

**COURSE: PROBABILITY AND STATISTICS****Experiment: 3**

Name: Riya Dhakaan	Roll no: A044	Batch B-Tech IT
Year II Sem IV	Date 28.01.2026	Grade:

**Objective:**

- Understand and apply discrete probability distributions using R.
- Learn to use R functions to compute probabilities, cumulative probabilities, and simulate random values for discrete distributions.

Code for installing packages:

install.packages("e1071")

Code to load and use package:

library("e1071")

library("distr")

Code to remove package after use:

detach("stats", unload = TRUE)

Syntax for working with discrete data is as follows:

```
ddiscrete(x, probs, values =
  1:length(probs)) pdiscrete(q, probs, values
  = 1:length(probs)) qdiscrete(p, probs,
  values = 1:length(probs)) rdiscrete(n,
  probs, values = 1:length(probs), ...)
```

where

x, q vector or array of quantiles.

p vector or array of probabilities. n number of observations.

Probs probabilities of the distribution. values values of the distribution.

... ignored (only there for backwards compatibility)

These functions provide information about the discrete distribution where the

**COURSE: PROBABILITY AND STATISTICS**

probability of the elements of values is proportional to the values given in probs, which are normalized to sum up to 1. ddiscrete gives the density, pdiscrete gives the distribution function, qdiscrete gives the quantile function and rdiscrete generates random deviates.

```
X=c(0,1,2,3,4) P=c(0.1,0.15,0.2,0.55) XP=X*P
data.frame(X,P,XP)
```

```
mean=sum(XP)
```

**TO FIND THE MISSING VALUE IN A PROBABILITY DISTRIBUTION**

convert the question of solving the equation to finding the root of a function and use the following steps: ]

For example to find root for  $0.6+6x=1$ , use the following code: f <- function(x) (0.6+6\*x-1)  
`uniroot(f, lower=0, upper=1)$root`

**TO FIND THE DISTRIBUTION OF A NEW VARIABLE GIVEN AS A FUNCTION OF RANDOM VARIABLE:**

```
x=c(-1,0,1,2)
```

```
y=x^*
```

```
x+1
```

```
y
```

```
prob = rep(1/4,4)
```

```
tapply(prob,y,sum)
```

---

```
#Install package
install.packages("stats")
```

**COURSE: PROBABILITY AND STATISTICS**

```
#To remove package after
use

detach("stats", unload = TRUE)

#Load library to load and use
package library(e1071)

library(distr)

#Prefix used



- p for "probability", the cumulative distribution function (c. d. f.)
- q for "quantile", the inverse c. d. f.
- 1. d for "density", the density function (p. f. or p. d. f.)
- r for "random", a random variable having the specified distribution



#Frequency table

random=sample(1:10, size=1000, replace = TRUE)
t=table(random)

barplot(t)

#How to enter data

rdiscrete( 30, c('0.2','0.5','0.3') )

rdiscrete( 100, c('0.2','0.5','0.3'), c("A","B","C"))

#Example

y= rdiscrete( 100, c(1/4,2/4,1/4), c(0,1,2))
factor(y)

levels(factor(y))
table((factor(y)))

#To find probability associated to any random variable for
example x=1

ddiscrete(1, c(1/4,2/4,1/4), c(0,1,2))

#Example of rolling of die

# generate the vector of probabilities

probability <- rep(1/6, 6)

# plot the probabilities

barplot(probability, xlab = "outcomes", main = "Probability
Distribution")

# generate the vector of cumulative probabilities
```

**COURSE: PROBABILITY AND STATISTICS**

```

cum_probability <- cumsum(probability)

# plot the probabilities

barplot(cum_probability, xlab = "outcomes", main = "Cumulative
Probability Distribution")

Note: Plots must be customized by using the knowledge of
Practical 2.

#Mean and variance
X=c(0,1,2,3,4)
P=c(0.1,0.15,0.2,0.55)
XP=X*P

data.frame(X,P,XP)
mean=sum(XP)

#Find unknown for 0.6+6x=1

f <- function(x) (0.6+6*x-1)
uniroot(f, lower=0, upper=1)$root

```

**Task 1:**

PDF of random variable X is:

X	1	2	3	4	5	6	7
P(X)	k	2k	3k	$k^2$	$k^2+k$	$2k^2$	$4k^2$

Find  $k$ ,  $P(X < 5)$ ,  $P(1 \leq X \leq 5)$ ,  $E(x)$  and  $V(x)$

Write an R program for the above problem. Also, write an R program to plot the probability distribution and cumulative distribution function.

**Hint:** For creating a vector with cumulative sum of another vector's elements in R use:  
**cumsum(vectorname)**

**Task 1 – Mean (Expected Value) of X**

**If I repeat this experiment many times, what value of X do I expect on average?**

$$k + 2k + 3k + k^2 + (k^2 + k) + 2k^2 + 4k^2 = 1$$

$$8k^2 + 7k - 1 = 0$$

**COURSE: PROBABILITY AND STATISTICS**

$$K = 1/8$$

$$P(x < 5) = 58/64$$

$$\text{Mean} = 201/64$$

$$k = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$k = \frac{-7 \pm \sqrt{49 + 32}}{16}$$

$$k = \frac{-7 \pm \sqrt{81}}{16}$$

$$k = \frac{-7 \pm 9}{16}$$

$$k = \frac{1}{8}$$

Mean:

$$E(X) = \sum x \cdot P(X = x)$$

$$\frac{8 + 32 + 72 + 4 + 45 + 12 + 28}{64} = \frac{201}{64}$$

**COURSE: PROBABILITY AND STATISTICS**

$$E(X) = \frac{201}{64} \approx 3.14$$

**Variance=10607/4096**

**Variance tells us how spread out the values are from the mean**

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

$$E(X^2) = \frac{8 + 64 + 216 + 16 + 225 + 72 + 196}{64} = \frac{797}{64}$$

$$[E(X)]^2 = \left(\frac{201}{64}\right)^2 = \frac{40401}{4096}$$

$$E(X^2) = \frac{797}{64} = \frac{51008}{4096}$$

$$\text{Var}(X) = \frac{51008 - 40401}{4096} = \frac{10607}{4096}$$

$$\boxed{\text{Var}(X) = \frac{10607}{4096} \approx 2.59}$$

**CDF (Cumulative Distribution Function)**

**What is the probability that X is less than or equal to a value?**

**COURSE: PROBABILITY AND STATISTICS**

$$F(x) = P(X \leq x)$$

```
install.packages("e1071")
```

```
library ("e1071")
```

```
library ("distr")
```

Code to remove package after use:

```
detach("stats", unload = TRUE)
```

```
x <- 1:7
```

```
k <- 1/8
```

```
p <- c(k,2*k,3*k,k^2,k^2+k,2*k^2,4*k^2)
```

```
EX <- sum(x*p)
```

```
VX <- sum(x^2*p)-EX^2
```

```
EX
```

```
VX
```

```
sum(p[x<5])
```

```
sum(p[x>=1 & x<=5])
```

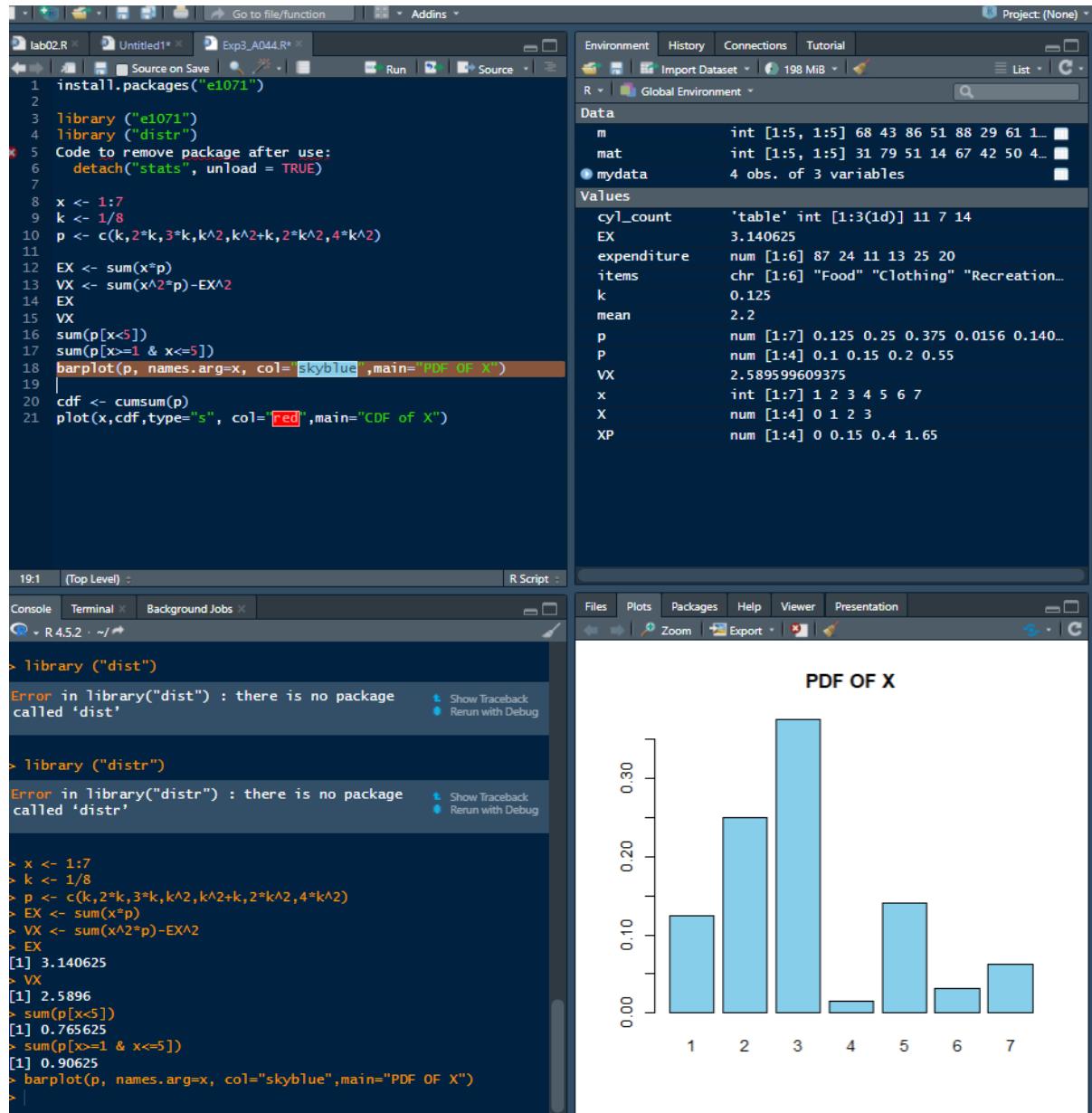
```
barplot(p, names.arg=x, col="skyblue",main="PDF OF X")
```

```
cdf <- cumsum(p)
```

```
plot(x,cdf,type="s", col="red",main="CDF of X")
```

## Mukesh Patel School of Technology Management &amp; Engineering

## COURSE: PROBABILITY AND STATISTICS



## COURSE: PROBABILITY AND STATISTICS

The screenshot shows the RStudio interface with the following components:

- Code Editor:** Contains R code for calculating the PDF and CDF of a random variable X. The code includes library imports, package removal, and statistical calculations.
- Data View:** Shows the Global Environment with variables like m, mat, mydata, cdf, cyl\_count, EX, expenditure, items, k, mean, p, p\_bar, VX, x, X, and XP.
- Console:** Displays errors related to the 'dist' package and the calculated values for various parameters.
- Plots:** A plot titled "CDF of X" showing a step function that increases from 0.1 at x=1 to 1.0 at x=7.

**Task 2:**

A random variable X has the following pdf:

X	-2	-1	0	1	2	3
P(X)	0.1	k	0.2	2k	0.3	3k

Find k, P(X < 2), cdf, mean and variance of X.

Write an R program for the above problem. Also, write an R program to plot the probability distribution and cumulative distribution function.

**COURSE: PROBABILITY AND STATISTICS**

$$\sum P(X = x) = 1$$

$$0.1 + k + 0.2 + 2k + 0.3 + 3k = 1$$

$$k = \frac{1}{15}$$

Find  $P(X < 2)$

$$P(X < 2) = 0.1 + \frac{1}{15} + 0.2 + \frac{2}{15}$$

$$P(X < 2) = 0.3 + 0.2 = 0.5$$

**CDF (Cumulative Distribution Function)**

$$F(x) = P(X \leq x)$$

<b>X</b>	-2	-1	0	1	2	3
<b>F(X)</b>	0.1	1/6	11/30	1/2	0.8	1

**Mean (Expected Value)**

$$E(X) = \sum x \cdot P(X)$$

**COURSE: PROBABILITY AND STATISTICS**

$$E(X) = \frac{20}{15} - \frac{4}{15} = \frac{16}{15}$$

$$E(X) = \frac{16}{15}$$

**Variance of X**

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

$$E(X^2) = \sum x^2 P(X)$$

$$\begin{aligned} E(X^2) &= \frac{6}{15} + \frac{1}{15} + 0 + \frac{2}{15} + \frac{18}{15} + \frac{27}{15} \\ &= \frac{54}{15} \end{aligned}$$

$$[E(X)]^2 = \left(\frac{16}{15}\right)^2 = \frac{256}{225}$$

$$\frac{54}{15} = \frac{810}{225}$$

$$\text{Var}(X) = \frac{810}{225} - \frac{256}{225} = \frac{554}{225}$$

$$\text{Var}(X) = \frac{554}{225}$$

**COURSE: PROBABILITY AND STATISTICS**

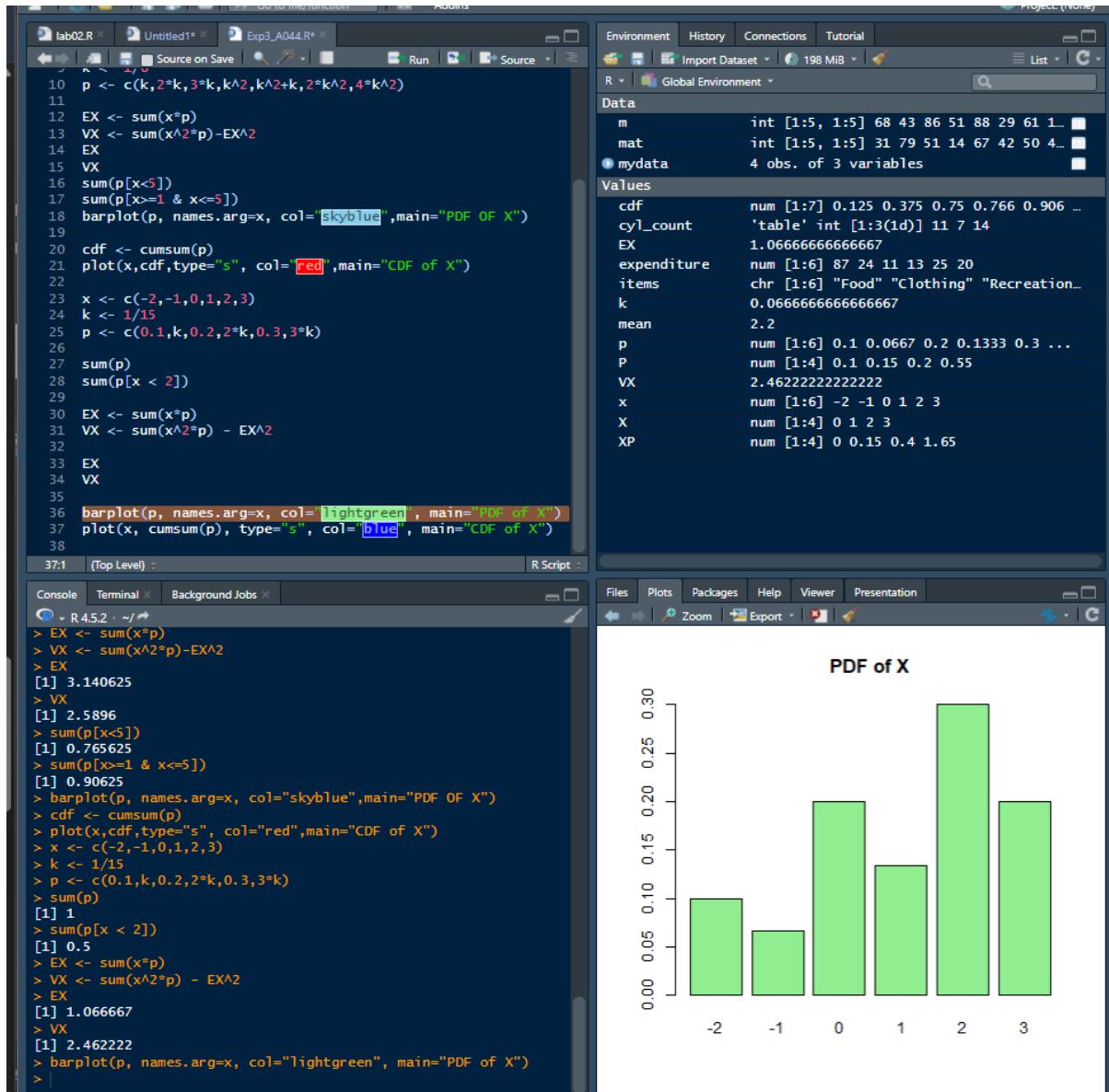
```
x <- c(-2,-1,0,1,2,3)
k <- 1/15
p <- c(0.1,k,0.2,2*k,0.3,3*k)
```

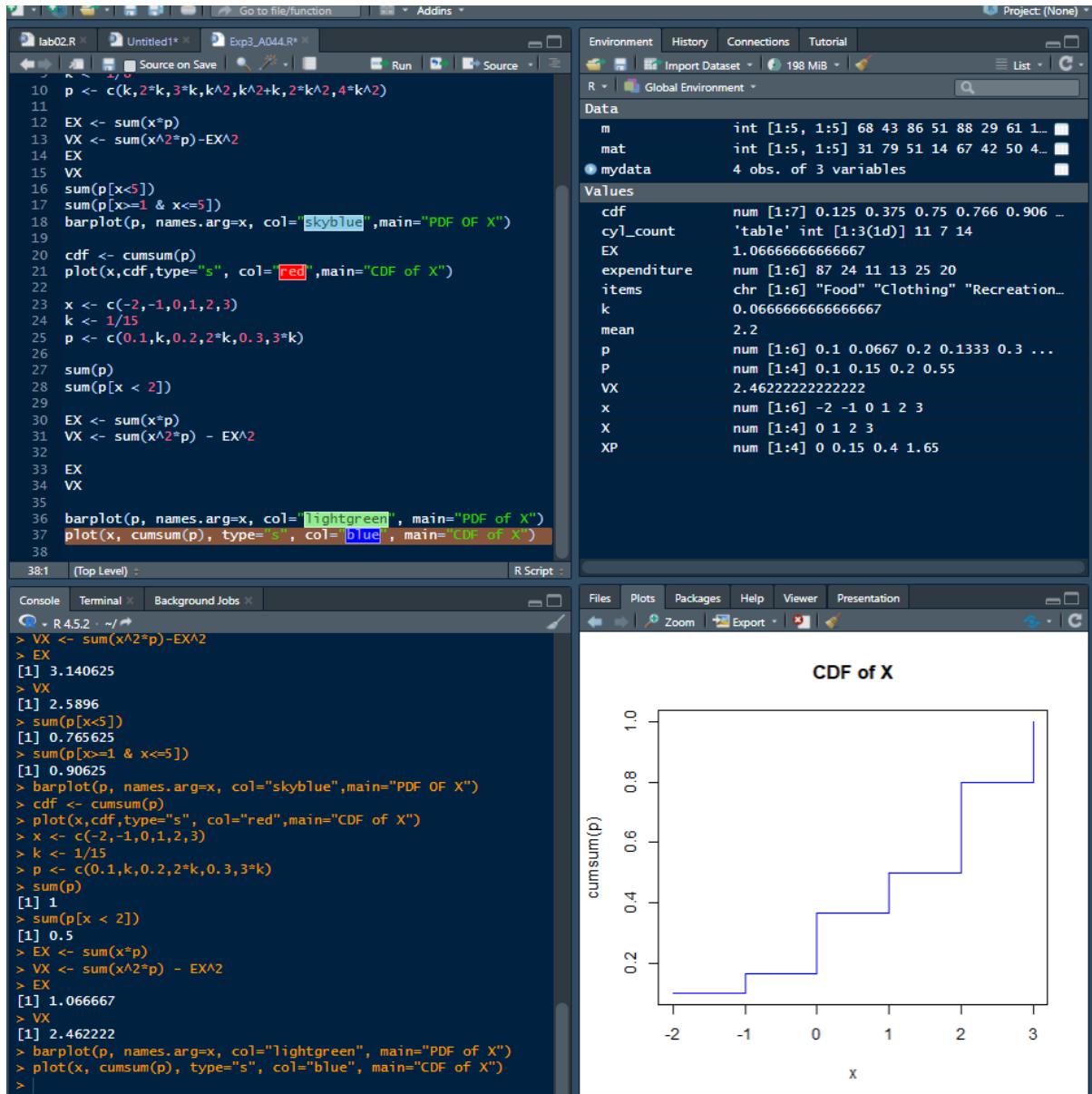
```
sum(p)
sum(p[x < 2])
```

```
EX <- sum(x*p)
VX <- sum(x^2*p) - EX^2
```

```
EX
VX
```

```
barplot(p, names.arg=x, col="lightgreen", main="PDF of X")
plot(x, cumsum(p), type="s", col="blue", main="CDF of X")
```

**COURSE: PROBABILITY AND STATISTICS**

**COURSE: PROBABILITY AND STATISTICS****Task 3:**

A RV X has the following probability distribution:

X	-2	-1	0	1	2
P(X=x)	1/5	1/5	2/5	2/15	1/15

Find the probability distribution of Y = X<sup>2</sup>+1. Also find mean and variance of Y.

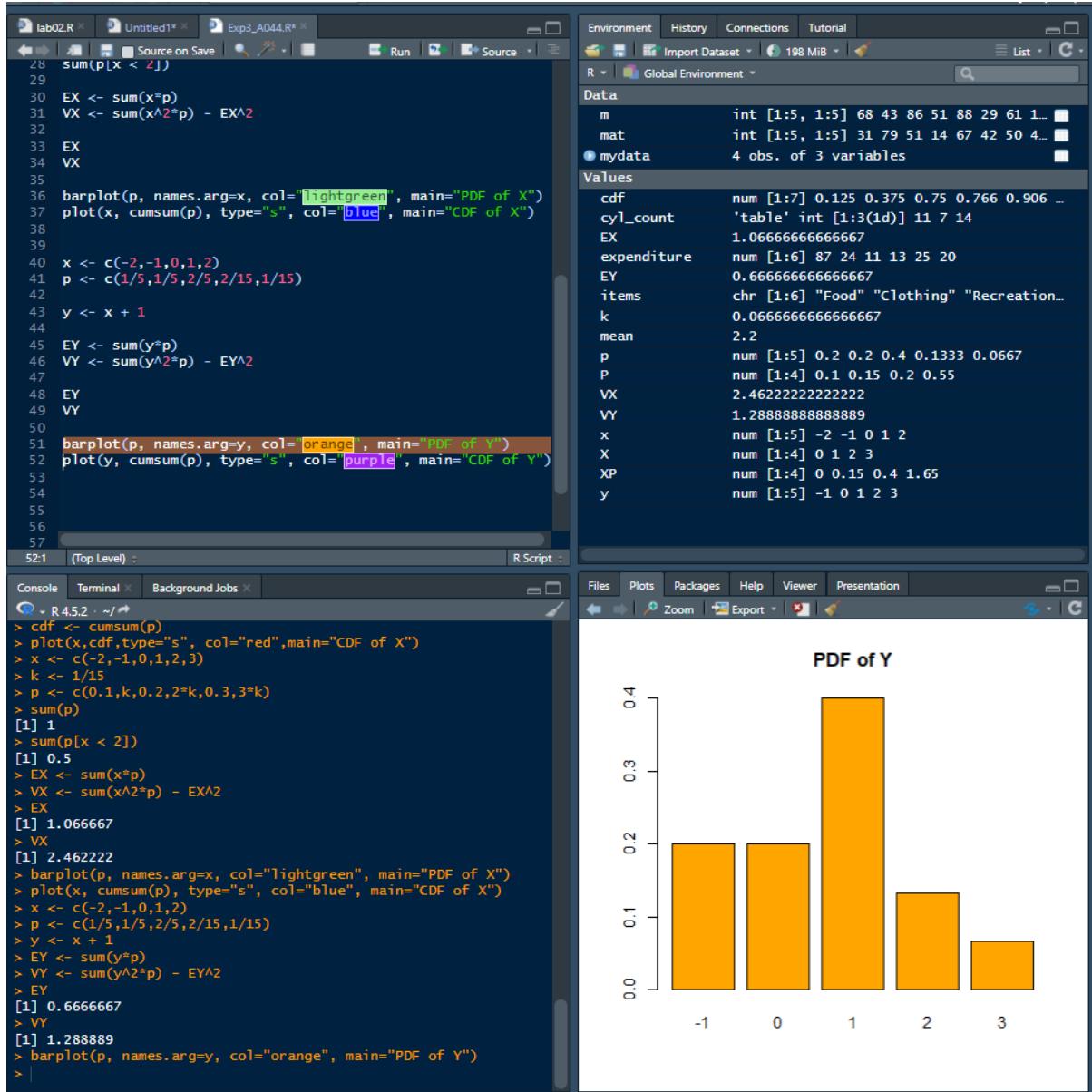
**COURSE: PROBABILITY AND STATISTICS**

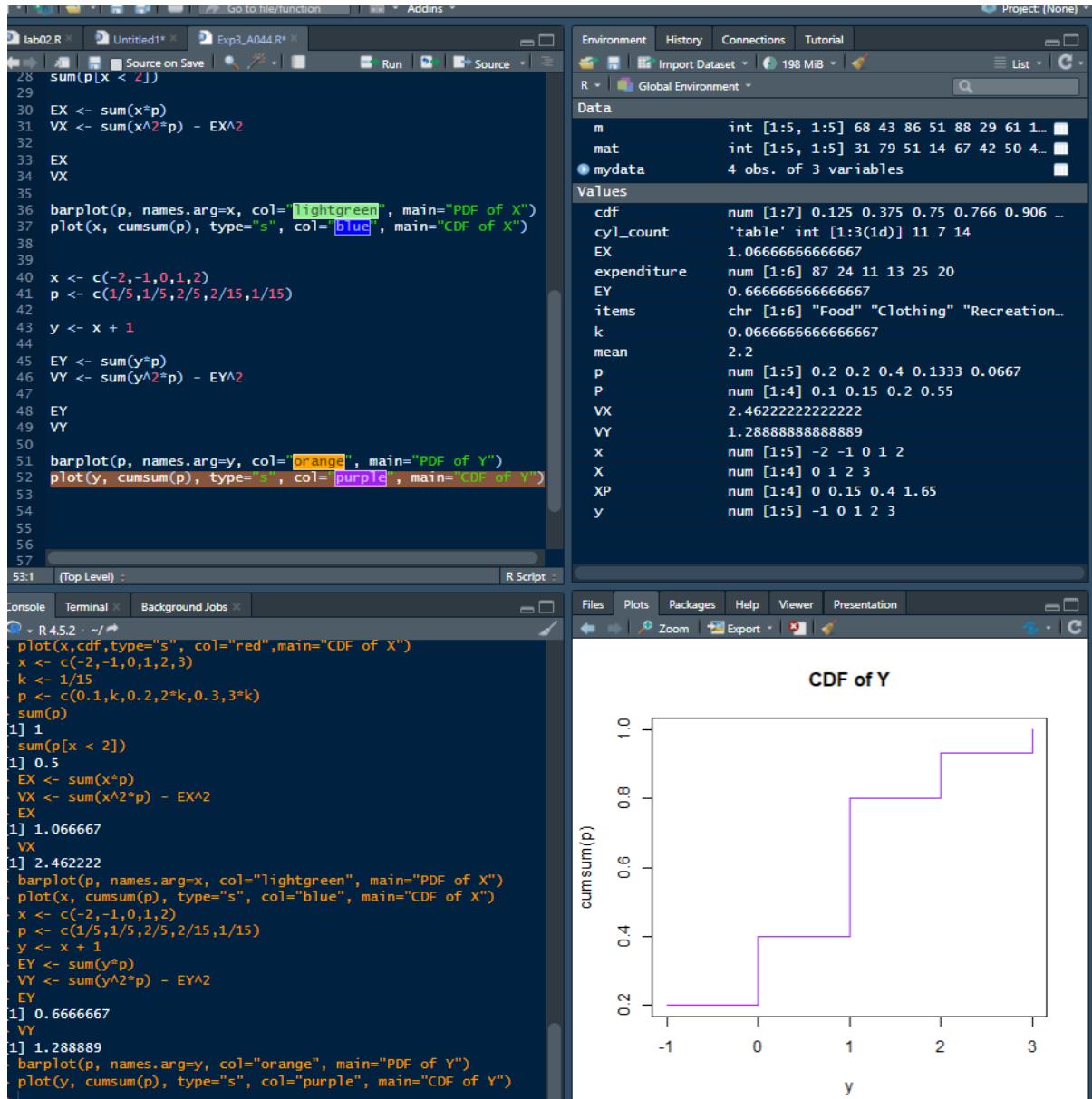
Write an R program for the above problem. Also, write an R program to plot the probability distribution and cumulative distribution function.

X	$X^2$	$Y = X^2 + 1$
-2	4	5
-1	1	2
0	0	1
1	1	2
2	4	5

## Mukesh Patel School of Technology Management & Engineering

### COURSE: PROBABILITY AND STATISTICS



**COURSE: PROBABILITY AND STATISTICS****Task 4:**

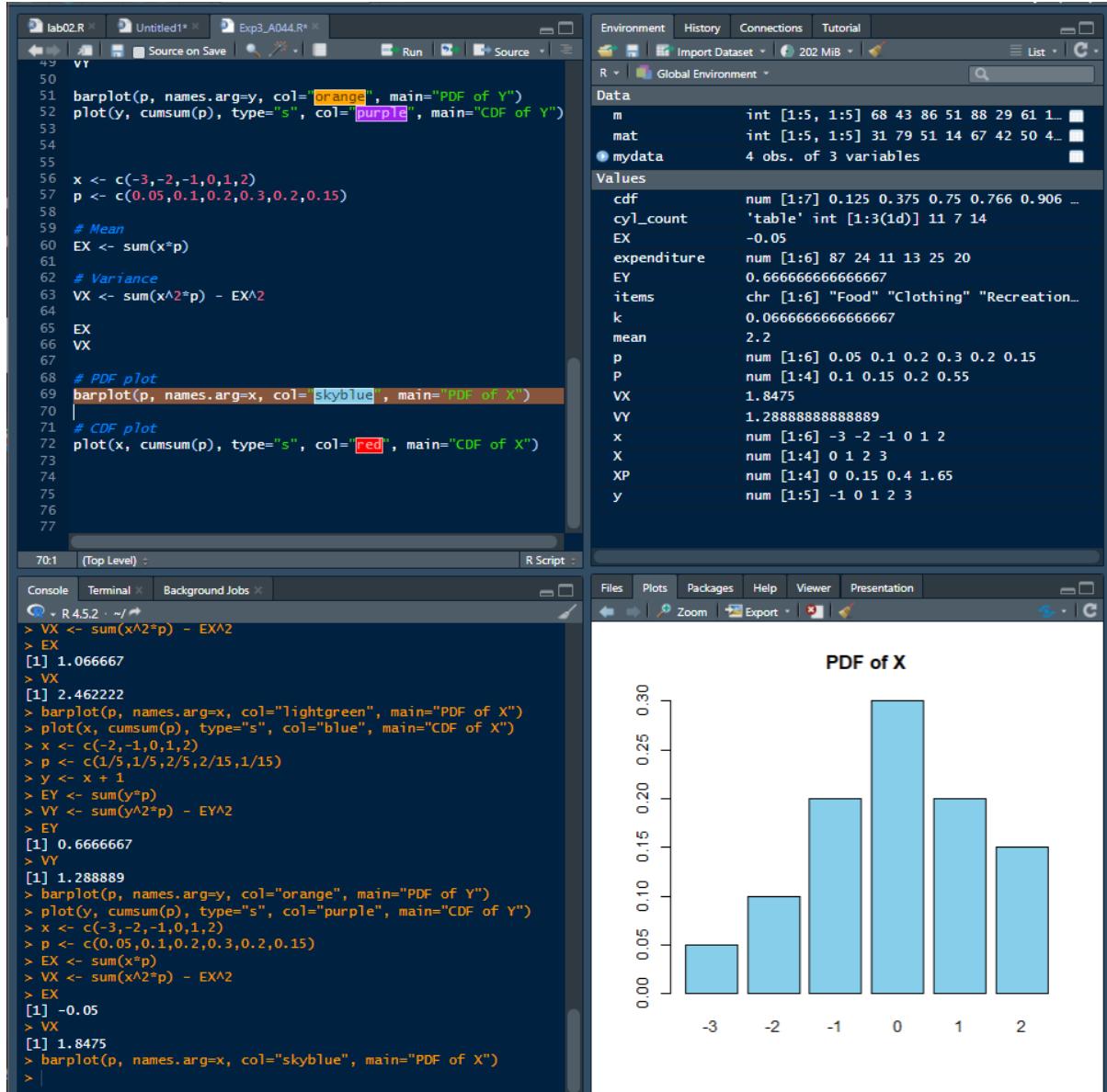
Given the following distribution:

X	-3	-2	-1	0	1	2
P(X=x)	0.05	0.1	0.2	0.3	0.2	0.15

Find Mean and Variance.

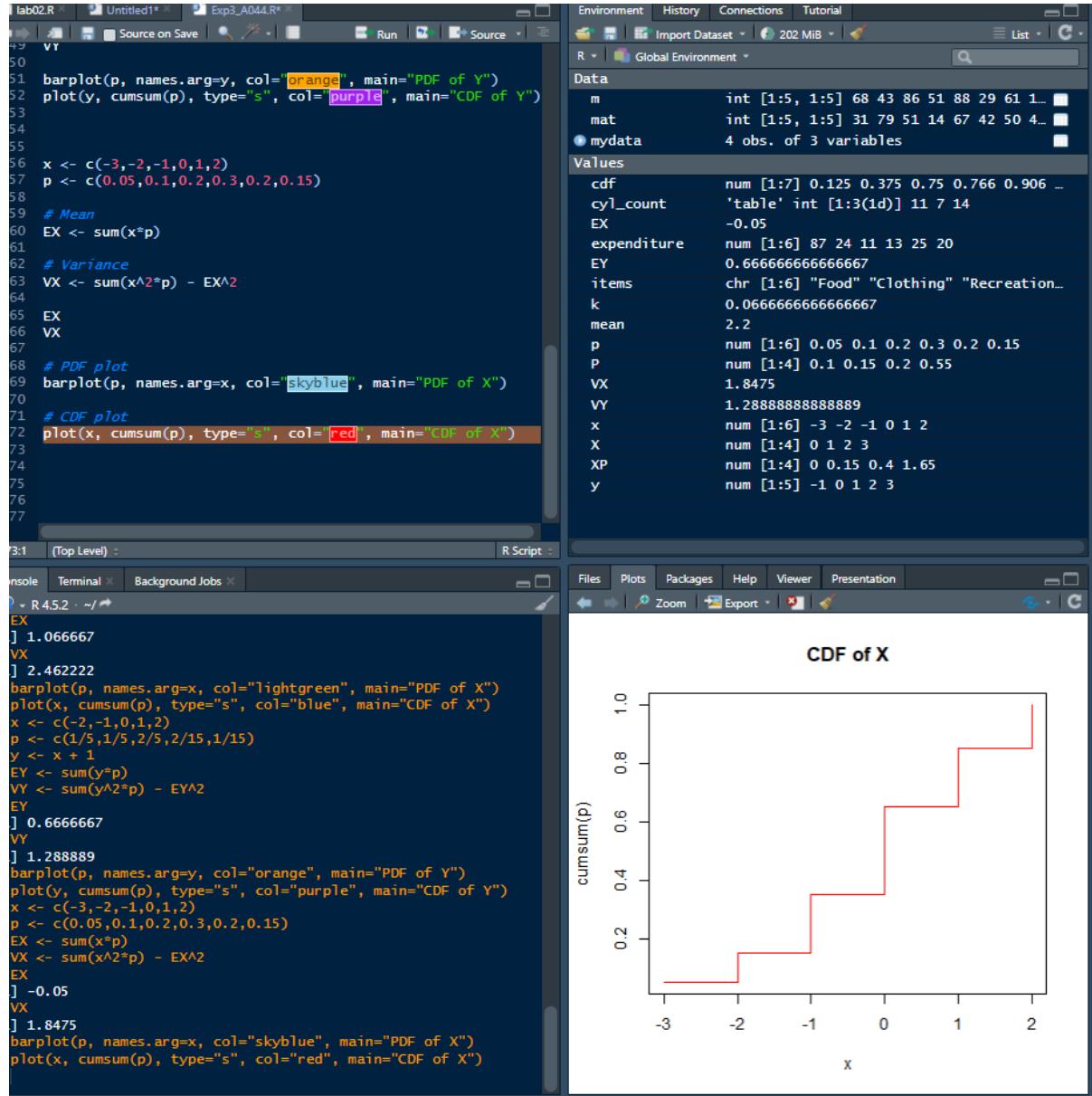
Write an R program for the above problem.

### COURSE: PROBABILITY AND STATISTICS



## Mukesh Patel School of Technology Management & Engineering

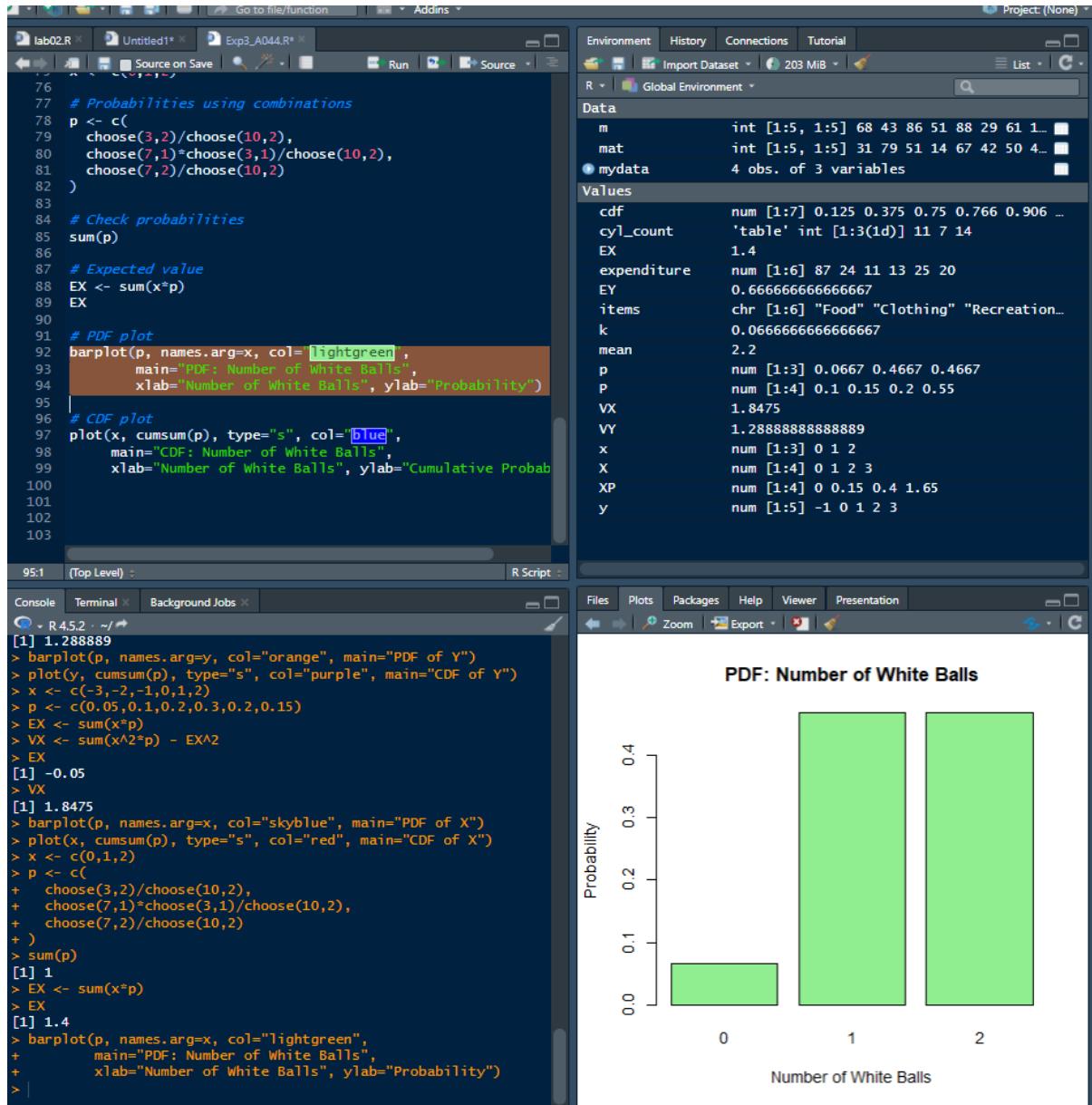
### COURSE: PROBABILITY AND STATISTICS



### Task 5:

An urn contains 7 white and 3 red balls. Two balls are drawn together, at random from this urn. Compute the expected number of white balls drawn.

Using an R program, find the expectation of the above problem. Also write a program to plot probability distribution and cumulative probability distribution. Hint: For calculating combination in R use: choose(n,r)

**COURSE: PROBABILITY AND STATISTICS**

## COURSE: PROBABILITY AND STATISTICS

The figure shows a screenshot of the RStudio interface. The left pane contains an R script named 'Exp3\_A044.R' with code related to probability calculations and plotting. The right pane shows the Global Environment and a plot titled 'CDF: Number of White Balls'. The plot displays a step function representing the cumulative distribution of the number of white balls.

**Script (Exp3\_A044.R):**

```
76  
77 # Probabilities using combinations  
78 p <- c(  
79   choose(3,2)/choose(10,2),  
80   choose(7,1)*choose(3,1)/choose(10,2),  
81   choose(7,2)/choose(10,2)  
82 )  
83  
84 # Check probabilities  
85 sum(p)  
86  
87 # Expected value  
88 EX <- sum(x*p)  
89 EX  
90  
91 # PDF plot  
92 barplot(p, names.arg=x, col="lightgreen",  
93   main="PDF: Number of White Balls",  
94   xlab="Number of White Balls", ylab="Probability")  
95  
96 # CDF plot  
97 plot(x, cumsum(p), type="s", col=blue,  
98   main="CDF: Number of White Balls",  
99   xlab="Number of White Balls", ylab="Cumulative Probability")  
00  
01  
02  
03
```

**Console:**

```
0:1 (Top Level) : R Script :  
sole Terminal X Background Jobs X  
R 4.5.2 · ~  
x <- c(0.05,0.1,0.2,0.3,0.2,0.15)  
EX <- sum(x*p)  
X <- sum(x^2*p) - EX^2  
EX  
-0.05  
X  
1.8475  
barplot(p, names.arg=x, col="skyblue", main="PDF of X")  
plot(x, cumsum(p), type="s", col="red", main="CDF of X")  
x <- c(0,1,2)  
y <- c(  
  choose(3,2)/choose(10,2),  
  choose(7,1)*choose(3,1)/choose(10,2),  
  choose(7,2)/choose(10,2),  
  1  
  EX <- sum(x*p)  
  EX  
  1.4  
barplot(p, names.arg=x, col="lightgreen",  
  main="PDF: Number of White Balls",  
  xlab="Number of White Balls", ylab="Probability")  
plot(x, cumsum(p), type="s", col="blue",  
  main="CDF: Number of White Balls",  
  xlab="Number of White Balls", ylab="Cumulative Probability")  
0:1
```

**Data View:**

Object	Type	Value
m	int	[1:5, 1:5] 68 43 86 51 88 29 61 1 ...
mat	int	[1:5, 1:5] 31 79 51 14 67 42 50 4 ...
mydata	4 obs. of 3 variables	

**Values View:**

Object	Type	Value
cdf	num	[1:7] 0.125 0.375 0.75 0.766 0.906 ...
cyl_count	'table'	int [1:3(1d)] 11 7 14
EX	1.4	
expenditure	num	[1:6] 87 24 11 13 25 20
EY	0.6666666666666667	
items	chr	[1:6] "Food" "Clothing" "Recreation..."
k	0.0666666666666667	
mean	2.2	
p	num	[1:3] 0.0667 0.4667 0.4667
P	num	[1:4] 0.1 0.15 0.2 0.55
VX	1.8475	
VY	1.28888888888889	
x	num	[1:3] 0 1 2
X	num	[1:4] 0 1 2 3
XP	num	[1:4] 0 0.15 0.4 1.65
y	num	[1:5] -1 0 1 2 3

**Plot:**

**CDF: Number of White Balls**

The plot shows the Cumulative Distribution Function (CDF) for the number of white balls. The x-axis is labeled 'Number of White Balls' and ranges from 0.0 to 2.0. The y-axis is labeled 'Cumulative Probability' and ranges from 0.0 to 1.0. The CDF is a step function that remains at 0.0 until x=1.0, then jumps to approximately 0.5, and finally reaches 1.0 at x=2.0.

## **Home Work Questions:**

Solve all the above questions manually.