IT2120 - Probability and Statistics

Lab Sheet 04

IT24104087-Rathnamalala R.I.B.T

```r
setwd("C:\\Users\\IT24104087\\Desktop\\IT24104087")
Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE)

colnames(Delivery_Times) <- "Delivery_Time"
head(Delivery_Times)
str(Delivery_Times)

breaks <- seq(20, 70, length.out = 10)
hist(Delivery_Times$Delivery_Time,right = FALSE,breaks = breaks,main = "Histogram of Delivery Times",xlab = "Delivery Tim

hist_data <- hist(Delivery_Times$Delivery_Time, breaks = breaks, right = FALSE, plot = FALSE)
frequencies <- hist_data$counts
cum_freq <- cumsum(frequencies)
print(frequencies)
print(cum_freq)

midpoints <- hist_data$mids
plot(midpoints, cum_freq, type = "b",
     main = "Cumulative Frequency Polygon for Delivery Time",
     xlab = "Delivery Time (minutes)",
     ylab = "Cumulative Frequency",
     pch = 16)
```

Comments on distribution

It is approximately symmetric with no strong left or right skew.

- The spread ranges from 20 to 67 minutes, with most values between 31.12 and 58.92 minutes.
- No extreme outliers, suggesting two subgroups of delivery times, possibly due to different delivery conditions or order types.