MODULE-2

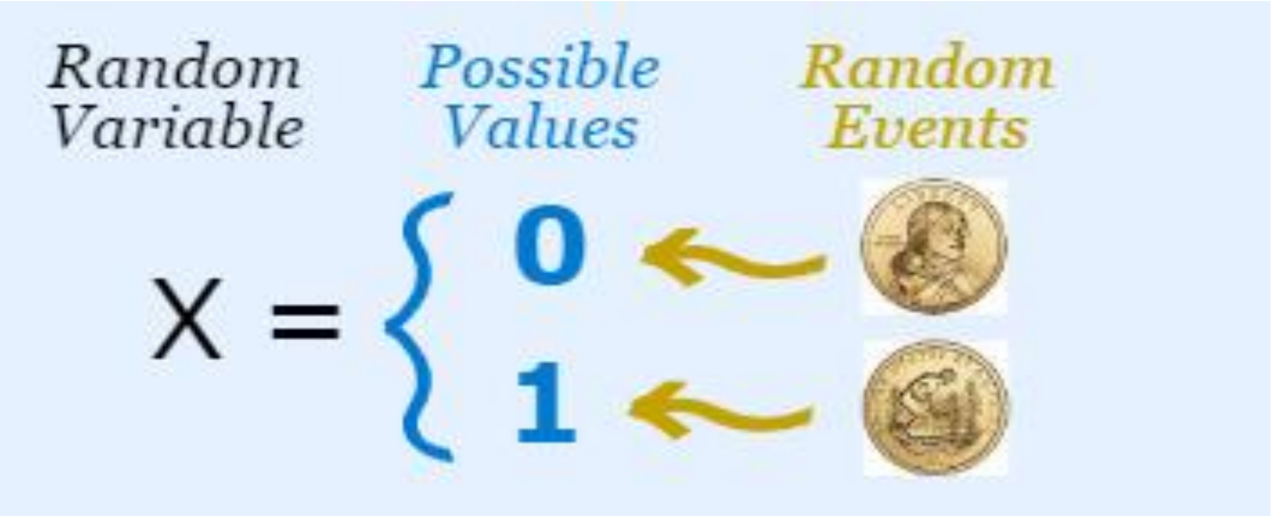
RANDOM VARIABLES

Random Variables

A Random Variable is a set of **possible values** from a random experiment.

Example: Tossing a coin: we could get Heads or Tails.

Let's give them the values **Heads=0** and **Tails=1** and we have a Random Variable "X“:



- In short:

$$X = \{0, 1\}$$

- Note: We could choose Heads=100 and Tails=150 or other values if we want! It is our choice.

- We have an **experiment** (such as tossing a coin)
- We give **values** to each event
- The **set of values** is a **Random Variable**

Not Like an Algebra Variable

In Algebra a variable, like **x,** is an unknown value:

Example: $x + 2 = 6$

In this case we can find that $x=4$

But a Random Variable is different ...

A Random Variable has a whole **set of values** ...

and it could take on **any** of those values, randomly.

EXAMPLES

Example: $X = \{0, 1, 2, 3\}$

- X could be 0, 1, 2, or 3 **randomly**.
- And they might each have a different probability.

NOTE: We use a capital letter, like **X** or **Y**, to avoid confusion with the Algebra type of variable.

Sample Space

A Random Variable's set of values is the Sample Space.

Example: Throw a die once

Random Variable **X** = "The score shown on the top face".

X could be 1, 2, 3, 4, 5 or 6

So the Sample Space is {1, 2, 3, 4, 5, 6}

DEFINATIONS

- **Random Experiment**: If in each trail an experiment conducted under identical conditions, the outcome is not unique, but may be one of the possible outcomes, then such an experiment is called a *random experiment*.

Example: Throwing a die, a pack of cards.

- ***Outcome:*** The result of a random experiment will be called an *outcome*.
- ***Trial***: Any particular performance of a random experiment is called a ***trial*** and outcome or combination of outcomes are termed as ***events***.

Example: Tossing of a coin: {H, T}

-

  Exhaustive events or cases: The total number of possible outcomes of a random experiments is known as the exhaustive events or cases.

- ***Favourable Events or Cases***: The number of cases favorable to an event in a trail is the number of outcomes which entail the happening of the event.

Example: In drawing a card from a pack of cards the number of cases
favorable to drawing of an ace is 4,
for drawing a spade card is 13 and
for drawing a red card is 26

Mutually Exclusive events: Events are said to be mutually exclusive or incompatible if the happening of any one of them precludes the happening of all the others, *i.e.,* no two or more of them can happen simultaneously in the same trail.

***Equally likely events*:** Outcomes of trail are said to be equally likely if taking into consideration all the relevant evidences, there is no reason to expect one in preference to the others.

Independent Events: Several events are said to be independent if the happening (non-happening) of an event is not affected by the supplementary knowledge concerning the occurrence of any number of remaining events.

Statistical Definition:

If a random experiment or trail results in 'n' exhaustive, mutually exclusive and equally likely outcomes(or cases) , out of which 'm' are favourable to the occurrence of an event $\boldsymbol{E}$, then the probability '**p**' of occurrence of E, denoted by $\boldsymbol{P(E)}$, is given by:

$$p = P(E) = \frac{Number\ of\ favourable\ cases}{Total\ number\ of\ exhaustive\ cases} = \frac{m}{n}$$

(i) Since $m \geq 0, n \geq 0\ and\ m \leq n$

$P(E) \geq 0\ and\ P(E) \leq 1 \quad \Rightarrow 0 \leq P(E) \leq 1$

(ii) $P(E) + P(\bar{E}) = 1$

Probability

- P(X = value) = probability of that value

Example (continued): Throw a die once

X = {1, 2, 3, 4, 5, 6}

In this case they are all equally likely, so the probability of any one is 1/6

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

Note that the sum of the probabilities = **1**, as it should be.

Example: How many heads when we toss 3 coins?

X = "The number of Heads" is the Random Variable.

In this case, there could be 0 Heads (if all the coins land Tails up), 1 Head, 2 Heads or 3 Heads.

So the Sample Space = {0, 1, 2, 3}

But this time the outcomes are NOT all equally likely.

The three coins can land in eight possible ways:

{HHH, HTH, THH, T TH, HHT, HT T, THT, T T T}

- $P(X = 3) = 1/8$
- $P(X = 2) = 3/8$
- $P(X = 1) = 3/8$
- $P(X = 0) = 1/8$

Example: Two dice are tossed.

The Random Variable is **X** = "The sum of the scores on the two dice".

Let's make a table of all possible values:

		1st Die					
		1	**2**	**3**	**4**	**5**	**6**
2nd Die	**1**	2	3	4	5	6	7
	2	3	4	5	6	7	8
	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

There are **6 × 6 = 36** possible outcomes, and the Sample Space (which is the sum of the scores on the two dice) is {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}

Sum of numbers (x)	*Favourable sample points*	*No. of favourable cases*	*Probability* $p(x)$
2	(1, 1)	1	$\frac{1}{36}$
3	(1, 2), (2, 1)	2	$\frac{2}{36}$
4	(1, 3), (3, 1), (2, 2)	3	$\frac{3}{36}$
5	(1, 4), (4, 1), (2, 3), (3, 2)	4	$\frac{4}{36}$
6	(1, 5), (5, 1), (2, 4), (4, 2), (3, 3)	5	$\frac{5}{36}$
7	(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3)	6	$\frac{6}{36}$
8	(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)	5	$\frac{5}{36}$
9	(3, 6), (6, 3), (4, 5), (5, 4)	4	$\frac{4}{36}$
10	(4, 6), (6, 4), (5, 5)	3	$\frac{3}{36}$
11	(5, 6), (6, 5)	2	$\frac{2}{36}$
12	(6, 6)	1	$\frac{1}{36}$

- 2 occurs just once, so $P(X = 2) = 1/36$
- 3 occurs twice, so $P(X = 3) = 2/36 = 1/18$
- 4 occurs three times, so $P(X = 4) = 3/36 = 1/12$
- 5 occurs four times, so $P(X = 5) = 4/36 = 1/9$
- 6 occurs five times, so $P(X = 6) = 5/36$
- 7 occurs six times, so $P(X = 7) = 6/36 = 1/6$
- 8 occurs five times, so $P(X = 8) = 5/36$
- 9 occurs four times, so $P(X = 9) = 4/36 = 1/9$
- 10 occurs three times, so $P(X = 10) = 3/36 = 1/12$
- 11 occurs twice, so $P(X = 11) = 2/36 = 1/18$
- 12 occurs just once, so $P(X = 12) = 1/36$

One dimensional random variable:

• Discrete random variable:

A random variable is called discrete random variable if its set of possible outcomes is countable.

Example: number of defectives in a sample of k items.

• Probability function (or) Probability Mass Function (p.m.f):

If X is a discrete R.V. which can take the values such that , then is called the probability mass function and it satisfies the following conditions:

$$(i)\ \ p_i \geq 0, \text{ for all } i, \text{ and}$$

$$(ii)\ \ \sum_i p_i = 1$$

- ***Probability distribution function:***

The collection of pairs is called the probability distribution of the R.V. X.

$X = x_i$	$P(X = x_i)$
x_1	p_1
x_2	p_2
$\vdots$	$\vdots$
x_r	p_r
$\vdots$	$\vdots$

Example 13·1. *A die is tossed twice. Getting 'an odd number' is termed as a success. Find the probability distribution of the number of successes.*

Solution. Since the cases favourable to getting an odd number in a throw of a die are (1, 3, 5), *i.e.*, 3 in all,

$$\text{Probability of success}(S) = \frac{3}{6} = \frac{1}{2} \quad ; \quad \text{Probability of failure } (F) = 1 - \frac{1}{2} = \frac{1}{2}\cdot$$

If X denotes the number of successes in two throws of a die, then X is a random variable which takes the values 0, 1, 2.

$$P(X = 0) = P[F \text{ in 1st throw and } F \text{ in 2nd throw}] = P(FF) = P(F) \times P(F) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

$$P(X = 1) = P(S \text{ and } F) + P(F \text{ and } S) = P(S)\,P(F) + P(F)\,P(S) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}$$

$$P(X = 2) = P(S \text{ and } S) = P(S)\,P(S) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

Hence the probability distribution of X is given by :

x	0	1	2
$p(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Example 13·2. *Two cards are drawn*

(*a*) *successively with replacement*

(*b*) *simultaneously* (*successively without replacement*),

from a well shuffled deck of 52 cards. Find the probability distribution of the number of aces.

Solution. Let X denote the number of aces obtained in a draw of two cards. Obviously, X is a random variable which can take the values 0, 1 or 2.

(*a*) Probability of drawing an ace $= \frac{4}{52} = \frac{1}{13}$ $\Rightarrow$ Probability of drawing a non-ace $= 1 - \frac{1}{13} = \frac{12}{13}$

Since the cards are drawn with replacement, all the draws are independent.

$$P(X = 2) = P(\text{Ace and Ace}) = P(\text{Ace}) \times P(\text{Ace}) = \frac{1}{13} \times \frac{1}{13} = \frac{1}{169} \cdot$$

$$\begin{aligned} P(X = 1) &= P(\text{Ace and Non-ace}) + P\,(\text{Non-ace and Ace}) \\ &= P(\text{Ace}) \times P(\text{Non-ace}) + P(\text{Non-ace}) \times P(\text{Ace}) \\ &= \frac{1}{13} \times \frac{12}{13} + \frac{12}{13} \times \frac{1}{13} = \frac{24}{169} \end{aligned}$$

$$P(X = 0) = P(\text{Non-ace and Non-ace}) = P(\text{Non-ace}) \times P(\text{Non-ace}) = \frac{12}{13} \times \frac{12}{13} = \frac{144}{169}$$

Hence, the probability distribution of X is :

x :	0	1	2
$p(x)$:	$\frac{144}{169}$	$\frac{24}{169}$	$\frac{1}{169}$

$b)$ If cards are drawn without replacement, then exhaustive number of cases of drawing 2 cards out of 52 cards is ${}^{52}C_2$.

$$\therefore \quad P(X=0) = P(\text{No ace}) = P(\text{Both cards are non-aces}) = \frac{{}^{48}C_2}{{}^{52}C_2} = \frac{48 \times 47}{52 \times 51} = \frac{188}{221}$$

$$P(X=1) = P(\text{one ace}) = P(\text{one ace and one non-ace}) = \frac{{}^{4}C_1 \times {}^{48}C_1}{{}^{52}C_2} = \frac{4 \times 48 \times 2}{52 \times 51} = \frac{32}{221}$$

$$P(X=2) = P(\text{both aces}) = \frac{{}^{4}C_2}{{}^{52}C_2} = \frac{4 \times 3}{52 \times 51} = \frac{1}{221}$$

Hence, the probability distribution of X becomes :

x :	0	1	2
$p(x)$:	$\frac{188}{221}$	$\frac{32}{221}$	$\frac{1}{221}$

Example 13·5. *Four bad apples are mixed accidentally with 20 good apples. Obtain the probability distribution of the number of bad apples in a draw of 2 apples at random.*

Solution. Let X denote the number of bad apples drawn. Then X is a random variable which can take the values 0, 1 or 2.

There are 4 + 20 = 24 apples, in all and the exhaustive number of cases of drawing two apples is ${}^{24}C_2$.

$$\therefore \qquad P(X=0) = \frac{{}^{20}C_2}{{}^{24}C_2} = \frac{20 \times 19}{24 \times 23} = \frac{95}{138} \quad ; \quad P(X=1) = \frac{{}^{4}C_1 \times {}^{20}C_1}{{}^{24}C_2} = \frac{2 \times 4 \times 20}{24 \times 23} = \frac{40}{138}$$

$$P(X=2) = \frac{{}^{4}C_2}{{}^{24}C_2} = \frac{4 \times 3}{24 \times 23} = \frac{3}{138}$$

Hence, the probability distribution of X is :

x :	0	1	2
$p(x)$:	$\frac{95}{138}$	$\frac{40}{138}$	$\frac{3}{138}$

Example 13·3. *Obtain the probability distribution of X, the number of heads in three tosses of a coin (or a simultaneous toss of three coins).*

Solution. Obviously, X is a random variable which can take the values 0, 1, 2 or 3. The sample space S consists of $2^3 = 8$ sample points, as given below :

$$\begin{aligned} S &= \{(H, T) \times (H, T) \times (H, T)\} \\ &= \{(HH, HT, TH, TT) \times (H, T)\} \\ &= \{HHH, HTH, THH, TTH, HHT, HTT, THT, TTT\} \end{aligned}$$

The probability distribution of X is given in Table 13·2.

TABLE 13·2 : **PROBABILITY DISTRIBUTION OF NUMBER OF HEADS IN 3 TOSSES OF A COIN**

No. of heads (x)	*Favourable events*	*No. of favourable cases*	*Probability* $p(x)$
0	$\{TTT\}$	1	$\frac{1}{8}$
1	$\{TTH, HTT, THT\}$	3	$\frac{3}{8}$
2	$\{HTH, THH, HHT\}$	3	$\frac{3}{8}$
3	$\{HHH\}$	1	$\frac{1}{8}$

Problem

1. If the random variable takes the values 1,2,3 and 4 such that $2P(X=1)=3P(X=2)=P(X=3)=5P(X=4)$, find the probability distribution and the cumulative distribution function of X.

Solution:

Let $2P(X=1)=3P(X=2)=P(X=3)=5P(X=4)=k$

That is, $P(X=1)=\frac{k}{2},\ P(X=2)=\frac{k}{3},\ P(X=3)=k\ ,\ P(X=4)=\frac{k}{5}$

We know that, $\sum_i p_i=1$

That is, $\frac{k}{2}+\frac{k}{3}+k+\frac{k}{5}=1 \Rightarrow k=\frac{30}{61}$

The probability distribution of X is given by

x_i	1	2	3	4
$P(x_i)$	$\frac{15}{61}$	$\frac{10}{61}$	$\frac{30}{61}$	$\frac{6}{61}$

The c.d.f. F(x) is defined as $F(x) = P(X \le x)$.

When $x < 1$, $F(x) = 0$

When $1 \le x < 2$, $F(x) = P(X = 1) = \dfrac{15}{61}$

When $2 \le x < 3$, $F(x) = P(X = 1) + P(X = 2) = \dfrac{25}{61}$

When $3 \le x < 4$, $F(x) = P(X = 1) + P(X = 2) + P(X = 3) = \dfrac{55}{61}$

When $x \ge 4$, $F(x) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = 1$

3. A random variable X has the following probability distribution,

Values of X	-2	-1	0	1	2	3
$P(X = x)$	0.1	K	0.2	2K	0.3	3K

i)Find K, (ii) Evaluate $P(X < 2)$ and $P(-2 < X < 2)$ (iii) find the c.d.f. of X

Solution: (i) $K = \frac{1}{15}$

(ii) $P(X < 2) = \frac{1}{2}$ and $P(-2 < X < 2) = \frac{2}{5}$

(iii) F(x) = 0, when $x < -2$

$= \frac{1}{10}$, when $-2 \le x < -1$,

$= \frac{1}{6}$, when $-1 \le x < 0$,

$= \frac{11}{30}$, when $0 \le x < 1$,

$= \frac{1}{2}$, when $1 \le x < 2$,

$= \frac{4}{5}$, when $2 \le x < 3$,

=1, when $3 \le x$

- ## ***Continuous random variable*** :

A random variable X is said to be a continuous random variable if it takes all possible values between certain limits or in an interval which may be finite or infinite.

Example: person's height, (within the range of human heights), not just certain fixed heights.

Probability density function (p.d.f):

In case of a continuous random variable, we do not talk of Probability at a particular point (which is always zero) but we always talk of probability in an interval. If $p(x)\,dx$ is the probability that the random variable X takes the value in a small interval of magnitude dx, *e.g.*, $(x, x + dx)$ or $(x - dx/2 , x + dx/2)$, then $p(x)$ is called the *probability density function* (*p.d.f.*) of the *r.v.* X.

Probability Density Function

If X is a continuous RV such that

$$P\left\{x-\frac{1}{2}\mathrm{d}x \le X \le x+\frac{1}{2}\mathrm{d}x\right\} = f(x)\mathrm{d}x$$

then $f(x)$ is called the *probability density function* (shortly denoted as pdf) of X, provided $f(x)$ satisfies the following conditions:

(i) $f(x) \ge 0$, for all $x \in R_x$, and

(ii) $\int_{R_X} f(x)\mathrm{d}x = 1$

Moreover, $P(a \le X \le b)$ or $P(a < X < b)$ is defined as

$$P(a \le X \le b) = \int_a^b f(x)\mathrm{d}x.$$

Note *When X is a continuous RV*

$$P(X = a) = P(a \le X \le a) = \int_a^a f(x)\,\mathrm{d}x = 0$$

This means that it is almost impossible that a continuous RV assumes a specific value. Hence $P(a \le X \le b) = P(a \le X < b) = P(a < X \le b) = P(a < X < b)$.

Distribution Function (or) Cumulative Distribution Function (cdf)

The cumulative distribution function $F(x)$ of a random variable X with probability distribution $P(x)$ and probability desity function $f(x)$ is

$$F(x) = P(X \leq x) = \begin{cases} \sum_{t \leq x} P(X = t), & \text{If } X \text{ is discrete} \\ \int_{-\infty}^{x} f(x)dx, & \text{If } X \text{ is continuous} \end{cases}$$

Let $F(x)$ be a cumulative distribution function of a continuous random variable X. Then $P(a < X < b) = F(b) - F(a)$ and $f(x) = \frac{d(F(x))}{dx}$, if the derivative exists.

Properties: 1. $0 \le F(x) \le 1 \; ; -\infty < x < \infty$

2. $F'(x) = \frac{d}{dx} F(x) = f(x) \ge 0$

$\Rightarrow$ $F(x)$ is non-decreasing function of x.

3. $F(-\infty) = \lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} \int_{-\infty}^{x} f(x)\, dx = \int_{-\infty}^{-\infty} f(x)\, dx = 0$

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

4. $F(x)$ is continuous function of x on the right.

5. The discontinuities of $F(x)$ are at the most countable.

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

If X takes integral values, *viz.*, 1, 2, 3, …then

$$F(x) = P(X = 1) + P(X = 2) + \ldots + P(X = x)$$

$$\Rightarrow \quad F(x) = p(1) + p(2) + p(3) + \ldots + p(x)$$

Remarks. 1. In the above case,

$$F(x-1) = p(1) + p(2) + \ldots + p(x-1)$$

$$\therefore \quad F(x) - F(x-1) = p(x) \quad \Rightarrow \quad p(x) = F(x) - F(x-1)$$

MOMENTS

If X is a discrete $r.v.$ with probability function $p(x)$ then :

$$\mu'_r = r\text{th moment about any arbitrary point 'A'} = \sum (x - A)^r \,.\, p(x)$$

$$\mu_r = r\text{th moment about mean } (\bar{x}) = \sum (x - \bar{x})^r \,.\, p(x)$$

In particular,

$$\text{Mean } (\bar{x}) = \text{First moment about origin} = \sum x \, p(x).$$

[Taking $A = 0$ and $r = 1$ in (13·6)].

$$\text{Variance } (x) = \mu_2 = \sum (x - \bar{x})^2 \,.\, p(x)$$

If $f(x)$ is the p.d.f of a random variable 'X' which defined in the interval (a, b) then (i) Mean = $\int_a^b xf(x)dx$

(ii) Variance = $\int_a^b (x - mean)^2 f(x)dx$

(iii) Moment about mean = $\int_a^b (x - mean)^r f(x)dx$

EXAMPLE

Suppose that the error in the reaction temperature, in C, for a controlled laboratory experiment is a continuous random variable X having the probability density function

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0. & \textit{elsewhere.} \end{cases}$$

1. Verify that $f(x)$ is a density function.
2. Find $P(0 < X \leq 1)$.

Solution

For(1),

Obviously, $f(x) \geq 0$. *To verify condition*

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

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)dx &= \int_{-1}^{-\infty} f(x)dx + \int_{-1}^{2} f(x)dx + \int_{2}^{\infty} f(x)dx \\
&= 0 + \int_{-1}^{2} \frac{x^2}{3} dx + 0 \\
&= \left[\frac{x^3}{3 \times 3}\right]_{-1}^{2} \\
&= \frac{8}{9} - \frac{(-1)}{9} \\
&= \frac{8}{9} + \frac{(1)}{9} \\
&= 1.
\end{aligned}$$

$$
\begin{aligned}
P(0 < X \leq 1) &= \int_0^1 f(x)dx \\
&= \int_0^1 \frac{x^2}{3} dx \\
&= \left[\frac{x^3}{3 \times 3} \right]_0^1 \\
&= \frac{1}{9} - 0 \\
&= \frac{1}{9}.
\end{aligned}
$$

Problem:

Is the function defined as follows a density function? $f(x) = \begin{cases} e^{-x} & ,x \geq 0 \\ 0 & ,x < 0 \end{cases}$

Solution:

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

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

$$= 0 + [-e^{-x}]_0^{\infty} \quad = -e^{-\infty} + 1 = 1$$

$\Rightarrow f(x)$ Satisfy the condition of the density function.

Problem:

If $f(x) = \begin{cases} xe^{-\frac{x^2}{2}} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$

Then show that $f(x)$ is a p.d.f. and find its c.d.f. F(x).

Solution: $\int_0^\infty f(x)dx = \int_0^\infty xe^{\frac{-x^2}{2}} dx = 1$

To find F(x): $F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{0} f(x)dx + \int_0^x f(x)dx = 0 + \int_0^x xe^{\frac{-x^2}{2}} dx$

$$= 1 - e^{\frac{x^2}{2}}$$

Problem:

If a random variable 'X' has the p.d.f. $f(x) = \begin{cases} \dfrac{(x+1)}{2} & ,-1 < x < 1 \\ 0 & ,otherwise \end{cases}$

Then find the mean and variance.

Solution:

We know that, Mean = $\int_{-1}^{1} xf(x)dx = \int_{-1}^{1} x\dfrac{(x+1)}{2}dx = \dfrac{1}{2}\int_{-1}^{1}(x^2+x)dx = \dfrac{1}{3}$

Variance = $\int_{-1}^{1}(x-mean)^2 f(x)dx = \int_{-1}^{1}(x-\dfrac{1}{3})^2 \dfrac{(x+1)}{2}dx$

$$= \int_{-1}^{1}(9x^2+1-6x)(x+1)dx = \frac{2}{9}$$

Problem:

Verify whether the following is a distribution function.

$$F(x)=\begin{cases}0, & x<-a\\ \frac{1}{2}\left(\frac{x}{a}+1\right) & ,-a<x<a\\ 1 & ,x>1\end{cases}$$

Solution:

We know that $f(x)=\dfrac{dF(x)}{dx}=\begin{cases}\dfrac{1}{2a} & ,-a<x<a\\ 0 & ,otherwise\end{cases}$

Also, $\int\limits_{-a}^{a} f(x)dx=\int\limits_{-a}^{a}\frac{1}{2a}dx=\frac{1}{2a}[x]_{-a}^{a}=\frac{1}{2a}(2a)=1$

Therefore, $f(x)$ is a p.d.f. and F(x) is a distribution function.

1. Define a random variable and its probability distribution. Explain by means of two examples.

2. State, with reasons, if the following probability distributions are admissible or not.

(*i*)

x :	0	1	2
$p(x)$:	0·3	0·2	0·5

(*ii*)

x :	–1	0	1
$p(x)$:	0·4	0·4	0·3

(*iii*)

x :	0	1	2	3
$p(x)$:	0·2	0·3	0·3	0·1

(*iv*)

x :	–2	–1	0	1	2
$p(x)$:	0·3	0·4	– 0·2	0·2	0·3

Ans. (*i*) Yes, (*ii*) No, since $\sum p(x) > 1$, (*iii*) No, since $\sum p(x) < 1$, (*iv*) No, since $p(0) = -0{\cdot}2$ which is not possible.

Problem:

A random variable X has the following probability mass function is given

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

1. Find k.
2. Evaluate $P(X \leq 2)$ and $P(-1 \leq X \leq 2)$.
3. Find the cumulative distribution function.

For(1), since $P(x)$ is p.m.f, and $\sum_x P(x) = 1.$, it follows that we get

$$\sum_x P(x) = 1$$

$$0.1 + k + 0.2 + 2k + 0.3 + k = 1$$

$$0.6 + 4k = 1$$

$$4k = 0.4$$

$$k = \frac{0.4}{4}$$

$$k = 0.1$$

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

For(2),

$$P(X \leq 2) = P(X = -2) + P(X = -1) + P(X = 0) + P(X = 1) + P(X = 2$$

$$= 0.1 + 0.1 + 0.2 + 0.2 + 0.3$$

$$= 0.9$$

$$\begin{aligned} P(X \leq 2) &= 1 - P(X > 2) \\ &= 1 - P(X = 3) \\ &= 1 - 0.1 \\ &= 0.9 \end{aligned}$$

Now,

$$\begin{aligned} P(-1 \leq X \leq 2) &= P(X = -1) + P(x = 0) + P(x = 1) + P(x = 2) \\ &= 0.1 + 0.2 + 0.2 + 0.3 \\ &= 0.8 \end{aligned}$$

$$F(x) = P(X \leq x) = \sum_{t \leq x} P(X = t).$$

$$F(x) = \begin{cases} P(X \leq -2) = 0.1, & \text{If } x = -2 \\ P(X \leq -1) = 0.2, & \text{If } x = -1 \\ P(X \leq 0) = 0.4, & \text{If } x = 0 \\ P(X \leq 1) = 0.6, & \text{If } x = 1 \\ P(X \leq 2) = 0.9, & \text{If } x = 2 \\ P(X \leq 3) = 1, & \text{If } x = 3. \end{cases}$$

Problem:

A random variable X has the following probability mass function is given

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

1. Find k.
2. Evaluate $P(X \leq 6)$, $P(X \geq 6)$ and $P((1.5 < X < 4.5)/(X > 2))$
3. Find the cumulative distribution function.

Problem:

Let X be a continuous random variable with probability density function (pdf) is

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

1. Find $P(X \leq 0.4)$, $P(X \geq \frac{3}{4})$ and $P(X \geq \frac{1}{2}.)$
2. Evaluate $P(\frac{1}{2} < X \leq \frac{3}{4})$ and $P(X > \frac{3}{4} / X > \frac{1}{2})$.
3. Find the cumulative distribution function.

Solution:

$$\text{For}(2), P\left(X > \frac{3}{4} / X > \frac{1}{2}\right) = \frac{P((X > \frac{3}{4}) \cap (X > \frac{1}{2}))}{P(X > \frac{1}{2})}$$

$$= \frac{P(X > \frac{3}{4})}{P(X > \frac{1}{2})}$$

$$= \frac{\int_{\frac{3}{4}}^{1} 2x dx}{\int_{\frac{1}{2}}^{1} 2x dx}$$

$$= \frac{\left[x^2\right]_{\frac{3}{4}}^{1}}{\left[x^2\right]_{\frac{1}{2}}^{1}} = \frac{\frac{7}{16}}{\frac{3}{4}} = \frac{7}{12}.$$

Problem:

The Department of Energy (DOE) puts projects out on bid and generally estimates what a reasonable bid should be. Call the estimate b. The DOE has determined that the density function of the winning (low) bid is

$$f(y) = \begin{cases} \frac{5}{8b}, & \frac{2}{5}b \leq y \leq 2b \\ 0. & \text{elsewhere.} \end{cases}$$

Find $F(y)$ and use it to determine the probability that the winning bid is less than the DOE's preliminary estimate b.

Solution:

To find $F(y)$, $F(y) = P(Y \le y) = \int_{-\infty}^{y} f(y)dy.$

For $\frac{2}{5}b \le y \le 2b$, then $P(Y \le y) = \int_{\frac{2}{5}b}^{2b} \frac{5}{8b}dy = \left[\frac{5y}{8b}\right]_{\frac{2}{5}b}^{2b} = \frac{5y}{8b} - \frac{1}{4}.$

Thus,

$$F(y) = \begin{cases} 0, & y \le \frac{2}{5}b \\ \frac{5y}{8b} - \frac{1}{4}, & \frac{2}{5}b \le y \le 2b \\ 1. & y \ge 2b. \end{cases}$$

Two-dimensional Random Variables

Definition: Let X and Y be two r.v. defined on the same sample space S, then the function (X, Y) that assigns a point in $R^2 (= R \times R)$, is called a two-dimensional random variable.

When (X, Y) is a two-dimensional discrete r.v., the possible values of (X, Y) may be represented as (x_i, y_j), $i = 1, 2, \ldots\ldots, n$; $j = 1, 2, \ldots\ldots, m$.

If (X, Y) can assume all values in a specified region R in the xy –plane, (X, Y) is called a two-dimensional continuous R.V.

Joint Probability Function

Let X and Y be random variables on a sample space S with respective image sets $X(S) = \{x_1, x_2, \ldots\ldots, x_n\}$ and $Y(S) = \{y_1, y_2, \ldots\ldots, y_m\}$.

$$X(S) \times Y(S) = \{x_1, x_2, \ldots\ldots, x_n\} \times \{y_1, y_2, \ldots\ldots, y_m\}$$

Into a probability space by defining the probability of the ordered pair to be $\boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x}_i, \boldsymbol{Y} = \boldsymbol{y}_j)$ or $\boldsymbol{p}(\boldsymbol{x}_i, \boldsymbol{y}_j)$.

The function $\boldsymbol{p}$ on $X(S) \times Y(S)$ is defined by:

$$p_{ij} = P\left(X = x_i \cap Y = y_j\right) = p\left(x_i, y_j\right) \text{ is called the}$$

joint probability function of $\boldsymbol{X}$ *and* $\boldsymbol{Y}$.

$X \backslash Y$	y_1	y_2	y_3	$\cdots$	y_j	$\cdots$	y_m	Total
x_1	p_{11}	p_{12}	p_{13}	$\cdots$	p_{1j}	$\cdots$	p_{1m}	$p_{1\cdot}$
x_2	p_{21}	p_{22}	p_{23}	$\cdots$	p_{2j}	$\cdots$	p_{2m}	$p_{2\cdot}$
x_3	p_{31}	p_{32}	p_{33}	$\cdots$	p_{3j}	$\cdots$	p_{3m}	$p_{3\cdot}$
.	.	.	.	$\cdots$	.	$\cdots$	.	
.	.	.	.	$\cdots$	.	$\cdots$	.	$\vdots$
.	.	.	.	$\cdots$	.	$\cdots$	.	
x_i	p_{i1}	p_{i2}	p_{i3}	$\cdots$	p_{ij}	$\cdots$	p_{im}	$p_{i\cdot}$
.	.	.	.	$\cdots$	.	$\cdots$	.	
.	.	.	.	$\cdots$	.	$\cdots$	.	$\vdots$
.	.	.	.	$\cdots$	.	$\cdots$	.	
x_n	p_{n1}	p_{n2}	p_{n3}	$\cdots$	p_{nj}	$\cdots$	p_{nm}	$p_{n\cdot}$
Total	$p_{\cdot 1}$	$p_{\cdot 2}$	$p_{\cdot 3}$	$\cdots$	$p_{\cdot j}$	$\cdots$	$p_{\cdot m}$	1

Joint Probability Mass Function:

If (X, Y) is a two-dimensional discrete random variable, then the joint discrete function of X, Y also called the joint probability mass function of (X, Y) denoted by $\boldsymbol{p_{XY}}$ is defined as:

$$p_{XY}(x_i, y_j) = P(X = x_i, Y = y_j) \text{ for a value of } (x_i, y_j) \text{ of } (X, Y)$$

$$p_{XY}(x_i, y_j) = 0 \text{ , otherwise.}$$

$$\sum\sum p_{XY}(x_i, y_j) = 1, \text{ where the summation is taken over all possible values of } (X, Y).$$

Marginal Probability Function:

$$p_X(x_i) = P(X = x_i)$$

$$= P(X = x_i \cap Y = y_1) + P(X = x_i \cap Y = y_2) + \ldots + P(X = x_i \cap Y = y_m)$$

$$p_X(x_i) = p_{i1} + p_{i2} + \ldots + p_{ij} + \cdots + p_{im} = \sum_{j=1}^{m} p_{ij} = p_{i\cdot}$$

Also, $\sum_{i=1}^{n} p_{i\cdot} = p_{1\cdot} + p_{2\cdot} + \ldots + \cdots + p_{n\cdot} = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1$

Similarly, $p_Y(y_j) = P(Y = y_j) = \sum_{i=1}^{n} p_{ij} = p_{\cdot j}$

which is the marginal probability function of Y.

Conditional Probability Function:

The conditional probability mass function of X, given $Y = y$ is defined as:

$$p_{X|Y}(x|y) = \frac{P(X=x,Y=y)}{P(Y=y)}, \text{ provided } P(Y = y) \neq 0$$

Similarly, $$p_{Y|X}(y|x) = \frac{P(X=x,Y=y)}{P(X=x)}$$

A necessary and sufficient condition for the discrete random variables X and Y to be independent is:

$P\left(X = x_i, Y = y_j\right) = P\left(X = x_i\right) P(Y = y_j)$ for all values (x_i, y_j) of (X,Y).

Joint Probability Density Function

If (X, Y) is two-dimensional continuous RV such that,

$$P\left(x - \frac{dx}{2} \leq X \leq x + \frac{dx}{2} \; and \; y - \frac{dy}{2} \leq Y \leq y + \frac{dy}{2}\right) = f(x,y)dxdy$$

Then $f(x,y)$ is called the joint pdf of (X, Y), provided $f(x,y)$ satisfies the following conditions:

(i) $f(x,y) \geq 0$, for all $(x,y) \in R$, where R is the range space.

(ii) $\iint_R f(x,y)\ dx\, dy = 1$.

Two-dimensional distribution function

Definition: The distribution function of the two-dimensional random variable (X, Y) is a real valued function F defined for all real x and y by the relation:

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) = \sum_{y_j \leq y} \sum_{x_i \leq x} p_{ij}$$

In the continuous case,

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\, dx\, dy$$

Properties: $\frac{\partial^2 F}{\partial x \partial y} = f(x, y)$

Marginal Distribution Functions:

$$F_X(x) = P(X \le x) = P(X \le x, Y < \infty) = \lim_{y \to \infty} F_{XY}(x, y) \ = \ F_{XY}(x, \infty)$$

Similarly, $F_Y(y) = P(Y \le y) = P(X < \infty, Y \le y) = \lim_{x \to \infty} F_{XY}(x, y) \ = \ F_{XY}(\infty, y)$

In case of joint discrete r.v., the marginal distribution functions are:

$$F_X(x) = \sum_y P(X \le x, Y = y) \text{ and } F_Y(y) = \sum_x P(X = x, Y \le y)$$

In case of Jointly continuous r.v.,

$$F_X(x) = \int_{-\infty}^{x} \{\int_{-\infty}^{\infty} f_{XY}(x, y)\, dy\} dx \quad F_Y(y) = \int_{-\infty}^{y} \{\int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\} dy$$

Marginal Density function:

$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$ - is the marginal density function of X.

$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$ - is the marginal density function of Y

$$P(a \le X \le b) = P(a \le X \le b, -\infty \le Y \le \infty)$$
$$= \int_a^b \{\int_{-\infty}^{\infty} f_{XY}(x,y)\, dy\}\, dx$$
$$= \int_a^b f_X(x)dx$$

$$P(a \le X \le b) = \int_c^d f_Y(y)dy$$

Independent Random Variables

If (X,Y) is two-dimensional independent RVs,

$$P(X = x,\ Y = y) = P(X = x) \cdot P(Y = y) \quad \text{- Discrete}$$

$$f(x, y) = f(x) \cdot f(y) \quad \text{- Continuous}$$

Example 10: For a bivariate probability distribution of (X, Y) given below,

X \ Y	1	2	3	4	5	6
0	0	0	1/32	2/32	2/32	3/32
1	1/16	1/16	1/8	1/8	1/8	1/8
2	1/32	1/32	1/64	1/64	0	2/64

Find $P(X \leq 1)$

$P(Y \leq 3)$

$P(X \leq 1, Y \leq 3)$

$P(X \leq 1 | Y \leq 3)$

$P(Y \leq 3 | X \leq 1)$

$P(X + Y \leq 4)$

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$= \sum_{j=1}^{6} P(X = 0, Y = j) + \sum_{j=1}^{6} P(X = 1, Y = j) = \frac{1}{4} + \frac{5}{8} = \frac{7}{8}$$

$$P(Y \leq 3) = P(Y = 1) + P(Y = 2) + P(Y = 3)$$

$$= \sum_{i=0}^{2} P(X = i, Y = 1) + \sum_{i=0}^{2} P(X = i, Y = 2) + \sum_{i=0}^{2} P(X = i, Y = 3)$$

$$= \frac{3}{32} + \frac{3}{32} + \frac{11}{64} = \frac{23}{64}$$

$$P(X \leq 1, Y \leq 3) = \sum_{j=1}^{3} P(X = 0, Y = j) + \sum_{j=1}^{3} P(X = 1, Y = j)$$

$$= \frac{1}{32} + \frac{1}{4} = \frac{9}{32}$$

$$P(X \leq 1 | Y \leq 3) = \frac{P(X\leq 1, Y\leq 3)}{P(Y\leq 3)} = \frac{9/32}{23/64} = \frac{18}{23}$$

$$P(Y \leq 3 | X \leq 1) = \frac{P(X\leq 1, Y\leq 3)}{P(X\leq 1)} = \frac{9/32}{7/8} = \frac{9}{28}$$

$$P(X + Y \leq 4) = \sum_{j=1}^{4} P(X = 0, Y = j) + \sum_{j=1}^{3} P(X = 1, Y = j) + \sum_{j=1}^{2} P(X = 2, Y = j)$$

$$= \frac{3}{32} + \frac{1}{4} + \frac{1}{16} = \frac{13}{32}$$

Example 11:

Three balls are drawn at random without replacement from a box containing 2 white, 3 red and 4 black balls. If ***X*** denotes the number of white balls drawn and ***Y*** denotes the number of red balls drawn. Find the joint probability distribution of (X, Y).

Solution:

$$P(X=0, Y=0) = \text{P(drawing 3 balls none of which is white or red)}$$

$$= \text{P(all the 3 balls are black)}$$

$$= \frac{{}^4C_3}{{}^9C_3} = \frac{1}{21}$$

$$P(X=0, Y=1) = \text{P(drawing 1 red and 2 black balls)} = \frac{3}{14}$$

X \ Y	0	1	2	3
0	$\frac{1}{21}$	$\frac{3}{14}$	$\frac{1}{7}$	$\frac{1}{84}$
1	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{1}{14}$	0
2	$\frac{1}{21}$	$\frac{1}{28}$	0	0

Example 13

If X and Y are two random variables having joint density function:

$$f(x,y) = \begin{cases} \frac{1}{8}(6-x-y) & ;\ 0 \le x \le 2,\ 2 \le y \le 4 \\ 0 & \end{cases}$$

Find the (i) Marginal density functions of X and Y

(ii) $P(X < 1 \cap Y < 3)$

(iii) $P(X + Y < 3)$ (iv) $P(X<1 \mid Y<3)$

Example 13

The joint pdf of a two-dimensional RV (X,Y) is given by

$$f(x,y) = xy^2 + \frac{x^2}{8} \;\; ; \; 0 \le x \le 2, \; 0 \le y \le 1$$

Compute $P(X > 1),\ P(Y < \frac{1}{2}),\ P(X > 1 | Y < \frac{1}{2}),\ P(X < Y)$ and $P(X + Y \le 1)$

(i) $P(X > 1) = \underset{\substack{R_1 \\ (x>1)}}{\int\int} f(x, y)\,\mathrm{d}x\,\mathrm{d}y$

$$= \int_0^1 \int_1^2 \left(xy^2 + \frac{x^2}{8} \right) \mathrm{d}x\mathrm{d}y = \frac{19}{24}$$

(ii) $P(Y < 1/2) = \underset{\substack{R_2 \\ \left(y<\frac{1}{2}\right)}}{\int} \left(xy^2 + \frac{x^2}{8} \right) \mathrm{d}x\,\mathrm{d}y$

$$= \int_0^{1/2} \int_0^2 \left(xy^2 + \frac{x^2}{8} \right) \mathrm{d}x\,\mathrm{d}y$$

$$= \frac{1}{4}$$

(iii) $P(X > 1, Y < 1/2) = \displaystyle\int\limits_{\substack{R_3 \\ \left(x>1\ \&\ y<\frac{1}{2}\right)}} \left(xy^2 + \frac{x^2}{8}\right) \mathrm{d}x\,\mathrm{d}y$

$$= \int_0^{1/2}\int_1^2 \left(xy^2 + \frac{x^2}{8}\right) \mathrm{d}x\,\mathrm{d}y = \frac{5}{24}$$

(iv) $P(X > 1/Y < \frac{1}{2}) = \dfrac{P\left(X>1, Y<\frac{1}{2}\right)}{P\left(Y<\frac{1}{2}\right)} = \dfrac{5/24}{1/4} = \dfrac{5}{6}$

(v) $P(Y < \frac{1}{2}/X > 1) = \dfrac{P\left(X>1, Y<\frac{1}{2}\right)}{P\left(X>1\right)} = \dfrac{5/24}{19/24} = \dfrac{5}{19}$

(vi) $P(X < Y) = \underset{\substack{R_4 \\ (x<y)}}{\int\int}\left(xy^2 + \frac{x^2}{8}\right) dx\, dy$

$$= \int_0^1 \int_0^y \left(xy^2 + \frac{x^2}{8}\right) dx\, dy = \frac{53}{480}$$

(vii) $P(X + Y \le 1) = \underset{\substack{R_5 \\ (x+y\le 1)}}{\int\int}\left(xy^2 + \frac{x^2}{8}\right) dx\, dy$

$$= \int_0^1 \int_0^{1-y} \left(xy^2 + \frac{x^2}{8}\right) dx\, dy = \frac{13}{480}$$

Solution:

$$P(X > 1) = \frac{19}{24}$$

$$P\left(Y < \frac{1}{2}\right) = \frac{1}{4}$$

$$P(X > 1, Y < \frac{1}{2}) = \frac{5}{24}$$

$$P(X > 1 | Y < \frac{1}{2}) = \frac{5/24}{1/4} = \frac{5}{6}$$

$$P(X < Y) = \frac{53}{480}$$

$$P(X + Y \leq 1) = \frac{13}{480}$$

Example 3 *The joint probability mass function of (X, Y) is given by* $p(x, y) = k(2x + 3y)$, $x = 0, 1, 2$; $y = 1, 2, 3$. *Find all the marginal and conditional probability distributions. Also find the probability distribution of* $(X + Y)$.

Solution The joint probability distribution of (X, Y) is given below. The relevant probabilities have been computed by using the given law.

X	Y		
	1	2	3
0	$3k$	$6k$	$9k$
1	$5k$	$8k$	$11k$
2	$7k$	$10k$	$13k$

$$\sum_{j=1}^{3} \sum_{i=0}^{2} p(x_i, y_j) = 1$$

i.e., the sum of all the probabilities in the table is equal to 1.
i.e., $72k = 1$.

$$\therefore \qquad k = \frac{1}{72}$$

Marginal Probability Distribution of X: $\{i, p_{i*}\}$

$X = i$	$p_{i*} = \sum_{j=1}^{3} p_{ij}$
0	$p_{01} + p_{02} + p_{03} = \frac{18}{72}$
1	$p_{11} + p_{12} + p_{13} = \frac{24}{72}$
2	$p_{21} + p_{22} + p_{23} = \frac{30}{72}$
	Total $= 1$

Marginal Probability Distribution of Y: $\{j, p_{*j}\}$

$Y = j$	$p_{*j} = \sum_{i=1}^{2} p_{ij}$
1	15/72
2	24/72
3	33/72
	Total = 1

Conditional distribution of X, given $Y = 1$, is given by $\{i, P(X = i/Y = 1)\} =$
$\{i, P(X = i, Y = 1)/P(Y = 1)\}$
$= \{i, p_{i1}/p_{*1}\}, i = 0, 1, 2.$
the tabular representation is given below:

$X = i$	p_{i1}/p_{*1}
0	$3k/15k = \frac{1}{5}$
1	$5k/15k = \frac{1}{3}$
2	$7k/15k = \frac{7}{15}$
	Total = 1

The other conditional distributions are given below:

C.P.D. of X, given $Y = 2$	
$X = i$	p_{i2}/p_{*2}
0	$\frac{6k}{24k} = \frac{1}{4}$
1	$\frac{8k}{24k} = \frac{1}{3}$
2	$\frac{10k}{24k} = \frac{5}{12}$
	Total = 1

C.P.D. of X, given $Y = 3$	
$X = i$	p_{i3}/p_{*3}
0	$\frac{9k}{33k} = \frac{3}{11}$
1	$\frac{11k}{33k} = \frac{1}{3}$
2	$\frac{13k}{33k} = \frac{13}{33}$
	Total = 1

C.P.D. of Y, given $X = 0$	
$Y = j$	p_{oj}/p_{o*}
1	$\frac{3k}{18k} = \frac{1}{6}$
2	$\frac{6k}{18k} = \frac{1}{3}$
3	$\frac{9k}{18k} = \frac{1}{2}$
	Total = 1

C.P.D. of Y, given $X = 1$	
$Y = j$	p_{1j}/p_{1*}
1	$\frac{5k}{24k} = \frac{5}{24}$
2	$\frac{8k}{24k} = \frac{1}{3}$
3	$\frac{11k}{24k} = \frac{11}{24}$
	Total = 1

C.P.D. of Y, given $X = 2$	
$Y = j$	p_{2j}/p_{2*}
1	$\frac{7k}{30k} = \frac{7}{30}$
2	$\frac{10k}{30k} = \frac{1}{3}$
3	$\frac{13k}{30k} = \frac{13}{30}$
	Total = 1

Probability distribution of $(X + Y)$	
$(X + Y)$	P
1	$p_{01} = \frac{3}{72}$
2	$p_{02} + p_{11} = \frac{11}{72}$
3	$p_{03} + p_{12} + p_{21} = \frac{24}{72}$
4	$p_{13} + p_{22} = \frac{21}{72}$
5	$p_{23} = \frac{13}{72}$
	Total = 1

Example 14

Joint density function of X and Y is given by:

$$f(x, y) = 4xye^{-(x^2+y^2)} \quad ; \quad x \geq 0, \; y \geq 0$$

Test whether X and Y are independent.

Find the conditional density function of X given $Y = y$.

Solution Here the range space is the entire first quadrant of the xy-plane. By the property of the joint pdf

$$\iint\limits_{x>0,\, y>0} kxy\, e^{-(x^2+y^2)}\, dx\, dy = 1$$

i.e., $$k\int_0^\infty ye^{-y^2}\, dy \int_0^\infty xe^{-x^2}\, dx = 1$$

i.e., $$\frac{k}{4} = 1$$

$\therefore$ $$k = 4$$

Now $f_x(x) = \displaystyle\int_0^\infty 4x\, e^{-x^2} \times e^{-y^2}\, dy = 2x\, e^{-x^2},\ x > 0$

Similarly, $f_y(y) = 2ye^{-y^2},\ y > 0.$

Now $f_x(x) \times f_y(y) = 4\, x\, y\, e^{-(x^2+y^2)} = f(x, y)$

$\therefore$ The RVs x and y are independent.

Mathematical Expectation

The *average* value of a random phenomenon is also termed as its ***mathematical expectation*** or ***expected value***.

The expected value of a discrete r.v. is a weighted average of all possible values of the r.v., where the weights are the probabilities associated with the corresponding values.

For a discrete r.v. $\boldsymbol{X}$ with p.m.f. $\boldsymbol{p(x)}$,

$$E(X) = \sum_x x\, p(x)$$

For a continuous r.v. $\boldsymbol{X}$ with p.d.f. $\boldsymbol{f(x)}$,

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

Expected Value of a function of a random variable

Consider a r.v. X with *p.d.f. (p.m.f.)* $f(x)$ and distribution function $F(x)$.
If $g(\cdot)$ is a function such that g(X) is a r.v. and $E[g(X)]$ exists, then

$$E[g(X)] = \sum_{x} g(x)\, p(x) \qquad \text{(for discrete r.v.)}$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\; dx \quad \text{(for continuous r.v.)}$$

Expected Value of Two-dimensional RV

If (X, Y) is a two-dimensional RV with joint $p.d.f.\ f(x, y)$ or

joint $p.m.f.\ p_{ij}$, then

$$E[g(X,Y)] = \sum_i \sum_j g\left(x_i, y_j\right) p_{ij} \quad \text{(for discrete r.v.)}$$

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y)\ dxdy \quad \text{(for continuous r.v.)}$$

Particular cases:

1. If we take $g(X) = X^r$, r - being a positive integer, then

$$E[X^r] = \int_{-\infty}^{\infty} x^r f(x)\ dx$$

$$E[X^r] = \sum_x x^r p(x)$$

Which is defined as $\boldsymbol{\mu'_r}$, the r^{th} **moment (about origin)** of the probability distribution. Thus μ'_r (about origin) = $E(X^r)$

In particular, if $r = 1$ μ'_1 (about origin) = $E(X)$ = $\bar{\mathbf{X}}$ = **Mean**

$r = 2$ μ'_2 (about origin) = $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\ dx$

$$\mu_2 = \mu'_2 - {\mu'_1}^2 = E(X^2) - [E(X)]^2$$

2. If $g(X) = E[X - E(X)]^r = E[X - \bar{x}]^r$, then

$$E[X - E(X)]^r = \int_{-\infty}^{\infty} [x - E(X)]^r f(x)\, dx = \int_{-\infty}^{\infty} [x - \bar{x}]^r f(x)\, dx$$

$$E[X - E(X)]^r = \sum_x [x - E(X)]^r p(x) = \sum_x [x - \bar{x}]^r p(x)$$

which is $\boldsymbol{\mu_r}$, the ***r***

which is $\boldsymbol{\mu_r}$, the $\boldsymbol{r}^{\text{th}}$ **moment about mean.**

if $r = 2$, we get $\mu_2 = E[X - E(X)]^2 = \int_{-\infty}^{\infty} [x - \bar{x}]^2 f(x)\, dx$ = **Variance**

3. If $g(x) = \text{constant} = c$,

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

$$E(c) = \sum_x c\, p(x) = c \sum_x p(x) = c$$

13·7. THEOREMS ON EXPECTATION

Theorem 13·1. *$E(c) = c$, where c is a constant.* ...

Theorem 13·2. *$E(cX) = c\, E(X)$, where c is a constant.* ...

Theorem 13·3. *$E(aX + b) = a\, E(X) + b$, where a and b are constants.* ...

Theorem 13·4. (Addition Theorem of Expectation). *If X and Y are random variables then*

$$E(X + Y) = E(X) + E(Y), \quad \ldots$$

i.e., Expected value of the sum of two random variables is equal to the sum of their expected values

The result can be generalised to n variables. *If $X_1, X_2, \ldots, X_n$ are n random variables, then*

$$E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n) \quad \ldots$$

i.e.,

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) \quad \text{or} \quad \text{simply} \quad E(\sum X) = \sum E(X) \quad \ldots$$

Corollary. $E(aX + bY) = a\, E(X) + bE(Y)$, where a and b are constants. ..

However, $$E\left(\frac{1}{X}\right) \neq \frac{1}{E(X)}$$

$$E(X^{1/2}) \neq [E(X)]^{1/2}$$

$$E(\log X) \neq \log E(X)$$

$$E(X^2) \neq [E(X)]^2$$

- **Theorem. (Multiplication Theorem of Expectation).** *If X and Y are independent random variables, then* **$E(XY) = E(X) \cdot E(Y)$**

- In general, if $X1$, $X2$, …, Xn *are n independent random variables, then* **$E(X1\ X2\ X3 \ldots Xn) = E(X1) \cdot E(X2) \ldots E(Xn)$**

- **Remark. T***he multiplication theorem of expectation holds only for independent events while no such condition on the variables is required for the addition theorem of expectation.*

VARIANCE OF X IN TERMS OF EXPECTATION

$$\sigma_x^2 = E[X - E(X)]^2$$

Also, we have a simplified expression for σ_x^2 given by

$$\Rightarrow \qquad \sigma_x^2 = E(X^2) - [E(X)]^2$$

For the probability distribution $\{x, p(x)\}$, we have :

$$\text{Mean} = E(X) = \sum x \times p(x)$$

$$\sigma_x^2 = E(X^2) - [E(X)]^2 = \sum x^2\, p(x) - [\sum x\, p(x)]^2$$

Theorem 13·6. $\mathrm{Var}\,(X \pm c) = \mathrm{Var}\,X$, where c is a constant.

This theorem states *that variance is independent of change of origin*.

Theorem 13·7. $\mathrm{Var}\,(aX) = a^2 \,.\, \mathrm{Var}\,(X)$, where a is a constant.

This theorem states that *variance is not independent of change of scale*.

Corollary. Combining the results of the above two theorems, we get :

$$\mathrm{Var}\,(aX \pm b) = \mathrm{Var}\,(aX) = a^2\,\mathrm{Var}\,(X)$$

Theorem 13·8. $\mathrm{Var}\,(c) = 0$, where c is constant.

Co-Variance

Let (x_i, y_i), $i = 1, 2, \ldots, n$ be n paired observations on two variables X and Y. Then,

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \quad ; \quad E(Y) = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y} \quad ; \quad E(XY) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i$$

If X and Y are two random variables, then the covariance between them is defined as:

$$\begin{aligned} Cov(X,Y) &= E[\{X - E(X)\}\{Y - E(Y)\}] \\ &= E[XY - X\,E(Y) - Y\,E(X) + E(X)E(Y)] \\ &= \mathrm{E}(XY) - E(X)\,E(Y) - E(Y)\,E(X) + E(X)E(Y) \\ Cov(X,Y) &= \mathrm{E}(XY) - E(X)\,E(Y) \end{aligned}$$

If X and Y are independent random variables, $\mathrm{E}(XY) = E(X)\,E(Y)$

and hence, $Cov(X,Y) = E(X)\,E(Y) - E(X)\,E(Y) = 0$

Properties:

1. $$\begin{aligned} Cov(aX, bY) &= E[\{aX - E(aX)\}\{bY - E(bY)\}] \\ &= E[a\{X - E(X)\}\, b\{Y - E(Y)\}] \\ &= ab\, E[\{X - E(X)\}\, \{Y - E(Y)\}] \\ Cov(X, Y) &= ab\, Cov(X, Y) \end{aligned}$$

2. $Cov(X + a, Y + b) = Cov(X, Y)$

3. $Cov\left(\frac{X-\bar{X}}{\sigma_X}, \frac{Y-\bar{Y}}{\sigma_Y}\right) = \frac{1}{\sigma_X\, \sigma_Y} Cov(X, Y)$

4. If X and Y are independent, $Cov(X, Y) = 0$, but the converse is not true.

5. $V(X_1 \pm X_2) = V(X_1) + \mathrm{V}(X_2) \pm 2Cov(X_1, X_2)$

6. $V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2\, V(X_i) + 2\, \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j\, Cov(X_i, X_j)$

Correlation Coefficient:

If X and Y are two random variables, then the covariance between them is denoted as $Cov(X,Y)$,

$$Cov(X,Y) = E(XY) - E(X)\,E(Y)$$

and the Co-efficient of correlation $(r_{XY}\ or\ \rho_{XY})$ between X and Y is defined as ***a numerical measure of linear relationship between them***,

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X\,\sigma_Y} = \frac{Cov(X,Y)}{\sqrt{Var(x)}\,\sqrt{Var(y)}}$$

Properties:

1. If X and Y are independent, $Cov(X, Y) = 0$ and $\rho_{XY} = 0$ but the converse is not true.

   When $\rho_{XY} = 0$, we say that X and Y are uncorrelated.

2. $|\rho_{XY}| \leq 1$

   i.e., if $\rho_{XY} = +1 \Rightarrow$ Positive Correlation

   $\rho_{XY} = -1 \Rightarrow$ Negative Correlation

   $\rho_{XY} = 0 \Rightarrow$ UnCorrelated random variables.

Example:

Find the expectation of the number on a die when thrown.

Solution. Let X denote the number obtained on the die. Then X is a random variable which can take any one of the values $1, 2, 3, \ldots, 6$ each with equal probability $1/6$ as given in the adjoining Table.

x	1	2	3	4	5	6
p	1/6	1/6	1/6	1/6	1/6	1/6

$$E(X) = \sum x \,.\, p(x) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6}$$

$$= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = \frac{7}{2} = 3{\cdot}5.$$

Example: ***What is the expected number of heads appearing when a fair coin is tossed three times ?***

Solution:

x	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

$$\therefore \quad E(X) = \sum x\, p(x)$$

$$= \frac{1}{8}\left(0 + 1 \times 3 + 2 \times 3 + 3 \times 1\right) = \frac{12}{8} = 1{\cdot}5$$

Example: ***A random variable X is defined as the sum of faces when a pair of dice is thrown. Find the expected value of X*.**

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

$$
\begin{aligned}
E(X) &= \sum x \,.\, p(x) \\
&= 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + 5 \times \frac{4}{36} + 6 \times \frac{5}{36} + 7 \times \frac{6}{36} + 8 \times \frac{5}{36} + 9 \times \frac{4}{36} + 10 \times \frac{3}{36} + 11 \times \frac{2}{36} + 12 \times \frac{1}{36} \\
&= \frac{1}{36}\left[2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12\right] = \frac{252}{36} = 7
\end{aligned}
$$

Example 13·12. *From a bag containing 4 white and 6 red balls, three balls are drawn at random .*

(i) *Find the expected number of white balls down.*

(ii) *If each white ball drawn carries a reward of Rs. 4 and each red ball Rs. 6, find the expected reward of the draw.*

Solution. There are 4 white and 6 red balls *i.e.*, 10 balls in the bag.

Three balls can be drawn out of 10 balls in ${}^{10}C_3$ ways, which gives the exhaustive number of cases.

$$p(x) = \text{Probability of drawing } x \text{ white balls} = \frac{{}^4C_x \times {}^6C_{3-x}}{{}^{10}C_3} \quad ; \quad x = 0, 1, 2, 3$$

$$p(0) = \frac{{}^4C_0 \times {}^6C_3}{{}^{10}C_3} = \frac{1 \times 6 \times 5 \times 4 \times 3\,!}{3\,! \times 10 \times 9 \times 8} = \frac{1}{6} = \frac{5}{30} \quad ; \quad p(1) = \frac{{}^4C_1 \times {}^6C_2}{{}^{10}C_3} = \frac{4 \times 15}{120} = \frac{15}{30} \; ;$$

$$p(2) = \frac{{}^4C_2 \times {}^6C_1}{{}^{10}C_3} = \frac{6 \times 6}{120} = \frac{9}{30} \quad ; \quad p(3) = \frac{{}^4C_3 \times {}^6C_0}{{}^{10}C_3} = \frac{4 \times 1}{120} = \frac{1}{30}$$

x	$p(x)$
0	5/30
1	15/30
2	9/30
3	1/30

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

$$= 0 \times \frac{5}{30} + 1 \times \frac{15}{30} + 2 \times \frac{9}{30} + 3 \times \frac{1}{30}$$

$$= \frac{1}{30}\left(15 + 18 + 3\right) = \frac{36}{30} = \frac{6}{5} = 1{\cdot}2$$

(ii)

x	$p(x)$	*Reward $R(x)$ in Rs.*
0	5/30	$4 \times 0 + 6 \times 3 = 18$
1	15/30	$4 \times 1 + 6 \times 2 = 16$
2	9/30	$4 \times 2 + 6 \times 1 = 14$
3	1/30	$4 \times 3 + 6 \times 0 = 12$

$\therefore$ Expected reward of the draw is :

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

$$= \text{Rs.} \left[\frac{5}{30} \times 18 + \frac{15}{30} \times 16 + \frac{9}{30} \times 14 + \frac{1}{30} \times 12\right]$$

$$= \text{Rs.}\ (3 + 8 + 4{\cdot}20 + 0{\cdot}40)$$

$$= \text{Rs.}\ 15{\cdot}60$$

Example 13·15. *A random variable X has the following probability distribution :*

x	:	4	5	6	8
p	:	0·1	0·3	0·4	0·2

Find $E[X - E(X)]^2$.

x	p	px	px^2
4	0·1	0·4	1·6
5	0·3	1·5	7·5
6	0·4	2·4	14·4
8	0·2	1·6	12·8
Total	1	$\sum px = 5{\cdot}9$	$\sum px^2 = 36{\cdot}3$

$$E[X - E(X)]^2 = \text{Var}(X) = E(X^2) - [E(X)]^2$$
$$= \sum px^2 - (\sum px)^2$$

$$\begin{aligned} E[X - E(X)]^2 &= \sum px^2 - (\sum px)^2 \\ &= 36{\cdot}3 - (5{\cdot}9)^2 \\ &= 36{\cdot}30 - 34{\cdot}81 \\ &= 1{\cdot}49 \end{aligned}$$

Example 13·18. *If* $V(X) = V(Y) = \frac{1}{4}$ *and* $V(X - Y) = \frac{1}{3}$, *the correlation coefficient between random variables X and Y is*

(*i*) $\frac{1}{3}$; (*ii*) $-\frac{1}{3}$; (*iii*) $\frac{2}{3}$; (*iv*) *none of these.*

Solution. We are given : $\sigma_x^2 = \sigma_y^2 = \frac{1}{4} \quad \Rightarrow \quad \sigma_x = \sigma_y = \frac{1}{2}$. ...(*)

and $\text{Var}(X - Y) = \frac{1}{3}$

$\Rightarrow$ $\text{Var}(X) + \text{Var}(Y) - 2\,\text{Cov}(X, Y) = \frac{1}{3}$ [From (13·32*a*)]

$\Rightarrow$ $\sigma_x^2 + \sigma_y^2 - 2r\,\sigma_x\,\sigma_y = \frac{1}{3}$ $\left(\because \frac{\text{Cov}(x, y)}{\sigma_x\,\sigma_y} = r\right)$

$\Rightarrow$ $\frac{1}{4} + \frac{1}{4} - 2r\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{3} \quad \Rightarrow \quad \frac{r}{2} = \frac{1}{2} - \frac{1}{3} = \frac{3-2}{6} = \frac{1}{6} \quad \Rightarrow \quad r = \frac{1}{3}$

$\therefore$ (*i*) is the correct answer.

Example 13·19. *Let X and Y be two random variables each taking three values –1, 0 and 1, and having the joint probability distribution as given in Table 13·5.*

Obtain the marginal probability distributions of X and Y and hence their expected values.

TABLE 13·5

Y \ X	–1	0	1
–1	0	·1	·1
0	·2	·2	·2
1	0	·1	·1

TABLE 13·6. **COMPUTATION OF MARGINAL PROBABILITY DISTRIBUTIONS**

$Y \downarrow$ \ $X \rightarrow$	-1	0	1	*Total* $g(y)$
-1	0	$\cdot 1$	$\cdot 1$	$\cdot 1 + \cdot 1 = 0\cdot 2$
0	$\cdot 2$	$\cdot 2$	$\cdot 2$	$\cdot 2 + \cdot 2 + \cdot 2 = 0\cdot 6$
1	0	$\cdot 1$	$\cdot 1$	$0 + \cdot 1 + \cdot 1 = 0\cdot 2$
Total $p(x)$	$0 + \cdot 2 + 0 = 0\cdot 2$	$\cdot 1 + \cdot 2 + \cdot 1 = 0\cdot 4$	$\cdot 1 + \cdot 2 + \cdot 1 = 0\cdot 4$	$\sum p(x) = \sum g(y) = 1$

TABLE 13·7. **MARGINAL PROBABILITY DISTRIBUTION OF *X***

x	-1	0	1	Total
$p(x)$	$0\cdot 2$	$0\cdot 4$	$0\cdot 4$	$1\cdot 0$

$$\begin{aligned} E(X) &= \sum x\, p(x) \\ &= -1 \times 0\cdot 2 + 0 \times 0\cdot 4 + 1 \times 0\cdot 4 \\ &= -0\cdot 2 + 0\cdot 4 = 0\cdot 2 \end{aligned}$$

TABLE 13·8. **MARGINAL PROBABILITY DISTRIBUTION OF Y**

y	–1	0	1	Total
$g(y)$	0·2	0·6	0·2	1·0

$$
\begin{aligned}
E(Y) &= \sum y g(y) \\
&= -1 \times 0{\cdot}2 + 0 \times 0{\cdot}6 + 1 \times 0{\cdot}2 \\
&= -0{\cdot}2 + 0{\cdot}2 = 0
\end{aligned}
$$

Example 13·20. If $p(x) = \begin{cases} 0{\cdot}1x, & x = 1,2,3,4 \\ 0\ , & \text{otherwise,} \end{cases}$

Find : (*i*) $P\{x = 1 \text{ or } 2\}$ (*ii*) $P\left\{\frac{1}{2} < x < \frac{5}{2} \;\middle|\; x > 1\right\}$

Solution (*i*) $P(X = 1 \text{ or } 2) = P(X = 1) + P(X = 2) = p(1) + p(2)$

$$= 0{\cdot}1 \times 1 + 0{\cdot}1 \times 2 = 0{\cdot}1 + 0{\cdot}2 = 0{\cdot}3$$

$$(ii)\ P\left\{\frac{1}{2} < X < \frac{5}{2} \;\middle|\; X > 1\right\} = \frac{P\left[\left(\frac{1}{2} < X < \frac{5}{2}\right) \cap X > 1\right]}{P(X > 1)} = \frac{P(X = 2)}{1 - P(X \le 1)}$$

$$= \frac{0{\cdot}1 \times 2}{1 - [p(0) + p(1)]} = \frac{0{\cdot}2}{1 - [0{\cdot}1 \times 0 + 0{\cdot}1 \times 1]} = \frac{0{\cdot}2}{0{\cdot}9} = \frac{2}{9}.$$

17. Let the random variable X assume the values x_1, x_2 and x_3. Then the function f denoted by $f(x_i) = P(X = x_i)$ is given by :

Values of X, (x)	:	–3	6	9
$P(X = x)$	:	1/6	1/2	1/3

Find $E(X)$, $E(X^2)$ and $E(2X + 1)^2$.

Ans. $E(X) = 5{\cdot}5$; $E(X^2) = 46{\cdot}5$; $E(2X + 1)^2 = 4E(X^2) + 4E(X) + 1 = 209$

19. A random variable X has the following probability distribution :

Value of X, (x)	:	0	1	2	3
$P(X = x)$	:	1/3	1/2	0	1/6

Find : (*i*) Var (X), (*ii*) Var(Y) where $Y = 2X - 1$.

Ans. Var$(X) = 1$; Var(Y) = Var$(2X - 1)$ = Var$(2X) = 2^2$ Var$(X) = 4 \times 1 = 4$.

Moment Generating Function (MGF)

If X is discrete, then $M_X(t) = E(e^{tx}) = \sum_{i=1}^{n} e^{tx_i} p(x_i)$,

where X takes the values $x_1, x_2, x_3, \ldots, x_n$ with probabilities

$p(x_1), p(x_2), p(x_3), \ldots p(x_n)$

If X is continuous, then $M_X(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$.

Properties of MGF:

- The r^{th} moment $\mu_r^{'}$ is the coefficient of $\frac{t^r}{r!}$ in the expansion of $M_X(t)$ in series of powers of t.

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

$$= 1 + tE(x) + \frac{t^2}{2!}E(x^2) + \cdots + \frac{t^r}{r!}E(x^r) + \dots$$

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

Therefore, $\mu_r^{'}$ is the coefficient of $\frac{t^r}{r!}$ in the expansion of $M_X(t)$ in series of powers of t.

- $$\mu_r' = E(X^r) = \left[\frac{d^r}{dt^r}(M_X(t))\right]_{t=0}.$$

  In particular, if $r = 1$ then $\mu_1' = \left[\frac{d(M_X(t))}{dt}\right]_{t=0}$ is the mean of X.

- $M_{cX}(t) = M_X(ct)$
- $M_{aX+b}(t) = e^{bt} M_X(at)$
- $M_{X+Y}(t) = M_X(t) M_Y(t)$

PROBLEMS:

Find the moment generating function of the random variable X whose probability mass function $P(X = x) = \frac{1}{2^x}, x = 1, 2, 3, \ldots,$ Deduce the mean and variance from moment generating function.

Answer: $M_X(t) = \frac{e^t}{2-e^t}$ and mean is 2.