Assignment

Ankush Sharma

17/03/2020

```r
setwd("/home/siddhant18241/Ankush/Asg5")
```

```r
library(dplyr)
library(plyr)
library(Rmisc)
```

```r
f<-read.csv("input.csv")
head(f)
```

```
##       X Patient_ID age gender SBP DBP
## 1  17          29  62      1 124  70
## 2  38          81  30      1 106  68
## 3 129         228  19      1 136  72
## 4 275         524  69      1 140 102
## 5 336         643  46      1 108  46
## 6 413         789  65      2 190  76
```

The file contains information regarding the patient age, gender, its unique id and systolic and diastolic blood pressure. The data can be summarized as:-

```r
summary(f)
```

```
##        X             Patient_ID          age            gender
##  Min.   :   17   Min.   :   29   Min.   :18.00   Min.   :1.000
##  1st Qu.:10257   1st Qu.:18822   1st Qu.:30.00   1st Qu.:1.000
##  Median :19350   Median :35482   Median :44.00   Median :2.000
##  Mean   :18596   Mean   :33609   Mean   :46.32   Mean   :1.521
##  3rd Qu.:27196   3rd Qu.:48992   3rd Qu.:62.00   3rd Qu.:2.000
##  Max.   :35232   Max.   :62003   Max.   :85.00   Max.   :2.000
##       SBP             DBP
##  Min.   : 82.0   Min.   :  0.0
##  1st Qu.:112.0   1st Qu.: 62.0
##  Median :122.0   Median : 70.0
##  Mean   :124.9   Mean   : 70.1
##  3rd Qu.:136.0   3rd Qu.: 78.0
##  Max.   :230.0   Max.   :128.0
```

Confidence Interval

The 95 % confidence interval for the overall population can be calculated for SBP as

```r
CI(f$SBP, ci=0.95)
```

```
##    upper     mean    lower
## 126.1807 124.9220 123.6633
```

The 95 % confidence interval for the overall population can be calculated for DBP as

```r
CI(f$DBP, ci=0.95)
```

```
##    upper     mean    lower
## 71.01097 70.09600 69.18103
```

Test

```r
male<-f[which(f$gender==1),]
female<-f[which(f$gender==2),]
```

Comparing SBP for male and female

Null hypothesis: No significant difference in the blood pressure between male and female Applying 2 sided t test

```r
t.test(male$SBP,female$SBP)
```

```
##
##  Welch Two Sample t-test
##
## data:  male$SBP and female$SBP
## t = 3.726, df = 991.73, p-value = 0.0002055
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.234984 7.208570
## sample estimates:
## mean of x mean of y
##  127.3820  122.6603
```

The result conclude that the SBP is significntly different between male and female

Comparing DBP for male and female

Null hypothesis: No significant difference in the blood pressure between male and female Applying 2 sided t test

```r
t.test(male$DBP,female$DBP)
```

```
##
##  Welch Two Sample t-test
##
## data:  male$DBP and female$DBP
## t = 6.9495, df = 983.72, p-value = 6.665e-12
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   4.554939 8.139582
## sample estimates:
## mean of x mean of y
##  73.40292  67.05566
```

The result conclude that the DBP is significntly different between male and female

Comparing BP difference in male >60 and <=60

Creating datasets

```
male_older_60<-male[which(male$age>60),]
male_younger_60<-male[which(male$age<=60),]
female_older_60<-female[which(female$age>60),]
female_younger_60<-female[which(female$age<=60),]
```

SBP in male

Null hypothesis: No significant difference in the blood pressure between male greater than 60 and younger

```
t.test(male_older_60$SBP,male_younger_60$SBP)
```

```
##
##  Welch Two Sample t-test
##
## data:  male_older_60$SBP and male_younger_60$SBP
## t = 6.0237, df = 164.36, p-value = 1.082e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##    8.871967 17.524326
## sample estimates:
## mean of x mean of y
##  137.1360  123.9379
```

result: significant difference in the blood pressure between male greater than 60 and younger

DBP in male

Null hypothesis: No significant difference in the blood pressure between male greater than 60 and younger

```
t.test(male_older_60$DBP,male_younger_60$DBP)
```

```
##
##  Welch Two Sample t-test
##
## data:  male_older_60$DBP and male_younger_60$DBP
## t = -2.385, df = 177.74, p-value = 0.01813
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -7.5044493 -0.7087032
## sample estimates:
## mean of x mean of y
##  70.36800  74.47458
```

result: significant difference in the blood pressure between male greater than 60 and younger

SBP in female

Null hypothesis: No significant difference in the blood pressure between female greater than 60 and younger

```
t.test(female_older_60$SBP,female_younger_60$SBP)
```

```
##
##  Welch Two Sample t-test
##
## data:  female_older_60$SBP and female_younger_60$SBP
## t = 11.646, df = 198.23, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  20.85493 29.35764
## sample estimates:
## mean of x mean of y
##  140.4901  115.3838
```

result: significant difference in the blood pressure between female greater than 60 and younger

DBP in female

Null hypothesis: No significant difference in the blood pressure between female greater than 60 and younger

```
t.test(female_older_60$DBP,female_younger_60$DBP)
```

```
##
##  Welch Two Sample t-test
##
## data:  female_older_60$DBP and female_younger_60$DBP
## t = -2.9104, df = 206.43, p-value = 0.004006
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
##  -7.780546 -1.496275
## sample estimates:
## mean of x mean of y
##  63.76159  68.40000
```

result: significant difference in the blood pressure between female greater than 60 and younger