**Assignment 2**

**Binomial and Normal distribution**

**Name: Motbhare Subodh Uddhav**

**Roll No: 92**

**Branch & Div: Computer Engineering & C**

**Gr No: 12120132**

**Source Code**

#Exercise-1

setwd("D:Data Science Program")

library(readxl)

travel = read\_excel("travelled abroad data.xlsx")

# Q.1)

p = length(travel$`Travelled abroad`[travel$`Travelled abroad`=="Y"])/nrow(travel) \* 100

p

# Q.2)

x0<-dbinom(0,10,p/100)

x0

x1 <- dbinom(1,10,p/100)

x1

x2<-dbinom(2,10,p/100)

x2

x3<-dbinom(3,10,p/100)

x3

x4<-dbinom(4,10,p/100)

x4

x5<-dbinom(5,10,p/100)

x5

x6<-dbinom(6,10,p/100)

x6

x7<-dbinom(7,10,p/100)

x7

x8<-dbinom(8,10,p/100)

x8

x9<-dbinom(10,10,p/100)

x9

# Q.3)

x<-c(x0,x1,x2,x3,x4,x5,x6,x7,x8,x9)

x<-signif(x,2)

x

barplot(x, names.arg=c(0,1,2,3,4,5,6,7,8,9), xlab='Number of people chosen for probability', ylab='Probabilties', main='Binomial distribution')

sum(dbinom(59:100,100,p/100))

#Exercise-2

setwd("D:Data Science Program")

hyp = read.csv("Hypothesis\_csv1.csv", header=TRUE)

hyp

# Q.1)

mean = mean(hyp$Life\_Hrs)

mean

std = sd(hyp$Life\_Hrs)

std

# Q.2)

SE = std/sqrt(nrow(hyp))

SE

# Q.3)

Z = (mean - 10000)/SE

Z

# Q.4)

p = pnorm(Z)

p

# Q.5)

# The p-value 0.0305 is less than the level of significance of 0.05 (5%). We reject the null hypothesis, Hence the claim made by the manufacturer is rejected.

# The p-value 0.0305 is greater than the level of significance of 0.01 (1%). We fail to reject the null hypothesis.

**Output:**

library(readxl)

> travel = read\_excel("travelled abroad data.xlsx")

>

>

>

>

> # Q.1)

>

> p = length(travel$`Travelled abroad`[travel$`Travelled abroad`=="Y"])/nrow(travel) \* 100

> p

[1] 56

>

> # Q.2)

> x0<-dbinom(0,10,p/100)

> x0

[1] 0.0002719736

>

> x1 <- dbinom(1,10,p/100)

> x1

[1] 0.003461482

>

> x2<-dbinom(2,10,p/100)

> x2

[1] 0.01982485

>

> x3<-dbinom(3,10,p/100)

> x3

[1] 0.06728435

>

> x4<-dbinom(4,10,p/100)

> x4

[1] 0.1498606

>

> x5<-dbinom(5,10,p/100)

> x5

[1] 0.228878

>

> x6<-dbinom(6,10,p/100)

> x6

[1] 0.2427494

>

> x7<-dbinom(7,10,p/100)

> x7

[1] 0.176545

>

> x8<-dbinom(8,10,p/100)

> x8

[1] 0.08426012

>

> x9<-dbinom(10,10,p/100)

> x9

[1] 0.003033055

>

> # Q.3)

> x<-c(x0,x1,x2,x3,x4,x5,x6,x7,x8,x9)

> x<-signif(x,2)

> x

[1] 0.00027 0.00350 0.02000 0.06700 0.15000 0.23000 0.24000 0.18000 0.08400 0.00300

> barplot(x, names.arg=c(0,1,2,3,4,5,6,7,8,9), xlab='Number of people chosen for probability', ylab='Probabilties', main='Binomial distribution')

>

> sum(dbinom(59:100,100,p/100))

[1] 0.3084356

>

>

> #Exercise-2

> setwd("D:Data Science Program")

Error in setwd("D:Data Science Program") :

cannot change working directory

>

> hyp = read.csv("Hypothesis\_csv1.csv", header=TRUE)

> hyp

Bearing.no. Life\_Hrs

1 1 9859.01

2 2 10111.94

3 3 10023.41

4 4 9770.15

5 5 10158.63

6 6 9639.62

7 7 10055.09

8 8 10095.38

9 9 9990.29

10 10 9883.86

11 11 10071.64

12 12 9685.03

13 13 9897.37

14 14 9918.80

15 15 9948.09

16 16 9963.22

17 17 9932.52

18 18 9775.70

19 19 10098.47

20 20 9870.34

21 21 10031.14

22 22 9992.01

23 23 10037.81

24 24 10065.94

25 25 9903.96

26 26 10134.47

27 27 10007.85

28 28 9810.02

29 29 10142.73

30 30 9824.97

31 31 10023.30

32 32 10109.40

33 33 9788.34

34 34 9833.43

35 35 9992.57

36 36 10020.62

37 37 10083.13

38 38 10022.64

39 39 9861.12

40 40 10001.23

41 41 9937.17

42 42 10084.14

43 43 9905.46

44 44 10065.08

45 45 10050.98

46 46 9641.05

47 47 9883.92

48 48 10199.71

49 49 10033.67

50 50 10013.62

>

> # Q.1)

> mean = mean(hyp$Life\_Hrs)

> mean

[1] 9965.001

> std = sd(hyp$Life\_Hrs)

> std

[1] 132.1739

>

> # Q.2)

> SE = std/sqrt(nrow(hyp))

> SE

[1] 18.69222

>

> # Q.3)

> Z = (mean - 10000)/SE

> Z

[1] -1.872394

>

> # Q.4)

> p = pnorm(Z)

> p

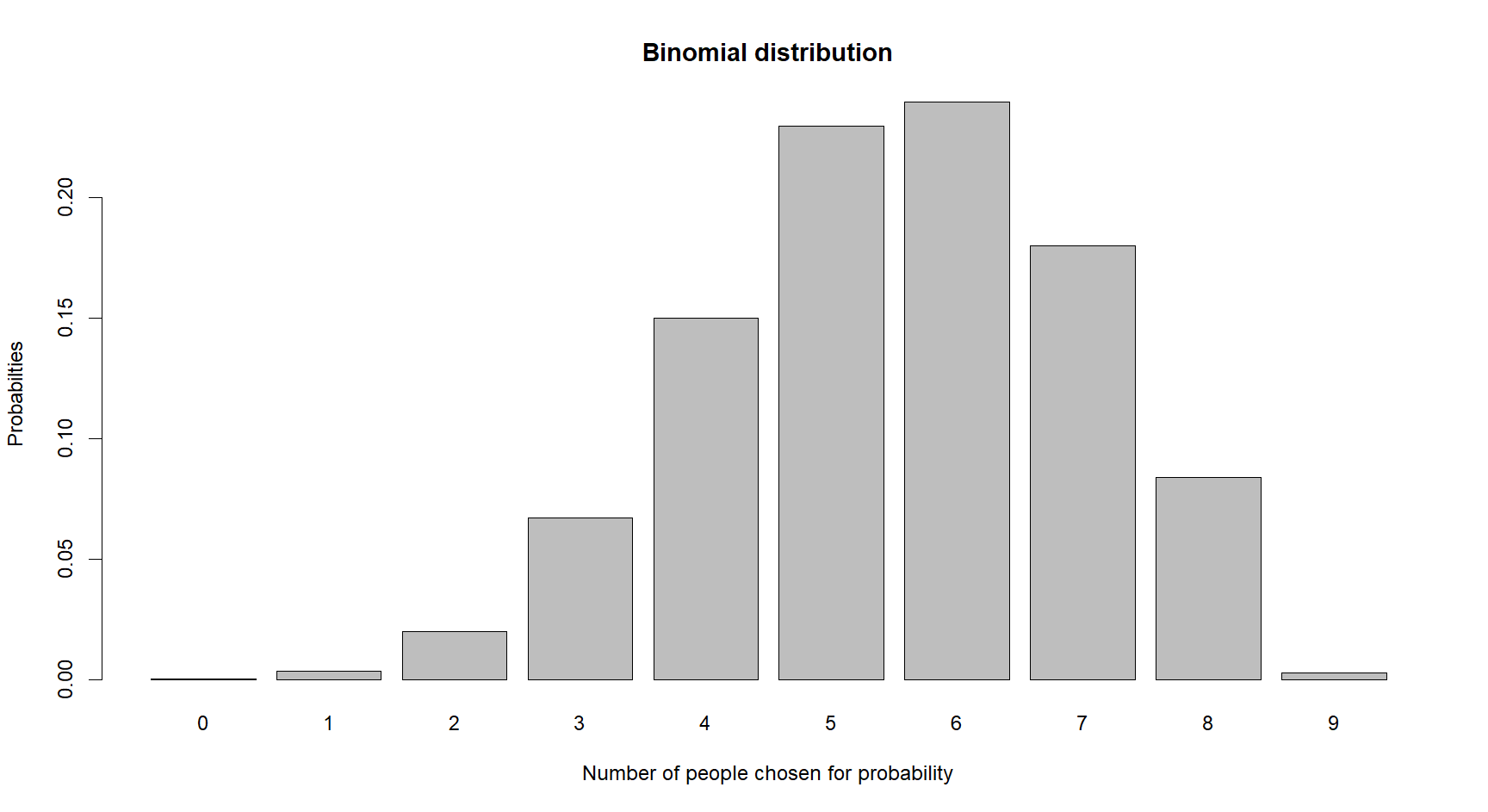
[1] 0.03057603

>

> # Q.5)

> # The p-value 0.0305 is less than the level of significance of 0.05 (5%). We reject the null hypothesis, Hence the claim made by the manufacturer is rejected.

> # The p-value 0.0305 is greater than the level of significance of 0.01 (1%). We fail to reject the null hypothesis.

****