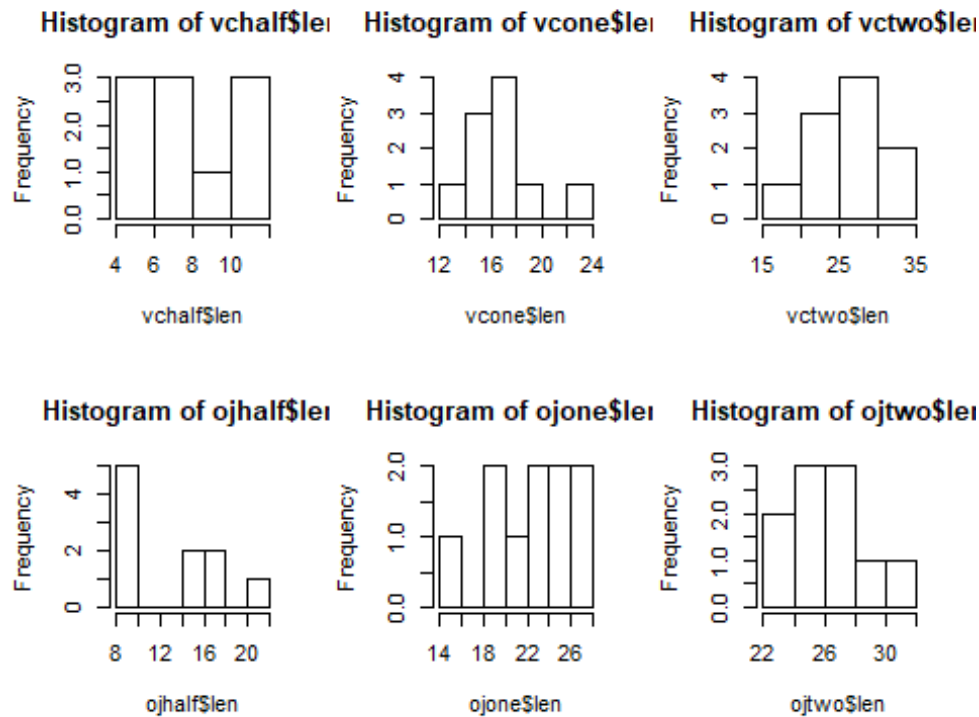# Statistical Inference Course Project Part 2

Pulan Yu

2/29/2020

Part 2: Basic Inferential Data Analysis

ToothGrowth dataset has three columns: len, supp and dos. The supp is a factor with two levels: VC and OJ. The dose has three values: 0.5, 1.0 and 2.0. The following code will do an exploeratory analysis

Supp has two values: VC and OJ respectively. The dose has five levels from 0.5,1.0,to 2.0.

```
par(mfrow=c(1,2))
vc <- ToothGrowth[ToothGrowth$supp =='VC',]
boxplot(vc$len~vc$dos,col=(c("red","gold","darkgreen")),xlab =
'dos',ylab='len',main="VC")
oj <- ToothGrowth[ToothGrowth$supp =='OJ',]

boxplot(oj$len~oj$dos,col=(c("red","gold","darkgreen")),xlab='dos',ylab='len'
,main='OJ')
```
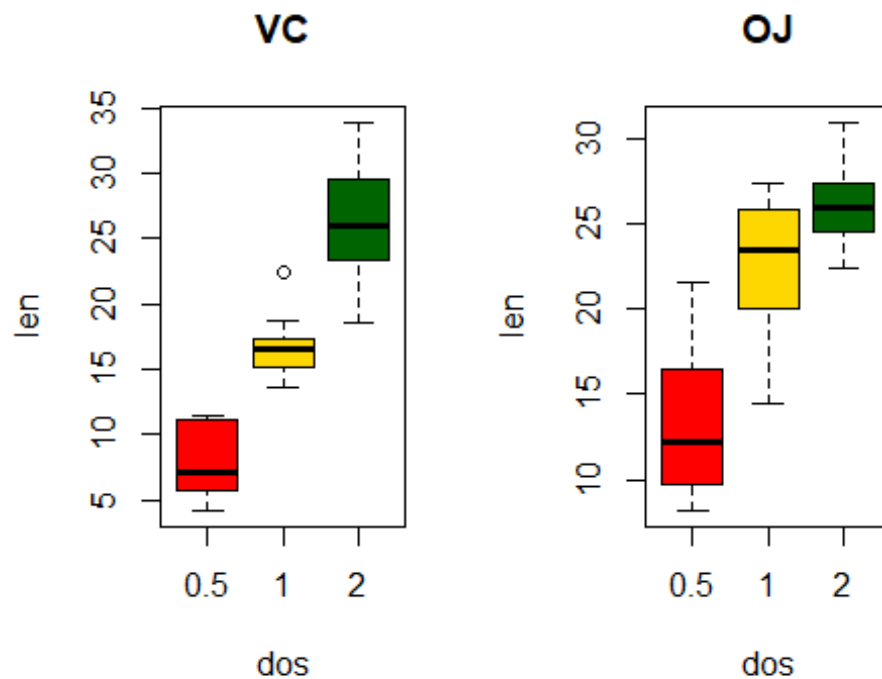
Since there are two variables: supp and dose, firstly, the effect of dose is investigated and then the effect of dose. The result format is confidence interval followed by P-Value

1. the 0.5 level of vc vs oj

```
vchalf <- vc[vc$dose==0.5,]
ojhalf <- oj[oj$dose==0.5,]
test <- t.test(vchalf$len,ojhalf$len,var.equal=TRUE)
res <- c(test$conf.int,test$p.value)
print(res)
```

```
## [1] -8.729738345 -1.770261655  0.005303661
```

2. the 1.0 level of vc vs oj

```
## [1] -9.0193080513 -2.8406919487  0.0007807262
```

3. the 2.0 level vc vs oj

```
## [1] -3.5629985  3.7229985  0.9637098
```

For VC supply

```
vchalf <- vc[vc$dose==0.5,]
vcone <- vc[vc$dose==1.0,]
vctwo <- vc[vc$dose==2.0,]
test <- t.test(vchalf$len,vcone$len,var.equal=TRUE)
res <- c(test$conf.int,test$p.value)
print('Dose 0.5 vs 1.0')
```

```
## [1] "Dose 0.5 vs 1.0"

  print(res)

## [1] -1.126435e+01 -6.315654e+00  6.492265e-07

  test <- t.test(vchalf$len,vctwo$len,var.equal=TRUE)
  res <- c(test$conf.int,test$p.value)
  print('Dose 0.5 vs 2.0')

## [1] "Dose 0.5 vs 2.0"

  print(res)

## [1] -2.183284e+01 -1.448716e+01  4.957286e-09

  test <- t.test(vctwo$len,vcone$len,var.equal=TRUE)
  res <- c(test$conf.int,test$p.value)
  print('Dose 2.0 vs 1.0')

## [1] "Dose 2.0 vs 1.0"

  print(res)

## [1] 5.771040e+00 1.296896e+01 3.397578e-05
```

For OJ supply

```
## [1] "Dose 0.5 vs 1.0"

## [1] -1.341081e+01 -5.529186e+00  8.357559e-05

## [1] "Dose 0.5 vs 2.0"

## [1] -1.627822e+01 -9.381777e+00  3.401859e-07

## [1] "Dose 2.0 vs 1.0"

## [1] 0.2194983 6.5005017 0.0373628
```

Conclusions:

For both OJ and VC supply, the level of difference has effect on the tooth length since the p-value of all t-test are less then 0.05 at 95% confidence level. The null hypothesis is rejected, the level has effect on the length and higher does is better for tooth growth.

For the same level, at low dose level 0.5 and 1.0, the supply has effect on the tooth length and OJ is better for tooth growth. Since the p-value of 0.5 and 1.0 level are 0.005304 and 0.0007807 respectively, so the null hypothesis is rejected and supply has impact on the length. For 2.0 level, the p-value is 0.9637, the null hypothese is failed to be rejected at 95% confidence level. It suggests that at high dose, the supply may not have statistical significant impact on tooth length.