Tooth growth length in response to orange juice and ascorbic acid in guinea pigs

Y.V. Wang, November 21, 2015

Overview

In this Project, I analyzed the ToothGrowth Data. This data documents the response of 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).

1. Summary of the data

This data frame with 60 observations on 3 variables: -len (a numeric factor), the Tooth length; -supp (a factor variable), which stands for Supplement type (VC-ascorbic acid or Orange juice); -dose (a numeric factor), which is the dose level (0.5,1, and 2mg) for each treatment in milligrams. I conducted the exploratory data analyses in Appendix 1, where supplementary figure 1 shows the boxplot of two different treatments and three dose levels. Also in Appendix 2 and printed table, I calculated the mean and standard deviation of each treatment and their associated dosages. From the preliminary analyses, it is quite clear that the mean of each treatment and dosage are quite different. To be sure, I perform statistical analyses below.

2. Confidence interval testing

According to preliminary analyses above, I tested following hypothesis individually (by only considering the techniques we learned from this class).

2a) Test A. Do teeth growths respond to different dosages of ascorbic acid (VC) differently? In another word, when given different dosages of VC, did the teeth growth length differ?

Assumptions for the test: i) these are independent samples for each group to be tested. ii). The sample for each group is normally distributed. iii). The population standard deviation is not the same (see appendix 2).

Hypothesis testing

Null hypothesis (H_0): the teeth growth length responded to 0.5 and 1 mg, 1 and 2mg, 0.5 and 2mg of VC the same.

Alternative hypothesis (H_a): the teeth growth length responded to different dosages of VC differently.

Given the sample sizes are small (10) for each dose and the population variance is unknown, t-test would be more appropriate).

Table 1. t- statistic for VC treatment (CI is 95%)

p.value	low_conf.int	upper_conf.int	compared.groups
6.81e-07	6.31	11.27	VC1~VC0.5
9.16e-05	5.69	13.05	VC2~VC1
4.68e-08	14.42	21.90	VC2~VC0.5

Interpretation and conclusion for table 1: Given the p-value for each compared pair are smaller than 0.05, this suggest that at 5% significance level, I conclude tooth growth length are not equal for 0.5, 1, and 2mg VC treatment. Similarly, looking at the confidence intervals (CI) for each compared group, I can see that all CI are greater than 0, and this confirms that the higher dosage is more effective than the lower dose VC treatment (please note the way I use t.test(), always in the order of high vs. low dose).

2b) Test B. Do teeth growths respond to different dosages of orange juice (OJ) differently? In another word, when given different dosages of OJ, did the teeth growth length differ?

Assumptions for the test: i) these are independent samples for each group; ii) The sample is normally distributed; iii) The population standard deviation is not the same (see appendix 2).
Hypothesis testing:
Null hypothesis (H_0): the teeth growth length responded to 0.5 and 1 mg, 1 and 2mg, 0.5 and 2mg of OJ the same.

Alternative hypothesis (H_a) : the teeth growth length responded to different dosages of OJ differently. Given the sample sizes are small (10) for each dose and the population variance is unknown, t-test would be more appropriate)(please note, that I compared the data in pairs)

For the hypothesis and confidence interval test code, see appendix 4
Table 2. t- statistic for OJ treatment (CI is 95%)

p.value	low_conf.int	upper_conf.int	compared.groups
8.79e-05	5.52	13.42	OJ1~OJ0.5
0.039	0.19	6.53	OJ2~OJ1
1.32e-06	9.32	16.34	OJ2~OJ0.5

Interpretation and conclusion for table 2: Similarly to VC treatment, the p-value for each compared group of OJ treatment are also smaller than 0.05, this suggest that at 5% significance level, tooth growth length are not equal for 0.5,1, and 2mg OJ treatment. Also, the confidence interval (CI) for each compared groups are greater than 0, that suggests the tooth growth of each compared group are different at 95% CI. Also, all CI are greater than 0, and this confirms **that the higher dosage is more effective than the lower dose OJ treatment** (please note the way I use t.test(), always in the order of high vs. low dose). Please also note the low end of CI (0.19) for OJ2 and OJ1 is quite close to 0, this is probably due to the outlier in the OJ1 group (see supplementary figure 1).

2c) Test C. Do teeth growth length is the same under the treatment (OJ and VC) if given the same dosage? in another word, do the same dosage of OJ and VC have the same effect on the tooth growth

Assumptions for the test: 1. These are independent samples for each group. 2. The sample is normally distributed. 3. The population standard deviation is not the same (see appendix 2).
Hypothesis testing:
Null hypothesis (H_0): the teeth growth length is the same when given the same

dosage of OJ and VC. Althernative hypothesis (H_a) : the teeth growth length is different when given the same dosages of VC and OJ differently. Given the sample sizes are small (10) for each dose and the population variance is unknown, t-test would be more appropriate).

For the hypothesis and confidence interval test code, see appendix 5

Table 3 t- statistic for same dose but two different treatment (VC and OJ) (CI is 95%)

p.value	low_conf.int	upper_conf.int	compared.groups
0.006	-8.78	-1.72	VC0.5~OJ0.5
0.001	-9.06	-2.80	VC1~OJ1
0.964	-3.64	3.80	VC2~OJ2

Interpretation and conclusion for table 3: Because the p-value for comparing the two group dose 0.5mg of VC and OJ and the p-valus for comparing 1mg of VC and OJ are smaller than 0.05, therefore, I conclude that the tooth growth length are different when give 0.5mg of OJ or VC. Same conclusion is also true also for 1 mg of OJ and VC. By looking at the 95% confidence interval (CI) for the these paired test, we can see the CI is entirely greater than 0, suggesting **that OJ treatment is slight more effective at dose level 0.5mg and 1mg**. However, the p-value (1.323784e-06) for comparing dose level 2mg of VC and OJ is greater than 0.05, suggesting that the tooth growth lengths at 2mg level of VC and OJ are the same. This conclusion is further supported by their 95% confidence interval (-3.64,3.80), which including 0.

3. Summary (please refer to table 1,2,3 as well as supplementary figure 1 and table1 in appendices)

When we assume the samples are independently taken and the each treated group has normal distribution, I draw following conclusion about the two treatments and 3 dose levels.

1. At 5% significance level, I am confident that the tooth growth lengths are greater at the higher dose than the lower dose when given that the OJ treatment.

2. At 5% significance level, I am confident that the tooth growth lengths are greater at the higher dose than the lower dose when given the VC treatment.

3. At 5% significance level, that data do not provide sufficient evidence to conclude that a difference exists between the guinea pigs' tooth growth lengths when given the higher dose (2mg) of OJ and VC.

4. At 5% significance level, I am confident that there is a difference between tooth growth lengths when given the 0.5 or 1mg doses of OJ and VC. The lower lose (ie, 0.5 and 1mg) of OJ is more effective in fostering tooth growth than the their respective dosages of VC.

Appendices

Supplementary Information for data exploration and innital research

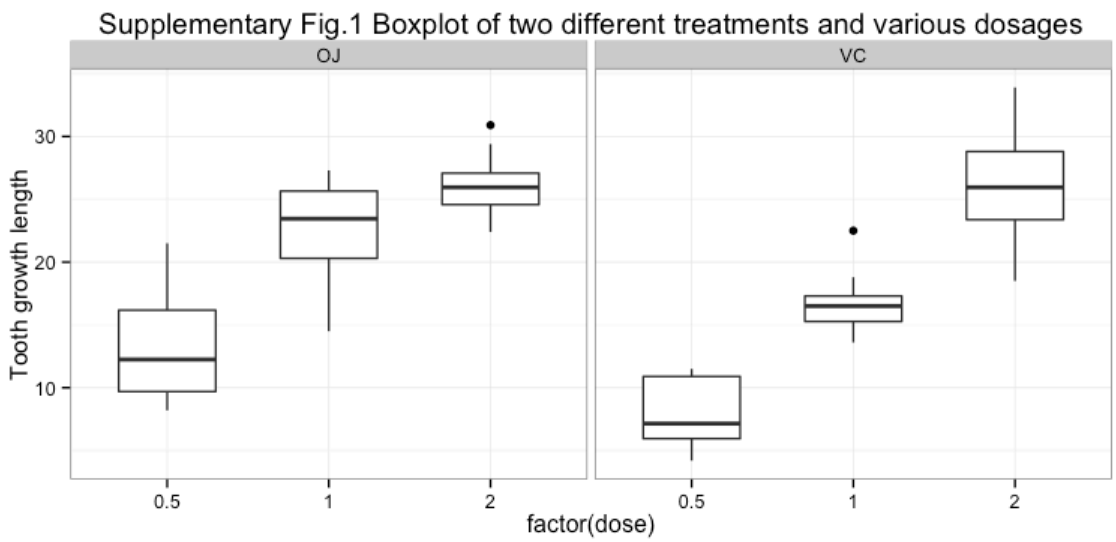
1. Data Summary and exploratory analyses

```
## '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 ...

# make a boxplot for the data for an innitial overview on the data
require(ggplot2)

## Loading required package: ggplot2

ggplot(ToothGrowth, aes(x=factor(dose),y=len))+
        geom_boxplot(outlier.size=2,outlier.shape=16,width=0.8)+
        facet_grid(.~supp)+theme_bw()+ylab("Tooth growth length")+
        ggtitle("Supplementary Fig.1 Boxplot of two different treatment
s and various dosages")
```



2. Calculate summary data (mean and sd) for tooth growth length. This step is to see whether the variance are similar for each group and treatment

```
attach(ToothGrowth)
ToothGrowth$dose<-as.factor(ToothGrowth$dose)
toothMean<-aggregate(len~supp+dose,data=ToothGrowth,mean) # calculate t
he mean
toothSd<-aggregate(len~supp+dose,data=ToothGrowth,sd) # calculate the s
d
toothSummary<-merge(toothMean,toothSd,by=c("supp","dose")) # merge the
mean and sd data
colnames(toothSummary)<-c("supp","dose","len.mean","len.SD") # set the
column name
toothSummary$len.mean<-round(toothSummary$len.mean,1)
toothSummary$len.SD<-round(toothSummary$len.SD,1)
print(toothSummary)
```

```
Supplementary table 1. The mean and sd for each dosage level and each t
reatment.

##   supp dose len.mean len.SD
## 1   OJ  0.5     13.2    4.5
## 2   OJ    1     22.7    3.9
## 3   OJ    2     26.1    2.7
## 4   VC  0.5      8.0    2.7
## 5   VC    1     16.8    2.5
## 6   VC    2     26.1    4.8

detach(ToothGrowth)
```

3. Code for Test A.

```
require(dplyr)
VC0.5<-ToothGrowth%>%filter(supp=="VC" & dose==0.5)%>%select(len) # gra
b data for VC dose0.5
VC1<-ToothGrowth%>%filter(supp=="VC" & dose==1)%>%select(len) #grab dat
a for VC dose 1
VC2<-ToothGrowth%>%filter(supp=="VC" & dose==2)%>%select(len) #grab dat
a for VC dose 2
VCpValues<-rbind(
  t.test(VC1, VC0.5, paired = FALSE, var.equal = FALSE)$p.value,
  t.test(VC2, VC1, paired = FALSE, var.equal = FALSE)$p.value,
  t.test(VC2, VC0.5, paired = FALSE, var.equal = FALSE)$p.value
) #compare p-values
VCconfidence<-rbind(
  t.test(VC1, VC0.5, paired = FALSE, var.equal = FALSE)$conf.int,
  t.test(VC2, VC1, paired = FALSE, var.equal = FALSE)$conf.int,
  t.test(VC2, VC0.5, paired = FALSE, var.equal = FALSE)$conf.int
) #compare 95% Confidence interval
VCsummary<-data.frame(cbind(VCpValues,VCconfidence))
colnames(VCsummary)<-c("p.value","low_conf.int","upper_conf.int")
rownames(VCsummary)<-c("VC1~VC0.5","VC2~VC1","VC2~VC0.5")
print(VCsummary) # table 1

##         p.value low_conf.int upper_conf.int compared.groups
## 1 6.811018e-07     6.314288       11.26571       VC1~VC0.5
## 2 9.155603e-05     5.685733       13.05427         VC2~VC1
## 3 4.681577e-08    14.418488       21.90151       VC2~VC0.5
```

4. Code for Test B.

```
require(dplyr)
OJ0.5<-ToothGrowth%>%filter(supp=="OJ" & dose==0.5)%>%select(len) #grab
data for OJ dose 0.5
OJ1<-ToothGrowth%>%filter(supp=="OJ" & dose==1)%>%select(len) #grab dat
a for OJ dose 1
OJ2<-ToothGrowth%>%filter(supp=="OJ" & dose==2)%>%select(len) #grab dat
a for OJ dose 2
OJpValues<-rbind(
```

```r
  t.test(OJ1, OJ0.5, paired = FALSE, var.equal = FALSE)$p.value,
  t.test(OJ2, OJ1, paired = FALSE, var.equal = FALSE)$p.value,
  t.test(OJ2, OJ0.5, paired = FALSE, var.equal = FALSE)$p.value
) #compare p-values
OJconfidence<-rbind(
  t.test(OJ1, OJ0.5, paired = FALSE, var.equal = FALSE)$conf.int,
  t.test(OJ2, OJ1, paired = FALSE, var.equal = FALSE)$conf.int,
  t.test(OJ2, OJ0.5, paired = FALSE, var.equal = FALSE)$conf.in
) #compare 95% Confidence interval
OJsummary<-data.frame(cbind(OJpValues,OJconfidence))
OJsummary$compared.groups<-factor(c("OJ1~OJ0.5","OJ2~OJ1","OJ2~OJ0.5"))
colnames(OJsummary)<-c("p.value","low_conf.int","upper_conf.int","compa
red.groups")
print(OJsummary) #table 2

##         p.value low_conf.int upper_conf.int compared.groups
## 1 8.784919e-05    5.5243656      13.415634       OJ1~OJ0.5
## 2 3.919514e-02    0.1885575       6.531443         OJ2~OJ1
## 3 1.323784e-06    9.3247594      16.335241       OJ2~OJ0.5
```

5. Code for Test C. Do teeth growth length is the same under the treatment (OJ and VC) if given the same dosage?

```r
Dose_pValues<-rbind(
  t.test(VC0.5, OJ0.5, paired = FALSE, var.equal = FALSE)$p.value,
  t.test(VC1, OJ1, paired = FALSE, var.equal = FALSE)$p.value,
  t.test(VC2, OJ2, paired = FALSE, var.equal = FALSE)$p.value
) #compare p-values
Dose_confidence<-rbind(
  t.test(VC0.5, OJ0.5, paired = FALSE, var.equal = FALSE)$conf.int,
  t.test(VC1, OJ1, paired = FALSE, var.equal = FALSE)$conf.int,
  t.test(VC2, OJ2, paired = FALSE, var.equal = FALSE)$conf.int
) #compare 95% Confidence interval
Dosage_summary<-data.frame(cbind(Dose_pValues,Dose_confidence))
Dosage_summary$compared.groups<-c("VC0.5~OJ0.5","VC1~OJ1","VC2~OJ2")
colnames(Dosage_summary)<-c("p.value","low_conf.int","upper_conf.int","
compared.groups")
print(Dosage_summary) #table 3

##       p.value low_conf.int upper_conf.int compared.groups
## 1 0.006358607    -8.780943      -1.719057     VC0.5~OJ0.5
## 2 0.001038376    -9.057852      -2.802148         VC1~OJ1
## 3 0.963851589    -3.638070       3.798070         VC2~OJ2
```