Statistical Inference course Project - Part 2

Guinea pig tooth growth experiment

Olivia Urbanowski

5/28/2020

Overview

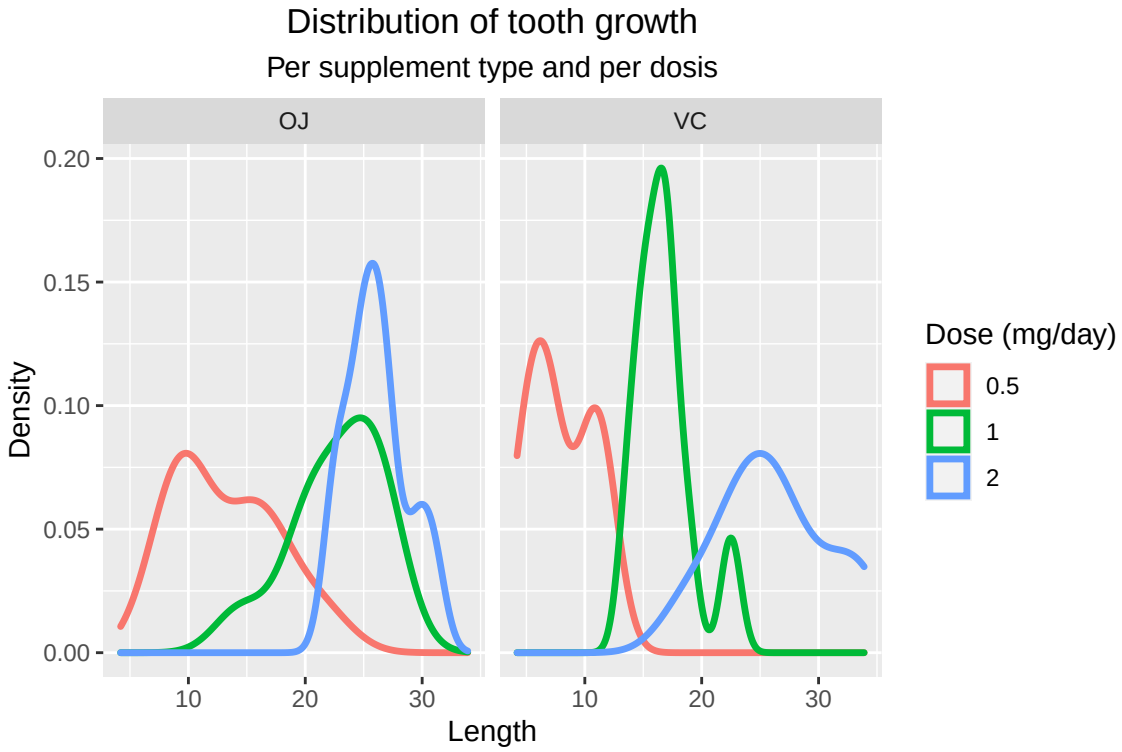
This report summarizes the analysis and results of an experiment of using of different tooth growth supplements, administrated at different dosis to guinea pigs.
The first section presents exploratory analysis of the collected data; the second section dives into the comparison of the effects of the supplements and the dosis. The third presents the conclusions.
The code for performing the analysis can be found in the appendix.

1. Data summary and exploratory analysis

The data we are analyzing presents 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 (OJ) or ascorbic acid (a form of vitamin C, VC).
Source: R Help

Table 1: Overview of the tooth growth data

Supp	Dose	Nb_obs	Min_g	Median_g	Max_g	Mean_g	SD_g
OJ	0.5	10	8.2	12.25	21.5	13.23	4.459708
OJ	1.0	10	14.5	23.45	27.3	22.70	3.910953
OJ	2.0	10	22.4	25.95	30.9	26.06	2.655058
VC	0.5	10	4.2	7.15	11.5	7.98	2.746634
VC	1.0	10	13.6	16.50	22.5	16.77	2.515309
VC	2.0	10	18.5	25.95	33.9	26.14	4.797731



A quick overview of the growth distribution highlights:

- For both supplements, the dosis appears to have played a key role in the total growth length (the higher the dose, the bigger the length)
- The choice of supplement used appears to have a mild impact on the growth, orange juice being more effective than vitamin C, but it's not conclusive
- In each of the six cases, the data appears to be more or less normally distributed

2. Analysis of tooth growth per type of supplement and dosis

As seen in the exploratory phase, for each supplement/dose combination, the observed growth presents a distribution which approximates that of a normal one. Thus, we will be using T testing methods to analyse
1. the impact of the supplement choice
2. the impact of the dose

2.1 Comparing the supplements

The code for this section can be found in the appendix

For this analysis we assume the variances of the data associated with OJ and VC respectively are different, based on the values presented in fig.1.

We performed a T test regarding the hypothesis H_0 : there is no true difference in the growth mean for both groups.

The p-value of this test is **0.061**, which means that if we want to have a 95% confidence, it is not relevant enough to reject the hypothesis.

Thus, **we cannot conclude that one of the supplements is more efficient than the other for tooth growth.**

2.2 Comparing the dosis

To compare the impact of the dose on the tooth growth per type of supplement, we performed t tests on each pair of dosis (per type of supplement), considering the hypothesis that there was no tooth growth, vs. the alternative hypothesis that the higher the dose, the bigger the growth.

Table 2: p-values of the T test results for difference in tooth growth

Supplement	Dose 0.5mg/day vs. 1mg/day	Dose 1mg/day vs. 2mg/day
Orange Juice	4.39245952758075e-05	0.0195975710231221
Vitamin C	3.40550885143253e-07	4.57780152831932e-05

The results of the performed t tests presented above all have p-values clearly lower than 5%. The results are thus significant: we can reject the initial hypothesis. We can conslude that **for both supplements, the higher the administrated dose, the bigger the growth, with a confidence interval of 95%.**

3.Conclusion

The performed analysis

- Do not allow us to conclude wether orange juice or ascorbic acid (vitamin C) is more efficient for tooth growth in guinea pig with a good enough confidence
- Show that we can conclude with a 95% confidence that increasing the dose also increases the growth length.

Code appendix

```
## this chunk serves as a purpose to set the environment, load the relevant
## packages and send all code to the code appendix
knitr::opts_chunk$set(echo = FALSE, warning = FALSE, message = FALSE,
  cache = TRUE, fig.align = "center")
library(dplyr)
library(ggplot2)
library(knitr)

set.seed(23)
## Loading the data and creating an overview
data("ToothGrowth")
growthdata <- group_by(ToothGrowth, supp, dose)
overview <- summarize(growthdata, length(len), min(len), median(len), max(len),
                      mean(len), sd(len))
colnames(overview) <- c("Supp", "Dose", "Nb_obs", "Min_g", "Median_g", "Max_g",
                        "Mean_g", "SD_g")
kable(overview, caption="Overview of the tooth growth data")
## Having a first look at how the data is distributed
g <- ggplot(data= growthdata, aes(x = len))
g + geom_density(size = 1.2, aes(color=as.factor(dose))) + facet_grid(.~supp
    ) + labs(title = "Distribution of tooth growth", subtitle =
    "Per supplement type and per dosis", x="Length", y="Density", color =
    "Dose (mg/day)") + theme(plot.title = element_text(hjust = 0.5),
    plot.subtitle = element_text(hjust = 0.5))
## Comparing the supplements: performing a t test
Orangejuice <- filter(growthdata, supp == "OJ")
Vitamin <- filter(growthdata, supp == "VC")
Testsupp <- t.test(Vitamin$len, Orangejuice$len,
                   paired = FALSE, var.equal = FALSE)
## Comparing the dosis
OJ05 <- filter(Orangejuice, dose == "0.5")
OJ1 <- filter(Orangejuice, dose == "1")
OJ2 <- filter(Orangejuice, dose == "2")
VC05 <- filter(Vitamin, dose == "0.5")
VC1 <- filter(Vitamin, dose == "1")
VC2 <- filter(Vitamin, dose == "2")

ptestVC1 <- t.test(VC05$len, VC1$len, alternative = "less", paired = FALSE,
                   var.equal = FALSE)$p.value
ptestVC2 <- t.test(VC1$len, VC2$len, alternative = "less", paired = FALSE,
                   var.equal = FALSE)$p.value
ptestOJ1 <- t.test(OJ05$len, OJ1$len, alternative = "less", paired = FALSE,
                   var.equal = FALSE)$p.value
ptestOJ2 <- t.test(OJ1$len, OJ2$len, alternative = "less", paired = FALSE,
                   var.equal = FALSE)$p.value
pvalues <- cbind(c("Orange Juice","Vitamin C"), c(ptestOJ1,ptestVC1),
                 c(ptestOJ2,ptestVC2))
colnames(pvalues) <- c("Supplement", "Dose 0.5mg/day vs. 1mg/day",
                       "Dose 1mg/day vs. 2mg/day")
kable(pvalues,
      caption = "p-values of the T test results for difference in tooth growth")
```