Introduction:

Suppose that to each point of a sample space we assign a number. We then have a function defined on the sample space. This function is called a random variable (or stochastic variable) or more precisely a random function (stochastic function). It is usually denoted by a capital letter such as X or Y. In general, a random variable has some specified physical, geometrical, or other significance.

Discrete and Continuous Random Variables:

Def. A random variable is called discrete if its range includes finite number of values or countable infinite number of values $x1, x2, \dots$.

Example: The number of members of a family, number of passengers on a plane, number of road accidents occurs in a day in a city …etc. Def. A random variable is called continuous if it can assume any value from a specified interval of the form [a, b]. That is; if for some $a < b$ any number x between a and b is possible.

Example: Heights of students in a class, temperature and barometric pressures of different cities, age of people or objects … etc.

Distribution Function:

With every random variable is associated another function known as its distribution function.

EXAMPLE 1: Suppose that a coin is tossed twice so that the sample space is S {HH, HT, TH, TT}. Let X represent the number of heads that can come up. With each sample point we can associate a number for X as shown in Table 2-1. Thus, for example, in the case of HH (i.e., 2 heads), X 2 while for TH (1 head), X 1. It follows that X is a random variable.

Table1

Sample Point	HH	HT	TH	TT
X	2	1	1	0

It should be noted that many other random variables could also be defined on this sample space, for example, the square of the number of heads or the number of heads minus the number of tails. A random variable that takes on a finite or countably infinite number of values is called a discrete random variable while one which takes on a noncountably infinite number of values is called a nondiscrete random variable.

Discrete Probability Distributions

Let X be a discrete random variable, and suppose that the possible values that it can assume are given by x1, x2, x 3, . . . , arranged in some order. Suppose also that these values are assumed with probabilities given by

$P(X = x_k) = f(x_k)$ $k = 1,2,3,\ldots\ldots$

It is convenient to introduce the probability function, also referred to as probability distribution, given by $P(X = x) = f(x)$

For x = xk, this reduces to (1) while for other values of x, $f(x)= 0$. In general, f(x) is a probability function

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

EXAMPLE 2:

Find the probability function corresponding to the random variable X of Example 2.1. Assuming that the coin is fair, we have

P(HH) = 1/4 P(HT) = 1/4 P(TH) = 1/4 P(TT) = 1/4 Then

P(X = 0) = P(TT) = 1/4

P(X = 1) P(HT **U** TH) = 1/4+1/4 = 1/2

P(X = 2) = P(HH) = 1/4

The probability function is thus given by table 2

Table 2

X	0	1	2
f(x)	1/4	1/2	1/4

Distribution Functions for Discrete Random Variables The distribution function for a discrete random variable X can be obtained from its probability function by noting that, for all x in (-∞,∞)

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u)$$

where the sum is taken over all values u taken on by X for which u<=x.

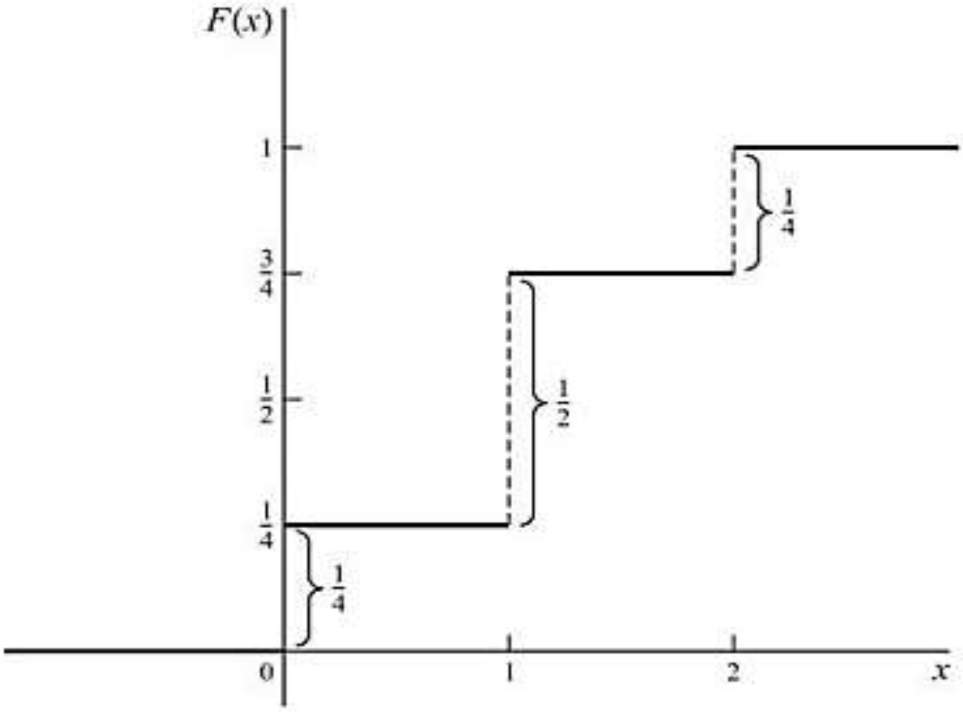
If X takes on only a finite number of values x1, x2, . . . , xn, then the distribution function is given by

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \leq x < x_2 \\ f(x_1) + f(x_2) & x_2 \leq x < x_3 \\ \vdots & \vdots \\ f(x_1) + \cdots + f(x_n) & x_n \leq x < \infty \end{cases}$$

EXAMPLE 2.3 (a) Find the distribution function for the random variable X of Example 2.2. (b) Obtain its graph. (a) The distribution function is

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \leq x < x_2 \\ f(x_1) + f(x_2) & x_2 \leq x < x_3 \\ \vdots & \vdots \\ f(x_1) + \cdots + f(x_n) & x_n \leq x < \infty \end{cases}$$

(b)The graph of F(x) is shown in the figure above is:



Example:

Three coins are tossed. How many fall "heads"? Let X be the number of heads which appear in the three tosses. Thus number may be 0, 1, 2, or 3. A sample space for the experiment and the number of heads for each sample point and the probability of each sample point are shown in the following table:

Sample point	*HHH*	*HHT*	*HTH*	*THH*	*HTT*	*THT*	*TTH*	*TTT*
Number of heads	3	2	2	2	1	1	1	0
Probability	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

$$f_X(x) = \begin{cases} \frac{1}{8} & x = 0 \\ \frac{3}{8} & x = 1 \\ \frac{3}{8} & x = 2 \\ \frac{1}{8} & x = 3 \\ 0 & o.w. \end{cases}$$

It is usually preferable to give a formula to express the probabilities by means of a function such that its values, fX(x), equal p(X=x) for each x. Observing that the numerators of these probabilities, 1, 3, 3, 1, are the binomial coefficients

$$C_0^3,\ C_1^3,\ C_2^3,\ C_3^3$$

we find that the formula for the probability distribution can be written as:

$$f_X(x) = \frac{C_x^3}{8} \qquad \text{for } x = 0, 1, 2, 3$$

Example:

Find the distribution function of the total number of heads obtained in four tosses of a coin.

Ω= {HHHH, HHHT, HHTH, HTHH, THHH, HHTT, HTHT, HTTH, THHT, THTH, TTHH, HTTT, THTT, TTHT, TTTH, TTTT}.

$$f(x) = \frac{C_x^4}{16} \qquad x = 0, 1, 2, 3, 4$$

$$F_X(0) = f_X(0) = \frac{1}{16}$$

$$F_X(1) = f_X(0) + f_X(1) = \frac{1}{16} + \frac{4}{16} = \frac{5}{16}$$

$$F_X(2) = f_X(0) + f_X(1) + f_X(2) = \frac{1}{16} + \frac{4}{16} + \frac{6}{16} = \frac{11}{16}$$

$$F_X(3) = f_X(0) + f_X(1) + f_X(2) + f_X(3) = \frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{16} = \frac{15}{16}$$

$$F_X(4) = \sum_{x \in S} f_X(x) = 1$$

Hence, the distribution function is given by:

$$F(x) = \begin{cases} 0 & for \quad x < 0 \\ \frac{1}{16} & for \quad 0 \le x < 1 \\ \frac{5}{16} & for \quad 1 \le x < 2 \\ \frac{11}{16} & for \quad 2 \le x < 3 \\ \frac{15}{16} & for \quad 3 \le x < 4 \\ 1 & for \quad x \ge 1 \end{cases}$$

Probability Distributions:

If X is a discrete random variable with distinct values *X1, X2, …, Xn , …*,and S is a countable set of real numbers, then the function fx: ℝ → ℝ given by

$$f_X(x) = \begin{cases} P(X = x) & if\ x \in S \\ 0 & if\ x \notin S \end{cases}$$

is called the probability distribution or the probability mass function of X, and the set S is called the support of X.

A function can serve as the probability distribution of a discrete random variable X if its values, fX(x), satisfy the following conditions:

1. $f_X(x) > 0\ \ \forall\, x \in S$
2. $f_X(x) = 0\ \ \forall\, x \notin S$
3. $\sum_{x\in S} f_X(x) = \sum_{x\in S} P[(X = x)] = P(\{X \in S\}) = 1$

Example:

$$\text{Let } f(x) = \begin{cases} \frac{x}{10} & x = 1, 2, 3, 4 \\ 0 & elsewhere \end{cases}$$

Is *f(x)* a probability mass function?

Note that *fX* is of a discrete type.

1. $p(X=1) = \frac{1}{10},\ p(X=2) = \frac{2}{10}\ ,\ p(X=3) = \frac{3}{10},\ p(X=4) = \frac{4}{10}$

2. $\sum_{x \in S} f(x) = \frac{1}{10} + \frac{2}{10} + \frac{3}{10} + \frac{4}{10} = 1$

$\therefore$ *f(x)* is a probability mass function.

Example:

Check whether the function:

$$f(x) = \frac{x+2}{25} \qquad \text{, for } x = 1, 2, 3, 4, 5$$

can serve as the probability distribution of a discrete random variable.

1) $P(X=1) = \frac{3}{25}$, $P(X=2) = \frac{4}{25}$, $P(X=3) = \frac{5}{25}$, $P(X=4) = \frac{6}{25}$,

$P(X=5) = \frac{7}{25}$

2) $\sum_{x \in S} f_X(x\ ) = 1$