# Statistical Inference Project Part 2

*Li Xiaowei*

# Overview

This report is the part 2 of the statistical inference project. We'll analyze the ToothGrowth data in R datasets.

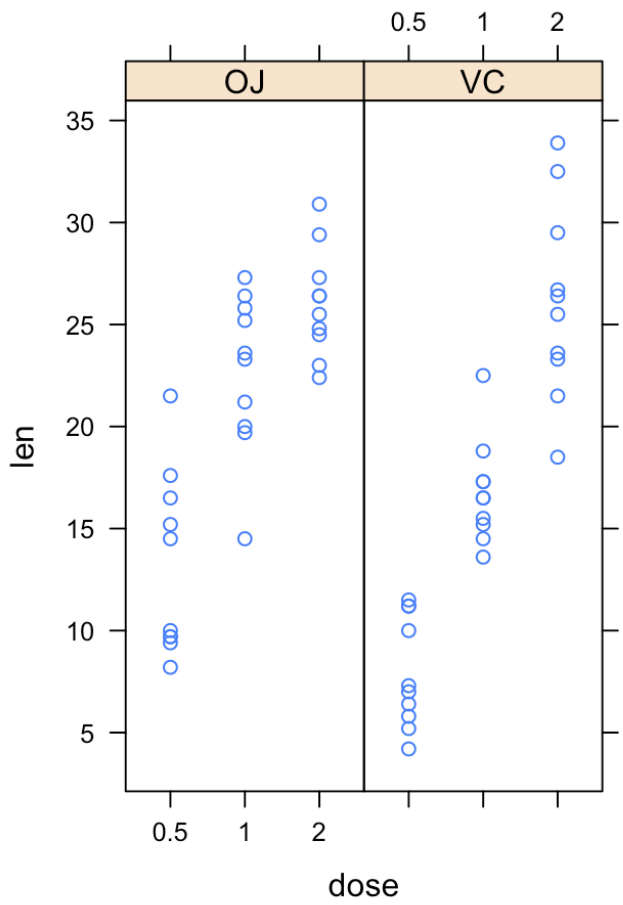
# Analysis of ToothGrowth

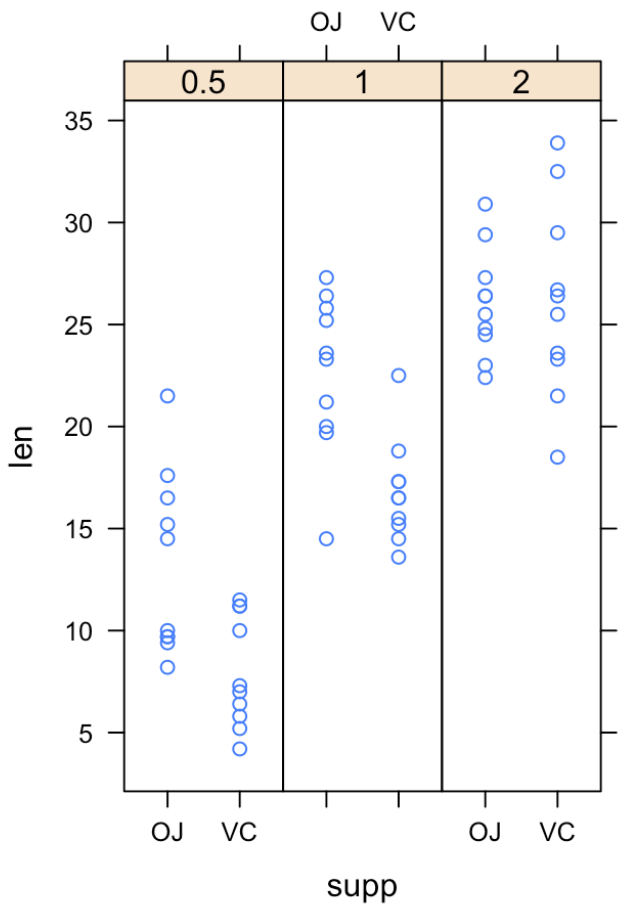
"ToothGrowth" dataset in R is a data frame with 60 observations on 3 variables. It's the length of teeth in each of 10 pigs at each of 3 dose levels of Vitamin C with 2 delivery methods(orange juice or ascorbic acid). Blow are 2 plots which addressing below 2 observations.

1. For the same delivery method, higher dose level is likely to be more effective than lower dose level.
2. For the same dose level, OJ(orange juice) is likely to be more effective than VC(ascorbic acid). It's quite obvious for dose level 0.5 and 1, but dose level 2 may tell a different story.

```
data("ToothGrowth")
library(lattice)
library(ggplot2)

tooth <- transform(ToothGrowth, dose = factor(dose))
plot1 <- xyplot(len ~ dose | supp, data = tooth, layout = c(2,1))
plot2 <- xyplot(len ~ supp | dose, data = tooth, layout = c(3,1))
print(plot1, position = c(0,0,0.5,1), more = TRUE)
print(plot2, position = c(0.5,0,1,1))
```




# Group the data

According to the documentation of ToothGrowth, the 60 observations come from 10 different pigs in 6 different testing scenarios. Hence, all the confidence intervals need to set the “paired” to “TRUE”.

There are 3 groups for each of the 2 delivery methods. 6 vectors created by below code.

```
gVC1 <- tooth[which(tooth$supp == "VC" & tooth$dose == 0.5), 1]
gVC2 <- tooth[which(tooth$supp == "VC" & tooth$dose == 1), 1]
gVC3 <- tooth[which(tooth$supp == "VC" & tooth$dose == 2), 1]
gOJ1 <- tooth[which(tooth$supp == "OJ" & tooth$dose == 0.5), 1]
gOJ2 <- tooth[which(tooth$supp == "OJ" & tooth$dose == 1), 1]
gOJ3 <- tooth[which(tooth$supp == "OJ" & tooth$dose == 2), 1]
```

# Confidence interval of the same delivery method

Let’s first explore the confidence interval of the same delivery method with different dose levels. We’ll compare dose levels of 0.5 to 1, and 1 to 2. The confidence levels are of 0.95 by default.

VC, dose level 0.5 vs 1

```
t.test(gVC1, gVC2, paired = TRUE)$conf[1:2]
```

```
## [1] -12.030399  -5.549601
```

VC, dose level 1 vs 2

```
t.test(gVC2, gVC3, paired = TRUE)$conf[1:2]
```

```
## [1] -13.334918  -5.405082
```

OJ, dose level 0.5 vs 1

```
t.test(gOJ1, gOJ2, paired = TRUE)$conf[1:2]
```

```
## [1] -14.615384  -4.324616
```

OJ, dose level 1 vs 2

```
t.test(gOJ2, gOJ3, paired = TRUE)$conf[1:2]
```

```
## [1] -7.2709376  0.5509376
```

The results are quite consistant that for the same delivery method a higher dose level always contributes a greater teeth length. Because the intervals are almost all below 0. This follows the observation from the earlier plot.

# Confidence interval of the same dose level

Then let's take a look at the confidence inteval of the same dose level but differenct delivery methods.

Dose level 0.5, VC vs OJ

```
t.test(gVC1, gOJ1, paired = TRUE)$conf[1:2]
```

```
## [1] -9.236542 -1.263458
```

Dose level 1, VC vs OJ

```
t.test(gVC2, gOJ2, paired = TRUE)$conf[1:2]
```

```
## [1] -9.908089 -1.951911
```

Dose level 2, VC vs OJ

```
t.test(gVC3, gOJ3, paired = TRUE)$conf[1:2]
```

```
## [1] -4.168976  4.328976
```

It's quite clear that for dose level 0.5 and 1 OJ always outperforms VC while it may not be true for dose level 2. The confidence interval of dose level 2 is almost symmetric of 0. This also follows the observation from earlier plot.