RANDOM VARIABLE (X)

A random variable (RV) X is a real valued function defined on the sample space, i.e., for each element of the sample space, the RV takes a unique real number.

This defines a one dimensional random variable.

Example: Let us consider throwing of two coins simultaneously. The sample space associated with this experiment shall be S = {HH, HT, TH, TT}. Here H represents the appearance of head and T represents the appearance of tail. We can define a random variable for this sample space as:

{(HH, 2), (HT, 1), (TH, 1), (TT, 0)}

This random variable can also be represented as:

$$X(\mathrm{HH}) = 2;\ X(\mathrm{HT}) = X(\mathrm{TH}) = 1;\ X(\mathrm{TT}) = 0.$$

Please note that this definition of RV defined on the given sample space is not unique.

We can define a number of other random variables on this given sample space.

PROBABILITY DISTRIBUTION

With a random variable *X*, we can associate a probability distribution. We will define the probability distributions of discrete and continuous random variables separately.

Discrete Random Variable

A set of numbers $\{p_n\}$ is the probability distribution of a random variable X taking values $x_1, x_2, \ldots, x_n, \ldots$ if it satisfies the following properties.

(i) $p_i \geq 0, i = 1, 2, \ldots$

(ii) $\sum_{i=1}^{\infty} p_i = 1$

We also write this as

$$p(x) = p_i, \text{ for } x = x_i;$$
$$= 0, \text{ otherwise}$$

This function $p(x)$ is also called the **probability mass function** (*pmf*) of the random variable X. This notation also means that,

$$P(X = x_i) = p_i, \quad i = 1, 2, \ldots$$

We also say that a function that satisfies (i) and (ii) above is a legitimate *pmf*.

There is a general convention of the notations here that random variables are represented by capital letters and their values are represented by small letters.

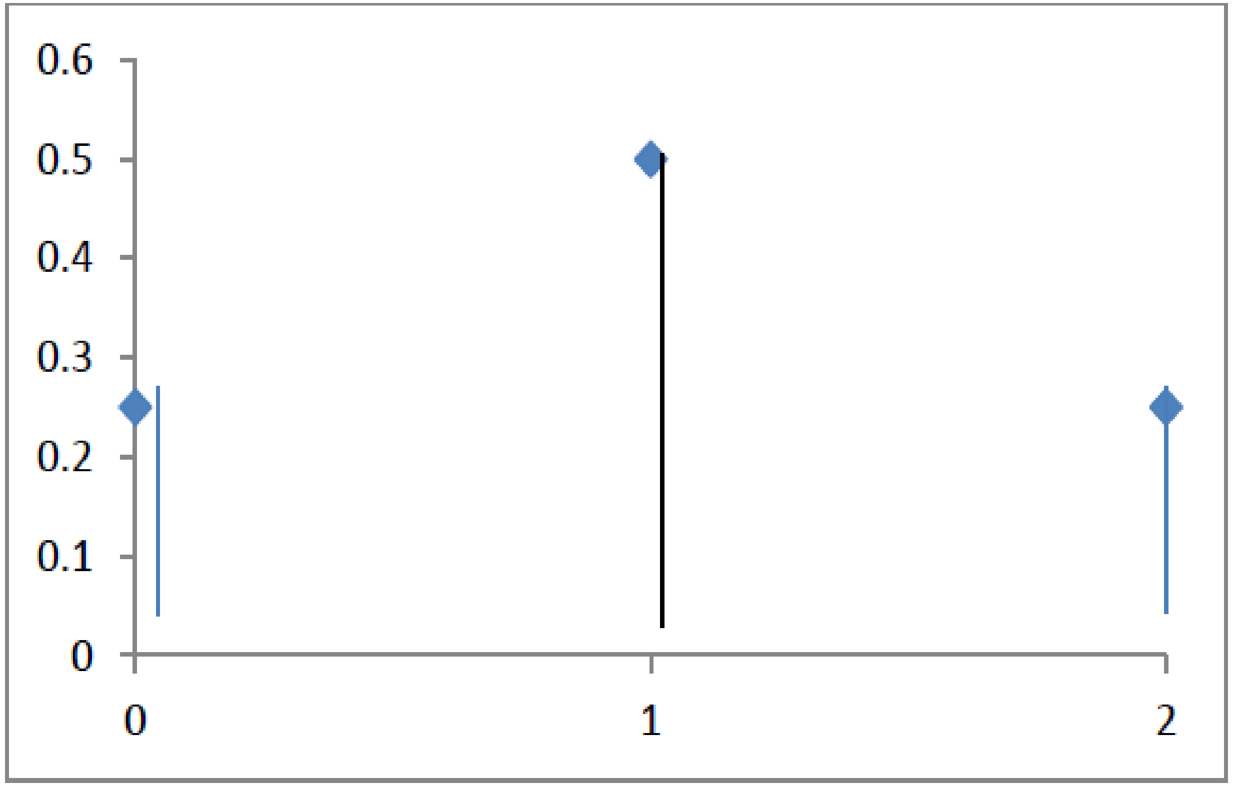
Example: we considered throwing two dice simultaneously in which the sample space is

$$S = \{HH, HT, TH, TT\}$$

Let us define here the random variable X as number of heads that appear in these two simultaneous throws. As such, X takes the values 0, 1 and 2. We can associate a *pmf* with this random variable as

$$P(X = 0) = 1/4,\ P(X = 1) = 1/2,\ P(X = 2) = 1/4.$$

This probability function is depicted in following Figure.



Probability Mass Function

Continuous Random Variable

When we deal with a continuous random variable, we do not have the values of the variable at discrete points rather the random variable takes the values from an interval. This interval can be considered as a subset of the real line. Here, instead of probability mass function, we define probability density function. Probability of an event depends upon the starting point and the length of the interval. Probability density function (*pdf*) is defined as the function that gives us the probability per unit interval. This can be illustrated with the following.

$$f(x) = \lim_{h \to 0} \frac{P(x < X < x + h)}{h}$$

As such, for a continuous random variable *X*, we define a *pdf*. A function *f(x)* is said to be a *pdf* if it satisfies the following properties.

(i) $f(x) \geq 0, \quad -\infty < x < \infty$

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

It means that ***pdf*** always takes nonnegative values and the integral of this function over the interval under consideration for the random variable should be 1. We also say that a function that satisfies (i) and (ii) above is a legitimate *pdf*.

Notes:

(i) Here, $P(a \leq X \leq b)$ is defined as $\int_a^b f(x)dx$.

(ii) For a continuous random variable, the probability that the variable takes a specific value is 0.

Example: Let us consider that *f(x)* = 2*x*, 0 < *x* < 1;

= 0, elsewhere

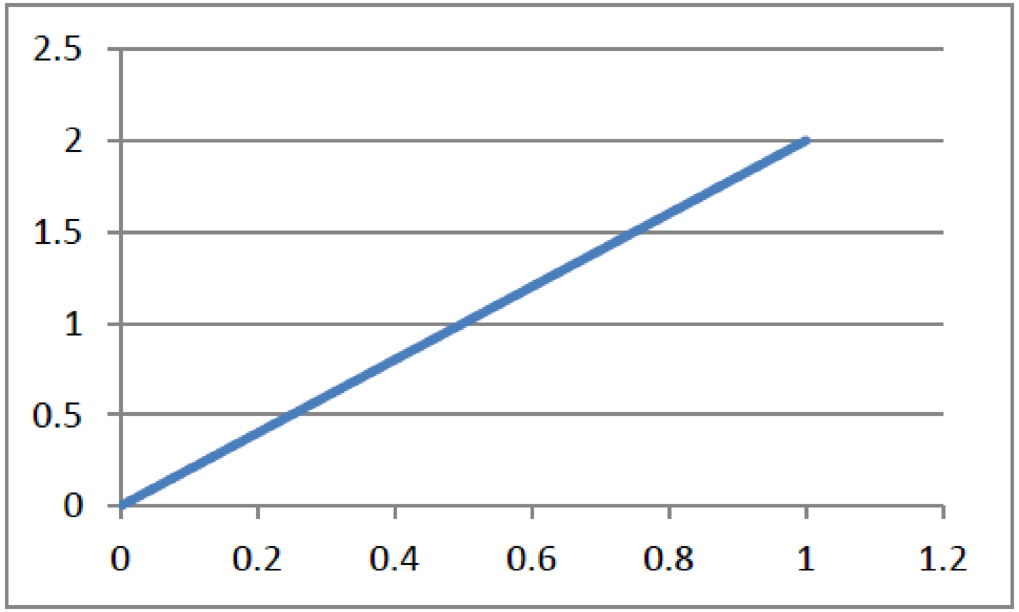
Show that this is a legitimate *pdf*.

(i) We can see that *f(x)* ≥ 0 for all *x* such that 0 < *x* < 1.

(ii) $\int_{-\infty}^{\infty} 2x\, dx = \int_{0}^{1} 2x\, dx = [x^2]_0^1 = 1.$

As such, this function is a legitimate *pdf*.

Graph of this function is given below. This can be noted that the area bounded by the function *f(x)*, line $x = 0$, line $x = 1$ and *x*-axis is 1 here.



Probability Density Function

Distribution Function (or, Cumulative Distribution Function)

The distribution function (*df*) of a random variable *X* is defined as,

$$F(t) = P(X \le t), \ -\infty < t < \infty$$

For a discrete random variable, we can see that this function will be of the following form,

$$F(t) = \sum_{x \le t} p(x), \ -\infty < t < \infty$$

And for a continuous random variable, this function shall be,

$$F(t) = P(X \le t) = \int_{-\infty}^{t} f(x)dx$$

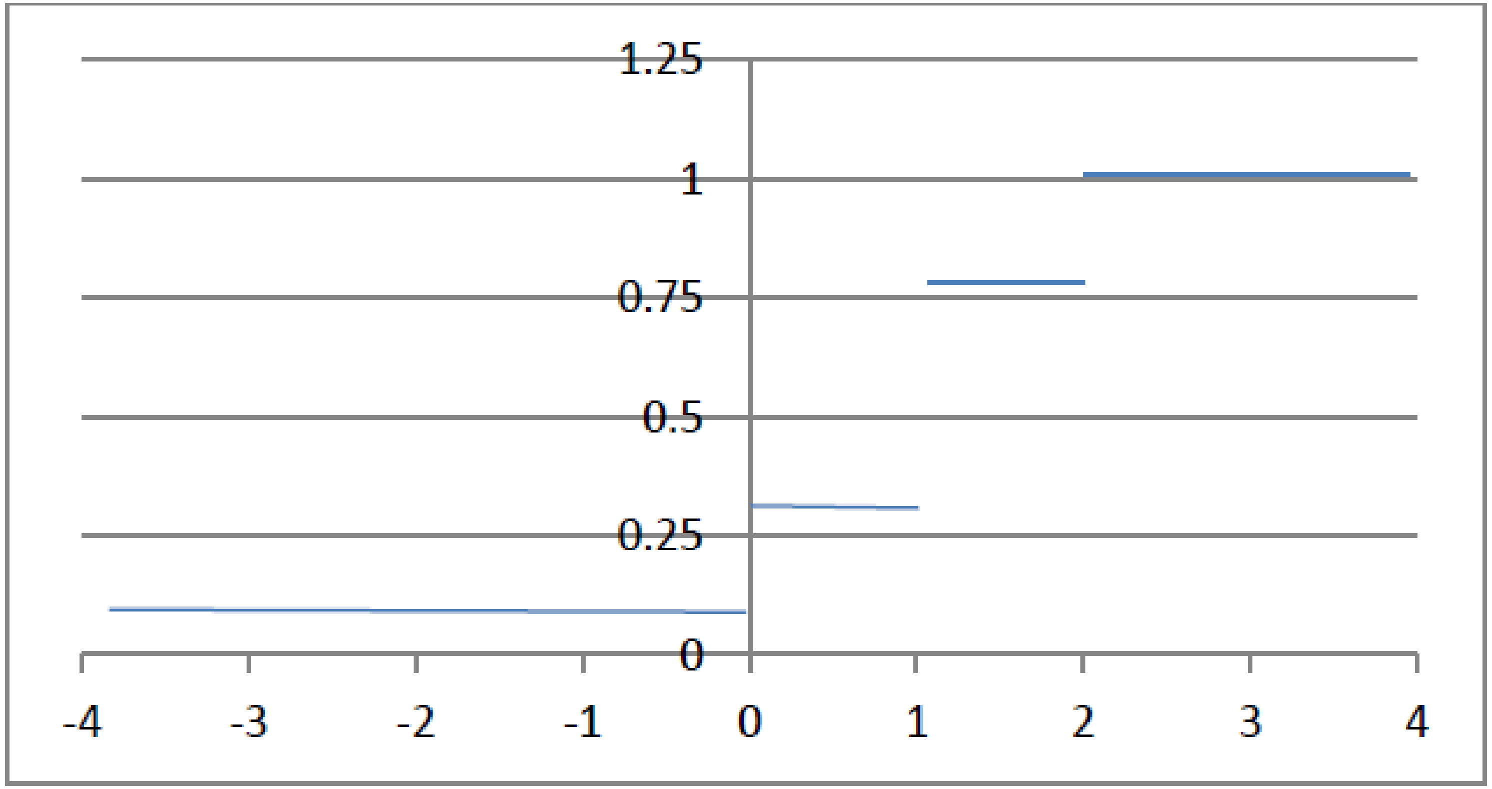
Example: Let us consider the *pmf* considered earlier and given as:

$$P(X = 0) = 1/4,\ P(X = 1) = 1/2,\ P(X = 2) = 1/4.$$

The distribution function for this random variable X can be given as below.

$$F(t) = \begin{cases} 0, & -\infty < t < 0 \\ 0.25, & 0 \leq t < 1 \\ 0.75, & 1 \leq t < 2 \\ 1, & t \geq 2 \end{cases}$$

Graph of this step function will look like below.



Distribution Function for a Discrete Variable

Let us now consider the *pdf*, $f(x) = 2x$, $0 < x < 1$;

$= 0$, elsewhere.

Distribution function for the random variable following this *pdf* shall be:

$$F(t) = \int_{-\infty}^{t} 2x\,dx = \int_{0}^{t} 2x\,dx = [x^2]_0^t = t^2.$$

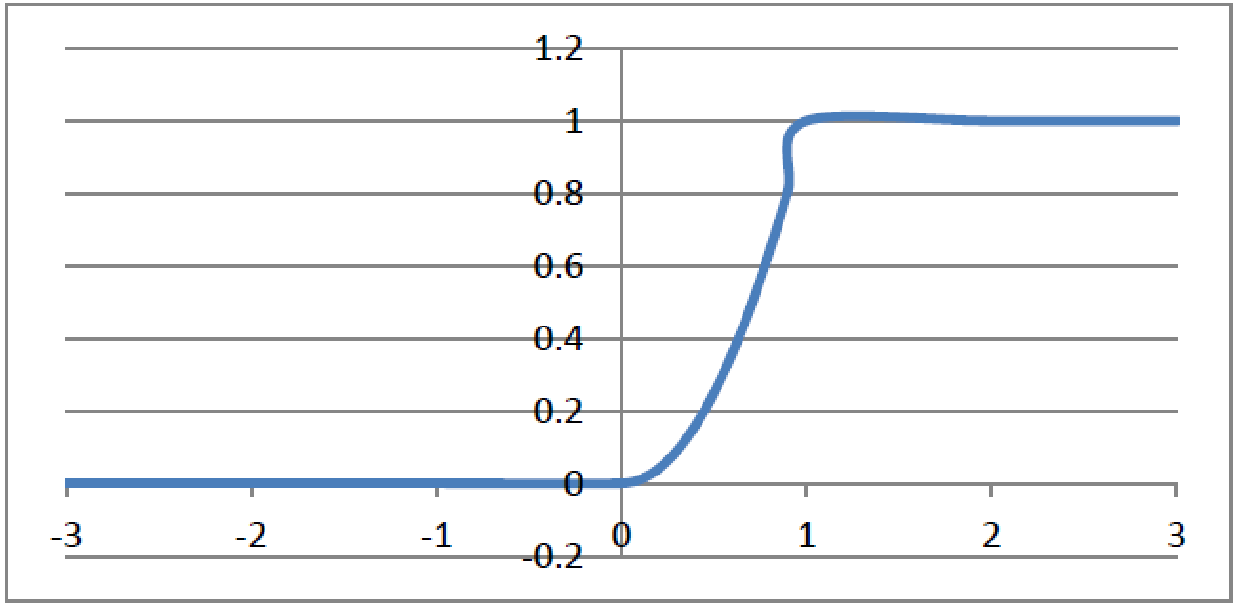
The distribution function shall thus be defined as,

$$F(t) = 0,\ t < 0$$

$$t^2, 0 \le t \le 1$$

$$1,\ t > 1$$

The graph of this function is:



Distribution Function for a Continuous Random Variable

One can note that this is a continuous function defined over the interval $(-\infty, \infty)$.

Example.

A random variable X has the following probability function:

Value of X, x :	0	1	2	3	4	5	6	7
$p(x)$:	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2+k$

(i) find k (ii) Evaluate $P(X<6)$ &

(iii) $P(X \geq 6)$ (iv) If $P(X \leq a) > \frac{1}{2}$ find a

(v) Determine the distribution function $F(x)$

Solution:

$$\text{Since} \sum_{x=0}^{7} p(x) = 1$$

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

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

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

$$\Rightarrow k = \frac{1}{10}, \quad \underline{k = -1} \rightarrow \text{rejected}$$

$$\Rightarrow k = \frac{1}{10}$$

(ii) $P(X<6) = P(X=0) + P(X=1) + \ldots P(X=5)$

$$= \frac{1}{10} + \frac{2}{10} + \frac{2}{10} + \frac{3}{10} + \frac{1}{100} = \frac{81}{100}$$

(iii) $\therefore P(X \geq 6) = 1 - P(X<6)$

$$= 1 - \frac{81}{100} = \frac{19}{100}$$

(iv) $P(X \leq a) > \frac{1}{2}$, by seeing distribution function

we have $\underline{a = 4}$

X	$F(x) = P(X \leq x)$
	0
0	$k = \frac{1}{10}$
1	$3k = \frac{3}{10}$
2	$5k = \frac{5}{10}$
3	$8k = \frac{8}{10} = \frac{4}{5}$
4	$8k + k^2 = 81/100$
5	$8k + 3k^2 = 83/100$
6	$9k + 10k^2 = 1$
7	

(v) → Table is distribution function.

Example 2:

If $p(x) = \begin{cases} \frac{x}{15}, & x = 1,2,3,4,5 \\ 0 & \text{otherwise} \end{cases}$

Find (i) $P(X = 1 \text{ or } 2)$, (ii) $P\left(\left(\frac{1}{2} < X < 5/2\right) \middle| X > 1\right)$

Soln $P(X = 1 \text{ or } X = 2) = P(X=1) + P(X=2)$

$$= \frac{1}{15} + \frac{2}{15} = \frac{3}{15} = \frac{1}{5}$$

iii) $P\left(\frac{1}{2} < X < 5/2 \mid X > 1\right) = \dfrac{P\left\{\left(\frac{1}{2} < X < 5/2\right) \cap (X > 1)\right\}}{P(X > 1)}$

$$= \frac{P\{(X = 1 \text{ or } 2) \cap (X > 1)\}}{P(X>1)}$$

$$= \frac{P(X = 2)}{1 - P(X=1)} = \frac{2/15}{14/15} = 1/7 \text{ Ans}$$

Ex: 3

Two dice are rolled. Let X denote the random variable which counts the number of points on the upturned faces, construct a table giving the non zero values of the probability mass function and draw the probability chart. Also find the distribution function of X.

$$p(2) = P(X=2) = P\{1,1)\} = \frac{1}{36}; \quad p(3) = P(X=3) = P\{(1,2),(2,1)\} = \frac{2}{36}$$

$$p(4) = P(X=4) = P\{(1,3),(2,2),(3,1)\} = \frac{3}{36}$$

$$p(5) = P(X=5) = P\{(1,4),(2,3)\ (3,2),(4,1)\} = \frac{4}{36}$$

$$\vdots \qquad \vdots$$

$$p(11) = P(X=11) = P\{(6,5),(5,6)\} = \frac{2}{36}; \quad p(12) = P(X=12) = P\{(6,6)\}$$

$=1/36$

Distribution Function.

$$F(1) = P(X \le 1) = 0,\ F(2) = P(X \le 2) = \frac{1}{36}$$

$$F(3) = P(X \le 3) = P(X = 2) + P(X = 3)$$

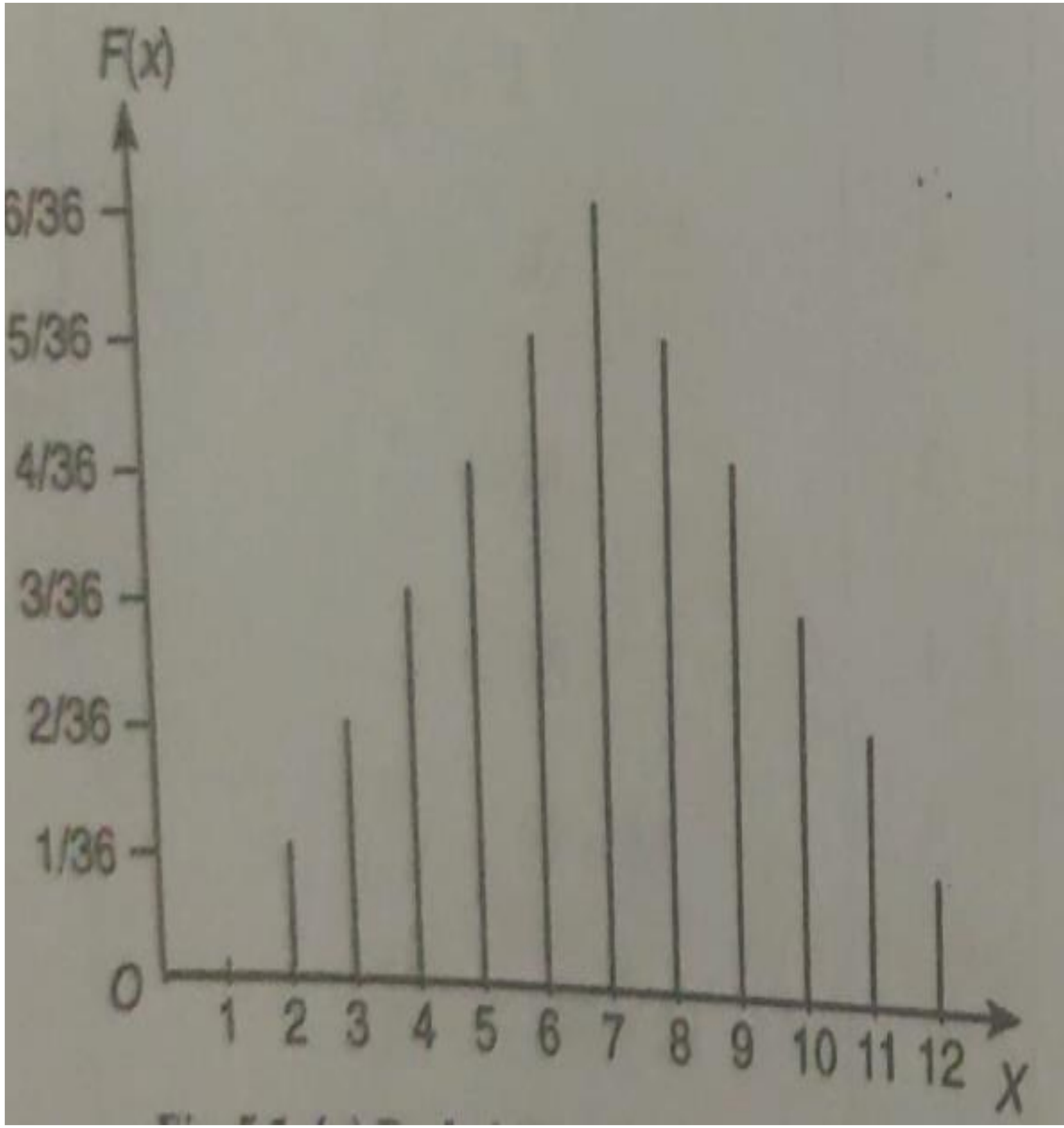
$$= p(2) + p(3) = \frac{1}{36} + \frac{2}{36} = \frac{3}{36}$$

$$F(4) = P(X \le 4) = P(X = 2) + P(X = 3) + P(X = 4)$$

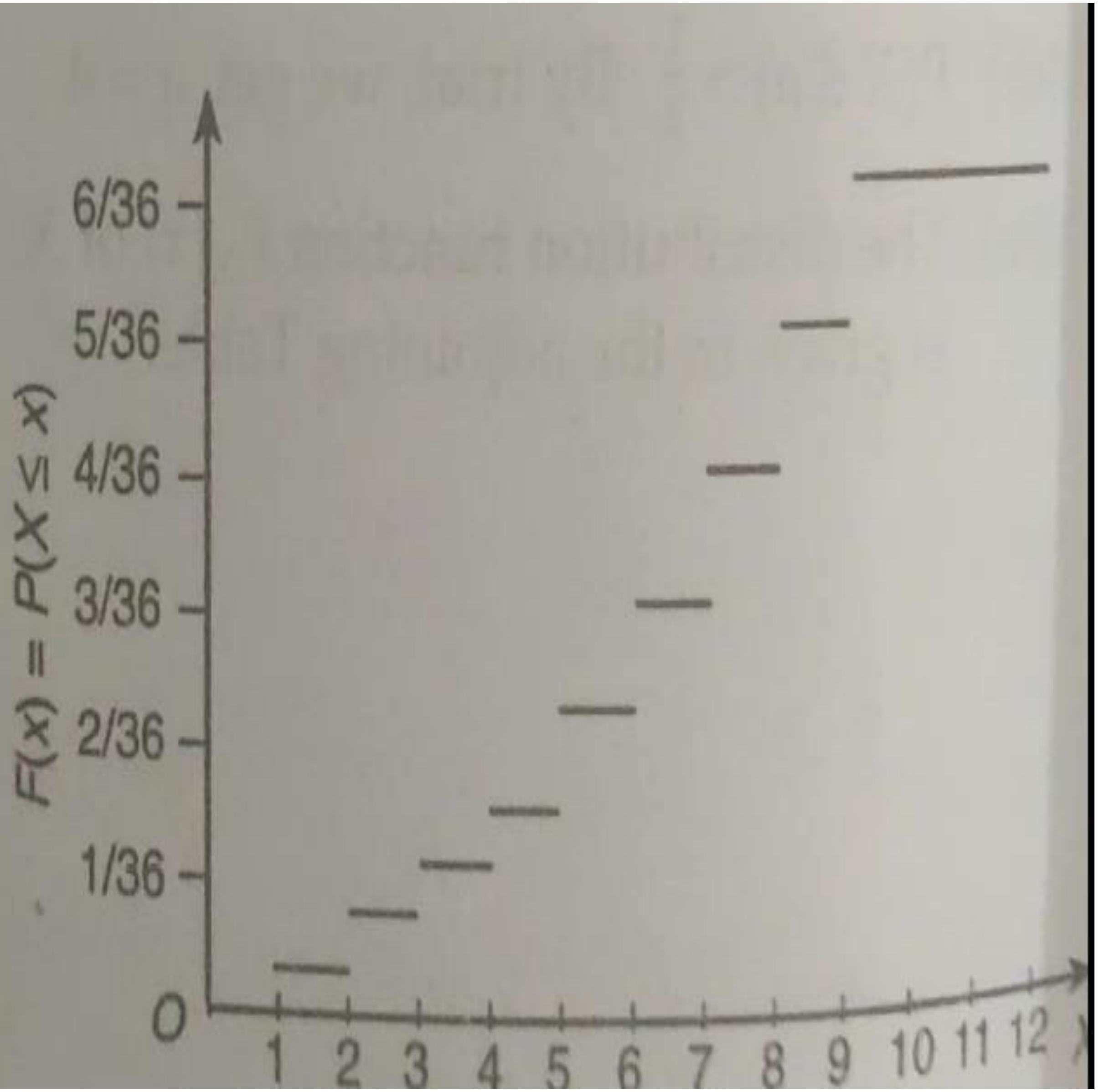
$$= \frac{1}{36} + \frac{2}{36} + \frac{3}{36} = \frac{6}{36}, \text{ and so on.}$$

The distribution function of X is shown in the adjointing Table.

$$F(x) = \begin{cases} 0, & \text{for } x < 2 \\ \frac{1}{36}, & \text{for } 2 \le x < 3 \\ \frac{3}{36}, & \text{for } 3 \le x < 4 \\ \frac{6}{36}, & \text{for } 4 \le x < 5 \\ \vdots & \vdots \\ \frac{35}{36}, & \text{for } 11 \le x < 12 \\ 1 & \text{for } x \ge 12 \end{cases}$$



Probability function



Distribution function

Example:

The diameter of an electric cable, say X is assumed to be a continuous random variable with p.d.f : $f(x) = 6x(1-x),\ 0 \leq x \leq 1$

(i) check that $f(x)$ is p.d.f.

(ii) Determine a number 'b' such that

$P(X < b) = P(X > b)$.

Solution:

For $0 \le x \le 1$, $f(x) \geq 0$,

$$\int_0^1 f(x)\, dx = 6\int_0^1 x(1-x)\, dx = 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1$$

$$= 1$$

Hence f(x is a p.d.f.

(ii)

$$\int_0^b f(x)\,dx = \int_b^1 f(x)\,dx.$$

$$\Rightarrow.\quad 6\int_0^b x(1-x)\,dx = 6\int_b^1 x(1-x)\,dx$$

$$= \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^b = \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_b^1$$

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

$$\Rightarrow (3b^2-2b^3)\frac{2}{6} = \frac{1}{6}$$

$$\Rightarrow 4b^3-6b^2+1 = 0$$

$$\Rightarrow (2b-1)(2b^2-2b-1)=0$$

$$\Rightarrow b=\frac{1}{2}, \quad b = \frac{2\pm\sqrt{12}}{4} = \frac{1\pm\sqrt{3}}{2}$$

So only real value between 0 & 1 is $\boxed{\frac{1}{2}}$

Example :- Let X be a continuous random variable with p.d.f

$$f(x) = \begin{cases} ax & 0 \le x < 1 \\ a & 1 \le x \le 2 \\ -ax + 3a & 2 \le x < 3 \\ 0 & \text{elsewhere} \end{cases}$$

(i) determine the constant a

(ii) compute $P(X \le 1.5)$

Solution :-

constant 'a' is determined from the consideration that total probability is unity i.e $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{2} f(x)\,dx + \int_{2}^{3} f(x)\,dx + \int_{3}^{\infty} f(x)$$

$=$

$$\Rightarrow \int_{0}^{1} ax\,dx + \int_{1}^{2} a\cdot dx + \int_{2}^{3} (-ax+3a)\,dx = 1$$

$$\Rightarrow a\left[\frac{x^2}{2}\right]_0^1 + a(x]_1^2 + \left\{-a\cdot\frac{x^2}{2} + 3ax\right]_2^3 = 1$$

$$\Rightarrow \frac{a}{2} + a + a\left[\left(-\frac{9}{2}+9\right) - (-2+6)\right\} = 1$$

$$\Rightarrow 2a = 1$$

$$\Rightarrow a = 1/2$$

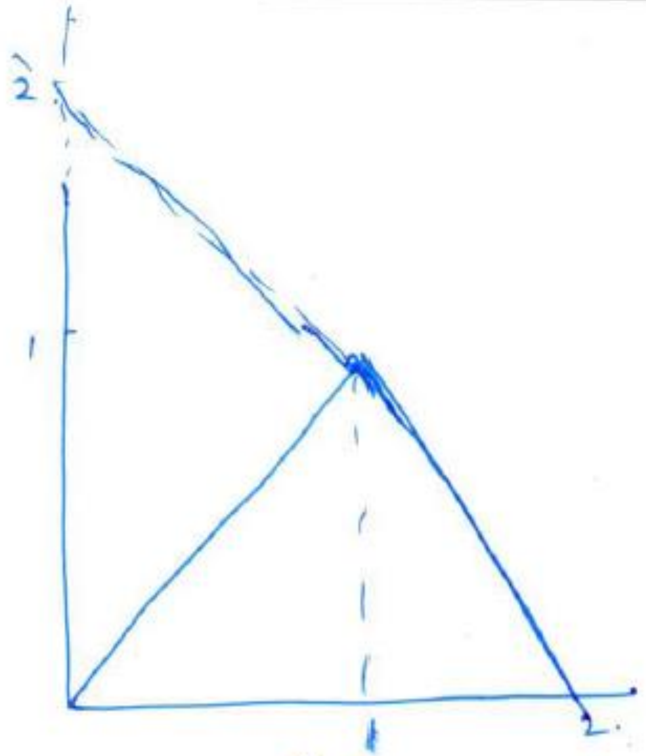
(ii) $P(X \leq 1.5) = \int_{-\infty}^{1.5} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_0^1 f(x)\,dx + \int_1^{1.5} f(x)\,dx$

$$= 0 + a\int_0^1 x\,dx + \int_1^{1.5} a\,dx$$

$$= \frac{a}{2} + 0.5a = a = 1/2$$

Ex:-

Given



$f(x)$

This is graph of $f(x)$; find the $f(x)$?
and then $F(x)$ (Distribution function). and calculate
the $P(0.3 < x \leq 1.5)$?

solution

$$f(x) = \begin{cases} x & 0 < x < 1 \\ 2-x & 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

check whether this is p.d.f. or not.

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

given

$$F(x) = \begin{cases} 0 & \text{if } x \le 0 \\ \int_0^x t\,dt = \frac{x^2}{2} & \text{if } 0 < x \le 1 \\ \int_0^1 t\,dt + \int_1^x (2-t)\,dt = 2x - \frac{x^2}{2} - 1 & \text{if } 1 < x \le 2 \end{cases}$$

Ex-2 for what value of k following function defines a PDF, also calculate

$P(X \geq 0.3)$, $f(x) = \begin{cases} kx(1-x), & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$

Solution $\int_0^1 f(x)\,dx = k/6$; so this determines

PDF for $k = 6$.

$$P(X > 0.3) = 1 - 6\int_0^{0.3} x(1-x)\,dx$$

$$= 0.784 \quad A_2.$$