Tooth Growth

Mao Soldevilla

Overview

In this project we can analyze the ToothGrowth data set included in R package

```
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 ...
```

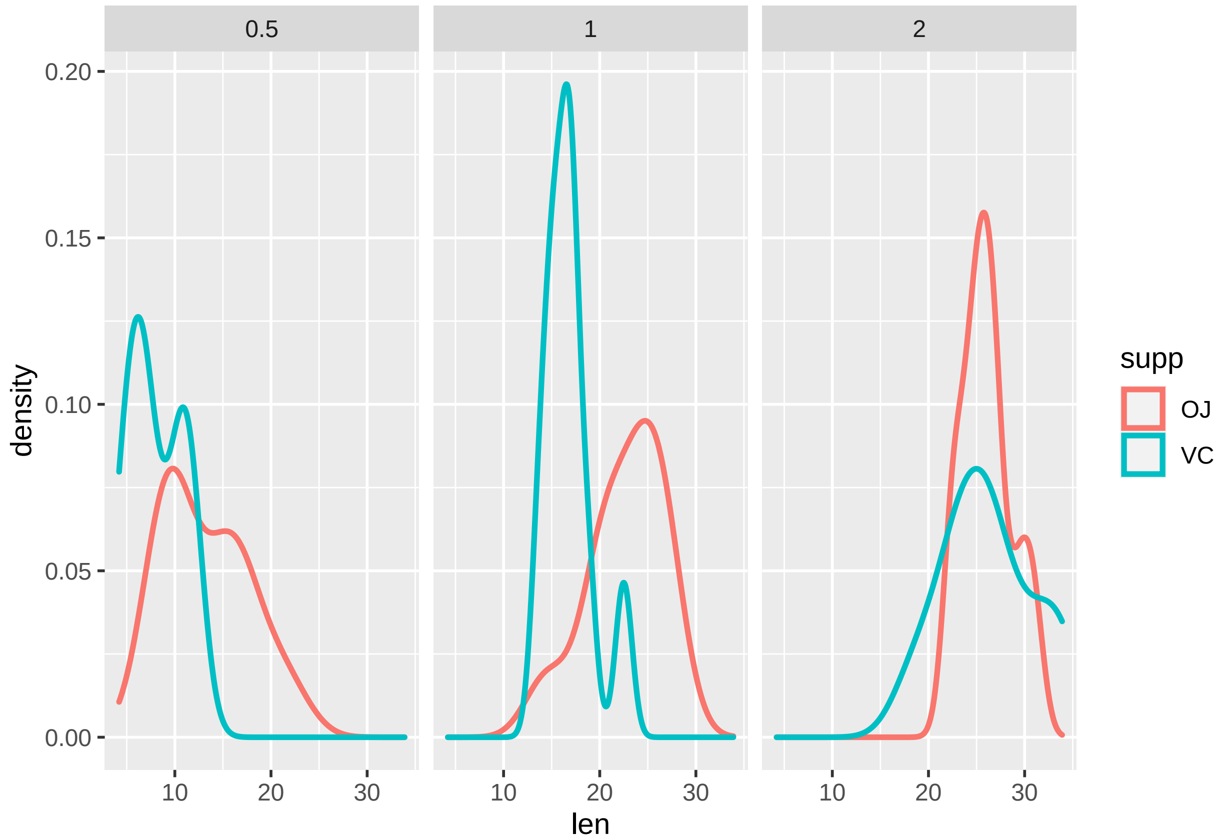
According our visualization our data frame have sixty obs of three variables. After that, we will check with summary what kind of data is in.

```
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
```

Plots for analyse

```
d <- ToothGrowth
g <- ggplot(d, aes(len, color = supp))
g <- g + facet_grid(.~dose)
g <- g + geom_density(size = 1)
print(g)
```



As we can see there are some differences in 0.5 and 1 doses instead of 2.

Now we see the data included in each plot.

Dose: 0.5

Analyzing the dose = 0.5

```
g1 <- d[d$supp == "OJ" & d$dose == 0.5, ]$len
g2 <- d[d$supp == "VC" & d$dose == 0.5, ]$len
difference <- g2 - g1
mn <- mean(difference)
s <- sd(difference)
n <- length(difference)
mn + c(-1, 1) * qt(.975, n - 1) * s / sqrt(n)
```

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

The T interval for dose = 0.5

Making a T test

```
t.test(difference)
```

```
##
##  One Sample t-test
```

```
##
## data:  difference
## t = -2.9791, df = 9, p-value = 0.01547
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -9.236542 -1.263458
## sample estimates:
## mean of x
##     -5.25
```

The P-value is 0.015472 which is < 0.05, and that means that the dosage supply is *acceptable*

Dose: 1

Analyzing the dose = 1

```r
g1 <- d[d$supp == "OJ" & d$dose == 1, ]$len
g2 <- d[d$supp == "VC" & d$dose == 1, ]$len
difference <- g2 - g1
mn <- mean(difference)
s <- sd(difference)
n <- length(difference)
mn + c(-1, 1) * qt(.975, n - 1) * s / sqrt(n)
```

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

The T interval for dose = 1

Making a T test

```r
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = -3.3721, df = 9, p-value = 0.008229
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -9.908089 -1.951911
## sample estimates:
## mean of x
##     -5.93
```

The P-value is 0.0082292 which is < 0.05, and that means that the dosage supply is *acceptable*

Dose: 2

Analyzing the dose = 2

```r
g1 <- d[d$supp == "OJ" & d$dose == 2, ]$len
g2 <- d[d$supp == "VC" & d$dose == 2, ]$len
difference <- g2 - g1
mn <- mean(difference)
s <- sd(difference)
n <- length(difference)
mn + c(-1, 1) * qt(.975, n - 1) * s / sqrt(n)
```

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

The T interval for dose = 2

Making a T test

```r
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 0.042592, df = 9, p-value = 0.967
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -4.168976  4.328976
## sample estimates:
## mean of x
##      0.08
```

The P-value is 0.9669567 which is > 0.05, and that means that the dosage supply is *not acceptable*