Probability R Assignment 3

Harshad Kumar Elangovan - 19200349

11/11/2019

Statistical Programming with R

Function pnorm can be used to calculate the areas under anormal distribution. This function takes three parameters 1.Observation value 2.Mean of the normal distribution 3.The standard deviation of the normal distribution

```
#Finding the area between 0.6 and 1.2 in a normal distribution with mean 0 and standard devia
tion 1

pnorm(1.2,mean=0,sd=1) #0.8849303
```

```
## [1] 0.8849303
```

```
pnorm(0.6,mean = 0,sd=1) #0.7257469
```

```
## [1] 0.7257469
```

```
# 0.8849303 - 0.7257469 = 0.1591834
```

Part A

Area left of 1.2:0.8849303 Area left of 0.6:0.7257469 Probability of Area between 0.6 and 1.2:0.1591834

This area can be calculated by R by assigning the result of the function to variables and subtracting them.

```
area1<-pnorm(1.2,mean=0,sd=1)

area2<-pnorm(0.6,mean = 0,sd=1)

totalarea<-area1-area2 #0.1591834
```

So, total area for the given range is 0.1591834.

Part B

The lengths of baby elephants' trunks follow a normal distribution with mean 1.8 metresand standard deviation 0.4 metres. What is the probability that a baby elephant will havea trunk between 1.6 and 2.1 metres long?

```
area1<-pnorm(2.1,mean = 1.8,sd=0.4) #0.7733726
area2<-pnorm(1.6,mean = 1.8,sd=0.4) #0.3085375
Probarea<-area1-area2 #0.4648351
```

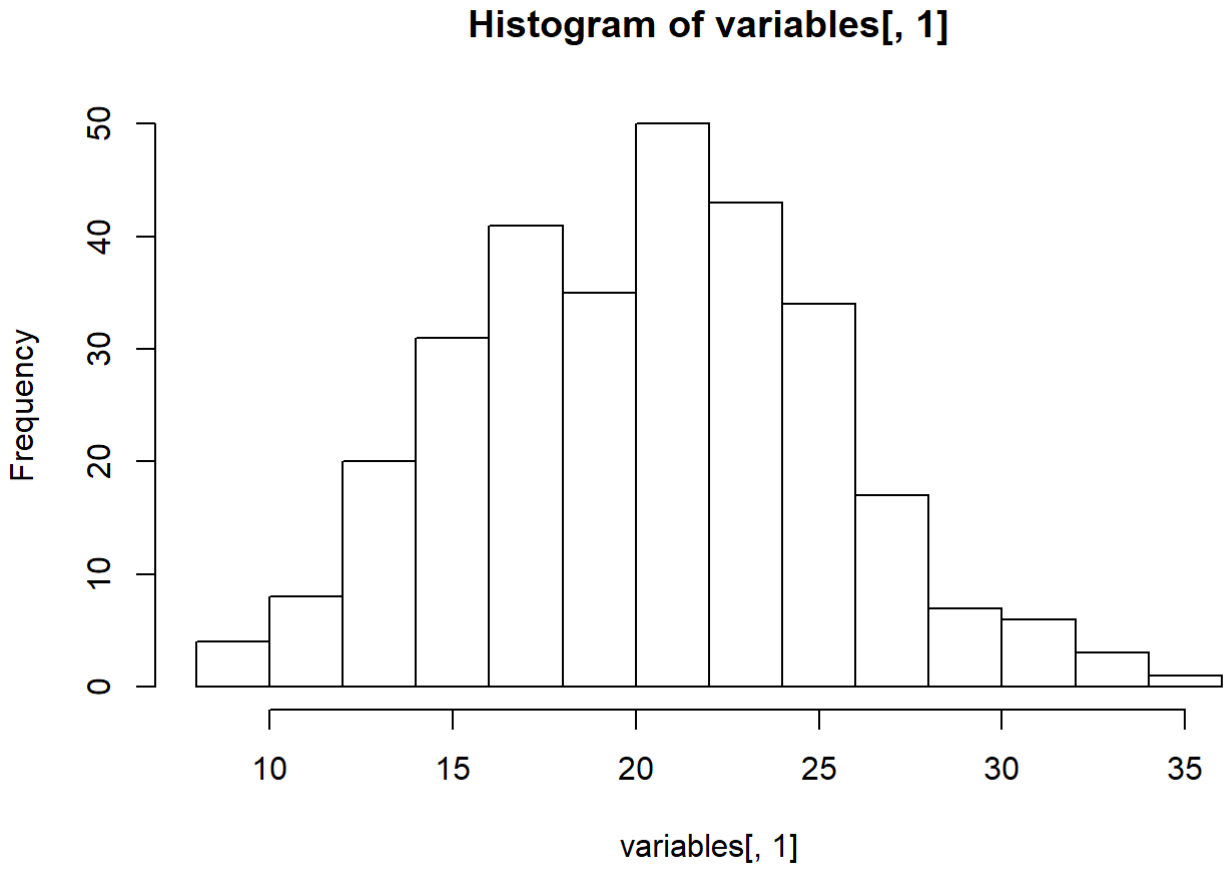
Probability that trunk is between 1.6 and 2.1 metres long: 0.4648351

Code used:

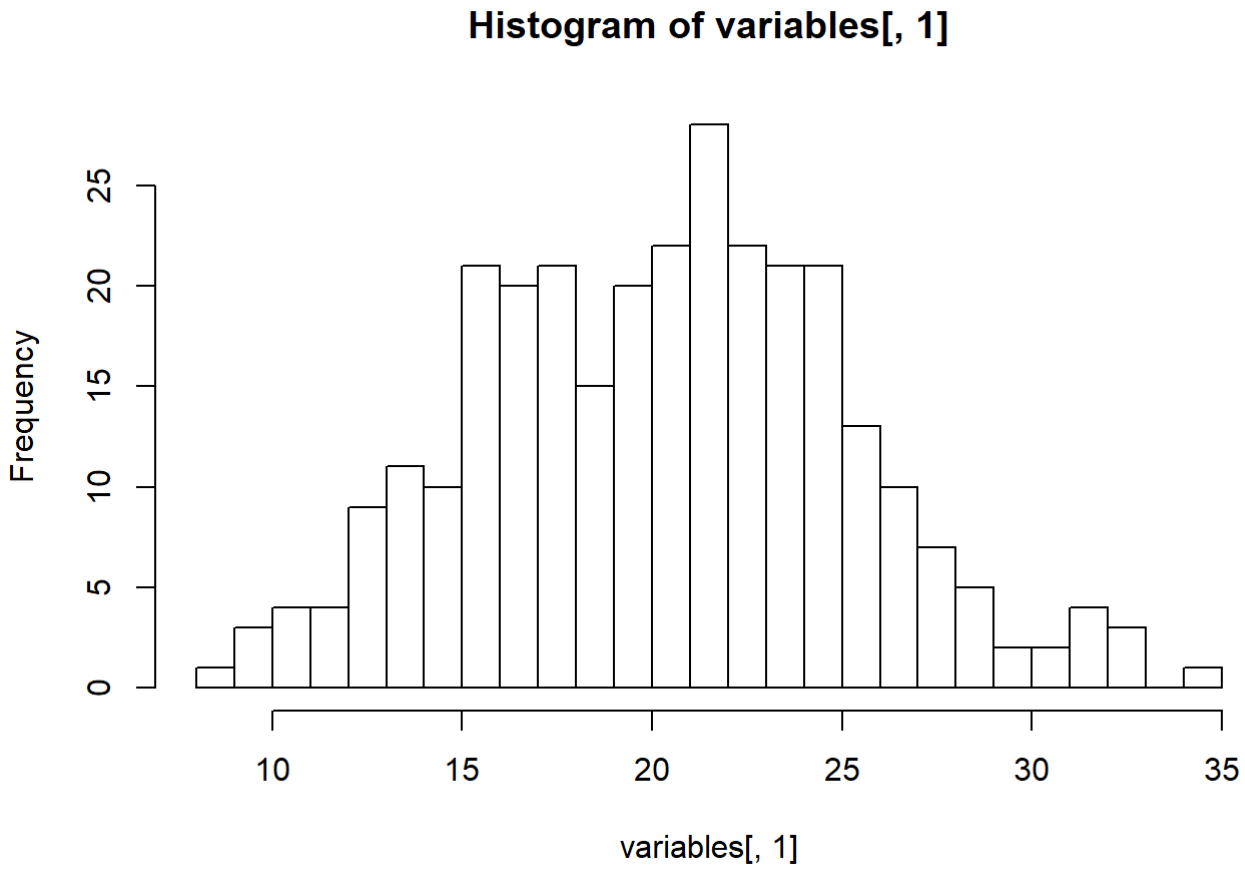
area1<-pnorm(2.1,mean = 1.8,sd=0.4)
area2<-pnorm(1.6,mean = 1.8,sd=0.4)
Probarea<-area1-area2

Part C

```
variables<-read.table("C:\\Users\\harshie\\Documents\\UCD\\Probability\\Lab3\\variables.txt",
header=TRUE)
variables <- data.frame(variables)
hist(variables[ , 1 ])
```



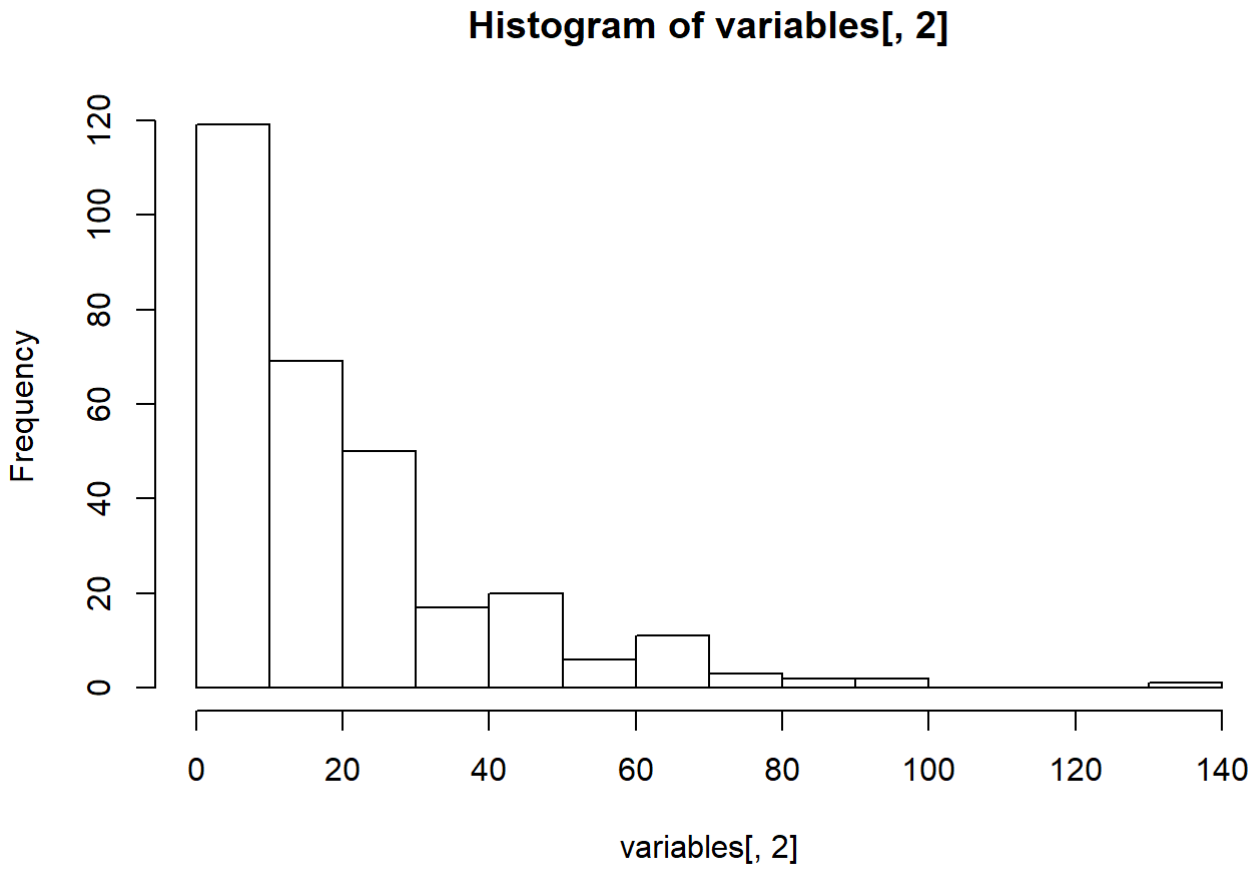

```
hist(variables[ , 1 ], breaks = 20)
```



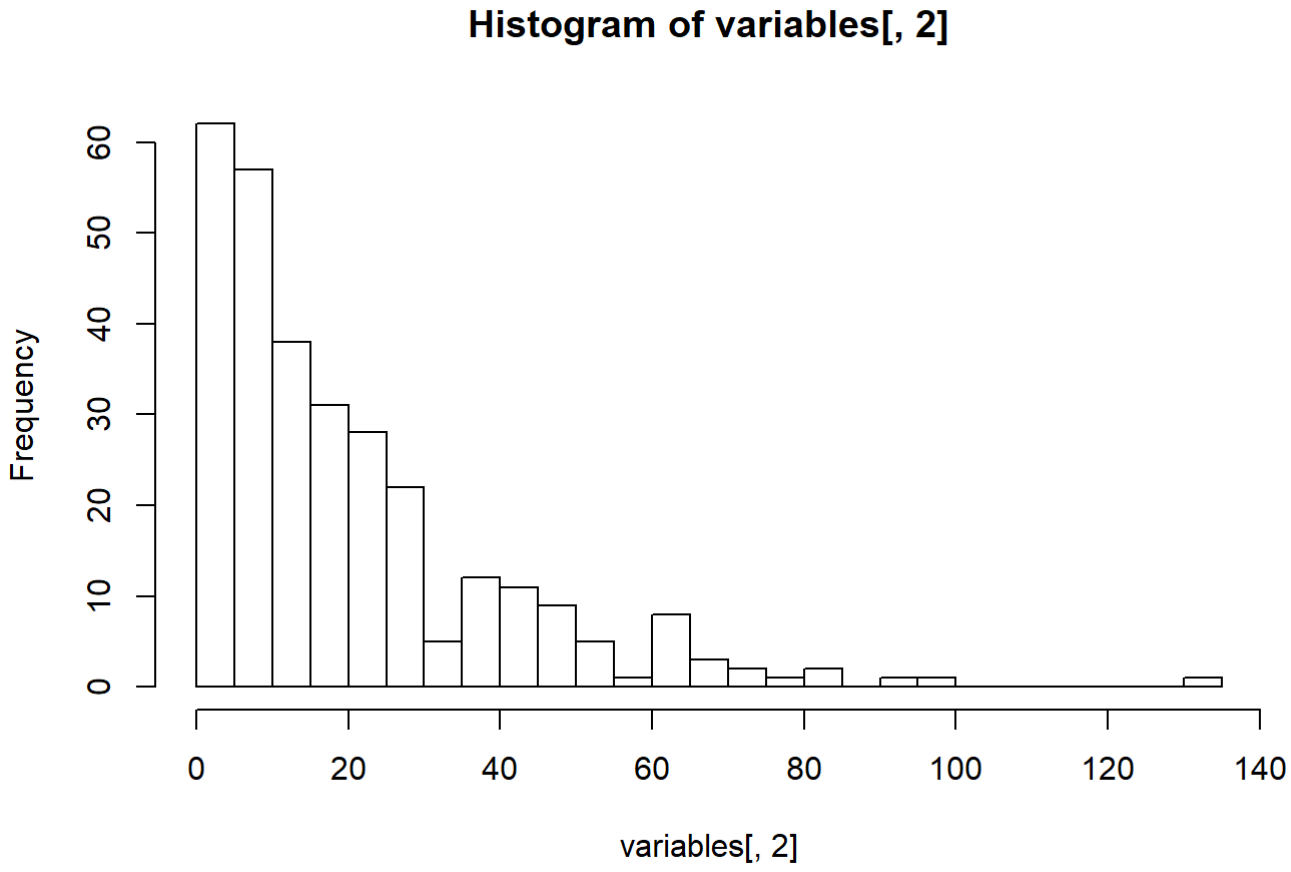

```
summary(variables[ , 1 ])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    8.27   16.56   20.47   20.29   23.59   34.60
```

```
hist(variables[ , 2 ])
```




```
hist(variables[ , 2 ], breaks = 20)
```



```
summary(variables[ , 2 ])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.055   5.832  13.511  20.078  27.245 131.378
```

Conclusion 1:

The data of the first dataset has a bell shaped curve near the mean and median and Mean is equal to the calculated Median. So this data proves the qualities of Normal distribution.

Code:

variables<-read.table("C:\Users\harshie\Documents\UCD\Probability\Lab3\variables.txt",header=TRUE)
variables <- data.frame(variables)
hist(variables[, 1])
hist(variables[, 1], breaks = 20)
summary(variables[, 1])

Conclusion 2:

The second dataset doesnot depicts the properties of Normal distribution since the curve is not bell shaped curve and the Mean and Median values vary each other(20.078 != 13.511 ie Mean !=Median).

Code:

hist(variables[, 2])
hist(variables[, 2], breaks = 20)
summary(variables[, 2])