Basic Inferential Data Analysis

Ismail HASSAN DJILAL

19 mars 2019

Part 2: Basic Inferential Data Analysis Instructions

In the second portion of the project, we're going to analyze the ToothGrowth data in the R datasets package.

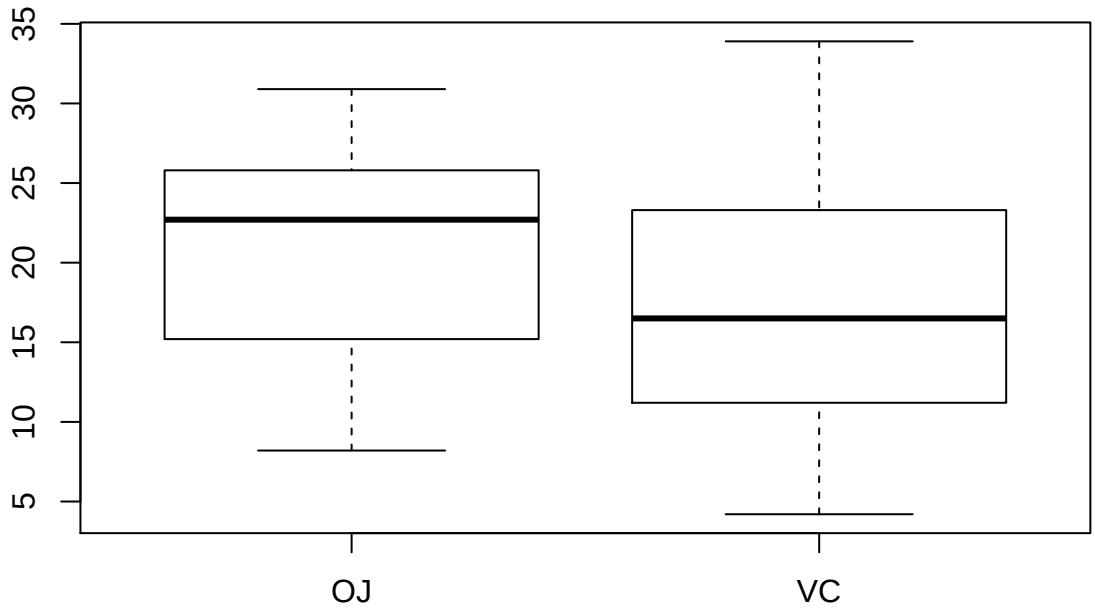
```
attach(ToothGrowth)
head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
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
```

```
boxplot(len~supp)
```



With the boxplot we can see the variation of the length of odontoblasts between guinea pigs according to the delivery methods OJ or VC .

```
library(dplyr)
```

```
## Warning: Installed Rcpp (0.12.18) different from Rcpp used to build dplyr (0.12.11).
## Please reinstall dplyr to avoid random crashes or undefined behavior.
```

```
##
## 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
```

```
t.test(len~supp, paired=FALSE, var.equal=F)
```

```
##
##  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
```

At 95% confidence level, there is no significant difference (p-value = 0.06>0.05) of the two means Furthermore 95% confidence interval ([-0.171, 7.5710156]) contain 0

Conclusion : we can not reject the null hypothese (The delivery method orange juice or ascorbic acid has no impact of the lenght of odontobasts) in favour of the alternative hypothesis

```
#subset of Toothgrowth, population of guinea pigs who received the levels of vitamin C 0.5 and 1

p<-ToothGrowth%>%filter(dose == 0.5  | dose == 1)
# t test of len according supp equal 0.5 or 1
t.test(len~dose, var.equal=F, data = p)
```

```
##
##  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
```

At 95% confidence level, there is a significant difference (p-value = 1.268e-07< 0.05) of the two means Furthermore 95% confidence interval ([-11.983781, -11.983781])

Conclusion : dose level change (0.5 and 1 dose levels) irrespective of delivery methods (OJ or VC) has an impact on the mean of len. We can reject the null hypothese.