Peer-graded Assignment: Statistical Inference Course Project Part 2

Kan Chuen LAM

January 3, 2018

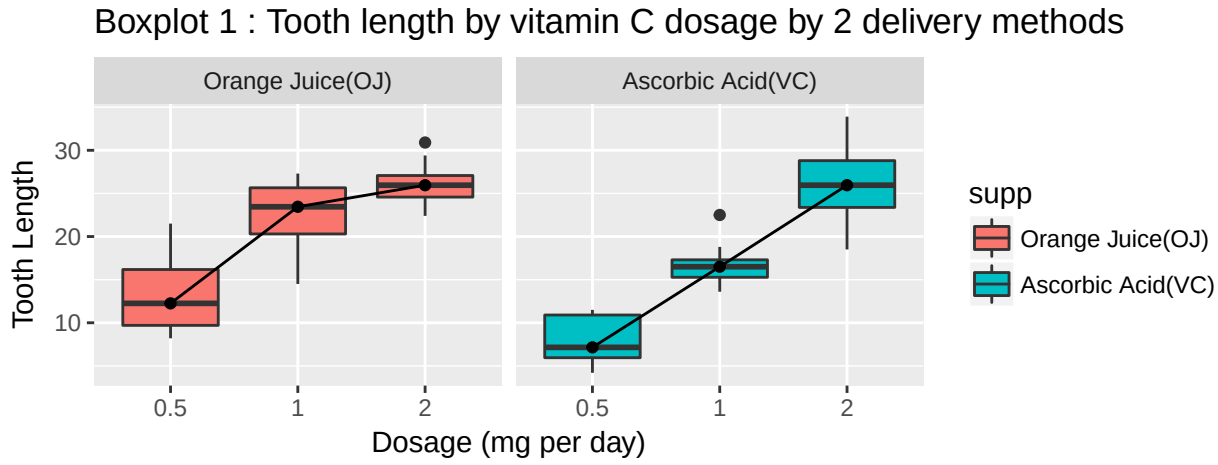
Overview

This is an analysis on the dataset ToothGrowth to investigate the relationship of the tooth length of Guinea Pigs under the impact of vitamin C at different dosage level using 2 different delivery supplements, orange juice and ascorbic acid. Let's look at the data summary :

```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

Exploratory Data Analysis

```
TG <- ToothGrowth
levels(TG$supp) <- c("Orange Juice(OJ)", "Ascorbic Acid(VC)")
ggplot(data=TG, aes(x=factor(dose),y=len)) + geom_boxplot(aes(fill=supp)) +
  stat_summary(fun.y=median, geom="line", aes(group=1)) +
  facet_grid(.~supp) + stat_summary(fun.y=median, geom="point") +
  labs(title="Boxplot 1 : Tooth length by vitamin C dosage by 2 delivery methods",
     x="Dosage (mg per day)", y="Tooth Length")
```



Observations :

- From Boxplot 1, we can see a general increase in tooth length with increase in dosage.
- OJ is a better method at low-mid level dosage, 0.5mg/day & 1.0mg/day.
- At high level dosage, both supplements demonstrate similar tooth length.

Hypothesis Testings

```
levels(TG$supp) <- c("OJ", "VC")
ci95 <- list()
ci95$OJ05 <- round(t.test(TG$len[TG$supp=="OJ" & TG$dose==0.5])$conf.int,2)
ci95$OJ10 <- round(t.test(TG$len[TG$supp=="OJ" & TG$dose==1.0])$conf.int,2)
ci95$OJ20 <- round(t.test(TG$len[TG$supp=="OJ" & TG$dose==2.0])$conf.int,2)
ci95$VC05 <- round(t.test(TG$len[TG$supp=="VC" & TG$dose==0.5])$conf.int,2)
ci95$VC10 <- round(t.test(TG$len[TG$supp=="VC" & TG$dose==1.0])$conf.int,2)
ci95$VC20 <- round(t.test(TG$len[TG$supp=="VC" & TG$dose==2.0])$conf.int,2)

ci95 <- t(as.data.frame(ci95))
colnames(ci95) <- c("2.5%", "97.5%")
rownames(ci95) <- c("Orange Juice(OJ) & 0.5mg/day",
                    "Orange Juice(OJ) & 1.0mg/day",
                    "Orange Juice(OJ) & 2.0mg/day",
                    "Ascorbic Acid(VC) & 0.5mg/day",
                    "Ascorbic Acid(VC) & 1.0mg/day",
                    "Ascorbic Acid(VC) & 2.0mg/day")
message("Table 95% Confidence Intervals for the\n
        tooth length at different supplement and dosage level")
```

```
## Table 95% Confidence Intervals for the
##
##         tooth length at different supplement and dosage level
```

```
ci95
```

```
##                               2.5% 97.5%
## Orange Juice(OJ) & 0.5mg/day  10.04 16.42
## Orange Juice(OJ) & 1.0mg/day  19.90 25.50
## Orange Juice(OJ) & 2.0mg/day  24.16 27.96
## Ascorbic Acid(VC) & 0.5mg/day  6.02  9.94
## Ascorbic Acid(VC) & 1.0mg/day 14.97 18.57
## Ascorbic Acid(VC) & 2.0mg/day 22.71 29.57
```

Observations :

- No overlapping of confidence levels at low level dosage in both OJ & VC supplements.
- No overlapping at all levels in VC supplements.
- There is overlapping at 1.0mg/day & 2.0mg/day in OJ supplements.

P-Values

```
pvalue <- list()
pvalue$OJ05v10 <- t.test(TG$len[TG$supp=="OJ" & TG$dose==0.5],
                         TG$len[TG$supp=="OJ" & TG$dose==1.0])$p.value
pvalue$OJ10v20 <- t.test(TG$len[TG$supp=="OJ" & TG$dose==1.0],
                         TG$len[TG$supp=="OJ" & TG$dose==2.0])$p.value
pvalue$VC05v10 <- t.test(TG$len[TG$supp=="VC" & TG$dose==0.5],
                         TG$len[TG$supp=="VC" & TG$dose==1.0])$p.value
pvalue$VC10v20 <- t.test(TG$len[TG$supp=="VC" & TG$dose==1.0],
                         TG$len[TG$supp=="VC" & TG$dose==2.0])$p.value
```

```
pvalue$OJ05VC05 <- t.test(TG$len[TG$supp=="OJ" & TG$dose==0.5],
                          TG$len[TG$supp=="VC" & TG$dose==0.5])$p.value
pvalue$OJ10VC10 <- t.test(TG$len[TG$supp=="OJ" & TG$dose==1.0],
                          TG$len[TG$supp=="VC" & TG$dose==1.0])$p.value
pvalue$OJ20VC20 <- t.test(TG$len[TG$supp=="OJ" & TG$dose==2.0],
                          TG$len[TG$supp=="VC" & TG$dose==2.0])$p.value


pvalue <- t(as.data.frame(pvalue))
colnames(pvalue) <- c("P-Value")
rownames(pvalue) <- c("Orange Juice(OJ) & Dosage 0.5mg/day Vs 1.0mg/day",
                      "Orange Juice(OJ) & Dosage 1.0mg/day Vs 2.0mg/day",
                      "Ascorbic Acid(VC) & Dosage 0.5mg/day Vs 1.0mg/day",
                      "Ascorbic Acid(VC) & Dosage 1.0mg/day Vs 2.0mg/day",
                      "Orange Juice(OJ) Vs Ascorbic Acid(VC), Dosage 0.5mg/day",
                      "Orange Juice(OJ) Vs Ascorbic Acid(VC), Dosage 1.0mg/day",
                      "Orange Juice(OJ) Vs Ascorbic Acid(VC), Dosage 2.0mg/day")
message(" P-Values for 2 samples T-test\n Tooth length for each dosage and supplment combination")
```

```
##  P-Values for 2 samples T-test
##  Tooth length for each dosage and supplment combination
```

```
pvalue
```

```
##                                                           P-Value
## Orange Juice(OJ) & Dosage 0.5mg/day Vs 1.0mg/day        8.784919e-05
## Orange Juice(OJ) & Dosage 1.0mg/day Vs 2.0mg/day        3.919514e-02
## Ascorbic Acid(VC) & Dosage 0.5mg/day Vs 1.0mg/day       6.811018e-07
## Ascorbic Acid(VC) & Dosage 1.0mg/day Vs 2.0mg/day       9.155603e-05
## Orange Juice(OJ) Vs Ascorbic Acid(VC), Dosage 0.5mg/day 6.358607e-03
## Orange Juice(OJ) Vs Ascorbic Acid(VC), Dosage 1.0mg/day 1.038376e-03
## Orange Juice(OJ) Vs Ascorbic Acid(VC), Dosage 2.0mg/day 9.638516e-01
```

Observations :

- The p-values align with our observations above that it fails to reject the null hypothesis that the mean difference of the tooth length Orange Juice(OJ) at 1.0mg/day and 2.0mg/day is the same.
- The p-values align with our observations above that it fails to reject the null hypothesis that the mean difference of the tooth length Orange Juice(OJ) and Ascorbic Acid(VC) at 2.0mg/day is the same.
- All other pvalues are below 0.05 threshold and hence corresponding null hypothesis of having the same means can be rejected.

Conclusion & Assumptions :

- Tooth lengths are t distributed with not equal variance among different combinations.
- The the tooth length generally increases as the increase in dosage level of Vitamin C per day.
- For Orange Juice(OJ), the impact is higher at low-mid level dosage, 0.5 & 1.0 mg/day, than Ascorbic Acid(VC).
- Both supplements demonstrate similar impact at high level, 2.0mg/day. Means of difference is not significant.
- The impact of dosage at 1.0 & 2.0 mg/day of Orange Juice(OJ) are similar. Means of difference is not significant.