Weilin Lu

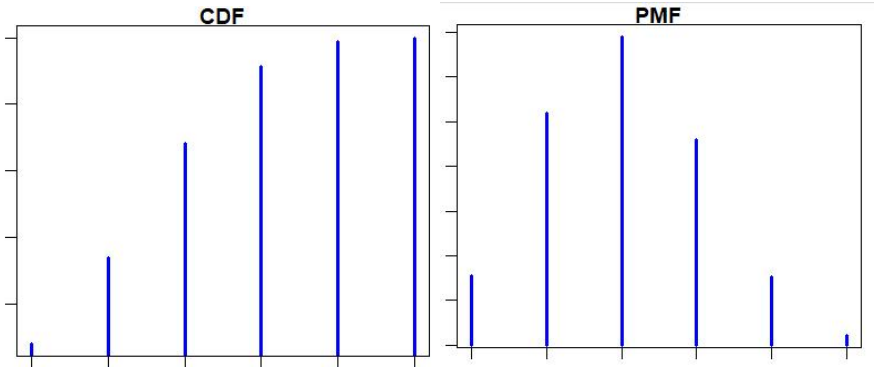
Assign 4

Part 1

a) :

```
#part 1
p_1 <- 0.4
n_1 <- 5
#a
successes <- 0:n_1
CDF1 <- pbinom(successes, size=n_1, prob=p_1)
print(CDF1)
PMF1 <- dbinom(successes, size=n_1, prob=p_1)
print(PMF1)
par(mfrow=c(1,1), mar = c(1,1,1,1))
plot(successes, CDF1 ,type='h',xlab="sucesses",ylab="probabilities",col="blue",lwd=3,main="CDF")
plot(successes, PMF1,type='h',xlab="sucesses",ylab="probabilities",col="blue",lwd=3,main="PMF")
```

```
> print(CDF1)
[1] 0.07776 0.33696 0.68256 0.91296 0.98976 1.00000
> print(CDF1)
[1] 0.07776 0.33696 0.68256 0.91296 0.98976 1.00000
> PMF1 <- dbinom(successes, size=n_1, prob=p_1)
> print(PMF1)
[1] 0.07776 0.25920 0.34560 0.23040 0.07680 0.01024
```



b) :

```
#b
p_1_b <- dbinom(2,n_1,p_1)
print(p_1_b)
```

```
> p_1_b <- dbinom(2,n_1,p_1)
> print(p_1_b)
[1] 0.3456
```

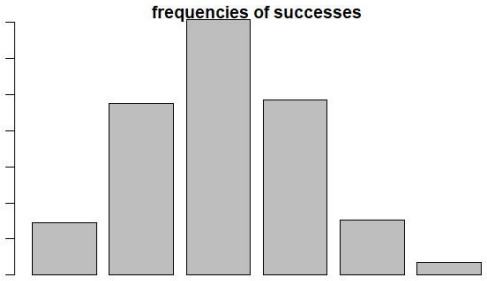
c) :

```
#c
p_1_c <- 1-pbinom(1,n_1,p_1)
print(p_1_c)
```

```
> print(p_1_c)
[1] 0.66304
```

d) :

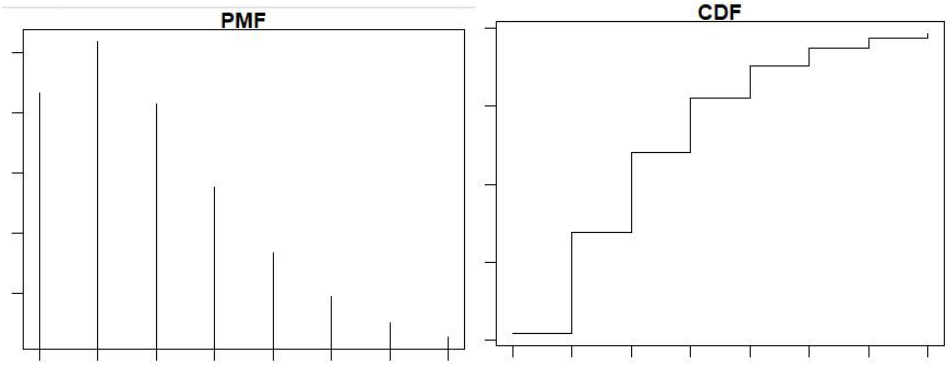
```
#d
set.seed(100)
perfect_scores_num <- rbinom(1000, size = n_1, prob = p_1)
barplot(table(perfect_scores_num), main = "frequencies of successes", xlab = "number of perfect scores", ylab = "frequency")
```



Part 2

a) :

```
#part 2
p_2 <- 0.6
r <- 3
n_2 <- 10
#a
x <- 0:(n_2 - r)
PMF2 <- dnbinom(x, size = r, prob = p_2)
plot(x, PMF2, type = "h", xlab = "number of failures", ylab = "probability", main = "PMF")
print(PMF2)
CDF2 <- pnbinom(x, size = r, prob = p_2)
Print(CDF2)
plot(x, CDF2, type = "s", xlab = "number of failures", ylab = "probability", main = "CDF")
```




```
> print(PMF2)
[1] 0.21600000 0.25920000 0.20736000 0.13824000 0.08294400 0.04644864 0.02477261 0.01274020
> CDF2 <- pnbinom(x, size = r, prob = p_2)
> print(CDF2)
[1] 0.2160000 0.4752000 0.6825600 0.8208000 0.9037440 0.9501926 0.9749652 0.9877054
```

b) :

```
#b
prob__of_4 <- dnbinom(4, size = r, prob = p_2)
prob__of_4
```

```
> prob__of_4
[1] 0.082944
```

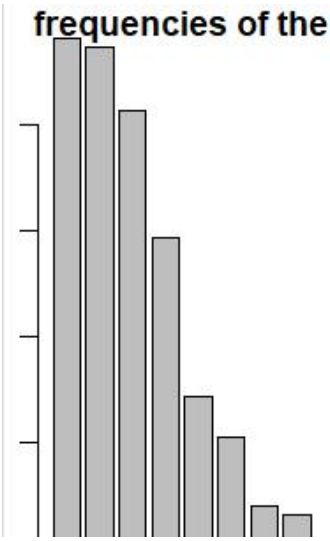
c) :

```
#c
atmost_4 <- pnbinom(4, size = r, prob = p_2)
atmost_4
```

```
> atmost_4
[1] 0.903744
```

d) :

```
set.seed(100)
failures <- rnbinom(1000, size = r, prob = p_2)
barplot(table(failures), main = "frequencies of the failures", xlab = "Number of failures", ylab = "Frequency")
```

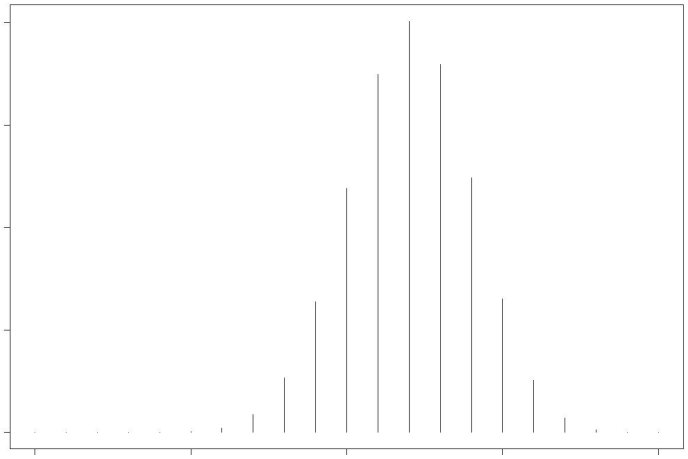


Part 3

a) :

```
#part 3
n_3 <- 20
p_3 <- 0.6
#a
question_num <- 0:20
probability <- dhyper(question_num,m=p_3*100,n=40,k=n_3)
print(probability)
plot(question_num, probability, type = "h", xlab = "Number of Multiple Choice Questions", ylab = "Probability")
```

```
> #a
> question_num <- 0:20
> probability <- dhyper(question_num,m=p_3*100,n=40,k=n_3)
> print(probability)
 [1] 2.571843e-10 1.469625e-08 3.744203e-07 5.665142e-06 5.718253e-05 4.098844e-04 2.167658e-03 8.670631e-03 2.666993e-02 6.376259e-02
[11] 1.192361e-01 1.748329e-01 2.007847e-01 1.797234e-01 1.242206e-01 6.530452e-02 2.550958e-02 7.137815e-03 1.346167e-03 1.526019e-04
[21] 7.820848e-06
> plot(question_num, probability, type = "h", xlab = "Number of Multiple Choice Questions", ylab = "Probability")
```



b) :

```
#b
exact10 <- dhyper(10,m=p_3*100,n=40,k=n_3)
exact10
```

```
> #b
> exact10 <- dhyper(10,m=p_3*100,n=40,k=n_3)
> exact10
[1] 0.1192361
```

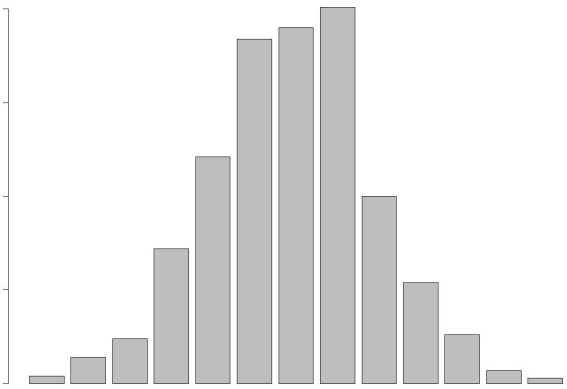
c) :

```
#c
least10 <- 1 - phyper(9,m=p_3*100,n=40,k=n_3)
least10
```

```
> #c
> least10 <- 1 - phyper(9,m=p_3*100,n=40,k=n_3)
> least10
[1] 0.8982561
```

d) :

```
#d
set.seed(100)
data_in_sim <- rhyper(1000,m=p_3*100,n=40,k=n_3)
print(data_in_sim)
barplot(table(data_in_sim), xlab = "Number of Multiple Choice Questions", ylab = "Frequency")
```



Part 4

a) :

```
#part 4
Question_num <- 10
#a
dpois(8, Question_num)
```

```
> dpois(8, Question_num)
[1] 0.112599
```

b) :

```
#b
ppois(8, Question_num)
```

```
> ppois(8, Question_num)
[1] 0.3328197
```

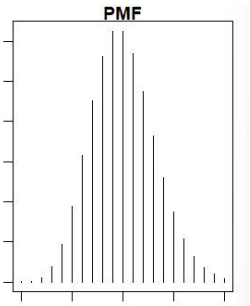
c) :

```
#c
ppois(12, Question_num) - ppois(5, Question_num)
```

```
> ppois(12, Question_num) - ppois(5, Question_num)
[1] 0.7244705
```

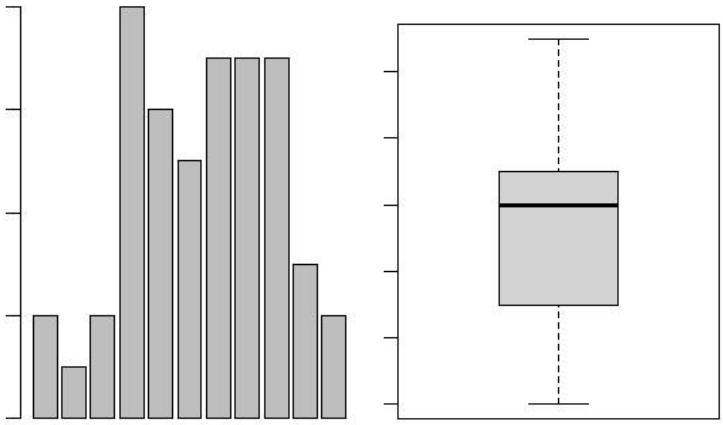
d) :

```
#d
first_20_question <- 0:20
prob_20 <- dpois(first_20_question, Question_num)
plot(first_20_question, prob_20, type = "h", xlab = "Number of Questions", ylab = "Probability", main = "PMF")
```



e) :

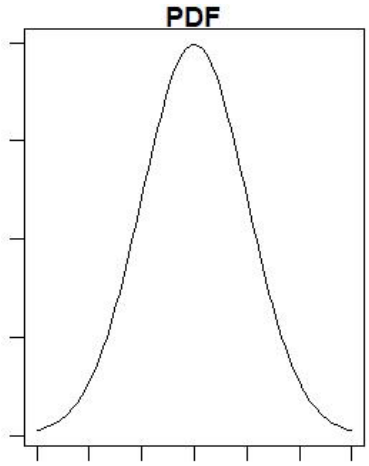
```
#e
set.seed(100)
data_in_sim <- rpois(50, Question_num)
barplot(table(data_in_sim), xlab = "Number of Questions", ylab = "Frequency")
boxplot(data_in_sim, xlab = "Number of Questions")
```



Part 5

a) :

```
#part 5
mean <- 100
stand_deviation <- 10
#a
x <- seq(mean - 3*stand_deviation, mean + 3*stand_deviation, length.out = 100)
y <- dnorm(x, mean = mean, sd = stand_deviation)
plot(x, y, type = "l", main = "PDF", xlab = "Money Spent ($)", ylab = "Density")
```



b) :

```
#b
1 - pnorm(120, mean = mean, sd = stand_deviation)
```

```
> 1 - pnorm(120, mean = mean, sd = stand_deviation)
[1] 0.02275013
```

c) :

```
#c
pnorm(90, mean = mean, sd = stand_deviation) - pnorm(80, mean = mean, sd = stand_deviation)
```

```
> pnorm(90, mean = mean, sd = stand_deviation) - pnorm(80, mean = mean, sd = stand_deviation)
[1] 0.1359051
```

d) :

```
1 - pnorm(mean + stand_deviation, mean = mean, sd = stand_deviation) + pnorm(mean - stand_deviation, mean = mean, sd = stand_deviation)
1 - pnorm(mean + 2*stand_deviation, mean = mean, sd = stand_deviation) + pnorm(mean - 2*stand_deviation, mean = mean, sd = stand_deviation)
1 - pnorm(mean + 3*stand_deviation, mean = mean, sd = stand_deviation) + pnorm(mean - 3*stand_deviation, mean = mean, sd = stand_deviation)
```

```
> 1 - pnorm(mean + stand_deviation, mean = mean, sd = stand_deviation) + pnorm(mean - stand_deviation, mean = mean, sd = stand_deviation)
[1] 0.3173105
> 1 - pnorm(mean + 2*stand_deviation, mean = mean, sd = stand_deviation) + pnorm(mean - 2*stand_deviation, mean = mean, sd = stand_deviation)
[1] 0.04550026
> 1 - pnorm(mean + 3*stand_deviation, mean = mean, sd = stand_deviation) + pnorm(mean - 3*stand_deviation, mean = mean, sd = stand_deviation)
[1] 0.002699796
```

e) :

```
#e
qnorm(0.1, mean = mean, sd = stand_deviation)
qnorm(0.9, mean = mean, sd = stand_deviation)
```

```
> qnorm(0.1, mean = mean, sd = stand_deviation)
[1] 87.18448
> qnorm(0.9, mean = mean, sd = stand_deviation)
[1] 112.8155
```

f) :

```
#f
qnorm(0.98, mean = mean, sd = stand_deviation)
```

```
> qnorm(0.98, mean = mean, sd = stand_deviation)
[1] 120.5375
```

g) :

```
#g
set.seed(100)
visitors <- rnorm(10000, mean = mean, sd = stand_deviation)
hist(visitors, main = "plot for 10,000 visitors using the above distribution",xlab = "Money Spent ($)", col = "blue", breaks = 20)
```