Vitamin C Effect on Tooth Growth in Guinea Pigs

Statistical Inference Coursera Project - Part 2

Magdalena Paszko

This paper is a part of the final project in the Coursera's Statistical Inference course (a part of Data Science specialization by Johns Hopkins University) in which we use hypothesis testing and confidence intervals to make inference on the effect of vitamin C on tooth growth in guinea pigs. The complete R code is available in the appendix.

Basic exploratory data analysis

The data analyzed below contains is **ToothGrowth** dataset from datasets package in R. It contains observations on tooth growth (**len** variable) of 60 guinea pigs that received one of three **dose** levels of vitamin C (0.5, 1 and 2 mg/day) by one of two supplements (**supp** variable): orange juice (OJ) or ascorbic acid (VC). Here is a basic summary of the data:

```
##       len         supp     dose
##  Min.   : 4.20   OJ:30   0.5:20
##  1st Qu.:13.07   VC:30   1  :20
##  Median :19.25           2  :20
##  Mean   :18.81
##  3rd Qu.:25.27
##  Max.   :33.90
```

Our focus in this analysis is whether there is a significant difference in tooth length between the different doses and supplements of vitamin C. In order to get a rough idea of possible differences we plot boxplots below.

Figure 1 Figure 2 Figure 3

In Figure 1 we plot a box-and-whisker of the tooth length by vitamin C supplement. Stating a difference between the two groups is not obvious: the median in the Orange Juice group is higher than that in the VC group, but the variability in the VC group is visibly higher - its range covers the OJ range entirely.

In Figure 2 we plot a box-and-whisker of the tooth length by vitamin C dose. There seem to be a considerable difference between the three groups: the higher the dose, the greater tooth length, although there is a quite pronounced dispersion in dose 1.

In Figure 3 a box-and-whisker of the tooth length by both factors. The differences between supplements seem to be more pronounced for the doses 0.5 and 1 than for dose 2. On the other hand in both supplement groups there is a visible growth of tooth length with increase of the dose.

Statistical inference on differences in tooth length

In order to make statistical inference about differences in tooth length we use hypothesis testing with Student's t-test and confidence intervals for means difference in each group.

Following Student's t-test assumptions we assume normality of the underlying data. Furthermore we assume unknown and unequal population variances, which gives Welch's test with the following statistic:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where $\bar{x_1}$ and $\bar{x}_2$ are sample group means and s_1^2 and s_2^2 are sample group variances.

We fix type I error at $\alpha = 5\%$ and, since we do not have any previous information about possible differences, we perform two-sided tests with the following hypotheses:

$H_0 : \mu_1 = \mu_2$

$H_A : \mu_1 \neq \mu_2$

Difference between supplements

We test a null hypothesis that tooth length means are equal in the two supplement groups (OJ and VC).

Table 1: t-test results for supplement groups

supp.pair	t	p.value	conf.int.min	conf.int.max
OJ-VC	1.92	0.0606345	-0.17	7.57

T statistic for the mean difference was 1.92. Under the null hypothesis the probability of obtaining a statistic at least as extreme is 0.06, which is greater than our fixed type I error. Also, the 95% confidence interval for the mean difference covers 0. We conclude that there is not enough evidence to reject the null hypothesis. The difference between means in two supplement groups seems to be statistically insignificant.

Difference between doses

We perform pairwise hypothesis tests that difference in means of the three dose groups are equal.

Table 2: t-test results for dose groups

dose.pair	t	p.value	conf.int.min	conf.int.max
0.5 - 1	-6.48	1.00e-07	-11.98	-6.28
0.5 - 2	-11.80	0.00e+00	-18.16	-12.83
1 - 2	-4.90	1.91e-05	-9.00	-3.73

For every dose pair under the null hypothesis the probability of obtaining t statistic as extreme as the calculated one is very low, considerably lower than the fixed type I error. Also, none of the calculated

95% confidence intervals for the mean difference covers 0. We conclude that for each dose pair we can reject the null hypothesis about equality of means. The difference between means in each dose group is statistically significant.

Difference between doses and supplements

Lastly we consider differences in means between doses and supplements. Given a relatively large number of pairwise comparisons to perform, we limit the analysis to plotting confidence intervals for each mean. Under the assumption of normality of the underlying data and unknown population variance, a $(1-\alpha)\%$ confidence interval is given by :

$$(\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}}; \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}})$$

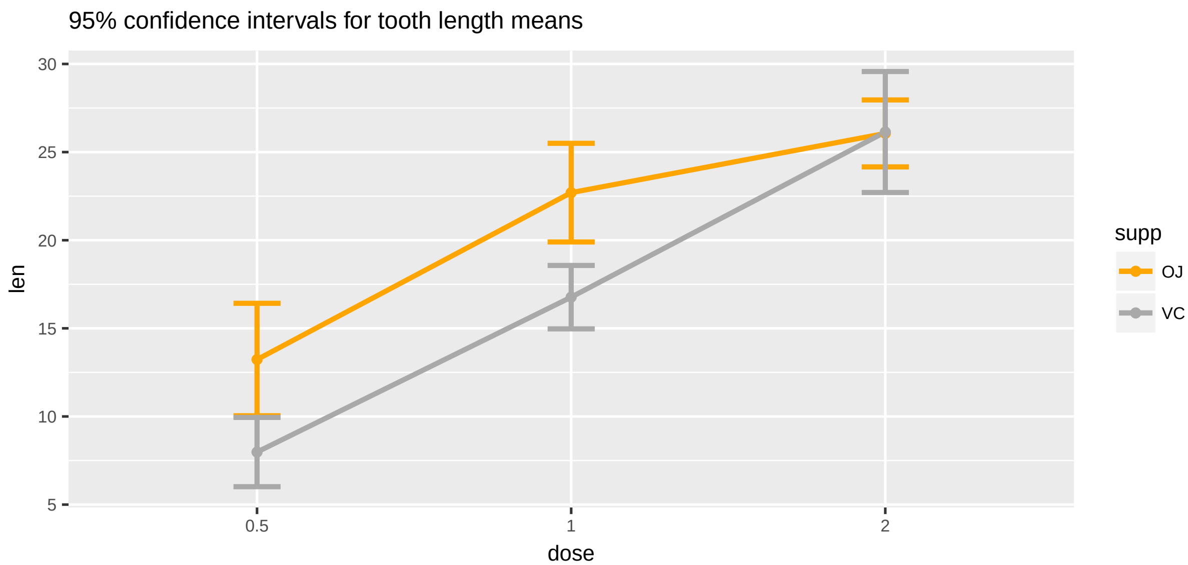

Figure 4

Figure 4 presents the results of calculating 95% confidence intervals for each of the six means.

For VC supplement none of the three dose mean confidence intervals overlap, we can thus state that in this group all dose means are significantly different. In the OJ group the situation is similar except that the intervals for dose 1 and 2 do overlap, thus we can state that in this group there is significant statistical difference between groups 0.5 and 1 and 0.5 and 2.

On the other hand, both for doses 0.5 and 1 the OJ and VC supplement mean intervals do not overlap and we can state that in those dose groups differences between supplements OJ and VC are statistically significant. Lastly, in the dose 2 group the two supplement means are almost identical and their respective confidence intervals overlap, which indicates that there is no evidence for significant difference between tooth length after receiving OJ and VC when applying dose 2.

Final Conclusions

Using two statistical methods: hypothesis testing and confidence intervals we found some significant effects of vitamin C supplement type and dose size on tooth length in guinea pigs, based on available data. Specifically we found that ascorbic acid gives significantly larger growth with each dose increase, whereas orange juice gives this effect only for smaller doses. Also, we found that for small and medium doses orange juice gives significantly larger tooth growth than ascorbic acid, but for big doses this effect is no longer present and the two supplements behave in a similar way.

Appendix - R code

Data loading and summary:

```
library(datasets)
df <- ToothGrowth
df$dose <- factor(df$dose, levels=c(0.5,1,2))
summary(df)
```

Exploratory data analysis - Figures 1, 2 and 3

```
library(ggplot2)
library(gridExtra)

g1 <- ggplot(df, aes(x=supp, y=len)) +
  geom_boxplot() +
  labs(title="Tooth length by supplement",
       caption="Figure 1") +
  theme(text = element_text(size=9),
        plot.title = element_text(size = 10))

g2 <- ggplot(df, aes(x=dose, y=len)) +
  geom_boxplot() +
  labs(title="Tooth length by dose",
    caption="Figure 2") +
  theme(text = element_text(size=9),
        plot.title = element_text(size = 10))

g3 <- ggplot(df, aes(x=dose, y=len, color=supp)) +
  geom_boxplot() +
  scale_color_manual(values=c("orange", "darkgray")) +
  labs(title="Tooth length by supplement \n and dose",
       caption="Figure 3")+
  theme(text = element_text(size=9),
        plot.title = element_text(size = 10),
        legend.position="bottom")
grid.arrange(g1, g2, g3, ncol=3)
```

T-test results for supplement groups - Table 1:

```
library(knitr)
t0 <- t.test(len ~ supp, data=df, conf.interval=0.95, var.equal=FALSE)
T0 <- data.frame("supp.pair"="OJ-VC", "t"=round(t0$statistic,2), "p.value"=t0$p.value,
                 "conf.int.min"=round(t0$conf.int[1],2),"conf.int.max"=round(t0$conf.int[2],2),
                 row.names = NULL)
kable(T0, caption="t-test results for supplement groups")
```

T-test results for dose groups - Table 2:

```
T <- data.frame("dose.pair" = c("0.5 - 1", "0.5 - 2", "1 - 2"))
t1 <- t.test(len ~ dose, data=df[df$dose %in% c(0.5, 1),], conf.interval=0.95, var.equal=FALSE)
t2 <- t.test(len ~ dose, data=df[df$dose %in% c(0.5, 2),], conf.interval=0.95, var.equal=FALSE)
t3 <- t.test(len ~ dose, data=df[df$dose %in% c(1, 2),], conf.interval=0.95, var.equal=FALSE)
T$t <- c(round(t1$statistic,2), round(t2$statistic,2), round(t3$statistic,2))
T$p.value <- c(t1$p.value, t2$p.value, t3$p.value)
T$conf.int.min <- c(round(t1$conf.int[1],2), round(t2$conf.int[1],2), round(t3$conf.int[1],2))
T$conf.int.max <- c(round(t1$conf.int[2],2), round(t2$conf.int[2],2), round(t3$conf.int[2],2))
kable(T, caption="t-test results for dose groups")
```

95% confidence intervals calculation and plotting - Figure 4:

```
library(Rmisc)
dfs <- summarySE(df, measurevar="len", groupvars=c("supp","dose"))
ggplot(dfs, aes(x=dose, y=len, group=supp, color=supp)) +
    scale_color_manual(values=c("orange", "darkgray")) +
    geom_line(size=1) +
    geom_point() +
    geom_errorbar(aes(ymin=len-ci, ymax=len+ci), width=0.15, size=1) +
    labs(title="95% confidence intervals for tooth length means",
         caption="Figure 4") +
    theme(text = element_text(size=9),
        plot.title = element_text(size = 10))
```