinea Pigs". The data is described as: *The response is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).*

2 Load the ToothGrowth data and perform some basic exploratory data analyses

```
data("ToothGrowth")
```

3 Provide a basic summary of the data

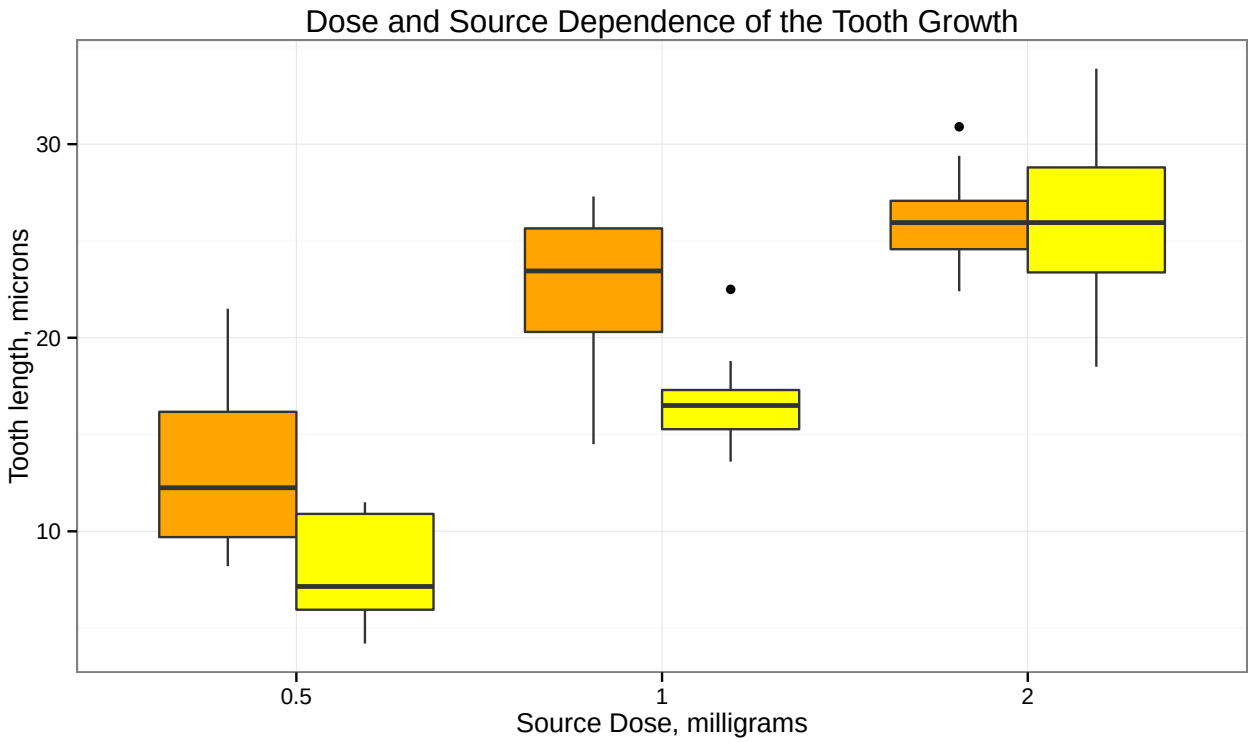
Well, the data is presented as a `data.frame` of 60 samples in 3 rows: len, supp, dose; The data is summarized in the Table 1 below (see **Code Block 1**).

- **len** is the tooth length (microns)
- **supp** is the supplement type (VC or OJ)
- **dose** is the dose (milligrams)

Table 1: Summary of the "ToothGrowth" Data (continued below)

supp	dose	min.len	max.len	mean.len	sd.len
OJ	0.5	8.2	21.5	13.23	4.46
OJ	1	14.5	27.3	22.7	3.911
OJ	2	22.4	30.9	26.06	2.655
VC	0.5	4.2	11.5	7.98	2.747
VC	1	13.6	22.5	16.77	2.515
VC	2	18.5	33.9	26.14	4.798

q25	median.len	q75
9.7	12.25	16.18
20.3	23.45	25.65
24.57	25.95	27.08
5.95	7.15	10.9
15.27	16.5	17.3
23.38	25.95	28.8



Plot 1. Vitamin C Sources Orange Juice Ascorbic Acid

There obviously exists a positive and proportionate effect of orange juice as compared with the ascorbic acid, however, this effect can only be observed at dose levels below 2 mg. We can also demonstrate this in numeric values (see below, see **Code Block 2**)

Table 3: Orange Juice vs Ascorbic Acid at 0.5 mg Dose

Test statistic	df	P value	Alternative hypothesis
3.17	14.97	*0.006359* * *	two.sided

Table 4: Orange Juice vs Ascorbic Acid at 1.0 mg Dose

Test statistic	df	P value	Alternative hypothesis
4.033	15.36	*0.001038* * *	two.sided

Table 5: Orange Juice vs Ascorbic Acid at 2.0 mg Dose

Test statistic	df	P value	Alternative hypothesis
-0.04614	14.04	*0.9639*	two.sided

Tables 3-5 confirm the data in the Plot 1 with a *t*-test (see **Code Block 3**. While the tests for 0.5- and 1.0-dose samples show significance with p-values **0.0063586** and **0.0010384** respectively, there seems to be no significant difference between the 2.0-dose level with p-value **0.9638516**.

4 Confidence intervals

Test	CI lower bound	CI upper bound	p-value
Dose 0.5 mg	1.719	8.781	0.006359
Dose 1.0 mg	2.802	9.058	0.001038
Dose 2.0 mg	-3.798	3.638	0.963852

Table 6 above provides upper and lower confidence interval bound for all three dose level comparisons (see **Code Block 4**).

5 Conclusions

Based on the data analysis discussed above, we can conclude that regardless of the source, tooth grows in guinea pigs is faster for those animals, that receive vitamin C supplemented ration. As for the sources of vitamin C, at lower doses (0.5-1 mg), orange juice favors tooth growth significantly more than ascorbic acid supplement. However, this effect is no longer significant at 2 mg dose.

6 References

[1] C. I. Bliss (1952) The Statistics of Bioassay.Academic Press. "The Effect of Vitamin C on Tooth Growth in Guinea Pigs."

7 Related R Code

7.1 Code Block 0

```
library(datasets)
library(ggplot2)
library(dplyr)
library(pander)
data("ToothGrowth")
```

7.2 Code Block 1

```
ToothGrowth %>% mutate(dose=factor(dose)) %>%
        group_by(supp, dose) %>%
        summarize(
                min.len=min(len),
                max.len=max(len),
                mean.len=mean(len),
                sd.len=sd(len),
                q25=quantile(len,probs =0.25),
                median.len=median(len),
                q75=quantile(len,probs =0.75)) %>%
        pander(caption = "Summary of the \"ToothGrowth\" Data")
```

7.3 Code Block 2

```
str(ToothGrowth)
ggplot(data=ToothGrowth, mapping=aes(x=factor(dose), y=len, fill=supp)) +
        geom_boxplot() +
        theme_bw()+
        labs(title="Dose and Source Dependence of the Tooth Growth") +
        theme(legend.key = element_rect(colour = NA), legend.position="bottom") +
        xlab("Source Dose, milligrams") +ylab("Tooth length, microns") +
        scale_fill_manual(name="Plot 1.   Vitamin C Sources",
                          labels=c("Orange Juice", "Ascorbic Acid"),
                          values=c("orange1", "yellow"))
```

7.4 Code Block 3

```
tooth.grows.list<- list()
ToothGrowth %>%
        group_by(supp, dose) %>%
        do({
                tooth.grows.list[[length(tooth.grows.list)+1]]<<- as.data.frame(.)
                names(tooth.grows.list)[length(tooth.grows.list)]<<-paste(
                        .$supp[1],.$dose[1],sep="_")
                return(.)
        }) -> ToothGrowth
#do t-tests
dose.0.5.t.test<- t.test(tooth.grows.list[["OJ_0.5"]]$len, tooth.grows.list[["VC_0.5"]]$len)
dose.1.0.t.test<- t.test(tooth.grows.list[["OJ_1"]]$len, tooth.grows.list[["VC_1"]]$len)
dose.2.0.t.test<- t.test(tooth.grows.list[["OJ_2"]]$len, tooth.grows.list[["VC_2"]]$len)

pander(dose.0.5.t.test, caption = "Orange Juice vs Ascorbic Acid at 0.5 mg Dose")
pander(dose.1.0.t.test, caption = "Orange Juice vs Ascorbic Acid at 1.0 mg Dose")
pander(dose.2.0.t.test, caption = "Orange Juice vs Ascorbic Acid at 2.0 mg Dose")
```

7.5 Code Block 4

```
ci.data<- rbind(
                c("Dose 0.5 mg",
                  round(dose.0.5.t.test$conf.int[1],3),
                  round(dose.0.5.t.test$conf.int[2],3),
                  round(dose.0.5.t.test$p.value,6)),
                c("Dose 1.0 mg",
                  round(dose.1.0.t.test$conf.int[1],3),
                  round(dose.1.0.t.test$conf.int[2],3),
                  round(dose.1.0.t.test$p.value,6)),
                c("Dose 2.0 mg",
                  round(dose.2.0.t.test$conf.int[1],3),
                  round(dose.2.0.t.test$conf.int[2],3),
                  round(dose.2.0.t.test$p.value,6))
                ) %>% as.data.frame()
colnames(ci.data)<- c("Test", "CI lower bound", "CI upper bound", "p-value")
pander(ci.data)
```