Random Variables:

Definition: A real valued function defined on the sample space (of a random experiment) is called a random variable. (r.v)

Example: 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 come up. Thus we can write

Sample point	HH	HT	TH	TT
X	2	1	1	0

Definition: A r.v is said to be **discrete** if it takes only a finite or countably infinite number of values.

Example: Suppose a die is tossed and X denotes the outcome (i.e. the number on the face that comes up) then X is a discrete r.v taking values 1,2,3,4,5,6.

Other examples: (i) Number of telephone calls received per unit time in a telephone exchange, (ii) the number of III Year students of the college who have got above 90% marks in *Computational Mathematics* subject, (iii) the number of alpha particles emitted by a radioactive source in a given time interval.

Definition: A r.v is said to be **continuous** if it takes (all) values in an interval. In other words, the values taken by a continuous r.v cannot be put in a 1-1 correspondence with the set of positive integers. (There are an uncountable number of values)

Example: Suppose X denotes the height of students (in cm) of a class. Then X can take any value in the interval (say) [0,400] and therefore X is a continuous r.v.

Probability mass function and Probability density function

Definition: Suppose X is a discrete r.v taking values $x_1, x_2, \dots$ Let $P(X = x_i) = p_i,\ i = 1,2,\dots$ Then $p_i,\ i = 1,2,\dots$ is called the **Probability mass function (p.m.f)** or **Probability function** if

$(i)\quad p_i \geq 0 \quad \forall i,$

$(ii)\ \sum_i p_i = 1$

The collection of pairs $\{x_i, p_i\}, i = 1,2,3,\dots$ is called the **probability distribution** of the r.v. X.

Remark: The probability distribution of the r.v. X is usually displayed in the form of a table given below:

X	$x_1 \quad x_2 \quad x_3 \quad \dots \quad x_i \quad \dots$
$P(X=x)$	$p_1 \quad p_2 \quad p_3 \quad \dots \quad p_i \quad \dots$

Example:
Suppose 2 fair dice are thrown and the sum (X) on the dice is noted. Obtain the probability distribution of X.

Solution: When 2 fair dice are thrown, the sample space is given by

$$S = \left\{ \begin{matrix} (1,1), (1,2), (1,3), (1,4), (1,5), (1,6) \\ (2,1), (2,2), (2,3), (2,4), (2,5), (2,6) \\ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6) \\ (4,1), (4,2), (4,3), (4,4), (4,5), (4,6) \\ (5,1), (5,2), (5,3), (5,4), (5,5), (5,6) \\ (6,1), (6,2), (6,3), (6,4), (6,5), (6,6) \end{matrix} \right\}$$

Therefore, if X denotes the sum on the 2 dice, then X takes the values 2,3,4,5,6,7,8,9,10,11,12.
The probability distribution of X is given by:

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

Definition: Suppose X is a continuous r.v taking values in an interval (a,b) (The interval can be open, half open, half closed or closed with $-\infty < a < b < \infty$). If there exists a function $f(x)$ such that

$$P(x-\frac{\delta x}{2} \le X \le x+\frac{\delta x}{2}) = f(x)\delta x$$

then *f(x)* is called the probability density function (p.d.f) of *X*, provided: $(i)\quad f(x) \ge 0\ \forall x \in R$ and $(ii)\ \int_{-\infty}^{\infty} f(x)dx = 1$

Examples:

1. A random variable X has the following Probability distribution:

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

(i) Find k (ii) Evaluate P(X<2) , P(X≥2), P(-2 < X < 2)

Solution: (i) To find k. We have

$$\sum_X p_X = 1$$
$$\Rightarrow 0.1 + k + 0.2 + 2k + 0.3 + 3k = 1$$
$$\Rightarrow 6k = 1 - 0.6$$
$$\Rightarrow k = \frac{0.4}{6}$$
$$\Rightarrow k = \frac{1}{15}$$

Therefore the probability distribution of X is given by

X	-2	-1	0	1	2	3
P(X=x)	0.1	1/15	0.2	2/15	0.3	3/15

$$Now,\ P(X<2) = P(X=-2)+P(X=-1)+P(X=0)+P(X=1)$$
$$=0.1+1/15+0.2+2/15$$
$$=0.3+3/15$$
$$=0.5$$
$$And\ P(X \ge 2) = P(X=2)+P(X=3)$$
$$=0.3+3/15$$
$$=0.5$$
$$And\ P(-2<X<2) = P(X=-1)+P(X=0)+P(X=1)$$
$$=1/15+0.2+2/15$$
$$=0.4$$

2. Suppose that in a certain region the daily rainfall (in inches) is a continuous r.v. X with p.d.f f(x) given by $f(x)=(3/4)(2x-x^2)\ for\ 0<x<2, and\ f(x)=0\ elsewhere$, find the probability that on a given day in this region, the rainfall is :
   (i) not more that 1 inch (ii) greater that 1.5 inches (iii) between 1 and 1.5 inches

Solution:

(i) P(rainfall is not more than 1 inch) $= P(X \le 1) = \int_{-\infty}^{1} f(x)dx$

$$= \int_0^1 \frac{3}{4}(2x - x^2)\, dx$$

$$= \frac{3}{4}\left(\frac{2x^2}{2} - \frac{x^3}{3}\right)\Bigg]_0^1$$

$$= \frac{3}{4}\left(1 - \frac{1}{3}\right) = \frac{1}{2}$$

(ii) P(rainfall is greater than 1.5 inches) $= P(X > 1.5) = \int_{1.5}^{\infty} f(x)dx$

$$= \int_{1.5}^{2} \frac{3}{4}(2x - x^2)\, dx$$

$$= \frac{3}{4}\left(\frac{2x^2}{2} - \frac{x^3}{3}\right)\Bigg]_{1.5}^{2}$$

$$= \frac{3}{4}\left((4 - (1.5)^2) - \frac{1}{3}(8 - (1.5)^3\right)\Bigg]_{1.5}^{2}$$

$$= 0.1562$$

(iii) P(rainfall is between 1 inch and 1.5 inches) $= P(1 < X < 1.5)$

$$= \int_1^{1.5} f(x)dx$$

$$= \int_1^{1.5} \frac{3}{4}(2x - x^2)\, dx$$

$$= \frac{3}{4}\left(\frac{2x^2}{2} - \frac{x^3}{3}\right)\Bigg]_1^{1.5}$$

$$= \frac{3}{4}\left[((1.5)^2 - 1) - \frac{1}{3}((1.5)^3 - 1)\right]$$

$$= 0.3515$$

Cumulative distribution function (c.d.f) or Distribution function

If X is a r.v, discrete or continuous, then the function $F(x) = P(X \le x)$ is called the cumulative distribution function (c.d.f) or distribution function of X.

If X is discrete, then

$F(x) = P(X \le x) = \sum_{x_i \le x} p_i$ where $p_i = P(X = x_i)$ is the probability mass function of X.

If X is continuous, then

$F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)dx$ where $f(x)$ is the probability density function of X.

Examples:

1. Suppose X is a discrete r.v with probability distribution

X	1	2	3
P(X=x)	1/4	1/2	1/4

Obtain the distribution function of X.

Solution: We have the distribution function of X to be given by

$F(x) = P(X \le x) = \sum_{x_i \le x} p_i$ if X is discrete. Therefore,

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

2. Suppose X is a continuous r.v. with $f(x) = \frac{1}{2}, 0 \le x \le 2.$ Obtain the distribution function of X.

Solution: We have the distribution function of X to be given by

$F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)dx$ if X is continuous. Therefore,

$$F(x) = \int_0^x \frac{1}{2}dx \quad = \frac{1}{2}x\Big|_0^x$$

$$= \frac{1}{2}x$$

$$\therefore F(x) = \begin{cases} \frac{1}{2}x, & 0 \le x \le 2 \\ 1, & x \ge 2 \end{cases}$$

3. Let X be a random variable denoting the number of points appearing in a throw of a fair die. Determine the distribution function $F(x)$ and draw its graph.

Solution: Here X takes values 1, 2,3,4,5 and 6 with equal probabilities. Hence the probability distribution of X is given by:

X	1	2	3	4	5	6
P(X=x)	1/6	1/6	1/6	1/6	1/6	1/6

Hence the distribution function of X is given by:

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

$$= \begin{cases} 0, & x < 1 \\ 1/6, & 1 \le x < 2 \\ 2/6, & 2 \le x < 3 \\ 3/6,, & 3 \le x < 4 \\ 4/6,, & 4 \le x < 5 \\ 5/6,, & 5 \le x < 6 \\ 1, & x \ge 6 \end{cases}$$

4. A continuous random variable has the following probability law:

$f(x) = kx^2$, $0 \le x \le 2$. Determine k and find P(0.1≤ X ≤0.4),P(0.2≤X<3)

Solution: Since f(x) is a pdf, we have,

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

$$\Rightarrow \int_0^2 kx^2\, dx = 1$$

$$\Rightarrow k\left[\frac{x^3}{3}\right]_0^2 = 1$$

$$\Rightarrow k = 3/8$$

$$(i)\ P(0.1 \le X \le 0.4) = \int_{0.1}^{0.4} \frac{3}{8}x^2 dx = \frac{3}{8}\left[\frac{x^3}{3}\right]_{0.1}^{0.4}$$

$$= 0.007875$$

$$(ii)\ P(0.2 \le X < 3) = \int_{0.2}^{3} \frac{3}{8}x^2 dx$$

$$= \int_{0.2}^{2} \frac{3}{8}x^2 dx \quad (\because x \le 2)$$

$$= \frac{3}{8}\left[\frac{x^3}{3}\right]_{0.2}^{2}$$

$$= 0.999$$

5. A random variable X has the following Probability distribution:

X	0	1	2	3	4	5	6	7
P(X=x)	0	k	2k	2k	3k	K^2	$2k^2$	$7k^2+k$

Evaluate (i) k (ii) P(X<6) , P(X≥6), P(0 < X < 5) . Find the minimum value of 'a' such that P(X≤a)>1/2. Also find the distribution function of X.

Solution: (i) **To find k**: We have

$$\sum_x p_x = 1$$

$$\Rightarrow 0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$$

$$\Rightarrow 10k^2 + 9k = 1$$

$$\Rightarrow 10k^2 + 9k - 1 = 0 \Rightarrow (10k-1)(k+1) = 0$$

$$\Rightarrow k = \frac{1}{10} \quad or \quad k = -1$$

But k=-1 is inadmissible, as the probabilities should be ≥0. Hence, $k = \frac{1}{10}$ Therefore the probability distribution of X is given by

X	0	1	2	3	4	5	6	7
P(X=x)	0	1/10	2/10	2/10	3/10	1/100	2/100	17/100

$$Now, \ P(X<6) = P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)+P(X=5)$$
$$=0+1/10+2/10+2/10+3/10+1/100$$
$$=81/100$$
$$[or \quad P(X<6) = 1-P(X\geq 6) = 1-\{P(X=6)+P(X=7)\}]$$
$$And \quad P(X\geq 6) = P(X=6)+P(X=7)$$
$$=2/100+17/100$$
$$=19/100$$
$$And \quad P(0<X<5) = P(X=1)+P(X=2)+P(X=3)+P(X=4)$$
$$=1/10+2/10+2/10+3/10$$
$$=8/10$$

By trial and error we find that

$$P(X\leq 3) = P(X=0)+P(X=1)+P(X=2)+P(X=3)$$
$$=0+1/10+2/10+2/10$$
$$=5/10 = 1/2$$
$$And \ P(X\leq 4) = P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)$$
$$=0+1/10+2/10+2/10+3/10$$
$$>1/2$$

Therefore, the required number ‘a’ equals 4. That is, a=4.
The distribution function is given by

$$F(x) = P(X\leq x)$$
$$= \begin{cases} 0, & x<1 \\ 1/10, & 1\leq x<2 \\ 3/10, & 2\leq x<3 \\ 5/10,, & 3\leq x<4 \\ 8/10,, & 4\leq x<5 \\ 81/100,, & 5\leq x<6 \\ 83/100, & 6\leq x<7 \\ 1, & x\geq 7 \end{cases}$$