5

Chapter Five Random Variables and Their Distribution Functions

Chapter Five

Random Variables and Their Distribution Functions

5.1 - The Concept of Random Variable

The reader may be familiar with mathematical variables. In this section we will define another type of variable called a random variable.

Definition 5.1.1.

Given a random experiment with sample space S . A *random variable* (r.v.) X is a function which assigns to each element $w \in S$ a real number X (w) in the set E.

Example 5.1.1.

a. Suppose that an experiment consists of tossing 2 fair coins. If we let X denote the number of heads appearing, then X is a random variable which takes on one of the values 0,1 and 2.

   Here S = {HH, HT, TH, TT}, then
   X (HH) = 2 ; X (HT) = X (TH) = 1 and
   X (TT) = 0 (see Fig. 5.1)

b. Suppose an experiment consists of tossing a fair die until number 6 appears. If we let Y denote the number of tosses required, then Y is a r.v. takes on one of the values 1,2,3, ...

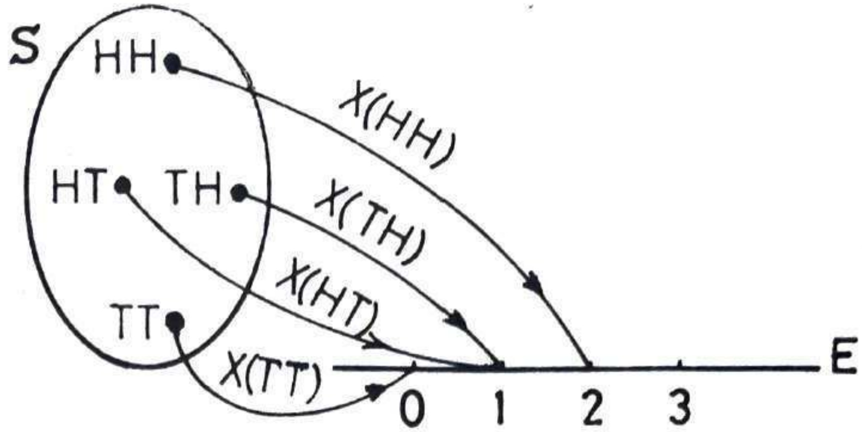

Fig. 5.1

Because the value of a random variable is determined by the outcome of the experiment, we may assign probabilities to the possible values of the random variable. For Example 5.1.1. (a), we have

$$P\{X = 2\} = P\{(HH)\} = \frac{1}{4}$$

$$P\{X = 1\} = P\{(HT), (TH)\} = \frac{1}{2}$$

$$P\{X = 0\} = P\{(TT)\} = \frac{1}{4}$$

Since X takes on one of the values 0, 1 and 2, we must have

$$\sum_{i=0}^{2} P(X = i) = 1$$

Notations :

If x is a real number, the set of all $w \in S$, such that $X(w) = x$, is denoted by $\{X = x\}$. Thus

$$P\{X = x\} = P\{w : X(w) = x\}.$$

Similarly, if $a, b \in R$, then

$$P\{X \leq a\} = P\{w : X(w) \leq a\},$$

$$P\{a \leq X \leq b\} = P\{w : a \leq X(w) \leq b\}$$

and so on.

Example 5.1.2.

Let an experiment consist of tossing a pair of unbiased dice. Let X be a random variable indicating the sum of the two faces, then X takes the values 2,3, ..., 12 with respective Probabilities.

$X = x$	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Definition 5.1.2.

A random variable X is called *discrete* if the set E is finite or countably infinite . If E is the set of the real numbers R or any subset of R, then X is called a *continuous* random variable.

In Examples 5.1.1. and 5.1.2, the random variables X and Y are discrete.

Example 5.1.3.

Let an experiment consist of measuring the lifetimes of 5 electric bulbs. The sample space S is the set of all 5-tuples $\{ w_1, w_2, w_3, w_4, w_5 \}$ where $w_i \geq 0$ for i = 1,2, ... 5. Let X be a r.v. representing the average lifetimes of the 5 bulbs, i.e.

$$X = \sum_{i=1}^{5} w_i / 5$$

Here X takes value on $[0, \infty)$, then X is a continuous r.v.

Remarks

1. If X and Y are two random variables defined on the same sample space S, then X+Y, aX, X/Y, aX +bY, XY are also random variables on S (where $a,b \in R$).

2. Also max (X,Y) and min (X,Y) are random variables, where max and min mean the maximum and the minimum of X and Y, respectively.

5.2. Distribution Function

Defintion 5.2.1

If X is a r.v. defined on the sample space S, the *cumulative distribution function* (c.d. f), or simply the *distribution function* F of the r.v. X, is the function from R to [0,1], and is defined for all $b \in R$, by

$F(b) = P\{X \leq b\}$ .... (5.2.1)

In other word, F(b) denotes the probability that the r.v. X takes on a value that is less than or equal to b. Some properties of the distribution function F are

1. F is a non decreasing function, that is, if $a < b$, then $F(a) \leq F(b)$

2. $F(\infty) = \lim_{b \to \infty} F(b) = 1$ and $F(-\infty) = \lim_{b \to -\infty} F(b) = 0$.

3. F(b) is continuous from the right, that is, $\lim_{h \to 0} F(b+h) = F(b)$ for $h > 0$.

For each random variable X there is one and only one distribution function. If it is known all probabilities asked about X can be answered in term of F. For example,

$P(a < X \leq b) = F(b) - F(a)$ for all $a < b$ .... (5.2.2)

$P(X > b) = 1 - F(b)$ for all $b \in R$ .... (5.2.3)

If we want to find P(X < b), we can apply the continuity property to obtain

$$
\begin{aligned}
P(X < b) &= P\left(\lim_{n \to \infty} \left\{X \leq b - \frac{1}{n}\right\}\right) \\
&= \lim_{n \to \infty} P\left\{X \leq b - \frac{1}{n}\right\} \\
&= \lim_{n \to \infty} F\left(b - \frac{1}{n}\right)
\end{aligned}
$$

It is not necessary that P(X < b) is equal to F(b), because F(b) includes P(X=b).

Example 5.2.1.

The distribution function of the r.v. X is given by

$$
F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} & 0 \leq x < 1 \\ \frac{3}{4} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}
$$

A graph of F(x) is presented in Fig. 5.2

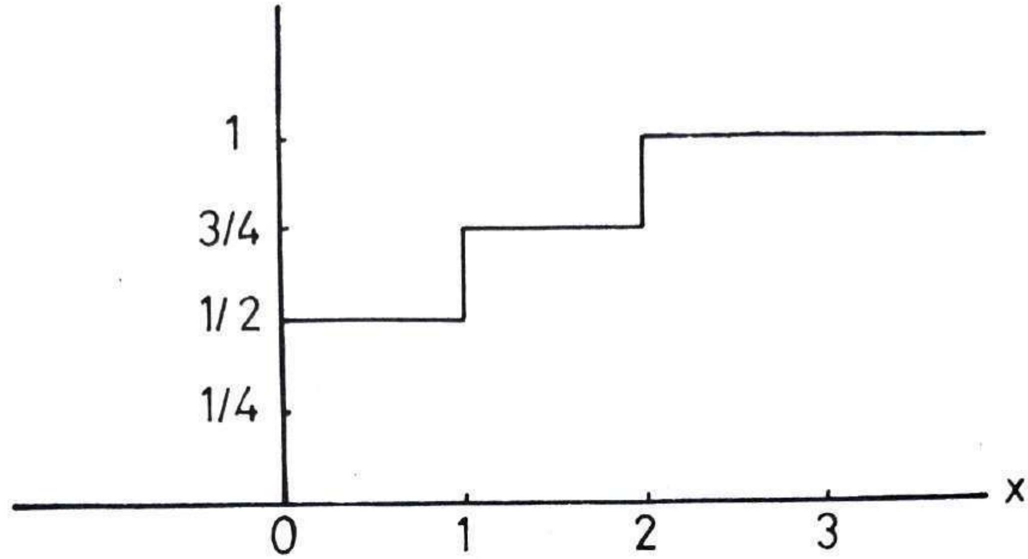

Fig. 5.2 Graph of F(x)

To calculate (a) $P(X < 2)$, (b) $P(X=1)$, (c) $P(X > \frac{3}{2})$ and (d) $P(1 < X \leq 4)$ ·

Solution : we have

a. $$P(X < 2) = \lim_{n \to \infty} P\left[X \leq 2 - \frac{1}{n} \right] = \lim_{n \to \infty} F\left(2 - \frac{1}{n} \right) = \frac{3}{4}.$$

b. $$P(X = 1) = P(X \leq 1) - P(X < 1) = F(1) - \lim_{n \to \infty} F\left(1 - \frac{1}{n} \right)$$

$$= \frac{3}{4} - \frac{1}{2} = \frac{1}{4}.$$

c. $$P\left(X > \frac{3}{2} \right) = 1 - P\left(X \leq \frac{3}{2} \right) = 1 - F\left(\frac{3}{2} \right) = 1 - \frac{3}{4} = \frac{1}{4}$$

d. $$P(0{\cdot}5 < X \leq 4) = F(4) - F(0{\cdot}5)$$

$$= 1 - \frac{1}{2} = \frac{1}{2}.$$

5.3 Discrete Random Variables

As we mentioned in Definition 5.1.2, a discrete random variable can take on at most a countable number of possible values $x_1, x_2, x_3, \ldots$.

Definition 5.3.1

The *probability mass function* (p.m.f.) P(a) of a discrete r.v.X is a function from E to [0,1] and is defined as

$$P(a) = P\{X=a\}, \ a \in E$$

The probability mass function P(a) is positive for at most a countable number of values of a. That is, if X must assume one of the values $x_1, x_2, \ldots\ldots$, then

1. $P(x_i) \geq 0$ for $i=1,2,3 \ldots$
2. $\sum_{i=1}^{\infty} P(x_i) = 1$

The distribution function of X is given by

$$F(x) = \sum_{x_i \leq x} P(x_i)$$

When F(x) is represented graphically, its type is the step function. If $x_1 < x_2 < x_3 < \ldots$, the distribution function has jumps at the points $x_1, x_2, x_3, \ldots$ and stays constant in the interval (x_i, x_{i+1}). The saltus at x_i is $P(x_i)$ (see Fig 5.3 below).

x

Example 5.3.1

Two fair dice are thrown once, Let X denote the maximum of the two numbers shown by two dice. Then the p. m. f. of X is given by

Values of X , x	1	2	3	4	5	6
P (x)	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

The distribution function of X is given by

$$F(x) = \begin{cases} 0 & \text{if} \quad x < 1 \\ \frac{1}{36} & 1 \leq x < 2 \\ \frac{4}{36} & 2 \leq x < 3 \\ \frac{9}{36} & 3 \leq x < 4 \\ \frac{16}{36} & 4 \leq x < 5 \\ \frac{25}{36} & 5 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$

The graph of F(x) is given below

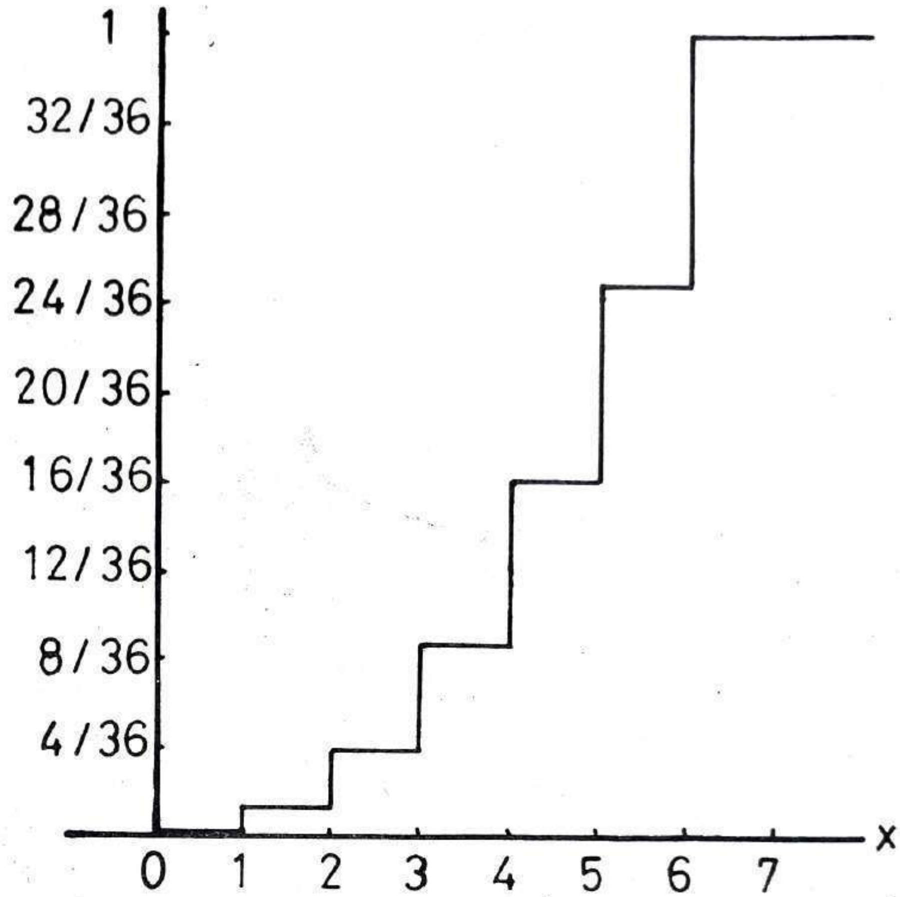

Fig 5 ..3 – The graph of F (x)

Example 5.3.2.

A box contains 6 red and 4 white balls. Two balls are drawn at random. Let Y denote the number of red balls that can be drawn. Then Y must assume one of the values 0,1 and 2 with respective probabilities 2/15, 8/15 and 5/15. Therefore, the p.m.f. of Y is given by

values of Y,y	0	1	2
P(y)	2/15	8/15	5/15

The distribution function of Y is given by

$$F(y) = \begin{cases} 0 & x < 0 \\ 2/15 & 0 \le x < 1 \\ 10/15 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

Example 5.3.3.

A r.v. X has the following p.m.f.

values of X,x	-2	-1	0	1	2	3
P(x)	k	2k	3k	5k	7k	2k

1- Find the value of K.

2- Find $P(X < 1)$, $P(X \ge 0)$, $P(-1 \le X < 2)$

1. Since $\Sigma\ P(x) = 1$, we have
$k + 2k+3k+5k+7k+2k = 20k = 1$

$$k = \frac{1}{20}$$

2. $P(X < 1) = P(X = -2) + P(X = -1) + P(X = 0)$

$$= \frac{1}{20} + \frac{2}{20} + \frac{3}{20} = \frac{6}{20} = \frac{3}{10}$$

$P(X \ge 0) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$

$$= \frac{3}{20} + \frac{5}{20} + \frac{7}{20} + \frac{2}{20} = \frac{17}{20}$$

and

$$P(-1 \leq X < 2) = P(X = -1) + P(X = 0) + P(X = 1)$$

$$= \frac{2}{20} + \frac{3}{20} + \frac{5}{20} = \frac{10}{20} = \frac{1}{2}$$

5.4. Special Univariate Discrete Distributions

5.4.1. Bernoulli Distribution

Suppose that a trial whose outcome can be classified as either a "success" or as a "failure" is performed. let X be a r. v. taking value 1 if the outcome is success and 0 if it is failure, then X is said to be a Bernoulli random variable and its p.m.f. is given by

$$P(X = i) = p^{i}(1 - p)^{1-i}, i = 0, 1 \qquad \ldots (5{\cdot}4{\cdot}1)$$

5.4.2 Binomial Distribution

Consider a series of n independent Bernoulli: trials (n is a finite integer) in which the probability of success p in each trial is constant. Then q=1-p is the probability of failure in each trial. Let X denote the number of successes in n trials. Then X can take the values 0,1,2, ... , n with non-zero probabilities, and X is said to be a binomial random variable with parameters n and p. Its p.m.f. is given by

$$P(X = k) = C(n, k)\, p^{k} q^{n-k}, k = 0, 1, \ldots n \qquad \ldots (5{\cdot}4{\cdot}2)$$

By applying Theorem 2.3.2 we can get

$$\sum_{k=0}^{n} P(X = k) = (p + q)^{n} = 1$$