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 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.

Note:Prefix used

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

To get Frequency table by random generation of data:

Sample() function is used to generate the random elements from the given data with or without replacement.

Syntax: sample(data, size, replace = FALSE, prob = NULL)

where

- data can be a vector or a dataframe
- size represents the size of the sample
- replace is used to set the values again repeated if it is set to true
- prob: a vector of probability weights for obtaining the elements of the vector being sampled

Examples:

1. random=sample(1:10, size=1000, replace = TRUE)
   t=table(random)
   barplot(t)
2. Also, we can use rdiscrete: a vector of length 30 whose elements are 1 with probability 0.2 and 2 with probability 0.8.
   rdiscrete (30, c(0.2, 0.8))
3. a vector of length 100 whose elements are A, B, C, D.The probabilities of the four values have the relation 1:2:3:3
   rdiscrete (100, c(1,2,3,3), c("A","B","C","D"))
4. rdiscrete(30, c(‘0.2’,’0.5’,’0.3’))
5. rdiscrete(100, c(‘0.2’,’0.5’,’0.3’), c(“A”,”B”,”C”))
6. y= rdiscrete(100, c(1/4,2/4,1/4), c(0,1,2))

   factor(y)

   levels(factor(y))

   table((factor(y)))

7. To find probability associated to any random variable for example x=1
8. ddiscrete(1, c(1/4,2/4,1/4), c(0,1,2))
9. Example of rolling of die

   probability <- **rep**(1/6, 6) *# generate the vector of probabilities*

PLOT THE PROBABILITIES

- **barplot**(probability, xlab = "outcomes", main = "Probability Distribution")
- *generate the vector of cumulative probabilities*

  cum_probability <- **cumsum**(probability)

- plot the probabilites

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

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

TO FIND MEAN AND VARIANCE OF A DISCRETE DATA

```
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)
```

EXERCISE (Programing and problem solving)

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 $k, P(X<5), P(1\le X\le 5)$

Write a R program for the above problem. Write R program to plot probability distribution and cumulative distribution.

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), c.d.f.

Write a R program for the above problem. Also write a R program to plot cumulative distribution function.

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 $V = X^2 + 1$.

Write a R program for the above problem. Write R program to plot probability distribution and cumulative distribution.

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 a R program for the above problem. Write R program to plot probability distribution and cumulative distribution.

5.

The pmf of a RV X is zero except at the points X=0, 1, 2. At these points $P(0) = 3c^3, P(1) = 4c - 10c^2, P(2) = 5c - 1$. Determine

i. c
ii. Find $P(X < 1), P(1 < X \leq 2), P(0 < X \leq 2)$

1/3,1/9,2/3,8/9