project_tooth

lipnerova

December 21, 2015

Report

Overview

This study analyse the ToothGrowth data in the R datasets package, aiming for comparing benefits of two diets on growth of teeth. The goals are to provide simple exploratory data analysis, compare two diets and draw a conclusion based on performed tests.

Loading the data

```
# Load libraries
library(dplyr); library(ggplot2); library(gridExtra); library(knitr)

# Load dataset
library(datasets)
data(ToothGrowth)
a<-(ToothGrowth) #give it easily typed name
```

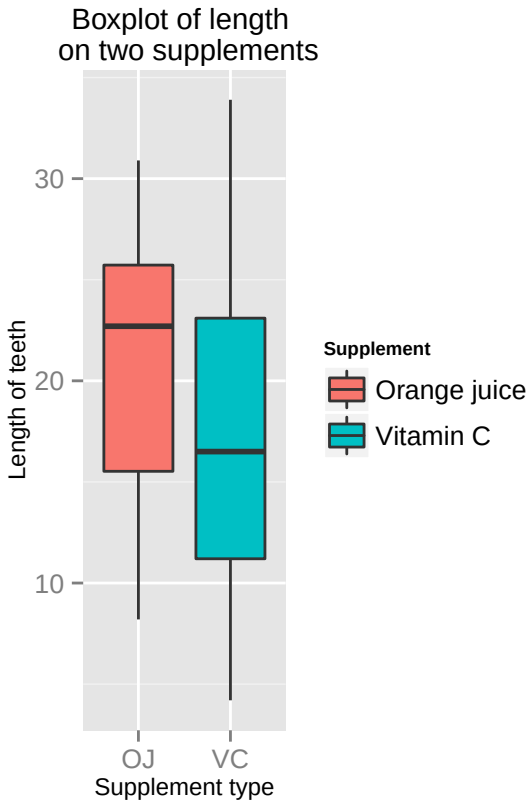
Basic xploratory data analyses

The data contains measurements of the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Animals were divided to 6 groups per 10 pigs, each group was fed with one of two supplements, orange juice (OJ) or ascorbic acid (vitami C, VC), in one of three doses (0.5, 1, or 2 mg/day). The output data has format of 60 observations and 3 variables - **len** for length of odontoblasts, **sup** for supplement type, and **dose** for dose recieved.

```
str(a) #structure of data
```

```
## '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: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

Boxplots of the data showing the differences and distributions of lengths of odontoblasts for each supplement on the left and for each dose on the right (see Appendix for code).



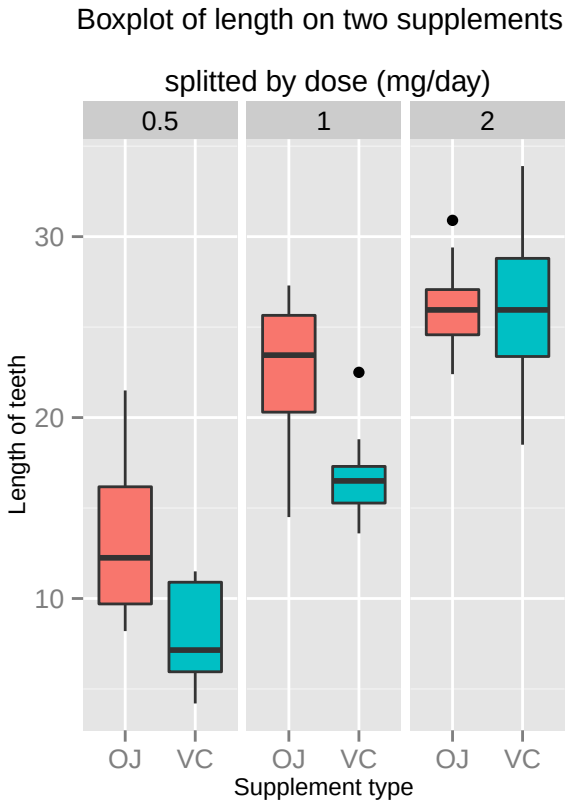


Table of means of length of odontoblasts for each dose of both supplements.

```
a_tab<-group_by(a, dose, supp) %>% summarise(mean(len))
kable(x=a_tab, digits=1, align="c")
```

dose	supp	mean(len)
0.5	OJ	13.2
0.5	VC	8.0
1.0	OJ	22.7
1.0	VC	16.8
2.0	OJ	26.1
2.0	VC	26.1

Testing

I decide to test the difference in teeth growth between supplementing orange juice (OJ) and vitamin C (VC). I use R function *t-test*. Based on EDA I suppose that orange juice gives better results.

Assumptions for testing: * Guinea pigs were selected randomly and each combination of dose and supplement fed to different group of pigs, thus they form independent groups (*paired=FALSE* as a default of t.test) * Lengths of odontoblasts are normal for tested sets of data (see Appendix for testing normality) * No info about variance, thus for sake of correctnes t.tests default *var.equal=FALSE* is used

Overall test:

H_0: there is no difference between OJ and VC

H_a: the OJ has bigger effect than VC

Partial tests for each dose:

H_0: there is no difference between OJ and VC for set dose level (e.g., 0.5, 1, and 2 mg/day).

H_a: the OJ has bigger effect than VC for set dose level

```
# Prepare table of results:
test_res <- NULL

# Test OJ vs VC overall:
test_res<-cbind(t.test(len ~ factor(supp), data=a, paired=F)$p.value,
    t.test(len ~ factor(supp), data=a, paired=F)$conf.int[1],
    t.test(len ~ factor(supp), data=a, paired=F)$conf.int[2])

# Test OJ vs VC for each dose:
for (i in c(0.5, 1, 2)) {
  test_res <- rbind(test_res,
              cbind(t.test(len ~ factor(supp), data=a[a$dose== i, ], paired=F)$p.value,
                    t.test(len ~ factor(supp), data=a[a$dose== i, ], paired=F)$conf.int[1],
                    t.test(len ~ factor(supp), data=a[a$dose== i, ], paired=F)$conf.int[2]
                    ))
  }

# Format table of results:
test_res<-as.data.frame(test_res)
row.names(test_res)<-c("OJ vs VC overall", "OJ vs VC - dose 0.5",
              "OJ vs VC - dose 1.0", "OJ vs VC - dose 2.0")
colnames(test_res)<-c("p-value", "95% CI (-)", "95% CI (+)")

# Print table of results:
kable(test_res, digits=3, row.names=T, align="c",
      caption="Summary table of test results")
```

Table 2: Summary table of test results

	p-value	95% CI (-)	95% CI (+)
OJ vs VC overall	0.061	-0.171	7.571
OJ vs VC - dose 0.5	0.006	1.719	8.781
OJ vs VC - dose 1.0	0.001	2.802	9.058
OJ vs VC - dose 2.0	0.964	-3.798	3.638

Conclusion

There is no statistically significant overall difference between orange juice and vitamin C, however for lower doses (0.5 mg/day and 1 mg/day) the orange juice has more prominent effect on growt of guinea pigs odontoblasts. High dose of 2 mg/day is sufficient to achieve maximum odontoblasts growth no matter the supplement. It also suggests that such a dose of vitamin C is high enough to compensate for whichever advantage the orange juice has at lower dose levels.

Appendix

```
# Overal boxplot:
p1 <- ggplot(a, aes(x=supp, y=len)) + geom_boxplot(aes(fill=supp)) +
  labs(title="Boxplot of length \n on two supplements",
       x="Supplement type",
       y="Length of teeth") +
  scale_fill_discrete(name="Supplement",
                      labels=c("Orange juice", "Vitamin C")) +
  theme(title = element_text(size = rel(0.75), hjust = 0.5))

# Dose boxplot:
p2 <- ggplot(a, aes(y=len, x=as.factor(supp))) +
  geom_boxplot(aes(fill=supp)) +
  guides(fill=FALSE) +
  facet_grid(.~dose) +
  labs(title="Boxplot of length on two supplements
       \n splitted by dose (mg/day)",
       x="Supplement type",
       y="Length of teeth") +
  theme(title = element_text(size = rel(0.75), hjust = 0.5))

# Arrange them:
grid.arrange(p1, p2, ncol=2, nrow=1)
```

```
shapiro.test(a$len[a$supp=="OJ" & a$dose == 0.5]) #norm
shapiro.test(a$len[a$supp=="OJ" & a$dose == 1]) #norm
shapiro.test(a$len[a$supp=="OJ" & a$dose == 2]) #norm
shapiro.test(a$len[a$supp=="VC" & a$dose == 0.5]) #norm
shapiro.test(a$len[a$supp=="VC" & a$dose == 1]) #norm
shapiro.test(a$len[a$supp=="VC" & a$dose == 2]) #norm
shapiro.test(a$len[a$supp=="OJ"]) #norm
shapiro.test(a$len[a$supp=="VC"]) #norm
shapiro.test(a$len) #norm
```