Sri Lanka Institute of Information Technology

Lab Submission
<Worksheet 5>

<IT24102257>

<Senadheera P.V.S.T>

Probability and statictics | IT2120

Exercise

Instructions: Create a folder in your desktop with your registration number (Eg: "IT....."). You need to save the R script file and take screenshots of the command prompt with answers and save it in a word document inside the folder. Save both R script file and word document with your registration number (Eg: "IT....."). After you finish the exercise, zip the folder and upload the zip file to the submission link.

1. Import the dataset ('Exercise – Lab 05.txt') into R and store it in a data frame called "Delivery_Times".

```
> #Exercise
> #(1)
> Delivery_Times <- read.table("Exercise - Lab 05.txt", header = TRUE, sep = ",")
```

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.

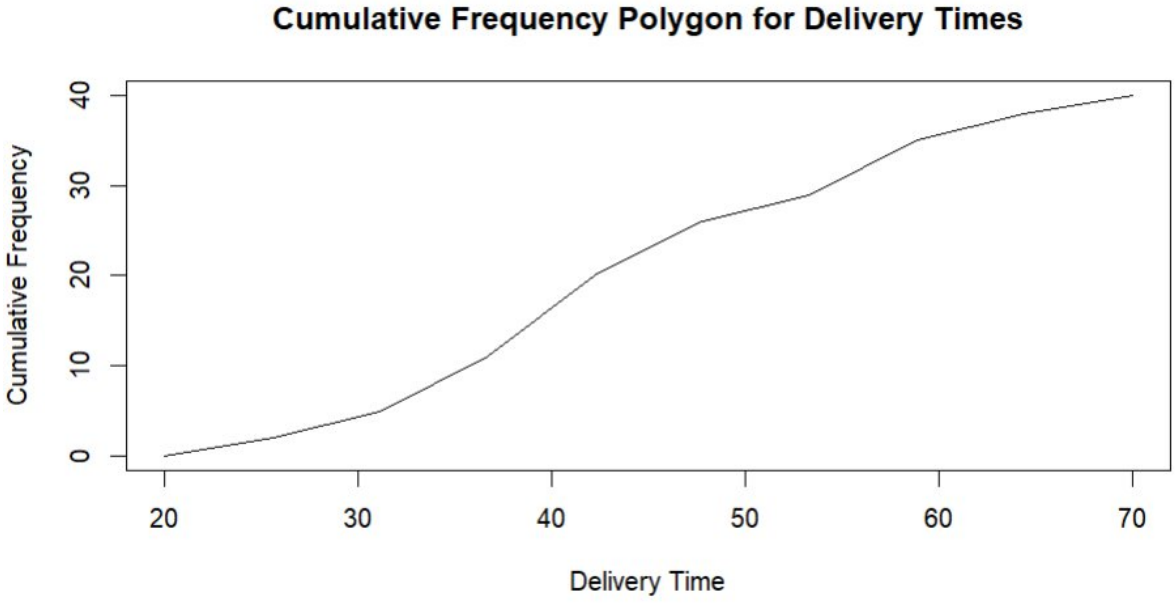
```
> #(2)
> hist(Delivery_Times$Delivery_Time_.minutes., main = "Histogram for Delivery Times", breaks = seq(20, 70, length = 10), right = FALSE)
```

3. Comment on the shape of the distribution.

```
> #(3)
> summary(Delivery_Times$Delivery_Time_.minutes.)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  20.00   36.00   42.50   43.75   54.00   67.00
> mean(Delivery_Times$Delivery_Time_.minutes.)
[1] 43.75
> median(Delivery_Times$Delivery_Time_.minutes.)
[1] 42.5
> #The distribution of delivery times is approximately symmetric,
> #but with a slight positive skew (a longer right tail).
```

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

```
#(4)
delivery_hist <- hist(Delivery_Times$Delivery_Time_.minutes., breaks = seq(20, 70, length = 10), right = FALSE, plot = FALSE)
delivery_breaks <- delivery_hist$breaks
delivery_freq <- delivery_hist$counts
delivery_cumfreq <- cumsum(delivery_freq)
delivery_new <- c()
for (i in 1:length(delivery_breaks)) {
  if (i == 1) {
    delivery_new[i] = 0
  } else {
    delivery_new[i] = delivery_cumfreq[i-1]
  }
}
plot(delivery_breaks, delivery_new, type = 'l', main = "Cumulative Frequency Polygon for Delivery Times",
     xlab = "Delivery Time", ylab = "Cumulative Frequency", ylim = c(0, max(delivery_cumfreq)))
```



B.Sc. (Hons) in Information Technology