MD2201: Data Science

Name of the student: Rajeev Tapadia **Roll No. : 67**

Div:C **Batch: 3**

Date of performance:

Experiment No.2

Title: Laboratory on Statistics

Aim: To find the probability for Binomial distribution and Normal distribution and verify the Normal approximation of Binomial distribution.

Software used: Programming language R.

Code Statement:

Write a single R code to answer the following questions.

Data Set: Travelled Abroad

1. Find out the % of Indians in the sample who have travelled abroad using the data source.

```
# 1.  Find out the % of Indians in the sample who have travelled abroad
p_ <- mean(dataset$Travelledabroad == "Y")
```

```
> # 1.  Find out the % of Indians in th
> mean(dataset$Travelledabroad == "Y")
[1] 0.56
```

4. Treating this value as 'p', calculate the following probabilities –
    a. What is the probability that in a randomly chosen sample of 10 persons, no one has travelled abroad?
    b. What is the probability that in a randomly chosen sample of 10 persons, exactly one has travelled abroad?
    c. What is the probability that in a randomly chosen sample of 10 persons, exactly two persons have travelled abroad?
    d. What is the probability that in a randomly chosen sample of 10 persons, exactly three persons have travelled abroad?
    e. What is the probability that in a randomly chosen sample of 10 persons, exactly four persons have travelled abroad?
    f. What is the probability that in a randomly chosen sample of 10 persons, exactly five persons have travelled abroad.
    g. What is the probability that in a randomly chosen sample of 10 persons, exactly six persons have travelled abroad?
    h. What is the probability that in a randomly chosen sample of 10 persons, exactly seven persons have travelled abroad?

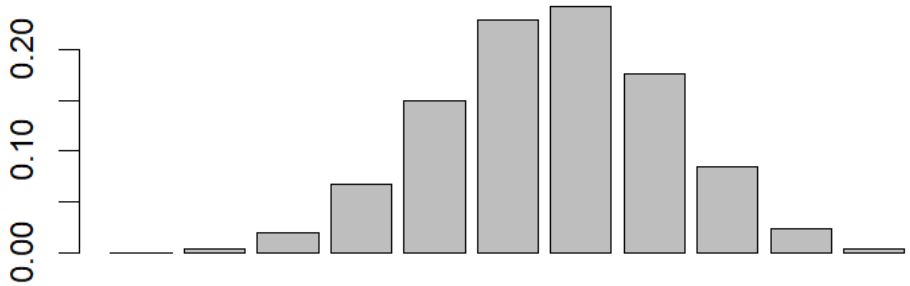
i. What is the probability that in a randomly chosen sample of 10 persons, exactly eight persons have travelled abroad?
j. What is the probability that in a randomly chosen sample of 10 persons, exactly nine persons have travelled abroad?
k. What is the probability that in a randomly chosen sample of 10 persons, all 10 persons have travelled abroad?

```
# b. exactly one person travelled abroad
dbinom(1, size = 10, prob = p)
# Probability that exactly 2 persons have traveled abroad
dbinom(2, size = 10, prob = p)
# Probability that exactly 3 persons have traveled abroad
dbinom(3, size = 10, prob = p)
# Probability that exactly 4 persons have traveled abroad
dbinom(4, size = 10, prob = p)
# Probability that exactly 5 persons have traveled abroad
dbinom(5, size = 10, prob = p)
# Probability that exactly 6 persons have traveled abroad
dbinom(6, size = 10, prob = p)
# Probability that exactly 7 persons have traveled abroad
dbinom(7, size = 10, prob = p)
# Probability that exactly 8 persons have traveled abroad
dbinom(8, size = 10, prob = p)
# Probability that exactly 9 persons have traveled abroad
dbinom(9, size = 10, prob = p)
# Probability that exactly 10 persons have traveled abroad
dbinom(10, size = 10, prob = p)
```

```
> # a. no one has travelled abroad
> dbinom(0, size = 10, prob = p)
[1] 0.0002719736
> # b. exactly one person travelled abroad
> dbinom(1, size = 10, prob = p)
[1] 0.003461482
> # Probability that exactly 2 persons have traveled abroad
> dbinom(2, size = 10, prob = p)
[1] 0.01982485
> # Probability that exactly 3 persons have traveled abroad
> dbinom(3, size = 10, prob = p)
[1] 0.06728435
> # Probability that exactly 4 persons have traveled abroad
> dbinom(4, size = 10, prob = p)
[1] 0.1498606
> # Probability that exactly 5 persons have traveled abroad
> dbinom(5, size = 10, prob = p)
[1] 0.228878
```

```
> # Probability that exactly 6 persons have traveled abroad
> dbinom(6, size = 10, prob = p)
[1] 0.2427494
> # Probability that exactly 7 persons have traveled abroad
> dbinom(7, size = 10, prob = p)
[1] 0.176545
> # Probability that exactly 8 persons have traveled abroad
> dbinom(8, size = 10, prob = p)
[1] 0.08426012
> # Probability that exactly 9 persons have traveled abroad
> dbinom(9, size = 10, prob = p)
[1] 0.02383115
> # Probability that exactly 10 persons have traveled abroad
> dbinom(10, size = 10, prob = p)
[1] 0.003033055
```

25. Plot the probability values as a Table / Bar graph/plot and interpret plot.



26. What is the probability that in the randomly chosen sample of 100 persons at least 59 have travelled abroad?

Hint: Expected to perform Normal approximation for the binary distribution.

```
no_sample <- 100
mu <- no_sample * p
sigma <- sqrt(no_sample * p * (1-p))
z <- (59 - mu)/sigma
1 - pnorm(z)
```

```
> 1 - pnorm(z)
[1] 0.2727998
>
```

Conclusion:

The analysis will provide insights into travel patterns and demonstrate how the Normal distribution can be used to approximate the Binomial distribution under certain conditions.