

Name – Atharva Sanjay shevate

Roll no -2

Div -E(1)

Srn – 202201727

INTRODUCTION

- Binomial distribution is a statistical probability distribution that summarizes the likelihood that a value will take one of two independent values under a given set of parameters or assumptions.
- The underlying assumptions of binomial distribution are that there is only one outcome for each trial, that each trial has the same probability of success, and that each trial is mutually exclusive or independent of one another.
- Binomial distribution is a common discrete distribution used in statistics, as opposed to a continuous distribution, such as normal distribution.

Data statistics model for Binomial Distribution in R programming.

Data Source: The cricket match data is loaded from the CSV file "vechical.csv."

Data Format: The dataset includes information about Vehical, such as

Scooter Owner, Car Owner And Bike Owner

code

Implementation # Load

the ggplot2 library

```
library(ggplot2)
```

Step 1: Load the Data data

```
<- read.csv(file.choose())
```

Step 2: Convert Data to Binary Format data\$car

```
<- ifelse(data$car == "Yes", 1, 0) data$bike <-
```

```
ifelse(data$bike == "Yes", 1, 0) data$scooter <-
```

```
ifelse(data$scooter == "Yes", 1, 0)
```

Step 3: Binomial Distribution Analysis for Car Users

```
car_prob <- dbinom(0:nrow(data), size = nrow(data),  
prob = mean(data$car))
```

```
# Create a data frame for the results
```

```
result_df <- data.frame(Owners = 0:nrow(data),  
Probability = car_prob)
```

```
# Step 4: Create a Beautiful Plot
```

```
car_plot <- ggplot(result_df, aes(x = Owners, y =  
Probability)) + geom_bar(stat = "identity", fill =  
"skyblue", color =  
"blue") + labs(title = "Binomial Distribution of Car  
Owners", x = "Number of Car Owners", y =  
"Probability") + theme_minimal() +  
theme(plot.title = element_text(hjust = 0.5))
```

```
# Print the plot print(car_plot)
```

```
# Step 3: Binomial Distribution
```

```
Analysis for Bike Users
```

```
bike_prob <-  
dbinom(0:nrow(data), size =  
nrow(data), prob =  
mean(data$bike))
```

```
# Step 4: Binomial Distribution Analysis for Scooter  
Users scooter_prob <- dbinom(0:nrow(data), size =  
nrow(data), prob = mean(data$scooter))
```

```
# Create data frames for the results
```

```
bike_result_df <- data.frame(Owners = 0:nrow(data),  
Probability = bike_prob) scooter_result_df <-  
data.frame(Owners = 0:nrow(data), Probability =  
scooter_prob)
```

```
# Step 5: Create Beautiful Plots
```

```
bike_plot <- ggplot(bike_result_df, aes(x = Owners, y =  
Probability)) + geom_bar(stat = "identity", fill =  
"skyblue", color =  
"blue") + labs(title = "Binomial Distribution of Bike  
Owners", x = "Number of Bike Owners",  
y = "Probability") + theme_minimal() +  
theme(plot.title = element_text(hjust = 0.5))
```

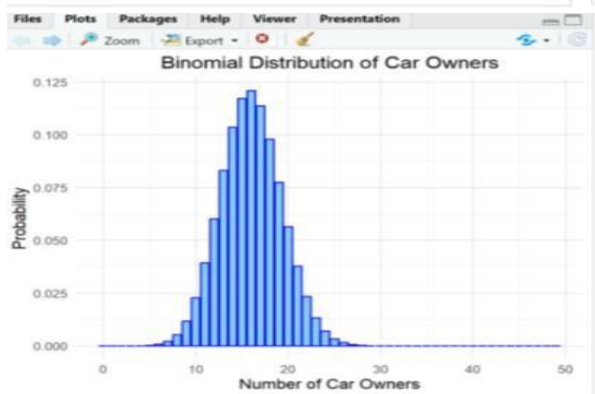
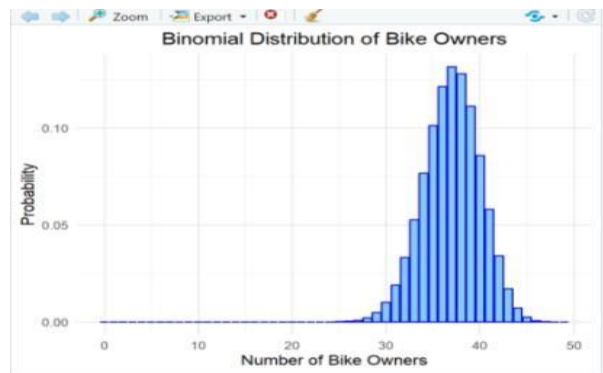
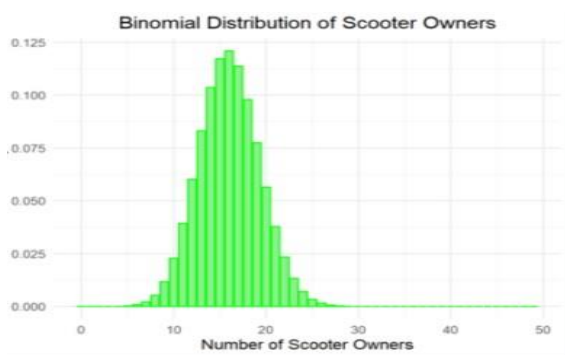
```
scooter_plot <- ggplot(scooter_result_df, aes(x =  
Owners, y = Probability)) + geom_bar(stat = "identity",  
fill = "lightgreen", color =
```

```

"green") +
  labs(title = "Binomial Distribution of Scooter
Owners",    x = "Number of Scooter
Owners",    y = "Probability") +
  theme_minimal() + theme(plot.title =
element_text(hjust = 0.5))

# Print the plots print(bike_plot)
print(scooter_plot)

```



Result Analysis

The result of this script is a set of three bar plots, each representing the binomial distribution of ownership for the respective type of vehicle (car, bike, or scooter). These plots show the probability of having a certain number of owners for each type of vehicle, given the dataset provided. The plots can be used to visualize the distribution of ownership within the dataset.

Conclusion

This script provides a data-driven analysis and visual representation of the ownership distribution for different types of vehicles in the dataset. It helps you understand the likelihood of having a specific number of owners for each type of vehicle, which can be valuable for making data-driven decisions or understanding patterns within the dataset.

