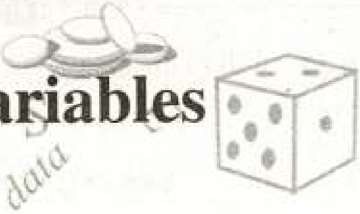
7 Random Variables



$0 \leq P(A) \leq 1$

7.1. Introduction

Every random experiment results in two or more outcomes and usually the interest is in a particular aspect of the outcomes of the experiment. For example, when a pair of dice is thrown, the interest may be in the total of upturned dots on both dice. In case of this experiment, the total may be 2 (one on each die), 3 (one on the first die and two on the other) and so on. It may be 4, 5, 6, 7, 8, 9, 10, 11 or 12. In the language of probability, these values associated with outcomes are the values of a so-called random variable. An other example is the total number of children in each of the fifty randomly chosen families. If no family has more than 5 children then the values of this random variable i.e., number of children in each family, would be 0,1,2,3,4,5 i.e., no child, one child, two children, 3 children, 4 children and five children respectively.

A variable whose values depend upon the outcomes of a random experiment is called a **random variable**. We will denote the random variable by the capital letters X, Y or Z and their values by the corresponding small letters x, y or z.

Example 7.1 : Let a pair of dice be thrown and Y denote the random variable that is the sum of upturned values on the two dice. There are 36 outcomes and Y assigns to the outcome (1, 1) the real number $1 + 1 = 2$. It assigns to the outcome (2, 1), the real number $2 + 1=3$ and so on uptil the outcome (6, 6), the real number $6 + 6 = 12$ so, the values assigned are 2,3,4,5,6,7,8,9,10,11 and 12.

7.2 Random numbers and their generation

Random numbers are a sequence of digits from the set $\{0, 1, 2, ..., 9\}$. So that, at each position in the sequence, each digit has the same probability 0.1 of being selected irrespective of the actual sequence, so far constructed. The probability is 0.1

because out of ten digits {0, 1, 2, ..., 9} each digit has equal probability i.e., 1/10 or 0.1. These are also known as random digits.

The simplest ways of achieving such numbers are games of chance such as dice, coins, cards or by repeatedly drawing numbered slips out of a hat. These are usually grouped purely for convenience of reading but this would become very tedious for long runs of each digits. Fortunately tables of random digits are now widely available (see table 7.1).

For implementation on computers to provide sequences of such digits easily and quickly, the most common methods are called **Pseudo random techniques**. Here, digits will re-appear in the same order (i.e., cycle) eventually but for a good technique the cycle might be tens of thousands of digits long. Of course the **Pseudo random digits** as the title says, are not truly random. In fact, they are completely deterministic but they do exhibit most of the properties of random digits.

Generally, these methods involve the recursive formula as

$$x_{n+1} = ax_n + b \pmod{m};\ n = 0,1,2,3, \ldots \qquad \ldots. \qquad (7.1)$$

Here a, b and m are suitably chosen integer constants and the seed x_1 (a starting number) is an integer. By mod m we means that if the answer is greater than m, then divide it by m and keep the remainder as a random number. Use of this formula gives rise to a sequence of integers each of which is in the range 0 to m-1. We simply run these together to give our sequence of pseudo random digits. Clearly this to be of any value; m, a and / or b should be large.

Example 7.2: Let $a = 13$, $b = 0$ and $m = 16$. Generate 4 random numbers.

Solution: According to the relation

$$x_{n+1} = ax_n + b \pmod{16} \text{ for } n = 0, 1, 2, \ldots$$

Let a seed x_0 be 5, then for $n = 0$, we have

$$\begin{aligned} x_1 &= 13x_o + b \pmod{16} \\ &= 13(5) + 0 \pmod{16} \\ &= 65 \pmod{16} \\ &= 1 \text{(dividing 65 by 16, the remainder is 1)} \end{aligned}$$

For $n = 1$, we have,

$$\begin{aligned} x_2 &= 13(1) + 0 \pmod{16} \\ &= 13 \end{aligned}$$

For $n = 2$, we have,

$$\begin{aligned} x_3 &= 13(13) + 0 \pmod{16} \\ &= 9 \end{aligned}$$

Similarly, for $n = 3$, we have,

$$\begin{aligned} x_4 &= 13(9) + 0 \pmod{16} \\ &= 5 \end{aligned}$$

So, the random numbers are 1, 13, 9, 5.

7.3 Application of random numbers

The random numbers have widely applicability in the simulation techniques (also called Monte Carlo Methods) which have been applied to many problems in the various sciences and are useful in the situations where direct experimentation is not possible, the cost of conducting an experiment is very high or the experiment takes too much time.

The random number tables are constructed from the random numbers where, the random numbers are grouped for the purposes of reading. The groups may consist of 2-digit, 3-digit, 4-digit or 5-digit sequence of random numbers (0, 1, 2, 3, 4, 5, 6, 7, 8, 9). A five digit sequence of these are given in table 7.1. We may use this table for 1-digit random numbers, 2-digit random numbers and so on uptil 5-digit random numbers depending upon the situation by using it row-wise or column-wise.

Table 7.1: Random Numbers

3 3 6 8 1	7 5 0 2 9	3 0 0 7 8	0 6 5 5 8	2 4 3 6 6
3 2 4 8 9	1 2 8 9 5	5 8 6 2 3	2 2 5 7 6	2 9 6 8 3
6 8 0 7 8	9 4 2 0 0	2 5 4 8 8	5 0 2 2 7	2 5 2 9 6
4 7 6 3 8	9 3 1 2 1	5 7 8 7 7	0 5 3 7 2	0 3 4 8 8
8 0 6 4 5	8 1 4 7 4	6 7 7 3 4	1 1 4 2 1	0 7 3 4 0
5 5 2 7 1	8 0 3 5 4	7 3 9 6 0	4 4 1 8 7	2 8 2 9 2
7 3 4 5 7	8 1 9 0 5	1 7 2 0 8	8 3 3 3 4	2 7 4 9 7
7 8 2 1 2	3 8 0 7 5	1 8 5 6 2	7 8 9 7 0	9 9 9 2 1
2 0 3 9 3	4 1 2 4 7	5 9 7 7 3	9 2 4 3 7	7 6 1 2 6
5 9 2 5 3	4 7 6 5 3	6 6 1 1 9	9 7 6 8 4	2 1 4 4 0
0 9 3 2 9	7 9 7 6 8	7 2 9 6 8	5 6 2 9 7	1 7 5 6 1
1 2 8 6 1	6 2 8 6 4	9 9 8 3 2	6 3 5 4 3	9 5 0 0 5
5 8 3 5 3	4 3 9 4 4	2 9 6 8 3	1 2 2 9 3	2 5 1 2 5
9 1 5 5 8	2 9 1 2 5	2 7 8 5 2	7 9 3 3 4	3 9 7 6 8
4 0 3 8 0	4 6 3 4 0	1 4 7 0 6	4 7 2 9 3	7 5 4 7 9
2 9 1 1 9	8 8 3 4 5	2 7 2 5 7	9 9 3 5 5	5 9 6 3 3
6 2 0 8 1	2 0 7 6 2	8 8 5 4 3	9 5 2 3 9	9 7 7 2 9
1 5 2 2 8	4 5 8 8 4	9 3 1 1 9	3 0 9 9 2	6 8 4 8 3
2 9 2 2 8	1 4 7 0 4	4 2 1 6 7	4 2 0 2 0	4 7 8 3 8
8 5 9 6 6	0 9 4 8 3	2 0 8 0 4	6 7 6 0 9	9 0 2 8 6
5 5 6 1 2	4 0 2 8 2	9 3 9 5 6	6 6 2 7 8	1 1 2 3 8
2 2 6 5 2	5 9 4 1 1	0 6 0 6 8	3 9 0 6 1	7 8 0 1 9
6 1 5 2 1	1 7 4 9 6	7 4 3 2 8	7 8 2 1 3	6 5 4 3 9
2 4 3 2 4	7 3 7 6 9	3 9 1 1 6	3 4 8 6 5	4 3 5 5 7
9 4 8 1 4	1 6 8 1 7	5 2 6 0 7	7 7 0 8 2	8 4 4 9 0
9 1 3 3 6	0 7 1 6 7	0 9 0 9 9	0 2 7 3 1	3 4 5 6 3
3 3 3 4 8	4 1 1 0 8	2 0 1 3 6	3 7 1 5 6	4 5 7 4 6
8 4 2 3 6	8 5 4 0 9	0 6 9 8 7	4 7 4 1 5	9 1 6 0 4
9 0 0 4 5	6 8 3 2 7	2 3 3 9 6	5 1 3 0 8	8 2 3 1 9
7 9 9 7 8	9 8 6 0 0	3 4 2 6 0	4 8 6 3 2	8 0 9 2 7
7 8 4 0 8	2 4 7 4 9	5 0 4 7 8	7 0 8 0 7	5 4 4 4 2
6 2 9 4 0	4 3 7 8 2	0 8 3 1 3	2 8 6 9 1	8 5 0 5 9
5 2 0 5 2	6 4 1 7 9	7 4 4 5 4	3 5 1 8 9	5 3 8 7 3
2 5 0 3 4	6 6 6 9 2	3 0 9 7 5	7 3 2 4 2	3 8 4 0 8
3 6 4 5 3	5 7 2 5 0	6 5 3 8 6	8 0 3 3 9	9 0 9 2 5
7 7 3 7 0	7 5 7 3 1	6 5 0 0 1	8 5 9 5 0	9 4 5 0 0
6 3 3 7 1	3 6 5 4 6	8 2 9 3 6	0 1 7 7 5	6 0 4 3 6
6 1 5 5 9	2 3 9 7 0	0 5 6 8 1	8 4 0 7 8	0 4 0 8 7
9 9 4 3 3	4 9 2 2 0	5 8 1 9 3	6 7 3 8 2	4 8 3 2 7
6 3 8 0 4	5 2 0 0 3	2 9 0 6 6	7 2 3 8 0	5 6 3 7 8
7 8 5 4 3	6 6 1 2 9	5 3 1 3 8	3 6 2 4 2	7 2 1 2 1
7 4 9 6 3	8 9 8 5 0	8 7 4 9 6	9 4 7 6 2	5 2 0 9 9
8 5 8 2 2	9 9 4 1 9	1 8 6 3 6	6 7 8 1 2	7 4 3 7 2
2 2 2 7 0	0 1 3 5 1	7 9 2 3 8	7 3 6 6 3	6 0 5 8 4
1 5 9 2 3	5 2 4 6 5	9 5 4 7 6	7 6 7 8 5	3 8 5 9 1
1 8 5 2 6	3 3 1 6 8	5 6 9 6 8	1 2 1 8 5	3 3 3 5 0
8 2 5 9 0	0 3 9 7 8	6 4 5 4 7	2 9 5 1 6	2 8 3 8 2
0 3 8 9 8	9 9 6 7 5	7 0 5 0 9	7 2 4 6 2	5 2 6 6 6
4 8 1 5 3	8 8 3 0 6	6 5 8 5 7	0 5 1 3 6	9 9 8 2 2
9 6 9 5 5	6 1 5 9 3	2 0 4 9 8	2 9 4 9 7	9 1 3 5 1
1 2 0 0 6	9 3 3 0 9	6 4 9 9 6	4 8 8 6 8	3 7 2 2 3
5 6 9 9 7	9 8 9 7 8	3 8 7 3 0	0 6 6 9 2	4 2 4 6 9
9 3 9 5 3	1 6 9 7 9	9 2 4 0 8	5 5 6 9 7	3 0 6 5 8
5 4 8 7 7	5 8 4 9 6	4 1 8 8 0	6 7 2 0 5	9 3 0 2 2
2 1 7 7 5	8 6 0 3 2	4 4 9 9 0	6 5 5 0 1	9 9 9 0 2

To explain the use of a random number table, consider the following example 7.3.

Example 7.3: Count the number of heads and tails when a single coin is thrown 12 times without throwing a coin.

Solution: The first step is to number both the heads and tails staying within 0 to 9. Let 0, 2, 4, 6, 8 (even numbers including zero) indicate heads and the odd numbers 1, 3, 5, 7, 9 represent tails.

The second step is to open the random number table given in table 7.1 and select arbitrarily a column, say column 1 and select arbitrarily a row say row 4 and start reading a set of 1-digit i.e., 4, 8, 5, 7, 7, 2, 5, 0, 1, 5, 9, 4. Third step is to interpret them as H, H, T, T, T, T, H, T, H, T, T, T, H as was decided head for 0, 2, 4, 6, 8 and tail for 1, 3, 5, 7, 9. So, there are 5 heads and 7 tails. These are given in the following frequency table.

Outcome	**frequency**
Head (one head)	5
Tail (0 head)	7

Similarly. to count the number of heads when two coins are thrown 20 times, the first step is to take two-digit random numbers. If both are even including zero, then both are heads; if both are odd then both are tails (0 head), and if one is even and one is odd then there is one head. The second step is to open the random number table and select arbitrarily a column say column 2 and select arbitrarily a row say row 2 and start reading two digit numbers i.e., 24, 80, 76, 06, 52, 34, 92, 03, 92, 93, 28, 83, 15, 03, 91, 20, 52, 92, 59, 56. Third step is to interpret them as HH, HH, TH, HH, TH, TH, TH, HT, TH, TT, HH, HT, TT, HT, TT, HH, TH, TH, TT, TH. So the frequency table would be

Outcome	f
0 head	4
1 head	11
2 head	5

Another method known as probability proportional to size (PPS) is used when we have information about the probabilities of the outcomes.

Example 7.4: If a coin is thrown thrice or three coins are tossed once) then the number of heads may be 0 (all tails) 1, 2 and 3 and the corresponding relative frequencies would be as given in adjoining table.

Find the sequence of the number of heads 20 times without throwing a coin?

Solution: Here we use 3-digits random numbers because probabilities are given to three decimals.

Number of heads	Probability
0	0.125
1	0.375
2	0.375
3	0.125

The first step is to make a cumulative probabilities (c.p) column.

The second step is to assign random numbers (*r*.no) from 0 to 999 because we have 3-digit probabilities. Note that in each class the random number assigned is one less than the number formed by the corresponding cumulative frequency. The reason is that the random numbers are from 0 to 999 not from 1 to 1000. These are shown in the table 7.2.

Table 7.2: cumulative probabilities and range.

Heads	Prob.	c.p.	r.no.
0	0.125	0.125	000-124
1	0.375	0.500	125-499
2	0.375	0.875	500-874
3	0.125	1.000	875-999

The third step is to open the random number table 7.1 and select arbitrarily a column, say column 4, and select arbitrarily a row, say row 2, the 3-digit random numbers are 441, 833, 789, 924, 976, 562, 635, 122, 793, 472, 993, 952, 309, 420, 676, 662, 390, 782, 348, 773 and the corresponding number of heads are 1,2, 2, 3, 3, 2, 2, 0, 2, 1, 3, 3, 1, 1, 2, 2, 1, 2, 1 and 2.

441 indicates 1 head because it is in the class corresponding to 1 head.

833 indicates 2 heads because it is in the class corresponding to 2 heads.

789 indicates 2 heads because it is in the class corresponding to 2 heads.

924 indicates 3 heads because it is in the class corresponding to 3 heads and so on. The corresponding frequency table is as given in table 7.3.

Table 7.3: Frequency Distribution

Heads	0	1	2	3
f	1	6	9	4

7.4 Concept of random variables and their construction from different fields.

As we are familiar by now that random variables arise from the outcomes of random experiments by associating a value to each outcome. The following examples may help explain them in detail.

Example 7.5: Consider an experiment in which three students in a class are asked to take one of the two courses Biology (B) or Computer Science (C).

Solution: Define the random variable Y by

Y = The number of students taking computer science.

The possible values are

No. student takes computer science so, $y = 0$

One student takes computer science so, $y = 1$

Two students take computer science so, $y = 2$

Three students take computer science so, $y = 3$

The possible outcomes of the experiment are:

BBB first takes biology, second takes biology, third takes biology.

CBB first takes computer science, second takes biology third takes biology.

BCB first takes biology, second takes computer science, third takes biology.

BBC first takes biology, second takes biology, third takes computer science.

CCB first takes computer science, second takes computer science, third takes biology.

CBC first takes computer science, second takes biology, third takes computer science.

BCC first takes biology, second takes computer science, third takes computer science.

CCC first takes computer science, second takes computer science, third takes computer science.

These are represented as follows:

BBB	CBB	BCB	BBC	CCB	CBC	BCC	CCC
0	1	1	1	2	2	2	3

Note that the values of the random variable Y are isolated points 0, 1, 2 and 3.

Example 7.6 : Consider an experiment of recording the time (minutes) taken by the customers to wait for its turn in a utility store while standing in a queue.

Define the random variable Y where Y is the time taken by the customers.

Solution: The first customer may take 5.0 minutes, the second may takes 6.0 minutes, the third may take 5.30 minutes, fourth may takes 12.0 minutes and so on.

It may be noted that Y may take any value in an interval on a number line.

7.5 Discrete and continuous random variables

Discrete random variable

A random variable is called discrete if the set of values it takes is a collection of isolated points on a real line i.e., the sample space S is a discrete sample space.

The outcomes of an experiment are noted and the value of the random variable is a number appropriately assigned to each outcome by some rule.

Example 7.7: Three coins are tossed and let the random variable Y denote the number of heads. Then the outcomes and the values of Y are given in adjoining table.

Outcomes	Value of Y
HHH	3
HHT	2
HTH	2
HTT	1
THH	2
THT	1
TTH	1
TTT	0

The values of Y are whole numbers. So, the sample space is discrete. Thus Y is a discrete random variable.

Example 7.5 above is also another example of a discrete random variable.

Continuous random variable: A random variable is called continuous if the set of values it takes is an entire interval on the number line i.e., the sample space S is continuous. The outcomes of an experiment are represented by the points on a line and the value of the random variable is a number appropriately assigned to each point by some rule.

Example 7.8: Consider the experiment of measuring the height of students in a statistics class. If the minimum height of a student is 5.0 feet and the maximum height is 5.8 feet, then the variable Y, the height of students, takes values between 5.0 and 5.8 feet i.e., in the interval 5.0 – 5.8 feet.

Example 7.6 is also an example of a continuous random variable.

Exercise 7

Ans on Page 257

7.1 i) Define random variable and give an example to explain it.

ii) What are random numbers and how these are generated. Also give an example to explain their application.

iii) Classify each of the following random variables as either discrete or continuous.

a) The number of pages in a book.

b) The number of questions asked in an oral examination.

c) The life time of a light bulb.

d) The amount of rainfall at a particular location during different months of 1996.

7.2 Generate the first 6 random digits using the pseudo-random number generator with

$m = 100,\ a = 21, b = 7, x_0 = 10$

7.3 Count the number of heads and tails when a single coin is tossed 10 times without throwing a coin.

7.4 Two coins are tossed. Let Y denotes the number of heads, then the possible number of heads and their corresponding probabilities are given in the adjoining table

Y	$p(Y)$
0	0.25
1	0.50
2	0.25

Find the sequence of number of heads 20 times without throwing a coin?

7.5 Let the digits 0, 1, 2, 3, 4 represent head and 5, 6, 7, 8, 9 represent tail, use random numbers to simulate 20 flips of a coin.

7.6 Two students in a class are asked to take one of the two courses Mathematics (M) or Biology (B). Define the random variable Y as number of students taking Biology. Write down the possible outcomes and the values assigned by the random variable Y.

7.7 Two coins are tossed and let the random variable y denote the number of heads. Write down the possible outcomes and the values assigned to the random variable.

7.8 There are three children in a family. Let the random variable denote the number of boys in a family. Write down the possible outcomes and the values assigned by the random variable assuming equal chances for boys and girls.

7.9 Four balls are drawn from a bag containing 5 white and 3 black balls. If X denotes the number of white balls drawn, then write down the possible values of the random variable X.

7.10 Differentiate between discrete and continuous random variable and give an example of each.

7.11 Fill in the blanks:

i) Random numbers are _________ by some _________ process.

ii) Random numbers are obtained in such a way that each digit has _________ probability.

iii) Random variable is also called _________ variable.

iv) A random variable assuming only a finite number of values is called_________ random variable.

v) A random variable assuming all possible values in a _________ is called _________ random variable.

vi) The sum of probabilities of events of a sample space is always _______________.

vii) A probability function is _________ function.

viii) The mean of probability distribution is called _________.

ix) $E(X) = \sum Xf(X)$ if it _________ absolutely.

x) A random variable may be _________ or _________.