Random Variable

A Random Variable X is a symbol that represents the outcome of an experiment.

Example - consider an experiment of tossing a pair of identical coins and let the random variable X be the number of heads obtained. The sample space S is follows.

Outcome	TT	HT	TH	HH
Value of X	0	1	1	2

$P(X=0) = \frac{1}{4}$, $P(X=1) = \frac{2}{4} = \frac{1}{2}$ and $P(X=2) = \frac{1}{4}$

Random Variable

Discrete   continuous

→ **Discrete Random Variable**

A discrete variable is a variable which can only take a countable numbers of values.

Eg. (0, 1, 2, 3),

→ **Continuous Random Variable**

A continuous random variable is random variable where the data can take infinitely many values

Example [0, 1], [a, b]

Distribution function

The cumulative distribution function (CDF also cumulative density funsity) of a real valued random variable X or just distribution function of X. evaluated at x is the probability that X will take a value less than or equal to x.

$$F_X(x) = P_X(X \leq x) \quad \forall x$$

Remark Domein of Distribution function is $(-\infty, \infty)$ and Range is $[0,1]$

→ Discrete Probability Distribution

The set of values x_i with their probability P_i constitute a discrete probability distribution of the discrete variate X

$p(x) = P(X = x)$ and if $x_1, x_2 \ldots$ are the possible values of X.

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

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

→ Continuous Probability Distribution

if $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx$

or $P\left(x - \frac{1}{2}dx \leq X \leq x + \frac{1}{2}dx\right) = f(x)\,dx$

where $F(x)$ is defined as the cumulative distribution function or simply the distribution function of the continuous variate X.

It is the probability that the value of the variate X will be $\leq x$.

→ The distribution function $F(x)$ has the following Properties

(i) $F'(x) = f(x) \geq 0$

So that $F(x)$ is a non-decreasing function.

(ii) $F(-\infty) = 0$

(iii) $F(\infty) = 1$

(iv) $P(a \leq x \leq b) = \int_a^b f(x)\,dx = \int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx$

$= F(b) - F(a)$

Example 26.28. *A die is tossed thrice. A success is 'getting 1 or 6' on a toss. Find the mean and variance of the number of successes.* (V.T.U., 2011 S ; Rohtak, 2004)

Solution. Probability of a success $= \frac{2}{6} = \frac{1}{3}$, Probability of failures $= 1 - \frac{1}{3} = \frac{2}{3}$.

$\therefore$ prob. of no success = Prob. of all 3 failures $= \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{8}{27}$

Probability of one successes and 2 failures $= 3c_1 \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{4}{9}$

Probability of two successes and one failure $= 3c_2 \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = \frac{2}{9}$

Probability of three successes $= \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{27}$

Now

$x_i = 0$	1	2	3
$p_i = 8/27$	4/9	2/9	1/27

$\therefore$ mean $\qquad \mu = \Sigma p_i x_i = 0 + \frac{4}{9} + \frac{4}{9} + \frac{1}{9} = 1.$

Also $\qquad \Sigma p_i x_i^2 = 0 + \frac{4}{9} + \frac{8}{9} + \frac{9}{27} = \frac{5}{3}$

$\therefore$ variance $\qquad \sigma^2 = \Sigma p_i x_i^2 - \mu^2 = \frac{5}{3} - 1 = \frac{2}{3}.$

Example 26.29. *The probability density function of a variate X is*

X :	0	1	2	3	4	5	6
$p(X)$:	k	$3k$	$5k$	$7k$	$9k$	$11k$	$13k$

(i) Find $P(X < 4)$, $P(X \geq 5)$, $P(3 < X \leq 6)$. (V.T.U., 2013)

(ii) What will be the minimum value of k so that $P(X \leq 2) > 3$.

Solution. (*i*) If X is a random variable, then

$$\sum_{i=0}^{6} p(x_i) = 1 \text{ i.e., } k + 3k + 5k + 7k + 9k + 11k + 13k = 1 \text{ or } k = 1/49.$$

$\therefore \qquad P(X < 4) = k + 3k + 5k + 7k = 16k = 16/49.$

$P(X \geq 5) = 11k + 13k = 24k = 24/49.$

$P(3 < X \leq 6) = 9k + 11k + 13k = 33k = 33/49.$

(*ii*) $\qquad P(X \leq 2) = k + 3k + 5k = 9k > 0.3$ or $k > 1/30$

Thus minimum value of $k = 1/30$.

Example 26.31. (i) *Is the function defined as follows a density function?*

$$f(x) = e^{-x}, \quad x \geq 0$$
$$= 0, \quad x < 0.$$

(ii) *If so, determine the probability that the variate having this density will fall in the interval (1, 2)?*
(iii) *Also find the cumulative probability function F(2)?*

Solution. (i) $f(x)$ is clearly > 0 for every x in (1, 2) and

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} 0 \cdot dx + \int_{0}^{\infty} e^{-x}\,dx = 1$$

Hence the function $f(x)$ satisfies the requirements for a density function.

(ii) Required probability $= P(1 \leq x \leq 2) = \int_{1}^{2} e^{-x}\,dx = e^{-1} - e^{-2} = 0.368 - 0.135 = 0.233.$

This probability is equal to the shaded area in Fig. 26.3 (*a*).

(iii) Cumulative probability function $F(2)$

$$\int_{-\infty}^{2} f(x)\,dx = \int_{-\infty}^{0} 0 \cdot dx + \int_{0}^{2} e^{-x}\,dx = 1 - e^{-2} = 1 - 0.135 = 0.865$$

which is shown in Fig. 26.3 (*b*).

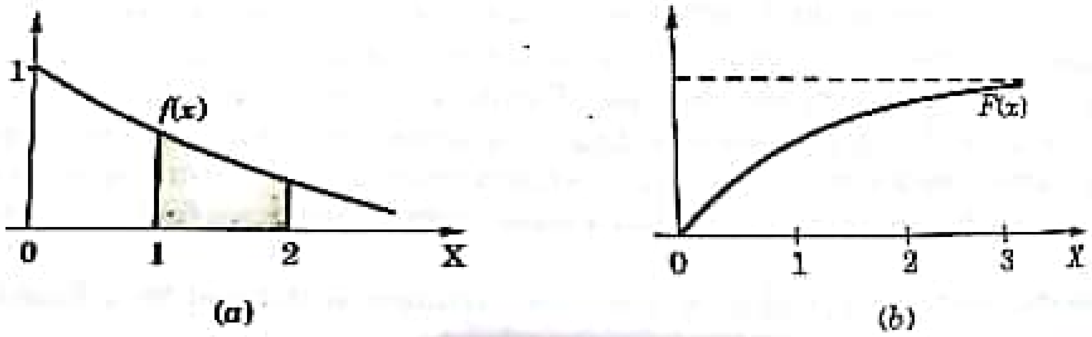

Fig. 26.3

26.10 (1) EXPECTATION

The mean value (μ) *of the probability distribution of a variate X is commonly known as its expectation and is denoted by E(X).* If $f(x)$ is the probability density function of the variate X, then

$$\sum_i x_i f(x_i) \quad \text{(discrete distribution)}$$

or

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx \quad \text{(continuous distribution)}$$

In general, expectation of any function $\phi(x)$ is given by

$$E[\phi(x)] = \sum_i \phi(x_i) f(x_i) \quad \text{(discrete distribution)}$$

or

$$E[\phi(x)] = \int_{-\infty}^{\infty} \phi(x) f(x)\,dx \quad \text{(continuous distribution)}$$

(2) **Variance** *of a distribution is given by*

$$\sigma^2 = \sum_i (x_i - \mu)^2 f(x_i) \quad \text{(discrete distribution)}$$

or

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\,dx \quad \text{(continuous distribution)}$$

where σ is the *standard deviation* of the distribution.

(3) *The* **rth moment** *about the mean* (denoted by μ_r) *is defined by*

$$\mu_r = \Sigma(x_i - \mu)^r f(x_i) \quad \text{(discrete distribution)}$$

or

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu)^r f(x)\,dx \quad \text{(continuous distribution)}$$

(4) **Mean deviation from the mean** *is given by*

$$\Sigma \,|\, x_i - \mu \,|\, f(x_i) \quad \text{(discrete distribution)}$$

or by

$$\int_{-\infty}^{\infty} |\, x - \mu \,|\, f(x)\,dx \quad \text{(continuous distribution)}$$

Example 26.32. *In a lottery, m tickets are drawn at a time out of n tickets numbered from 1 to n. Find the expected value of the sum of the numbers on the tickets drawn.* *(Rohtak, 2011)*

Scanned by CamScanner

Solution. Let $x_1, x_2, ..., x_n$ be the variables representing the numbers on the first, second, ..., nth ticket. The probability of drawing a ticket out of n tickets being in each case $1/n$, we have

$$E(x_i) = 1 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} + 3 \cdot \frac{1}{n} + ... + n \cdot \frac{1}{n} = \frac{1}{2}(n+1)$$

$\therefore$ expected value of the sum of the numbers on the tickets drawn

$$= E(x_1 + x_2 + ... + x_m) = E(x_1) + E(x_2) + ... + E(x_m)$$

$$= mE(x_i) = \frac{1}{2} m (n+1).$$

Example 26.33. *X is a continuous random variable with probability density function given by*

$$f(x) = kx \ (0 \le x < 2)$$
$$= 2k \ (2 \le x < 4)$$
$$= -kx + 6k \ (4 \le x < 6)$$

Find k and mean value of X. *(J.N.T.U., 2003)*

Solution. Since the total probability is unity

$\therefore$ $$\int_0^6 f(x)\, dx = 1$$

i.e., $$\int_0^2 kx\, dx + \int_2^4 2k dx + \int_4^6 (-kx + 6k)\, dx = 1$$

or $$k \left| x^2/2 \right|_0^2 + 2k \left| x \right|_2^4 + \left(-kx^2/2 + 6kx\right)_4^6 = 1$$

or $$2k + 4k + (-10k + 12k) = 1 \text{ i.e., } k = 1/8.$$

$$\text{Mean of } X = \int_0^6 x\, f(x)\, dx$$

$$= \int_0^2 kx^2\, dx + \int_2^4 2kx\, dx + \int_4^6 x(-kx + 6k) dx$$

$$= k \left| x^3/3 \right|_0^2 + 2k \left| x^2/2 \right|_2^4 + \left(-k \left| x^3/3 \right|_4^6 + 6k \left| x^2/2 \right|_4^6\right)$$

$$= k(8/3) + k(12) - k(152/3) + 3k(20) = \frac{1}{8}(24) = 3.$$

Example 26.34. *A variate X has the probability distribution*

x :	-3	6	9
$P(X = x)$:	$1/6$	$1/2$	$1/3$

Find E (X) and E (X²). Hence evaluate E (2X + 1)².

Solution. $$E(X) = -3 \times \frac{1}{6} + 6 \times \frac{1}{2} + 9 \times \frac{1}{3} = 11/2.$$

$$E(X)^2 = 9 \times \frac{1}{6} + 36 \times \frac{1}{2} + 81 \times \frac{1}{3} = 93/2$$

$\therefore$ $$E(2X+1)^2 = E(4X^2 + 4X + 1) = 4E(X^2) + 4E(X) + 1$$
$$= 4(93/2) + 4(11/2) + 1 = 209.$$

Example 26.35. *The frequency distribution of a measurable characteristic varying between 0 and 2 is as under*

$$f(x) = x^3, \quad 0 \le x \le 1$$
$$= (2-x)^3, \quad 1 \le x \le 2.$$

Calculate the standard deviation and also the mean deviation about the mean.

Solution. Total frequency $N = \int_0^1 x^3\, dx + \int_1^2 (2-x)^3\, dx = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$\therefore$ μ_1' (about the origin) $= \frac{1}{N}\left[\int_0^1 x \cdot x^3\, dx + \int_1^2 x(2-x)^3\, dx\right]$

$$= 2\left\{\left|\frac{x^5}{5}\right|_0^1 + \left|-x \cdot \frac{(2-x)^4}{4}\right|_1^2 - \left|\frac{(2-x)^5}{20}\right|_1^2\right\} = 2\left(\frac{1}{5} + \frac{1}{4} + \frac{1}{20}\right) = 1$$

μ_2' (about the origin) $= \frac{1}{N}\left[\int_0^1 x^2 \cdot x^3 dx + \int_1^2 x^2 (2-x)^3\, dx\right]$

$$= 2\left\{\left|\frac{x^6}{6}\right|_0^1 + \left|-x^2 \frac{(2-x)^4}{4}\right|_1^2 + \frac{1}{2}\int_1^2 x(2-x)^4\, dx\right\}$$

$$= 2\left\{\frac{1}{6} + \frac{1}{4} + \frac{1}{2}\left[\frac{1}{5} + \frac{1}{30}\right]\right\} = \frac{16}{15}$$

Hence $\sigma^2 = \mu_2 = \mu_2' - (\mu_1')^2 = \frac{1}{15}$

i.e., standard deviation $\sigma = \frac{1}{\sqrt{15}}$.

Mean deviation about the mean

$$= \frac{1}{N}\left\{\int_0^1 |x-1|\, x^3\, dx + \int_1^2 |x-1|\,(2-x)^3\, dx\right\}$$

$$= 2\left\{\int_0^1 (1-x)x^3\, dx + \int_1^2 (x-1)(2-x)^3\, dx\right\}$$

$$= 2\left\{\left(\frac{1}{4} - \frac{1}{5}\right) + \left(0 + \frac{1}{20}\right)\right\} = \frac{1}{5}.$$

GENERATING FUNCTION

COVARIANCE AND VARIANCE

Covariance: If $E(X) = \bar{X}$ and $E(Y) = \bar{Y}$, then the covariance of the random variables X and Y, denoted by cov (X, Y), is given by

$$\text{cov}(X, Y) = E[(X-\bar{X})(Y-\bar{Y})] = E(XY - X\bar{Y} - Y\bar{X} + \bar{X}\bar{Y})$$

$$= E(XY) - \bar{Y}E(X) - \bar{X}E(Y) + \bar{X}\bar{Y}$$

$$= E(XY) - \bar{Y}\bar{X} - \bar{X}\bar{Y} + \bar{X}\bar{Y}$$

or
$$\mathbf{cov}(X, Y) = E(XY) - \bar{X}\bar{Y}.$$

Note that, when X and Y are independent, we have

$$E(XY) = E(X)\,E(Y), \quad \text{then} \quad \text{cov}(X, Y) = 0.$$

Also we can easily establish that

1. cov $(aX, bY) = ab$ cov (X, Y)
2. cov $(X + a, Y + b) =$ cov (X, Y)
3. $\text{cov}\left(\frac{X-a}{h}, \frac{Y-b}{k}\right) = \frac{1}{hk}\text{cov}(X, Y)$
4. cov $(X, Y, Z) =$ cov $(X, Y) +$ cov (X, Z).

Variance: It is easy to verify the following properties:

1. Var $(X) =$ cov (X, X)
2. Var $(ax + b) = a^2$ Var (X)
3. Var $(aX + bY) = a^2$ Var $(X) + b^2$ Var $(Y) + 2ab$ cov (X, Y).

MOMENT GENERATING FUNCTION

The moment generating function (m.g.f.) of a random variable X about the origin is defined as

$$M_0(t) = E(e^{tX}) = \sum e^{tX} P(X).$$

This function $M_0(t)$ is called moment generating function because all the moments of X can be obtained from $M_0(t)$, as follows:

$$M_0(t) = \left[1 + tX + \frac{t^2}{2!}X^2 + \frac{t^3}{3!}X^3 + \ldots + \frac{t^r X^r}{r!} + \ldots\right] p(X)$$

$$= 1 + t\,E(X) + \frac{t^2}{2!}E(X^2) + \ldots + \frac{t^r}{r!}E(X^r) + \ldots$$

$$= 1 + t\,\mu_1' + \frac{t^2}{2!}\mu_2' + \ldots + \frac{t^r}{r!}\mu_r' + \ldots \quad \text{where} \quad \mu_r' = E(X^r).$$

Obviously, we can write here

$$\mu_r' = \frac{d^r}{dt^r}[M_0(t)].$$

As a general case, the moment generating function of the random variable X about an arbitrary point 'a' which can be $\bar{X}$ (or μ) also, is defined as

$$M_a(t) = E\left(e^{t(X-a)}\right) = e^{-at}E(e^{tX}).$$

Moment generating function of the sum of two independent variables is the product of their moment generating functions.

To establish it, consider two independent random variables X and Y whose m.g.fs. are $M_X(t)$ and $M_Y(t)$ respectively, then m.g.f. of $X + Y$ is given by

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E(e^{tX} \cdot e^{tY})$$

$$= E(e^{tX})\, E(e^{tY}) \qquad \text{(since } X \text{ and } Y \text{ are independent)}$$

Thus, $\quad \boldsymbol{M_{X+Y}(t) = M_X(t)\, M_Y(t).}$

Further, it is easy to verify that

1. $M_{cX}(t) = M_X(ct)$ where c is some constant

2. If $U = \dfrac{X-a}{h}$ then $M_U(t) = e^{at/h}\, M_X\left(\dfrac{t}{h}\right)$.

EXAMPLE 31.5. Find the probability distribution of the number of green balls drawn when three balls are drawn one by one without replacement from a bag containing three green and five white balls.

SOLUTION: Let X be the random variable which is the number of green balls drawn when three balls are drawn without replacement.

$$\therefore \quad P(X=0) = P \text{ (no green ball is drawn)} = P(W, W, W)$$

$$= \frac{5}{8}\cdot\frac{4}{7}\cdot\frac{3}{6}.$$

$$P(X=1) = P \text{ (one green ball is drawn)}$$

$$= P(G, W, W) + P(W, G, W) + P(W, W, G)$$

$$= \frac{3}{8}\cdot\frac{5}{7}\cdot\frac{4}{6}+\frac{5}{8}\cdot\frac{3}{7}\cdot\frac{4}{6}+\frac{5}{8}\cdot\frac{4}{7}\cdot\frac{3}{6}$$

$$P(X=2) = P(\text{two green balls are drawn})$$

$$= P(G, G, W) + P(G, W, G) + P(W, G, G)$$

$$= \frac{3}{8}\cdot\frac{2}{7}\cdot\frac{5}{6}+\frac{3}{8}\cdot\frac{5}{7}\cdot\frac{2}{6}+\frac{5}{8}\cdot\frac{3}{7}\cdot\frac{3}{6}$$

$$P(X=3) = P(G\,G\,G) = \frac{3}{8}\cdot\frac{2}{7}\cdot\frac{1}{6}.$$

Probability distribution is

X	0	1	2	3
$P(X)$	$\frac{5}{28}$	$\frac{15}{28}$	$\frac{15}{28}$	$\frac{1}{56}$

... 4 are known to be defective. A college

EXAMPLE 31.7. A random variate X has the following probability distribution

x	0	1	2	3	4
$p(x)$	k	$2k$	$2k$	k^2	$5k^2$

Find the value of k and obtain $P(X < 3)$ and $P(0 < X < 4)$. Determine the distribution function of X, also mean and variance of X.

SOLUTION: Since $\sum p(x) = 1$ we have $k + 2k + 2k + k^2 + 5k^2 = 1$ or $6k^2 + 4k + k - 1 = 0$ which gives $k = -1$ and $k = 1/6$. Therefore $k = 1/6$ as k cannot be negative. Ans.

Next, $$P(X < 3) = p(0) + p(1) + p(2) = \frac{1}{6} + \frac{2}{6} + \frac{2}{6} = \frac{5}{6}.$$

and $$P(0 < X < 4) = p(1) + p(2) + p(3) = \frac{2}{6} + \frac{2}{6} + \frac{1}{36} = \frac{25}{36}$$

Now, the probability distribution and the distribution function is as follows:

x	0	1	2	3	4
$p(x)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{1}{36}$	$\frac{5}{36}$
$F(x)$	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{5}{6}$	$\frac{31}{36}$	1

Then $$\text{mean} = \mu_x = \sum x_i p_i = 0 + \frac{2}{6} + \frac{4}{6} + \frac{3}{36} + \frac{20}{36} = \frac{59}{36}$$

$$\text{Variance} = \sigma_x^2 = \left[0\left(\frac{1}{6}\right) + 1\left(\frac{2}{6}\right) + 4\left(\frac{2}{6}\right) + 9\left(\frac{1}{36}\right) + 16\left(\frac{5}{36}\right)\right] - \left(\frac{59}{36}\right)^2$$

Two most important discrete probability distributions are – Binomial Distribution and Poisson distribution.