MATH1324 Assignment 1

Code

Modeling Body Measurements

Student Details

Yogesh Haresh Bojja (s3789918)

Problem Statement

Aim of this assignment is to visualize empirical data and normal distribution of BIACROMIAL DIAMETER from - “Heinz G, Peterson LJ, Johnson RW, Kerk CJ. 2003. Exploring Relationships in Body Dimensions. Journal of Statistics Education 11(2)” based on gender and determine whether the normal distribution fits the empirical data of the selected attribute.

Load Packages

Hide

```
# This is a chunk where you can load the necessary packages required to reproduce the
report
library(dplyr)
library(utils)
```

Data

Hide

```
# This is a chunk for your Data section.
df <- read.csv('bdims.csv')
df$sex <- factor(df$sex, levels = c(0,1), labels = c('Female', 'Male'))
t <- table(df$sex)
print(t)
```

```

Female   Male
   260    247
```

Hide

```
male_df = df %>% filter(df$sex=='Male')
female_df = filter(df, df$sex=='Female')
```

Summary Statistics

Hide

```
# This is a chunk for your Summary Statistics section.
a <- summary(female_df$bia.di)
print(paste('Minimum(Female) : ',a['Min.']))
```

```
[1] "Minimum(Female) :  32.4"
```

Hide

```
print(paste('Maximum(Female) : ',a['Max.']))
```

```
[1] "Maximum(Female) :  42.6"
```

Hide

```
print(paste('1st Quartile(Female) : ',a['1st Qu.']))
```

```
[1] "1st Quartile(Female) :  35.175"
```

Hide

```
print(paste('3rd Quartile(Female) : ',a['3rd Qu.']))
```

```
[1] "3rd Quartile(Female) :  37.8"
```

Hide

```
print(paste('Mean(Female) : ',round(a['Mean'], 2)))
```

```
[1] "Mean(Female) :  36.5"
```

Hide

```
print(paste('Median(Female) : ',a['Median']))
```

```
[1] "Median(Female) :  36.4"
```

Hide

```
print(paste('InterQuartile Range(Female) : ',round(IQR(female_df$bia.di), 2)))
```

```
[1] "InterQuartile Range(Female) :  2.62"
```

Hide

```
print(paste('Standard deviation(Female) : ',round(sd(female_df$bia.di), 2)))
```

```
[1] "Standard deviation(Female) :  1.78"
```

Hide

```
print(paste('Variance(Female) : ',round(var(female_df$bia.di), 2)))
```

```
[1] "Variance(Female) :  3.17"
```

Hide

```
a <- summary(male_df$bia.di)
print(paste('Minimum(Male) : ',a['Min.']))
```

```
[1] "Minimum(Male) :  34.1"
```

Hide

```
print(paste('Maximum(Male) : ',a['Max.']))
```

```
[1] "Maximum(Male) :  47.4"
```

Hide

```
print(paste('1st Quartile(Male) : ',a['1st Qu.']))
```

```
[1] "1st Quartile(Male) :  40"
```

Hide

```
print(paste('3rd Quartile(Male) : ',a['3rd Qu.']))
```

```
[1] "3rd Quartile(Male) :  42.6"
```

Hide

```
print(paste('Mean(Male) : ',round(a['Mean'], 2)))
```

```
[1] "Mean(Male) :  41.24"
```

Hide

```
print(paste('Median(Male) : ',a['Median']))
```

```
[1] "Median(Male) :  41.2"
```

Hide

```
print(paste('InterQuartile Range(Male) : ',round(IQR(male_df$bia.di), 2)))
```

```
[1] "InterQuartile Range(Male) :  2.6"
```

Hide

```
print(paste('Standard deviation(Male) : ',round(sd(male_df$bia.di), 2)))
```

```
[1] "Standard deviation(Male) :  2.09"
```

Hide

```
print(paste('Variance(Male) : ',round(var(male_df$bia.di), 2)))
```

```
[1] "Variance(Male) :  4.36"
```

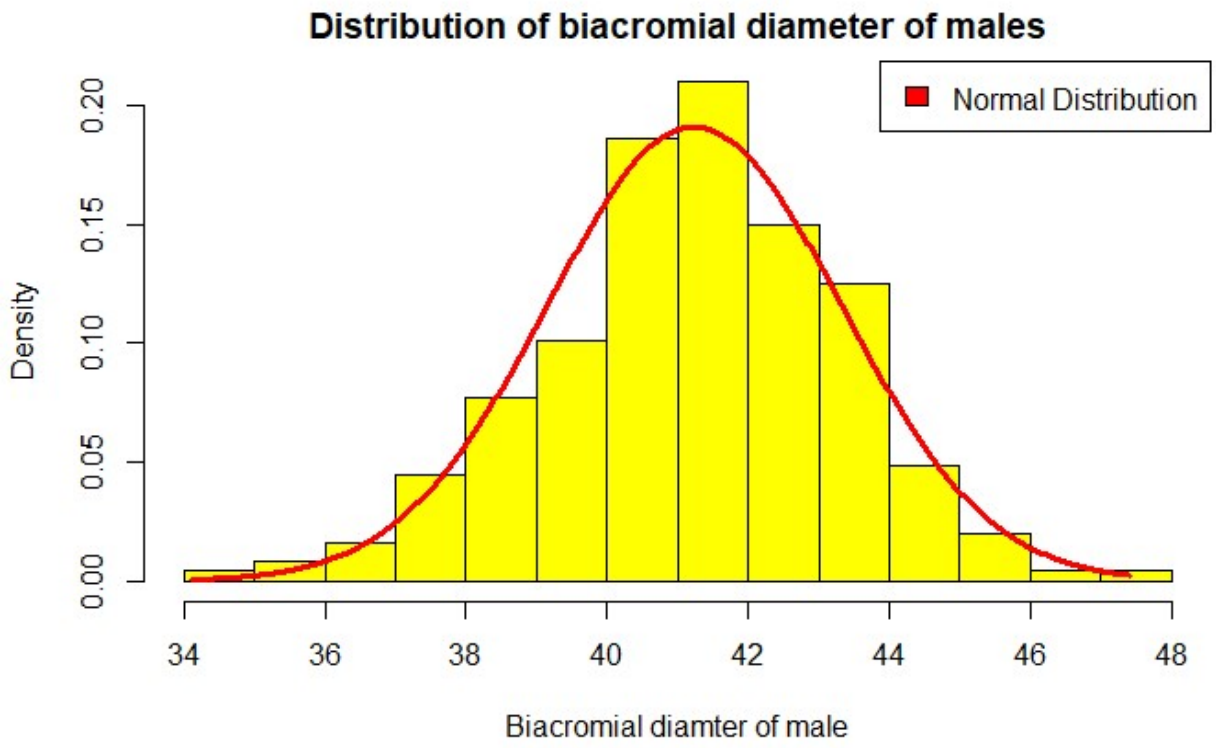
Distribution Fitting

Hide

```
# This is a chunk for your Distribution Fitting section.
hist(male_df$bia.di, col = 'yellow', xlab = 'Biacromial diamter of male', main = 'Dist
ribution of biacromial diameter of males', probability = TRUE, breaks = 15 )
x <- seq(from = min(male_df$bia.di), to = max(male_df$bia.di), length.out = 500)
y <- dnorm(x, mean = mean(male_df$bia.di), sd = sd(male_df$bia.di))
lines(x, y, col="red", lwd=3)
```

Hide

```
legend("topright", c("Normal Distribution"), fill=c("red"))
```

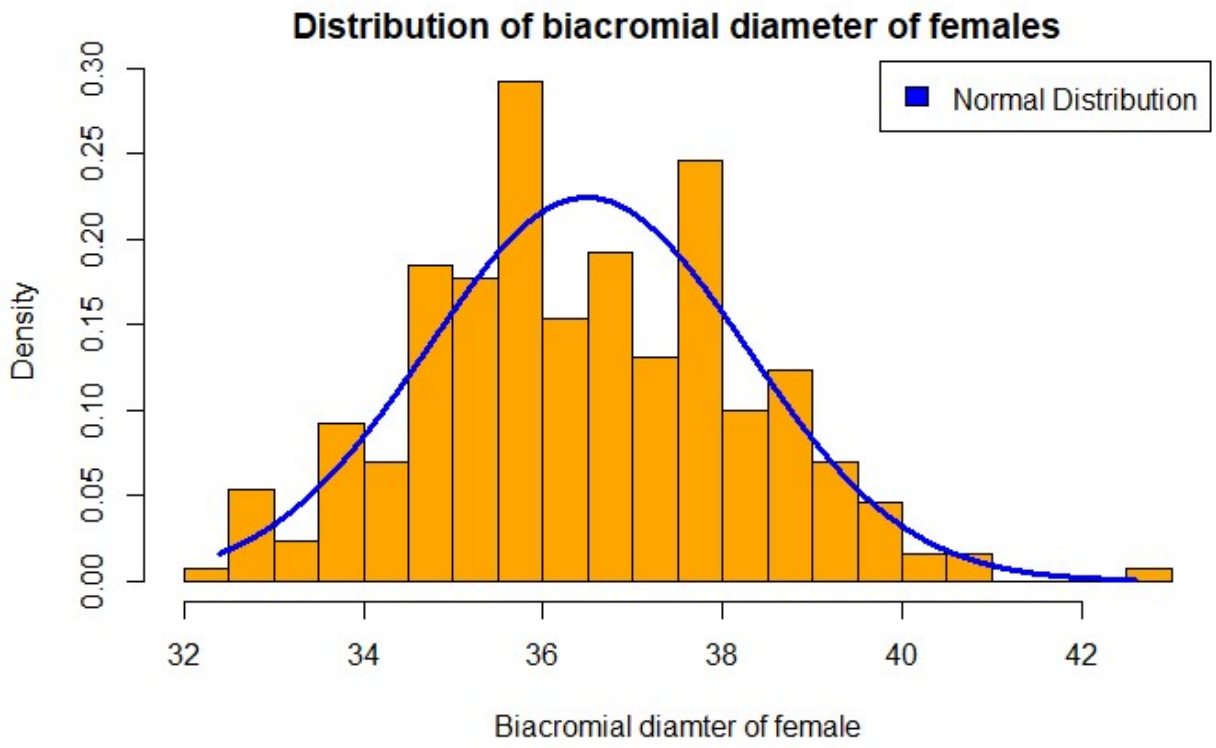


Hide

```
hist(female_df$bia.di, col = 'orange', xlab = 'Biacromial diamter of female', main =
'Distribution of biacromial diameter of females', probability = TRUE, breaks = 15)
x <- seq(from = min(female_df$bia.di), to = max(female_df$bia.di), length.out = 500)
y <- dnorm(x, mean = mean(female_df$bia.di), sd = sd(female_df$bia.di))
lines(x, y, col="blue", lwd=3)
```

Hide

```
legend("topright", c("Normal Distribution"), fill=c("blue"))
```



Interpretation

Interpretation of males biacromial diameter - Refering to the respective figure we can observe that density of the mean value of empirical data is more than that of the density of normal distributions mean. Mean of the normal distribution and empirical data is observed to be same. Normal distributions curve covers almost all the portion of empirical data.

Interpretation of females biacromial diameter - Refering to the respective figure we can observe that density of the mean value of empirical data is less than that of the density of normal distributions mean. Mean of the normal distribution and empirical data is observed to be same. Normal distributions curve covers majority of the portion of empirical data. Empirical data seems to have two peaks but still the curve mostly all the empirical data.

Finally after the investigation it can be concluded that both the normal distribution curves fit the empirical data well.