Sri Lanka Institute of Information Technology

Lab Submission
Lab sheet No 05

IT24101520

Athukorala A P H B

Probability and Statistics | IT2120

B.Sc. (Hons) in Information Technology

Exercise

1. Import the dataset (’Exercise – Lab 05.txt’) into R and store it in a data frame called” Delivery Times”.

```r
#Set working Directory
setwd("C:\\Users\\pamud\\Desktop\\IT24101520")
getwd()

#Import data set
DeliveryTimes <- read.table("Exercise - Lab 05.txt", header = TRUE)
colnames(DeliveryTimes) <- c("Time")

# Ensure it's numeric
DeliveryTimes$Time <- as.numeric(DeliveryTimes$Time)

```

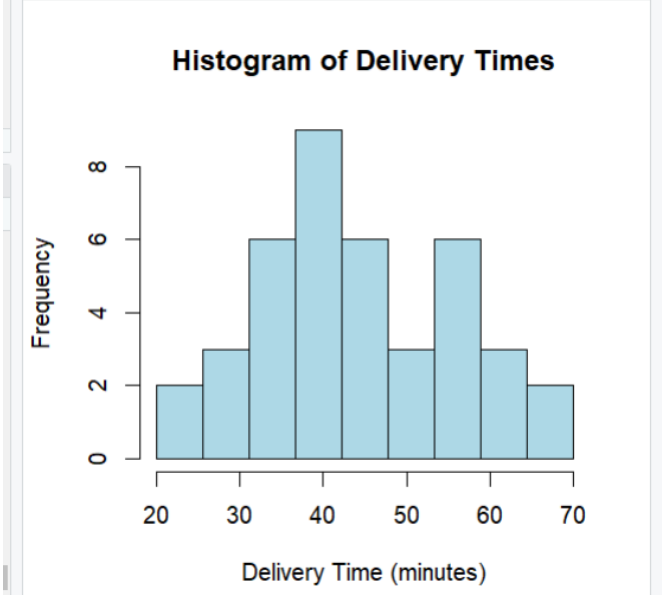
2. Draw a histogram for deliver times using nine class intervals where the lower limit is 20 and upper limit is 70. Use right open intervals.

```r
# Part 2: Histogram with 9 class intervals (20 to 70)
hist(DeliveryTimes$Time,
     breaks = seq(20, 70, length.out = 10),
     right = FALSE,
     main = "Histogram of Delivery Times",
     xlab = "Delivery Time (minutes)",
     ylab = "Frequency",
     col = "lightblue",
     border = "black")

# Define intervals
breaks <- seq(20, 70, length.out = 10)

# Create histogram object without plotting
freq_table <- hist(DeliveryTimes$Time,
                   breaks = breaks,
                   right = FALSE,
                   plot = FALSE)
```

```
> # Part 2: Histogram with 9 class intervals (20 to 70)
> hist(DeliveryTimes$Time,
+      breaks = seq(20, 70, length.out = 10),
+      right = FALSE,
+      main = "Histogram of Delivery Times",
+      xlab = "Delivery Time (minutes)",
+      ylab = "Frequency",
+      col = "lightblue",
+      border = "black")
>
>
> # Define intervals
> breaks <- seq(20, 70, length.out = 10)
```



3. Comment on the shape of the distribution.

The distribution of delivery times is roughly symmetric and unimodal, with most deliveries occurring around 35–45 minutes. There are fewer deliveries at the lower and higher extremes, indicating a fairly balanced spread around the central values.

4. Draw a cumulative frequency polygon (ogive) for the data in a separate plot.

```
# Create histogram object without plotting
freq_table <- hist(DeliveryTimes$Time,
                   breaks = breaks,
                   right = FALSE,
                   plot = FALSE)

# View frequencies
freq_table$counts   # frequency in each interval

# Optional: view the interval ranges
freq_table$breaks


# Cumulative frequency
cum_freq <- cumsum(freq_table$counts)
```

```
# Plot ogive
plot(breaks[-1], cum_freq, type = "o",
     main = "Cumulative Frequency Polygon (Ogive)",
     xlab = "Delivery Time",
     ylab = "Cumulative Frequency",
     col = "red")
```

```
> # Define intervals
> breaks <- seq(20, 70, length.out = 10)
> # Create histogram object without plotting
> freq_table <- hist(DeliveryTimes$Time,
+                    breaks = breaks,
+                    right = FALSE,
+                    plot = FALSE)
>
> # View frequencies
> freq_table$counts   # frequency in each interval
[1] 2 3 6 9 6 3 6 3 2
>
> # Optional: view the interval ranges
> freq_table$breaks
 [1] 20.00000 25.55556 31.11111 36.66667 42.22222 47.77778 53.33333 58.88889 64.44444 70.00000
>
>
> # Cumulative frequency
> cum_freq <- cumsum(freq_table$counts)
>
```

```
> # Cumulative frequency
> cum_freq <- cumsum(freq_table$counts)
> # Plot ogive
> plot(breaks[-1], cum_freq, type = "o",
+      main = "Cumulative Frequency Polygon (Ogive)",
+      xlab = "Delivery Time",
+      ylab = "Cumulative Frequency",
+      col = "red")
```