RANDOM VARIABLE

A **rule** that assigns a real number to each outcome is called **random variable**.

The rule is nothing but a function of the variable X that assigns a unique value to each outcome of the random experiment.

When a variable X takes the value x_i with he probability p_i (i=1,2,3,…,n) then X is called **random variable** or **stochastic variable** or **variate**.

There are two types of random variable : Discrete Random Variable and Continuous Random Variable.

DISCRETE RANDOM VARIABLE

A random variable X which can take only a finite number of values in an interval of the domain called discrete random variable.

Example :

- Number of mistakes in a page.
- Number appearing on the top of a die.

DISCRETE PROBABILITY DISTRIBUTION

If a random variable x can assume a discrete set of values say $x_1, x_2, \ldots, x_n$ with respect to probabilities $p_1, p_2, \ldots, p_n$ such that $p_1+p_2+\ldots+p_n=1$ then the occurrences of value x_i with respective probabilities p_i is called discrete probability distribution of X.

Example : Consider the experiment of throwing a pair of dice

Let X denotes the integer between 2 and 12

Then discrete probability distribution of X with probabilities $P(X)$ is given by

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

Probability Function or Probability Mass Function (pmf)

Probability Function or Probability Mass Function (pmf) of a random variable X is a function is a function *p(x)* which gives the probabilities corresponding to different possible discrete set of values say $x_1, x_2, \ldots, x_n$ of variable *x*.

$$p(x_i) = p(x = x_i) = \text{Probability that on variable } x \text{ assumes value } x_i$$

The function *p(x)* satisfies the condition

$(i) p(x_i) \geq 0$

$(ii) \sum p(x_i) = 1$

Cumulative Distribution Function (Distribution Function)

If X is a random variable then $P(X \leq x)$ is called the cumulative distribution function (cdf) or distribution function and is denoted by *F(x)*.

So, $F(x) = P(X \leq x)$

Expectation of a Discrete Random Variable

If x is a discrete random variable which assumes the discrete set of values $x_1, x_2, \dots, x_n$ with the respective probabilities $p_1, p_2, \dots, p_n$ then the expression or expected value of x is denoted by *E(X)* and defined as

$$E(X) = p_1x_1 + p_2x_2 + \cdots + p_nx_n = \sum_{i=1}^{n} p_ix_i$$

Similarly, the expected value of X^2 is defined as $E(X^2) = \sum_{i=1}^{n} p_ix_i^2$

Properties of Expectation

1. If X is a random variable and a be constant then

    i. $E(a) = a$

    ii. $E(aX) = aE(X)$

    iii. $E(X - \mu) = 0$

2. If x and y are two random variables then $E(X \pm Y) = E(X) \pm E(Y)$

3. $E(XY) = E(X)E(Y)$ if X and Y are two independent random variables.

4. If $y = ax + b$ where a and b are constants then $E(Y) = E(aX + b) = aE(X) + b$

A pair of coin is tossed , what is the expected value of getting head ?

Let X = number of heads

$X = 0, 1, 2$

Probability Distribution is given by

X	**0**	**1**	**2**
$P(X)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$$E(X) = (\frac{1}{4} \times 0) + (\frac{1}{2} \times 1) + (\frac{1}{4} \times 2) = 1$$

For the discrete probability distribution

x	0	1	2	3	4	5	6	7
p	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2+k$

Determine

(i) k (ii) mean (iii) variance

(iv) smallest value of k s.t. $P(X \leq x) > 1/2$

We know that,

$$\sum f(x) = 1$$

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

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

$$\Rightarrow k = -1 \text{ or } 1/10$$

$k = -1$ is not possible

$$\therefore k = 1/10$$

$$\text{Mean} = \sum x f(x) \qquad \Big|\ p(x) = f(x)$$

$$= 3.66$$

$$\text{Variance} = E[X^2] - (E[X])^2$$

$$= \left[0 + \left(1 \times \frac{1}{10}\right) + \left(2^2 \times \frac{2}{10}\right) + \left(3^2 \times \frac{2}{10}\right) + \left(4^2 \times \frac{3}{10}\right) + \left(5^2 \times \frac{1}{100}\right) + \left(6^2 \times \frac{2}{100}\right) + \left(7^2 \times \frac{77}{100}\right)\right] - (3.66)^2$$

$$= 37.7$$

$P(x \leq 0) = f(0) = 0$

$P(x \leq 1) = f(0) + f(1) = 0 + \frac{1}{10} = 0.1$

$P(X \leq 2) = f(0) + f(1) + f(2) = 0 + \frac{1}{10} + \frac{2}{10} = 0.03$

$P(X \leq 3) = f(0) + f(1) + f(2) + f(3) = 0.5$

$P(X \leq 4) = f(0) + f(1) + f(2) + f(3) + f(4)$

$= 0.8$ $\quad \left(\therefore \underline{x = 4} \right.$

$\therefore$ Smallest value of x s.t. $P(X \leq x) > \frac{1}{2}$ is 4

A fair coin is tossed until head or five tails occurs. Find expected no. of tosses of the coin.

Let, X = no. of tosses

X can take values $1, 2, 3, 4, 5, 6$

X	1	2	3	4	5	6
Outcome	H	TH	TTH	TTTH	TTTTH	TTTTT
P(X)	$\frac{1}{2}$	$(\frac{1}{2})^2$	$(\frac{1}{2})^3$	$(\frac{1}{2})^4$	$(\frac{1}{2})^5$	$(\frac{1}{2})^5$

$$\therefore E[X] = \sum x\,p(x) = 1.96$$

$\Rightarrow$ Expected no. of tosses ≈ 2

If X and Y are discrete random variables and k is a constant then prove that

(i) $E[X+k] = E[X] + k$

(ii) $E[X+Y] = E[X] + E[Y]$

X & Y are discrete R.V.
and k is constant.

$$E[X] = \sum_{i=1}^{n} x_i p_i(x_i)$$

$$E[X] = \sum x\, p(x)$$

$$\sum_{i=1}^{n} p_i = 1$$

$$E[X] = \sum x p$$

$$\sum p = 1 \quad \text{or} \quad \sum p(x) = 1$$

$$E[X+k] = \sum (x+k)\, p = \sum x p + \sum k p$$

$$= E[X] + k \sum p = E[X] + k$$

$$E[X+Y] = \sum (x+y)p$$

$$= \sum xp + \sum yp$$

$$= E[X] + E[Y]$$

A random variables X has the following probability distribution where k is some number

$$P(x) = \begin{cases} k, & x=0 \\ 2k, & x=1 \\ 3k, & x=2 \\ 0, & \text{otherwise} \end{cases}$$

(a) Determine value of k

(b) Find $P(x<2)$, $P(x \le 2)$, $P(x \geq 2)$

Independent Random Variable

$E[xy] = E[x]\,E[y]$

$$E[XY] = E[X]\,E[Y]$$

where X & Y are independent R.V.

Covariance

If x & y are two R.V. with mean $\bar{x}$ & $\bar{y}$ respectively then covariance between x & y is defined as

$$\operatorname{cov}(x,y) = E\left[(x-\bar{x})(y-\bar{y})\right]$$

$\Rightarrow$ The covariance of two independent random variable is zero.

Let x & y be two independent R.V.

$$\therefore E[xy] = E[x]\,E[y] \quad \text{--- (1)}$$

$$\mathrm{Cov}(x,y) = E[(x-\bar{x})(y-\bar{y})]$$

$$= E[(xy - x\bar{y} - \bar{x}y + \bar{x}\bar{y})]$$

$E[x+y]$

$= E[x] + E[y]$

$$= E[xy] - E[x\bar{y}] - E[\bar{x}y] + E[\bar{x}\bar{y}]$$

$$= E[x]E[y] - \bar{y}E[x] - \bar{x}E[y] + \bar{x}\bar{y}$$

(using 1)

$\bar{x}, \bar{y}$

const

$$= \bar{x}\bar{y} - \bar{y}\bar{x} - \bar{x}\bar{y} + \bar{x}\bar{y}$$

$$= 0$$

Correlation Coefficient

$$\rho = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x)}\,\sqrt{\text{Var}(y)}}$$

Properties of covariance

1. $\text{Cov}(X, X) = \text{Var}(X)$

2. If X & Y are two independent R.V. $\text{Cov}(X, Y) = 0$

③ $\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$

④ $\mathrm{cov}(aX, Y) = a\,\mathrm{cov}(X, Y)$, $a = \text{const}$

⑤ $\mathrm{cov}(X + c, Y) = \mathrm{cov}(X, Y)$, $c = \text{const}$

⑥ $\mathrm{cov}(X + Y, Z) = \mathrm{cov}(X, Z) + \mathrm{cov}(Y, Z)$

Properties of correlation

① $-1 \le \rho(X, Y) \le 1$

② If $\rho(X, Y) = 1$ then $Y = aX + b$ where $a > 0$

③ If $\rho(X, Y) = -1$ then $Y = aX + b$ where $a < 0$

④ $\rho(aX + b, cY + d) = \rho(X, Y)$ for $a, c > 0$

Moment of Random Variable

The moments of a random variable (or its distribution) are expected values of powers or related f^n of the random variable.

The rth moment of X is $\mu'_r = E(X^r)$

$= \sum x^r P(X = x)$

$= \sum x^r p(x)$

1st moment $\longrightarrow$ $\mu'_1 = E(X^1) = E(X)$

$= \sum x\, p(x)$

2nd moment $= \mu_2' = E[X^2] = \sum x^2 p(x)$

3rd moment $= \mu_3' = E[X^3] = \sum x^3 p(x)$

$\rightarrow$ r^{th} central moment $= \mu_r = E[X - \mu_x]^r$

2nd central moment $= \mu_2 = E(X - \mu_x)^2$

3rd " " $= \mu_3 = E[X - \mu_x]^3$

Q. The 2nd central moment is : (a) mean (b) variance (c) S.D. (d) none

The r^{th} central moment of X is $\mu_r = E(X - \mu_X)^r$

Q. Let X be a discrete random variable having probability mass f^n

$$P_X(x) = \begin{cases} 1/2 & , x = 1 \\ 1/3 & , x = 2 \\ 1/6 & , x = 3 \\ 0 & , \text{otherwise} \end{cases}$$

Find 3rd moment of X

3rd moment is given by $\mu_3' = E[X^3]$

$$= \sum x^3 p(x)$$

$$= \left\{(1)^3 \times \left(\frac{1}{2}\right)\right\} + \left\{(2)^3 \times \frac{1}{3}\right\} +$$

$$\left\{3^3 \times \frac{1}{6}\right\}$$

$$= 23/3$$

Let X be a discrete R.V. with p.m.f.

$$P_X(x) = \begin{cases} 3/4 & , \quad x=1 \\ 1/4 & ; \quad x=2 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find the 3rd central moment of X.

3rd central moment is given by $\mu_3 = E[X-\mu_X]^3$

Now,

$$\mu_X = E[X] = \sum x\,p(x) = \left(1\times\frac{3}{4}\right) + \left(2\times\frac{1}{4}\right) + 0$$

$$= \frac{3}{4} + \frac{2}{4} = \frac{5}{4}$$

$$\mu_3 = E\left[x - \frac{5}{4}\right]^3 = \sum \left(x - \frac{5}{4}\right)^3 p(x)$$

$$= \left(1 - \frac{5}{4}\right)^3 \left(\frac{3}{4}\right) + \left(2 - \frac{5}{4}\right)^3 \left(\frac{1}{4}\right)$$

$$= \frac{3}{32}$$

Moment Generating Function

The moment generating f^n (m.g.f.) of a R.V. X having the probability f^n $f(x)$ is given by

$$M_x(t) = E(e^{tx})$$

$$= \sum_x e^{tx} f(x) \longleftarrow \text{Discrete R.V.}$$

$$= \int e^{tx} f(x)\,dx \longleftarrow \text{cont. R.V.}$$

Here t is real const.

$$M_X(t) = E(e^{tX}) = E\left[1 + tX + \frac{(tX)^2}{2!} + \dots + \frac{(tX)^r}{r!} + \dots\right)$$

$$= E\left[1 + tX + \frac{t^2X^2}{2!} + \dots + \frac{t^rX^r}{r!} + \dots\right)$$

$$= 1 + tE[X] + \frac{t^2}{2!}E[X^2] + \dots + \frac{t^r}{r!}E[X^r) + \dots$$

$$1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \dots + \frac{t^r}{r!}\mu_r' + \dots$$

Coefficient of $\frac{t^r}{r!}$ will give rth moment about the origin.

Note ① $\frac{d^r}{dt^r}[M_X(t)] = \mu_r'$

3rd moment, $r = 3$

$\frac{d^3}{dt^3}(M_X(t)) = \mu_3'$ ← 3rd moment

Moment generating f^n of X about the point $x = a$

$$M_X(t) \text{ (about } x = a) = E\left[e^{t(X-a)}\right]$$

$$M_X(t) \text{ (about mean)} = E\left[e^{t(X-\bar{X})}\right]$$

Here, $\bar{X}$ = mean

Properties

① $M_{cX}(t) = M_X(ct)$

(2) $M_{X_1+X_2+\cdots+X_n}(t) = M_{X_1}(t)\cdot M_{X_2}(t)\cdots M_{X_n}(t)$

Q. Find m.g.f. of a random variable whose moments are

$$\mu_r' = (r+1)!\, 2^r$$

$$M_X(t) = \sum_{r=0}^{\infty}\frac{t^r}{r!}\mu_r' = \sum_{r=0}^{\infty}\frac{t^r}{r!}(r+1)!\, 2^r$$

$$= \sum_{r=0}^{\infty}\frac{t^r}{r!}(r+1)\cdot r!\, 2^r$$

$$= \sum_{r=0}^{\infty} t^r (r+1)\, 2^r = \sum_{r=0}^{\infty}(r+1)(2t)^r$$

$$M_X(t) = \left[(0+1)(2t)^0\right] + \left[(1+1)(2t)^1\right] + \left[(2+1)(2t)^2\right] + \cdots$$

$$= 1 + 2\cdot(2t) + 3(2t)^2 + \cdots$$

$$= (1-2t)^{-2}$$

Q. Show that mgf of a R.V. X having the probability density f^n

$$f(x) = \begin{cases} 1/3 & , -1 < x < 2 \\ 0 & , \text{elsewhere} \end{cases}$$

is $$M_X(t) = \begin{cases} \dfrac{e^{2t} - e^{-t}}{3t} & , t \neq 0 \\ 1 & , t = 0 \end{cases}$$

$$M_X(t) = E(e^{tX})$$

$$= \int_{-1}^{2} e^{tx} f(x)\, dx$$

$$= \int_{-1}^{2} e^{tx} \left(\frac{1}{3}\right) dx \qquad (1)$$

$$= \frac{1}{3} \int_{-1}^{2} e^{tx}\, dx$$

$$= \frac{1}{3} \left[\frac{e^{tx}}{t} \right]_{-1}^{2} = \frac{1}{3t} \left(e^{2t} - e^{-t} \right) \; ; t \neq 0$$

Sub. $t=0$ in (1) we get

$$M_X(t) = \int_{-1}^{2} e^{0.x} \left(\frac{1}{3}\right) dx$$

$$= \frac{1}{3}\int_{-1}^{2} dx$$

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

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

$$= \frac{3}{3} = 1$$

$$\therefore M_X(t) = \begin{cases} \dfrac{e^{2t}-e^{-t}}{3t}, & t \neq 0 \\ 1, & t = 0 \end{cases}$$

Q. Find the mgf of the R.V. X having the probability density f^n

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

Also find the mean and variance of X using mgf.

$$M_X(t) = E[e^{tX}]$$

$$= \int_0^1 e^{tx} f(x)\,dx + \int_1^2 e^{tx} f(x)\,dx$$

$$M_X(t) = \left[\frac{e^{tx} \cdot x}{t} \right]_0^1 - \int_0^1 \frac{e^{tx}}{t}(1)\,dx + \left[\frac{e^{tx}}{t}(2-x) \right]_1^2 - \int_1^2 \frac{e^{tx}}{t}(-1)\,dx$$

$$= \frac{e^t}{t} - \left[\frac{e^{tx}}{t^2} \right]_0^1 - \frac{e^t}{t} + \left[\frac{e^{tx}}{t^2} \right]_1^2$$

$$= \frac{e^t}{t} - \left[\frac{e^t}{t^2} - \frac{1}{t^2} \right] - \frac{e^t}{t} + \left[\frac{e^{2t}}{t^2} - \frac{e^t}{t^2} \right]$$

$$= \frac{e^{2t}}{t^2} - \frac{2e^t}{t^2} + \frac{1}{t^2}$$

1

$$= \frac{1}{t^2}\left(e^{2t} - 2e^t + 1\right) = \frac{(e^t-1)^2}{t^2}$$

11

Expanding $M_X(t)$ in (1) we get

$$M_X(t) = \frac{1}{t^2}\left[\left(1+2t+\frac{(2t)^2}{2!}+\frac{(2t)^3}{3!}+\cdots\right)-2\left(1+t+\frac{t^2}{2!}+\cdots\right)+1\right]$$

$$\frac{1}{t^2}\left(t^2+t^3+\frac{7}{12}t^4+\cdots\right)$$

$$= 1+t+\frac{7}{12}t^2+\cdots \qquad \text{(iii)}$$

Mean, $\mu_1' =$ coefficient of t in $M_X(t)$

$= 1$ (From (iii))

$\mu_2' =$ coefficient of $\frac{t^2}{2!}$ in $M_X(t)$ $= \frac{7}{12}\times 2! = \frac{7}{6}$

Variance $(\mu_2) = \mu_2' - (\mu_1')^2 = \frac{7}{6} - (1)^2 = \frac{1}{6}$