

# Statistical Inference Course Project Part 2

*Koji*

*2018/7/20*

## Overview

Analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Provide a basic summary of the data.
3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)
4. State your conclusions and the assumptions needed for your conclusions.

## Analysis

### ToothGrowth data

```
# Load the ToothGrowth data
data("ToothGrowth")
str(ToothGrowth)

## '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 ...
?ToothGrowth
```

### Description

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

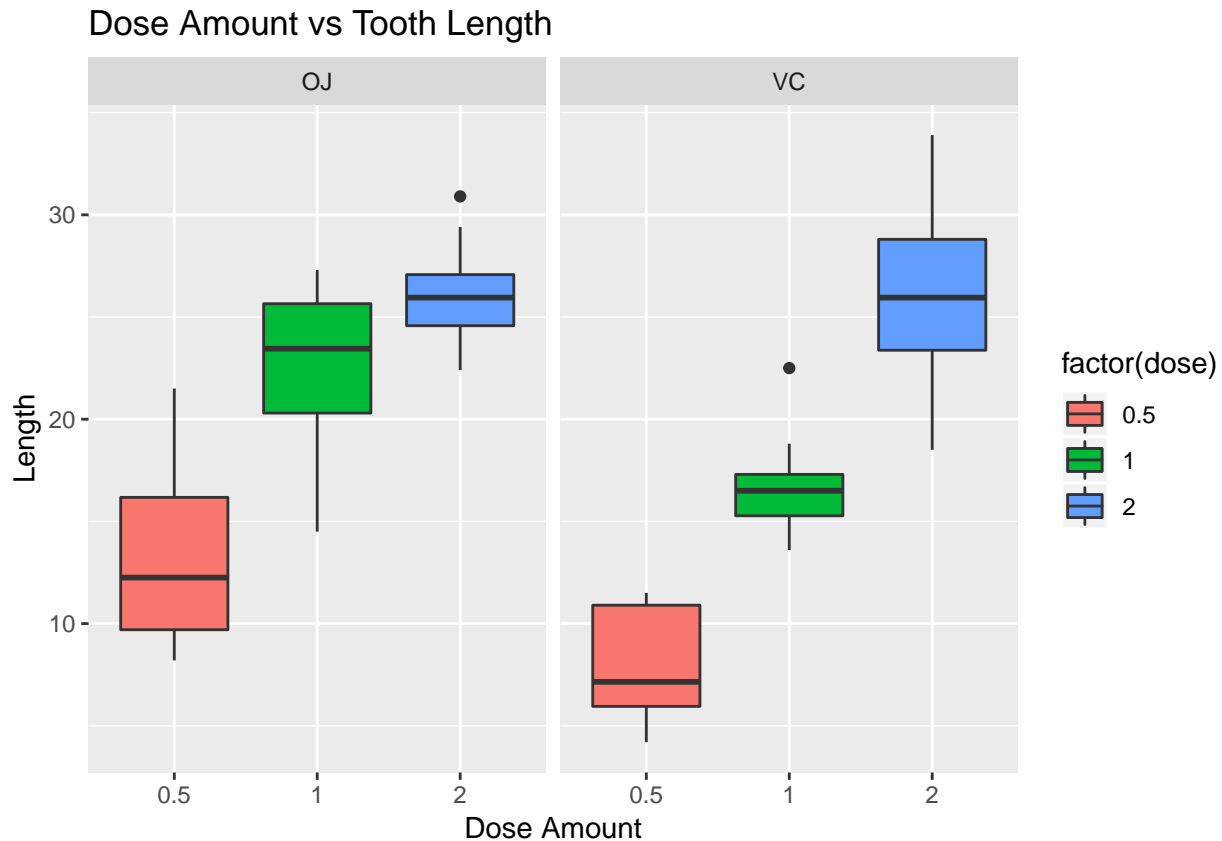
### Basic summary

```
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

g <- ggplot(ToothGrowth, aes(x = factor(dose), y = len))
g <- g + geom_boxplot(aes(fill = factor(dose)))
g <- g + facet_grid(. ~ supp)
```

```
g <- g + ggtitle("Dose Amount vs Tooth Length")
g <- g + xlab("Dose Amount") + ylab("Length")
g
```



## Hypothesis Test

Now we will compare tooth growth by supplement using a t-test.

```
t.test(len ~ supp, ToothGrowth)
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.1710156 7.5710156
## sample estimates:
## mean in group OJ mean in group VC
## 20.66333 16.96333
```

The p-value of this test was 0.06. Since the p-value is greater than 0.05 and the confidence interval of the test contains zero we can say that supplement types seems to have no impact on Tooth growth based on this test.

Now we'll compare tooth growth by dose, looking at the different pairs of dose values.

```
t.test(len ~ dose, ToothGrowth[ToothGrowth$dose %in% c(0.5, 1), ])
```

```
##
## Welch Two Sample t-test
##
## data: len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -11.983781 -6.276219
## sample estimates:
## mean in group 0.5 mean in group 1
## 10.605 19.735
```

```
t.test(len ~ dose, ToothGrowth[ToothGrowth$dose %in% c(1, 2), ])
```

```
##
## Welch Two Sample t-test
##
## data: len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
## 19.735 26.100
```

As can be seen, the p-value of each test was essentially zero and the confidence interval of each test does not cross over zero.

Based on this result we can assume that the average tooth length increases with an increasing dose, and therefore the null hypothesis can be rejected.

## Conclusions and Assumptions

In reviewing our t-test analysis from above, we can conclude that supplement type has no effect on tooth growth, and increasing the dose level leads to increased tooth growth..