Analyse ToothGrowth Data

Dexter Wang

22 May 2016

Synopsis

This is a practice of basic inferential analysis and hypothesis testing.

The data used 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 or ascorbic acid (a form of vitamin C and coded as VC).

Explore the data set

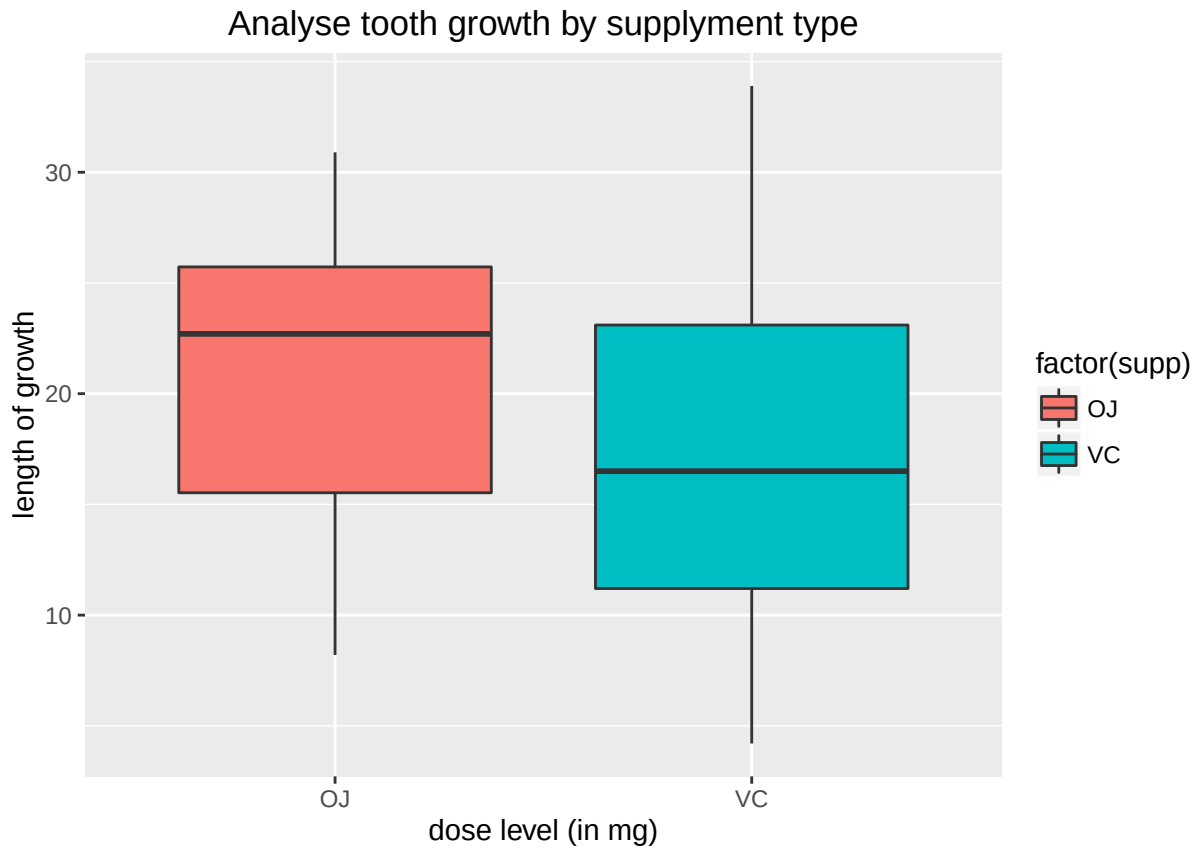
```
data(ToothGrowth)
head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

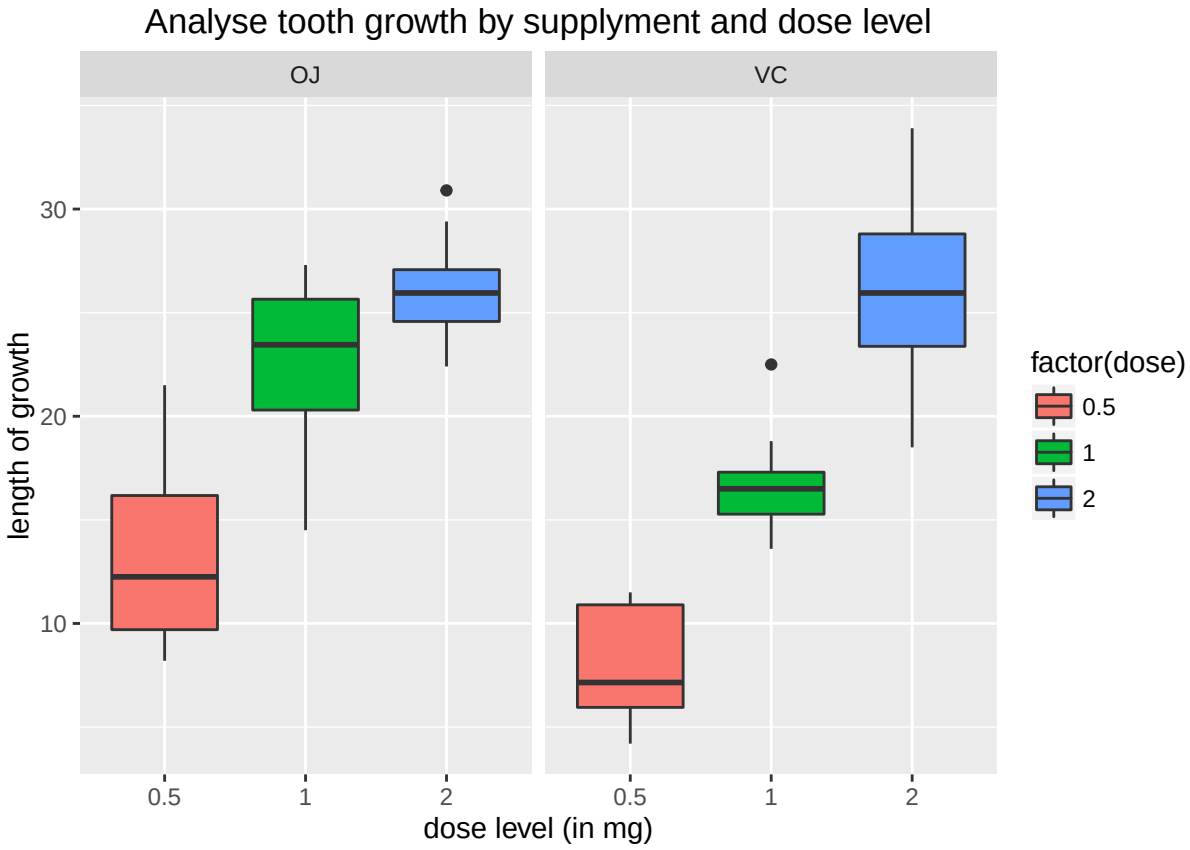
```
summary(ToothGrowth)
```

```
##       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
```

```
library(ggplot2)
title <- "Analyse tooth growth by supplyment type"
g <- ggplot(ToothGrowth, aes(x = factor(supp), y = len, fill = factor(supp))) +
    geom_boxplot() + labs(title = title, x = "dose level (in mg)",
    y = "length of growth")
g
```




```
title <- "Analyse tooth growth by supplyment and dose level"
g2 <- ggplot(ToothGrowth, aes(x = factor(dose), y = len, fill = factor(dose))) +
    geom_boxplot() + facet_grid(. ~ supp) + labs(title = title, x = "dose level (in mg)",
    y = "length of growth")
g2
```



From the graph, it can be seen that the tooth has higher rate of growth as dose level increases. The "OJ" treatment generally has more significant effect than "VC" to boost tooth growth. The exception occurs when giving 2.0 dose, where the mean growth is similar between the two but the "VC" treatment observes higher max growth.

Hypothesis testing

In this section, we try to find if dose level or treatment type has effect on tooth growth. We run several T test and capture p-value and confidence interval for each type of treatment overall and then between each corresponding dose level respectively.

The assumption is the variance between each group is not equal (var.equal=FALSE) in the T tests. The null hypothesis is that the effect on tooth growth in each group has no different than the other.

```
data <- ToothGrowth
data0.5 <- data[data$dose == 0.5, ]
data1 <- data[data$dose == 1, ]
data2 <- data[data$dose == 2, ]

t_supp <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = data)
t_0.5 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = data0.5)
t_1 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = data1)
```

```r
t_2 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = data2)

result <- data.frame(Test = c("OJ vs VC", "0.5 dose", "1 dose", "2 dose"),
    p.value = c(t_supp$p.value, t_0.5$p.value, t_1$p.value, t_2$p.value),
    ci.lower = c(t_supp$conf[1], t_0.5$conf[1], t_1$conf[1], t_2$conf[1]),
    ci.upper = c(t_supp$conf[2], t_0.5$conf[2], t_1$conf[2], t_2$conf[2]),
    `Mean OJ` = c(t_supp$estimate[1], t_0.5$estimate[1], t_1$estimate[1],
        t_2$estimate[1]), `Mean VC` = c(t_supp$estimate[2], t_0.5$estimate[2],
        t_1$estimate[2], t_2$estimate[2]))


result
```

```
##       Test     p.value   ci.lower ci.upper  Mean.OJ  Mean.VC
## 1 OJ vs VC 0.060634508 -0.1710156 7.571016 20.66333 16.96333
## 2 0.5 dose 0.006358607  1.7190573 8.780943 13.23000  7.98000
## 3   1 dose 0.001038376  2.8021482 9.057852 22.70000 16.77000
## 4   2 dose 0.963851589 -3.7980705 3.638070 26.06000 26.14000
```

From the result, it can be seen that for dose 0.5 and 1, treatment "OJ" has greater effect on tooth growth than "VC". Since the p values for these two tests are both less than the 0.05 bench mark. For dose 2, the null hypothesis cannot be rejected as the p value is greater than 0.05 and the confidence interval contains 0. Same to the test of "OJ vs VC", the null hypothesis cannot be rejected too.

To conclude, treatment "OJ" has greater effect on teeth growth than "VC" when giving low dose level between 0.5 and 1. It is uncertain which treatment has better effect on high dose level 2.0.

Thanks for viewing :)