The ToothGrowth data Analysis

Kiattisak Chaisomboon

07/06/2020

Overview

We are going to analyze the ToothGrowth data in the R datasets package. We will study the change of the tooth length in pigs each type of supplement and changing dose quantity.

Exploratory Data Analysis

First, loading the dataset and check its dimension

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
tg <- ToothGrowth
dim(tg)
```

```
## [1] 60  3
```

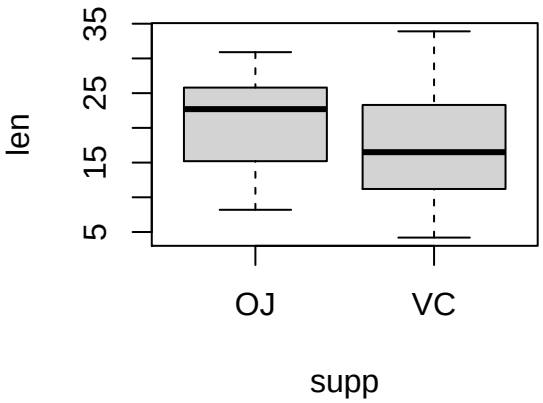
The dataset contains 60 rows and 3 columns

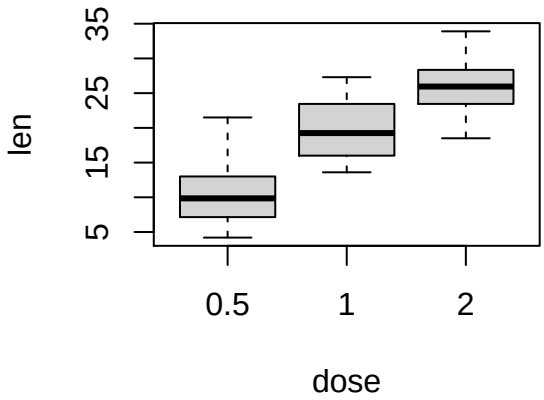
Next, summarizing the basic data from the dataset

```
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
```

```
par(mfrow=c(1,2))
boxplot(len ~ supp, data=tg)
boxplot(len ~ dose, data=tg)
```





It's seem giving orange joice is better than an acid and give more dose is better

Inferential Data Analysis

Create variables each type of supplement

```
OJ <- subset(tg, supp == 'OJ')
VC <- subset(tg, supp == 'VC')
```

Next,

```
t.test(OJ$len, VC$len, alternative='greater', paired = FALSE, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  OJ$len and VC$len
## t = 1.9153, df = 55.309, p-value = 0.03032
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4682687       Inf
## sample estimates:
## mean of x mean of y
##  20.66333  16.96333
```

```
half.dose <- subset(tg, dose == 0.5)
one.dose <- subset(tg, dose == 1)
two.dose <- subset(tg, dose == 2)
```

One dose effect greater than Half dose

```
t.test(one.dose$len, half.dose$len, alternative='greater', paired=FALSE, var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  one.dose$len and half.dose$len
## t = 6.4766, df = 37.986, p-value = 6.342e-08
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.753323      Inf
## sample estimates:
## mean of x mean of y
##    19.735    10.605
```

Two dose effect greater than One dose

```
t.test(two.dose$len, one.dose$len, alternative='greater', paired=FALSE, var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  two.dose$len and one.dose$len
## t = 4.9005, df = 37.101, p-value = 9.532e-06
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  4.17387     Inf
## sample estimates:
## mean of x mean of y
##    26.100    19.735
```