

7. A company has two plants to manufacture scooters. Plant *I* manufactures 70% of the scooters and Plant *II* manufactures 30%. At Plant *I*, 80% of the scooters produced are of standard quality and at Plant *II*, 90% of the scooters produced are of standard quality. A scooter is picked at random and found to be of standard quality. What is the chance that it has come from Plant *II*?

8. Suppose there is a chance for a newly constructed building to collapse, whether the design is faulty or not. The chance that the design is faulty is 10%. The chance that the building collapses is 95%. If the design is faulty and otherwise it is 45%. It is seen that the building collapsed. What is the probability that it is due to faulty design.

Answers:

1. $\frac{23}{45}$
2. a. 0.52 b. 0.035
3. (i) 0.682 , (ii) 0.318.
4. $\left(\frac{3}{8}\right) = 0.375$
5. a. $\frac{1}{5}$ b. $\frac{1}{5}$ c. $\frac{3}{5}$
6. 0.07
7. 0.03253
8. 0.19

Unit – 2

Probability distributions

2.1

Random Variable

While performing a random experiment we are mainly concerned with the assignment and computation of probabilities of events. In many experiments we are interested in some function of the outcomes of the experiment as opposed to the outcome itself. For instance, in tossing two dice we are interested in the sum of faces of the dice and are not really concerned about the actual outcome. That is, we may be interested in knowing that the sum is seven and not be concerned over whether actual outcome was (1, 6) or (2, 5) or (3, 4) or (4, 3) or (5, 2) or (6, 1). These quantities of interest or more formally these real valued function defined on the sample space are known as **random variables**.

Random variable (r. v): Let S be the sample space associated with a random experiment. Let $\mathbf{R}$ be the set of real numbers. If $X: S \rightarrow \mathbf{R}$, i.e., X is a real valued function defined on the sample space, then X is known as a **random variable**. In other words, random variable is a function which takes real values which are determined by the outcomes in the sample space.

The random variables are denoted by capital letters X, Y, Z ... etc.

Notation: Let $a, b \in \mathbf{R}$. The set of all ω in S such that $X(\omega) = a$ is denoted by $X = a$. That is, $X = a$ denotes the event $\{\omega \in S | X(\omega) = a\}$. Similarly $X \leq a$ denotes the event $\{\omega \in S | X(\omega) \leq a\}$ and $a < X \leq b$ denotes the event $\{\omega \in S | X(\omega) \in (a, b]\}$.

Let us consider a random experiment of three tosses of a coin. Then the sample space S consists of $2^3 = 8$ points as given below.

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

For each outcome ω in S define $X(\omega)$ as the number of heads in the outcome ω . Then X may take any one of the values 0, 1, 2 or 3. For each outcome in S , we have one value of X . Thus,

$$\begin{aligned} X(HHH) &= 3, \\ X(HTH) &= X(THH) = X(HHT) = 2 \\ X(TTH) &= X(HTT) = X(THT) = 1 \text{ and} \\ X(TTT) &= 0 \end{aligned}$$

This shows that X is a random variable.

Note that $X = 0, X = 1, X = 2$ and $X = 3$ respectively denote the events

$$\{TTT\}, \{TTH, HTT, THT\}, \{HTH, THH, HHT\} \text{ and } \{HHH\}$$

Discrete Random Variable (d. r. v): If the random variable assumes only a finite or countably infinite set of values, it is known as **discrete random variable**.

For example, the number of students attending the class, the number of defectives in a lot consisting of manufactured items and the number of accidents taking place on a busy road, etc., are all discrete random variables. In the above example X is a d.r.v.

Continuous Random Variable (c. r. v): If a random variable can assume uncountable set of values, it is said to a **continuous random variable**.

For example, the age, height or weight of the students in a class is all continuous random variables. In case of continuous random variable, we usually talk of the value in a particular interval and not at a point. Generally, discrete random

variable represents *count data* while continuous random variable represent *measured data*.

The probabilistic behavior of a d.r.v.X at each real point is described by a function called **probability mass function** and it is defined below:

Probability Mass Function (p.m.f): Let X be a discrete random variable with distinct values $x_1, x_2, \dots, x_n, \dots$. The function $p : \mathbf{R} \rightarrow \mathbf{R}$ defined as

$$p(x) = \begin{cases} P(X = x_i) = p_i & \text{if } x = x_i \\ 0 & \text{if } x \neq x_i, i = 1, 2, \dots, n, \dots \end{cases}$$

is called the **probability mass function** of r.v.X, if (i) $p(x) \geq 0 \forall x \in \mathbf{R}$ and (ii) $\sum_{x \in \mathbf{R}} p(x) = 1$

Probability Distribution: The set of all possible ordered pairs $(x_i, p(x_i)), i = 1, 2, \dots, n, \dots$ is called the **probability distribution** of the r.v.X.

In particular, if X takes the values $x_1, x_2, x_3, \dots, x_n$ then the probability of X is usually represented in a tabular form as given below:

Probability Distribution of r. v. x

x	x_1	x_2	x_3	$\dots$	x_n
$p(x)$	p_1	p_2	p_3	$\dots$	p_n

Note: The concept of probability distribution is analogous to that of frequency distribution. Just as frequency distribution tells us how the total frequency is distributed among different values (or classes) of the variable, similarly a probability distribution tells us how total probability 1 is distributed among the various values which the r. v. can take.

Example 1: Obtain the probability distribution of X , the number of heads in three tosses of a coin (or a simultaneous toss of three coins).

Solution:

The sample space S consists of $2^3 = 8$ sample points, as give below:

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

Obviously, X is a random variable which can take the values 0, 1, 2 or 3.

The probability distribution of X is computed as given below.

No. of heads a	$X = a$ $\{\omega \in S X(\omega) = a\}$	No. of favourable cases	$p(x) = P(X = a)$
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	$\frac{1}{8}$

Hence, the probability distribution of X is given by:

x	:	0	1	2	3
$p(x)$	:	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Probability Density Function (p. d. f): Let X be a continuous random variable defined on the sample space S . Let $f(x)$ be a real valued function defined on $\mathbf{R}$

such that, for any real numbers a and $b(a < b)$, $P(a \leq X \leq b) = \int_a^b f(x)dx$.

If the function $f(x)$ satisfies (i) $f(x) \geq 0 \quad \forall x \in \mathbf{R}$ and (ii) $\int_{-\infty}^{\infty} f(x)dx = 1$ then

$f(x)$ is known as probability density function (p.d.f) of X

Note:

1. If X is a c. r. v., then $P(X = a) = 0$ where a is some real number.
2. Unlike discrete probability distribution, a continuous probability distribution can't be presented in a tabular form.

Cumulative Distribution Function (c. d. f): The cumulative distribution function of a r. v. X is defined by

$$F(x) = P(X \leq x) = \begin{cases} \sum_{t \leq x} p(t) & \text{if } X \text{ is a d.r.v. with p.m.f } p(x) \\ \int_{-\infty}^x f(t) dt & \text{if } X \text{ is a c.r.v. with p.d.f } f(x) \end{cases}$$

Note: If X is a continuous random variable, then $\frac{d}{dx}F(x) = f(x)$

Properties of c.d.f.

1. If $a < b$, $P(a < X \leq b) = F(b) - F(a)$
2. $0 \leq F(x) \leq 1$ and $F(x) \leq F(y)$ if $x < y$
3. $F(-\infty) = 0$ and $F(\infty) = 1$
4. Discontinuities of $F(x)$ are at most countable.

Note: The c.d.f. is used to find the cumulative probabilities in a probability distribution.

Example 2:

- (i) Find the constant k such that

$$f(x) = \begin{cases} kx^2 & , \quad 0 < x < 3 \\ 0 & , \quad \text{otherwise} \end{cases}$$

is a p.d.f.

- (ii) Compute $P(1 < x < 2)$
- (iii) Find the c.d.f and use it to compute $P(1 < x \leq 2)$

Solution:

(i) $f(x)$ is a p.d.f if

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

$$\Rightarrow k \int_0^3 x^2 dx \Rightarrow k \left[\frac{x^3}{3} \right]_0^3 = 1 \Rightarrow k = \frac{1}{9}$$

$$f(x) = \begin{cases} \frac{1}{9}x^2 & , \quad 0 < x < 3 \\ 0 & , \quad \text{otherwise} \end{cases}$$

(ii)

$$\begin{aligned} P(1 < x < 2) &= \int_1^2 f(x) dx \\ &= \int_1^2 \frac{1}{9}x^2 dx = \frac{1}{9} \left[\frac{x^3}{3} \right]_1^2 = \frac{7}{27} \end{aligned}$$

(iii) We have,

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) du$$

If $x < 0$, then $F(x) = 0$. If $0 \leq x < 3$, then

$$F(x) = \int_{-\infty}^x f(u) du = \frac{1}{9} \int_0^x u^2 du = \frac{x^3}{27}$$

If $x \geq 3$, then

$$F(x) = \int_0^3 f(u) du + \int_3^x f(u) du = \frac{1}{9} \int_0^3 u^2 du + \int_3^x du = \frac{1}{9} \times 9 + 0 = 1$$

Thus, required c.d.f is

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x^3}{27} & , \quad 0 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

Hence $P(1 < x \leq 2) = P(x \leq 2) - P(x \leq 1)$

$$\begin{aligned} &= F(2) - F(1) \\ &= \frac{2^3}{27} - \frac{1^3}{27} = \frac{8}{27} - \frac{1}{27} = \frac{7}{27} \end{aligned}$$

Example 3: A die is tossed twice. Getting an *odd number* is termed as a success. Find the probability distribution and c.d.f of the number of successes.

Solution: Since the cases favorable to getting an odd number in a throw of a die are 1,3,5, i.e., 3 in all.

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

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.

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

$$= P(FF) = P(F) \times P(F) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$$

$$P(X = 1) = P(S \text{ and } F) + P(F \text{ and } S)$$

$$= P(S)P(F) + P(F)P(S)$$

$$= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}.$$

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

Hence the probability distribution of X is given by :

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

The c.d.f is given by

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{4} & \text{if } 0 \leq x < 1 \\ \frac{3}{4} & \text{if } 1 \leq x < 2 \\ 1 & \text{if } x \geq 2 \end{cases}$$

Example 4: 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 is $\frac{4}{52} = \frac{1}{13}$

$\Rightarrow$ Probability of drawing a non-ace is $1 - \frac{1}{13} = \frac{12}{13}$.

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

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

$$P(X = 1) = P(\text{Ace and Non-ace}) + P(\text{Non-ace and Ace})$$

$$= P(\text{Ace}) \times P(\text{Non-ace}) + P(\text{Non-ace}) \times P(\text{Ace})$$

$$= \frac{1}{13} \times \frac{12}{13} + \frac{12}{13} \times \frac{1}{13} = \frac{24}{169}.$$

$$P(X = 0) = P(\text{Non-ace and Non-ace})$$

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

Hence, the probability distribution of X is given by:

$x :$	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 ${}^{52}C_2$.

$$\therefore P(X = 0) = P(\text{No ace}) = P(\text{Both cards are non-aces})$$

$$= \frac{{}^{48}C_2}{{}^{52}C_2} = \frac{48 \times 47}{52 \times 51} = \frac{188}{221}$$

$$P(X = 1) = P(\text{one ace}) = P(\text{one ace and one non-ace})$$

$$= \frac{{}^4C_1 \times {}^{48}C_1}{{}^{52}C_2} = \frac{4 \times 48 \times 2}{52 \times 51} = \frac{32}{221}$$

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

Hence, the probability distribution of X is given by :

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

Example 5: If X is a continuous random variable with p.d.f

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

- (i) **Determine k .**
- (ii) **Compute $P(x \leq 1.5)$**

Solution:

- (i) Since $f(x)$ is the p.d.f, so we have

$$\begin{aligned}
& \int_{-\infty}^{\infty} f(x) dx = 1 \\
& \Rightarrow \int_0^1 f(x) dx + \int_1^2 f(x) dx + \int_2^3 f(x) dx = 1 \\
& \Rightarrow \int_0^1 kx dx + \int_1^2 k dx + \int_2^3 -k(x-3) dx = 1 \\
& \Rightarrow k \left[\frac{x^2}{2} \right]_0^1 + k[x]_1^2 - k \left[\frac{x^2}{2} - 3x \right]_2^3 = 1 \\
& \Rightarrow \frac{k}{2} + 2k - k - k \left[\left(\frac{9}{2} - 9 \right) - (2 - 6) \right] = 1 \\
& \Rightarrow k \left[\frac{1}{2} + 2 - 1 - \frac{9}{2} + 9 + 2 - 6 \right] = 1 \\
& \Rightarrow 2k = 1 \Rightarrow k = \frac{1}{2}
\end{aligned}$$

(ii)

$$\begin{aligned}
P(x \leq 1.5) &= \int_{-\infty}^{1.5} f(x) dx = \int_0^1 f(x) dx + \int_1^{1.5} f(x) dx \\
&= k \int_0^1 x dx + \int_1^{1.5} k dx = k \left[\frac{x^2}{2} \right]_0^1 + k[x]_1^{1.5} = k \left[\frac{1}{2} + \frac{1}{2} \right] = k = \frac{1}{2}
\end{aligned}$$

P1:

Two dice are rolled at random. Obtain the probability distribution of the sum of the numbers on them.

Solution:

When two dice are rolled, the sample space S consists of $6^2 = 36$, sample points as shown.

$$S = \{(1,1), (1,2), \dots, (1,6), (2,1), (2,2), \dots, (2,6), (3,1), (3,2), \dots, (3,6) \\ (4,1), (4,2), \dots, (4,6), (5,1), (5,2), \dots, (5,6), (6,1), (6,2), \dots, (6,6)\}$$

Let X denote the sum of the numbers on the two dice. Then X is a random variable which can take the values 2, 3, 4, ..., 12 with the probability distribution given by:

Sum of numbers (x)	Favourable sample points	No. of favourable cases	$p(x)$
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	$\frac{1}{36}$

Hence, the probability distribution of X is given by:

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

P2:

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 2 apples is ${}^{24}C_2$.

$$\therefore P(X = 0) = \frac{{}^{20}C_2}{{}^{24}C_2} = \frac{20 \times 19}{24 \times 23} = \frac{95}{138}$$

$$P(X = 1) = \frac{{}^4C_1 \times {}^{20}C_1}{{}^{24}C_2} = \frac{2 \times 4 \times 20}{24 \times 23} = \frac{40}{138}$$

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

Hence, the probability distribution of X is given by :

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

P3:

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

- i. Find k for which the above is a p.d.f.
- ii. Determine b such that $P(x < b) = P(x > b)$.

Solution:

- i. $f(x) = kx(1 - x)$, $0 \leq x \leq 1$ is the p.d.f. of a continuous random variable x if $\int_0^1 f(x) dx = 1$. That is $k \int_0^1 x(1 - x) dx = 1$

$$\Rightarrow k \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_0^1 = 1 \Rightarrow k = 6$$

- (i) Given $P(x < b) = P(x > b)$. That is,

$$\begin{aligned} \int_0^b f(x) dx &= \int_b^1 f(x) dx \\ \Rightarrow 6 \int_0^b x(1 - x) dx &= 6 \int_b^1 x(1 - x) dx \\ \Rightarrow \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_0^b &= \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_b^1 \\ \Rightarrow [3b^2 - 2b^3] &= [1 - 3b^2 + 2b^3] \\ \Rightarrow 4b^3 - 6b^2 + 1 &= 0 \\ \Rightarrow (2b - 1)(2b^2 - 2b - 1) &= 0 \\ \Rightarrow b = \frac{1}{2}, b = \frac{1 \pm \sqrt{3}}{2} \end{aligned}$$

Hence, $b = \frac{1}{2}$ is the only value lying in $[0, 1]$ and satisfying $P(x < b) = P(x > b)$

P4:

The time one has to spend in point of a bank counter is observed to be a random variable x with the p.d.f.

$$f(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{1}{9}(x+1) & , \quad 0 \leq x < 1 \\ \frac{4}{9}\left(x - \frac{1}{2}\right) & , \quad 1 \leq x < \frac{3}{2} \\ \frac{4}{9}\left(\frac{5}{2} - x\right) & , \quad \frac{3}{2} \leq x < 2 \\ \frac{1}{9}(4-x) & , \quad 2 \leq x < 3 \\ \frac{1}{9} & , \quad 3 \leq x < 6 \\ 0 & , \quad x \leq 6 \end{cases}$$

Let A, B be the events defined as

A : One waits between 0 and 2 minutes inclusive

B : One waits between 1 and 3 minutes inclusive

- i. Show that $P(B|A) = \frac{2}{3}$
- ii. Show that $P(\bar{A} \cap \bar{B}) = \frac{1}{3}$

Solution:

i.

$$\begin{aligned} P(A) &= \int_0^2 f(x) dx = \int_0^1 \frac{1}{9}(x+1) dx + \int_1^{\frac{3}{2}} \frac{4}{9}\left(x - \frac{1}{2}\right) dx + \int_{\frac{3}{2}}^2 \frac{4}{9}\left(\frac{5}{2} - x\right) dx \\ &= \frac{1}{9} \left[\frac{x^2}{2} + x \right]_0^1 + \frac{4}{9} \left[\frac{x^2}{2} - \frac{x}{2} \right]_1^{\frac{3}{2}} + \frac{4}{9} \left[\frac{5x}{2} - \frac{x^2}{2} \right]_{\frac{3}{2}}^2 \\ &= \frac{1}{9} \left[\left(\frac{1}{2} + 1 \right) \right] + \frac{4}{9} \left[\left(\frac{9}{8} - \frac{3}{4} \right) \right] + \frac{4}{9} \left[(5-2) - \left(\frac{15}{4} - \frac{9}{8} \right) \right] = \frac{1}{2} \end{aligned}$$

$$\begin{aligned}
 P(A \cap B) &= P(1 \leq x \leq 2) = \int_1^2 f(x) dx \\
 &= \int_1^{\frac{3}{2}} \frac{4}{9} \left(x - \frac{1}{2} \right) dx + \int_{\frac{3}{2}}^2 \frac{4}{9} \left(\frac{5}{2} - x \right) dx \\
 &= \frac{4}{9} \left[\frac{x^2}{2} - \frac{x}{2} \right]_1^{\frac{3}{2}} + \frac{4}{9} \left[\frac{5x}{2} - \frac{x^2}{2} \right]_{\frac{3}{2}}^2 \\
 &= \frac{4}{9} \left[\left(\frac{9}{8} - \frac{3}{4} \right) - \left(\frac{1}{2} - \frac{1}{2} \right) \right] + \frac{4}{9} \left[(5 - 2) - \left(\frac{15}{4} - \frac{9}{8} \right) \right] = \frac{1}{3}
 \end{aligned}$$

$$\begin{aligned}
 P(B) &= \int_1^3 f(x) dx \\
 &= \int_1^{\frac{3}{2}} \frac{4}{9} \left(x - \frac{1}{2} \right) dx + \int_{\frac{3}{2}}^2 \frac{4}{9} \left(\frac{5}{2} - x \right) dx + \int_2^3 \frac{1}{9} (4 - x) dx \\
 &= \frac{4}{9} \left[\frac{x^2}{2} - \frac{x}{2} \right]_1^{\frac{3}{2}} + \frac{4}{9} \left[\frac{5x}{2} - \frac{x^2}{2} \right]_{\frac{3}{2}}^2 + \frac{1}{9} \left[4x - \frac{x^2}{2} \right]_2^3 \\
 &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}
 \end{aligned}$$

$$\text{Thus } P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

$$\begin{aligned}
 \text{ii. } P(\overline{A} \cap \overline{B}) &= P(\overline{A \cup B}) = 1 - P(A \cup B) \\
 &= 1 - [P(A) + P(B) - P(A \cap B)] = 1 - \frac{2}{3} = \frac{1}{3}
 \end{aligned}$$

Alternatively, $\overline{A} \cap \overline{B}$ means the waiting time more than 3 minutes.

So,

$$\begin{aligned}
 P(\overline{A} \cap \overline{B}) &= P(x > 3) = \int_3^{\infty} f(x) dx = \int_3^6 f(x) dx + \int_6^{\infty} f(x) dx \\
 &= \frac{1}{9} \int_3^6 dx + 0 = \frac{1}{9} [6 - 3] = \frac{3}{9} = \frac{1}{3}
 \end{aligned}$$

2.1. Random Variables

Exercise

1. 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	2
$p(x):$	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

2. Two dice are thrown simultaneously and *getting a number less than 3* on a die is termed as a success. Obtain the probability distribution of the number of success.
3. Obtain the probability distribution of the number of sixes in two tosses of a die.
4. Obtain the probability distribution of number of heads of two tosses of a coin.
5. Three cards are drawn at random successively, with replacement, from a well shuffled pack of cards, getting 'a card of diamonds' is termed as a success. Obtain probability distribution of the number of successes.
6. Two cards are drawn without replacement, from a well shuffled pack of cards. Obtain the probability distribution of the number of face cards (Jack, Queen, King and Ace).

7. Five defective mangoes are accidentally mixed with twenty good ones and by looking at them it is not possible to difference between them. Four mangoes are drawn at random from the lot. Find the probability distribution of x , the number of defective mangoes.

8. Two bad eggs are mixed accidentally with 10 good ones and three are drawn at random from the lot. Obtain the probability distribution of the number of bad eggs drawn.

9. An urn contains 6 red and 4 white balls. Three balls are known at random. Obtain the probability distribution of the number of white balls drawn.

10. Suppose that the life in hours of a certain part of radio tube is a continuous random variable x with p.d. f given by

$$f(x) = \begin{cases} \frac{100}{x^2}, & \text{when } x \geq 100 \\ 0, & \text{otherwise} \end{cases}$$

- (i) What is the probability that all of three such tubes in a given radio set will have to be replaced during the first 150 hours of operation?
- (ii) What is the probability that none of three of the original tubes will have to be replaced during that first 150 hours of operation?
- (iii) What is the probability that a tube will last less than 200 hours if it is known that the tube still functioning after 150 hours of service.
- (iv) What is the maximum number of tubes that many be inserted into a set so that there is a probability of 0.5 that after 150 hours of services all of them are still functioning?

Answers:

1.

- (i) Yes
- (ii) No, since $\sum p(x) > 1$
- (iii) No, since $\sum p(x) < 1$
- (iv) No, since $p(0) = -0.2$ which is not possible.

2.

$x :$	0	1	2
$p(x) :$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{4}{9}$

3.

$x :$	0	1	2
$p(x) :$	$\frac{25}{36}$	$\frac{10}{36}$	$\frac{1}{36}$

4.

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

5.

$x:$	0	1	2	3
$p(x):$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

6.

$x:$	0	1	2
$p(x):$	$\frac{36C_2}{52C_2} = \frac{105}{221}$	$\frac{36C_1 \times 16C_1}{52C_2} = \frac{96}{221}$	$\frac{16C_2}{52C_2} = \frac{20}{221}$

7.

$x:$	0	1	2	3	4
$p(x):$	$\frac{20_{C_4}}{25_{C_4}} = \frac{969}{2530}$	$\frac{5_{C_1} \times 20_{C_3}}{25_{C_4}}$ $\frac{1140}{2530}$	$\frac{5_{C_2} \times 20_{C_2}}{25_{C_4}}$ $\frac{380}{2530}$	$\frac{5_{C_3} \times 20_{C_1}}{25_{C_4}}$ $\frac{40}{2530}$	$\frac{5_{C_1}}{25_{C_4}} = \frac{1}{2530}$

8.

$x:$	0	1	2	3
$p(x):$	$\frac{10_{C_3}}{12_{C_3}} = \frac{12}{22}$	$\frac{2_{C_1} \times 10_{C_2}}{12_{C_3}} = \frac{9}{22}$	$\frac{2_{C_2} \times 10_{C_1}}{12_{C_3}} = \frac{1}{22}$	0

9.

$x:$	0	1	2	3
$p(x):$	$\frac{5}{30}$	$\frac{15}{30}$	$\frac{9}{30}$	$\frac{1}{30}$

10.

- (i) $\frac{1}{3}$
- (ii) $\frac{2}{3}$
- (iii) $\frac{1}{4}$
- (iv) 1.7 (= 2 approximately)

2.4

Discrete Probability Distributions

Modules 2.1 and 2.2 deal with general properties of random variables. Random variables with special probability distributions are encountered in different fields of *science* and *engineering*. Some specific **discrete probability distributions** are discussed in this module and some specific **continuous probability distributions** are discussed in the next module 2.5.

Discrete Uniform Distribution: A r.v. X is said to have a **discrete uniform distribution** over the range $[1, n]$, if its p.m.f. is given by

$$p(x) = P(X = x) = \begin{cases} \frac{1}{n} & , \quad x = 1, 2, \dots, n \\ 0 & , \quad \text{otherwise} \end{cases}$$

Notation: $X \sim U(n)$, read as X follows discrete uniform distribution with parameter n .

Note: If all possible values of a r.v. are equally likely, then this distribution is used.

Example 1: If an unbiased coin is tossed once and X is equal to number of heads, then $X = 0, 1$ and

$$P(X = 0) = P(X = 1) = \frac{1}{2} \text{ and } X \sim U(2).$$

Example 2: If an unbiased die is thrown once and X is equal to number on the die, then $x = 1, 2, 3, 4, 5, 6$ and $P(X = i) = \frac{1}{6}$ for $i = 1, 2, 3, 4, 5, 6$ and $X \sim U(6)$.

Mean and Variance: We have $E(X) = \frac{1}{n} \sum_{i=1}^n i = \frac{n+1}{2}$

$$\text{and } E(X^2) = \frac{1}{n} \sum_{i=1}^n i^2 = \frac{(n+1)(2n+1)}{6}$$

$$\text{Thus } V(X) = E(X^2) - (E(X))^2 = \frac{(n+1)(n-1)}{12}$$

Bernoulli Experiment: A random experiment whose outcomes are of two types, **success (S)** and **failure (F)**, occurring with probabilities p and $q (= 1 - p)$ respectively, is called a **Bernoulli experiment**.

Conducting a Bernoulli experiment once is known as **Bernoulli trial**. Note that p and q are same in each trial and outcomes of different trials are independent.

Bernoulli distribution: In a Bernoulli experiment, if a r.v. X is defined such that it takes value 1 with probability p when S occurs and 0 with probability q when F occurs, then we say that X follows Bernoulli distribution and its p.m.f. is given by

$$p(x) = P(X = x) = \begin{cases} p^x q^{1-x} & , \quad x = 0, 1 \\ 0 & , \quad \text{otherwise} \end{cases}$$

Examples:

- 1) Tossing of a coin (results a head or tail)
- 2) Performance of a student in an examination (results pass or failure)
- 3) Sex of an unborn child (results female or male)

Mean and Variance:

$$\text{Mean} = \mu = E(X) = 0 \times q + 1 \times p = p$$

$$\text{and } E(X^2) = 0^2 \times q + 1^2 \times p = p$$

$$\therefore \text{Variance} = \sigma^2 = E(X^2) - (E(X))^2 = p - p^2 = p(1 - p) = pq$$

Binomial Distribution: Suppose we conduct n independent Bernoulli trials and we define

$$X = \text{number of successes in } n \text{ trials.}$$

Then X is a discrete random variable and it takes the values $0, 1, 2, \dots, n$.

Derivation of $P(X = x)$: Note that $X = x$ means that there are x successes and $(n - x)$ failures in n trials in a specified order (say) SSFSFFFS ... FSF.

Since outcomes of different trials are independent, by Multiplication Theorem, we have

$$\begin{aligned}
 P(SSFSFFFS \dots FSF) &= P(S) \cdot P(S) \cdot P(F) \cdot P(S) \cdot P(F) \cdot P(F) \cdot P(F) \cdot P(S) \cdot \\
 &\quad \dots P(F) \cdot P(S) \cdot P(F) \\
 &= p \cdot p \cdot q \cdot p \cdot q \cdot q \cdot q \cdot p \cdot \dots q \cdot p \cdot q \\
 &= \underbrace{p \cdot p \cdot \dots \cdot p}_{(x \text{ times})} \cdot \underbrace{q \cdot q \cdot \dots \cdot q}_{(n-x \text{ times})} = p^x q^{n-x}
 \end{aligned}$$

But x successes in n trials can occur in $\binom{n}{x}$ orders and the probability for each of these orders is same, viz., $p^x q^{n-x}$. Hence by addition theorem of probability

$$p(x) = P(X = x) = \binom{n}{x} p^x q^{n-x}$$

Definition: A r.v. X is said to follow a **binomial distribution** with parameters n and p if its p.m.f. is given by

$$p(x) = P(X = x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & , \quad x = 0, 1, 2, \dots, n, 0 < p < 1, q = 1 - p \\ 0 & , \quad \text{otherwise} \end{cases}$$

Notation: $X \sim B(n, p)$. Read as X follows binomial distribution with parameters n and p .

Real life examples:

- 1) Number of heads in n tosses of a coin
- 2) Number of boys in a family of n children
- 3) Number of times hitting a target in n attempts

Note:

1.

$$\sum_{x=0}^n p(x) = \sum_{x=0}^n \binom{n}{x} p^x q^{n-x} = (q + p)^n = 1$$

2. The c.d.f. of X is given by

$$F(x) = P(X \leq x) = \sum_{k=0}^x \binom{n}{k} p^k q^{n-k}, x = 0, 1, 2, \dots, n$$

Example 1: Four fair coins are tossed. If the outcomes are assumed to be independent, then find the p.m.f. and c.d.f. of the number of heads obtained.

Solution: Let X be the no. of heads in tossing 4 coins.

Then $X \sim B\left(4, \frac{1}{2}\right)$ where $p = P(\text{head}) = \frac{1}{2}$.

$$\begin{aligned} \text{Thus } p(x) = P(X = x) &= \binom{4}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{4-x} \\ &= \binom{4}{x} \left(\frac{1}{2}\right)^4 = \binom{4}{x} \left(\frac{1}{16}\right) \text{ for } x = 0, 1, 2, 3, 4. \end{aligned}$$

$$\text{Then } p(0) = \binom{4}{0} \left(\frac{1}{16}\right) = \frac{1}{16}$$

$$p(1) = \binom{4}{1} \left(\frac{1}{16}\right) = \frac{4}{16}$$

$$p(2) = \binom{4}{2} \left(\frac{1}{16}\right) = \frac{6}{16}$$

$$p(3) = \binom{4}{3} \left(\frac{1}{16}\right) = \frac{4}{16}$$

$$p(4) = \binom{4}{4} \left(\frac{1}{16}\right) = \frac{1}{16}$$

The p.m.f $p(x)$ and c.d.f $F(x)$ are given in the following table.

x	0	1	2	3	4
$p(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$
$F(x)$	$\frac{1}{16}$	$\frac{5}{16}$	$\frac{11}{16}$	$\frac{15}{16}$	1

Example 2: A and B play a game in which their chances of winning are in the ratio 3 : 2. Find A's chance of winning at least three games out of the five games played.

Solution:

Define X = No. of games A winning out of 5.

Here $p = P(A \text{ winning}) = \frac{3}{5}$ and $n = 5$ and $X \sim B\left(5, \frac{3}{5}\right)$. Thus,

$$p(x) = P(X = x) = \binom{5}{x} \left(\frac{3}{5}\right)^x \left(\frac{2}{5}\right)^{5-x} \text{ for } x = 0, 1, \dots, 5.$$

Required to find:

$$\begin{aligned} P(\text{A winning at least 3 out of 5 games}) &= P(X \geq 3) \\ &= P(X = 3) + P(X = 4) + P(X = 5) \\ &= \binom{5}{3} \left(\frac{3}{5}\right)^3 \left(\frac{2}{5}\right)^2 + \binom{5}{4} \left(\frac{3}{5}\right)^4 \left(\frac{2}{5}\right)^1 + \binom{5}{5} \left(\frac{3}{5}\right)^5 \\ &= \frac{3^3}{5^5} [10 \times 4 + 5 \times 3 \times 2 + 1 \times 9] \\ &= \frac{27 \times (40 + 30 + 9)}{3125} = 0.68 \end{aligned}$$

Example 3: The probability of a man hitting a target is $\frac{1}{4}$.

- (i) If he fires 7 times, what is the probability of his hitting the target at least twice?
- (ii) How many times must he fire so that the probability of his hitting the target at least once is greater than $\frac{2}{3}$?

Solution: Let X be the no. of times a man hitting the target in 7 fires. Here

$p = P(\text{man hitting the target}) = \frac{1}{4}$ and $n = 7$. Then $X \sim B\left(7, \frac{1}{4}\right)$ and

$$p(x) = \binom{7}{x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{7-x} \text{ for } x = 0, 1, 2, \dots, 7.$$

$$\begin{aligned}
 \text{(i)} \quad P(\text{at least two hits}) &= P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)] \\
 &= 1 - [p(0) + p(1)] \\
 &= 1 - \left\{ \binom{7}{0} \left(\frac{1}{4}\right)^0 \left(\frac{3}{4}\right)^7 + \binom{7}{1} \left(\frac{1}{4}\right) \left(\frac{3}{4}\right)^6 \right\} = \frac{4547}{8192} = 0.55
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad \text{Find } n \text{ such that } P(X \geq 1) &> \frac{2}{3} \\
 \Rightarrow 1 - P(X = 0) &> \frac{2}{3} \\
 \Rightarrow -1 + P(X = 0) &< -\frac{2}{3} \\
 \Rightarrow P(X = 0) &< \frac{1}{3} \\
 \Rightarrow \binom{n}{0} \left(\frac{1}{4}\right)^0 \left(\frac{3}{4}\right)^n &< \frac{1}{3} \\
 \Rightarrow \left(\frac{3}{4}\right)^n &< \frac{1}{3} \\
 \Rightarrow n[\log 3 - \log 4] &< \log 1 - \log 3 \\
 \Rightarrow n[\log 4 - \log 3] &> \log 3 \\
 \Rightarrow n > \frac{\log 3}{\log 4 - \log 3} &= 3.8, \text{ since } n \text{ cannot be fractional, the} \\
 &\text{required number of shots is 4.}
 \end{aligned}$$

Mean of Binomial Distribution:

$$\begin{aligned}
 \mu = E(X) &= \sum_{x=0}^n x p(x) = \sum_{x=0}^n x \binom{n}{x} p^x q^{n-x} \\
 &= \sum_{x=1}^n x \binom{n}{x} \binom{n-1}{x-1} p^x q^{n-x} = np \sum_{x=1}^n \binom{n-1}{x-1} p^{x-1} q^{n-x} \\
 &= np(q + p)^{n-1} = np
 \end{aligned}$$

$$\Rightarrow \mu = np$$

Variance of Binomial Distribution:

$$\sigma^2 = V(X) = npq \text{ (See } P_1 \text{ for proof)}$$

Example 4: One hundred balls are tossed into 50 boxes. What is the expected number of balls in the tenth box.

Solution: If we think of the balls tossed as Bernoulli trials in which a success is defined as getting a ball in the tenth box, then $p = \frac{1}{50}$. If X denotes the number of balls that go into the tenth box.

Then $X \sim B\left(100, \frac{1}{50}\right)$ and $E(X) = np = 100 \times \frac{1}{50} = 2$.

Example 5: The mean and variance of binomial distribution are 4 and $\frac{4}{3}$ respectively. Find $P(X \geq 1)$.

Solution: Here $X \sim B(n, p)$. But $np = 4$ and $np(1 - p) = \frac{4}{3}$.

Hence $4(1 - p) = \frac{4}{3} \Rightarrow 1 - p = \frac{1}{3} \Rightarrow p = \frac{2}{3}$ and $n = \frac{4}{p} = 4 \times \frac{3}{2} = 6$.

Thus, $X \sim B\left(6, \frac{1}{3}\right)$ and hence $P(X \geq 1) = 1 - P(X = 0)$

$$= 1 - \binom{6}{0} \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^6 = 1 - \left(\frac{1}{3}\right)^6$$

Poisson Distribution:

If $n \rightarrow \infty$ and $p \rightarrow 0$ such that $\lambda = np$ fixed, then $\binom{n}{x} p^x (1 - p)^{n-x} \rightarrow \frac{e^{-\lambda} \lambda^x}{x!}$ which is the p.m.f. of Poisson distribution (**See P_2**). Thus the p.m.f. of **Poisson distribution** is obtained as the limit of p.m.f. of **binomial distribution**.

Definition: A r.v. X is said to follow a **Poisson distribution** with parameter λ if its p.m.f. is given by

$$p(x) = P(X = x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!} & , \quad x = 0, 1, 2, \dots; \lambda > 0 \\ 0 & , \quad \text{otherwise} \end{cases}$$

Notation: Read $X \sim P(\lambda)$ as: X follows poisson distribution with parameter λ .

Real life examples

- 1) Number of defectives in a packet of 100 blades.
- 2) Number of telephone calls received at a particular telephone exchange in some unit of time.
- 3) Number of print mistakes in a page of a book.
- 4) The number of fragments received by a surface area 'A' from a fragment atom bomb.
- 5) The emission of radio active (alpha) particles.
- 6) Number of air accidents in some unit of time.

Note :

$$1. \sum_{x=0}^{\infty} p(x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1 \left(\because e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!} \right)$$

2. The c.d.f. of X is given by

$$F(x) = P(X \leq x) = \sum_{k=0}^x \frac{e^{-\lambda} \lambda^k}{k!}$$

Example 6: Messages arrive at a switchboard in a Poisson manner at an average rate of six per hour. Find the probability for each of the following events:

- a) Exactly two messages arrive within one hour.
- b) No message arrives within one hour.
- c) At least three messages arrive within one hour.

Solution: Let X be the r.v. that denotes the number of messages arriving at the switchboard within a one-hour interval. Then $X \sim P(6)$ and its p.m.f is given by

$$P(X = x) = p(x) = \frac{e^{-6} 6^x}{x!} \text{ for } x = 0, 1, 2, \dots$$

$$a) P(X = 2) = \frac{e^{-6}6^2}{2!} = \frac{36}{2}e^{-6} = 18e^{-6}.$$

$$b) P(X = 0) = \frac{e^{-6}6^0}{0!} = e^{-6}.$$

$$c) P(X \geq 3) = 1 - \{P(X = 0) + P(X = 1) + P(X = 2)\} \\ = 1 - \left\{ \frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!} \right\} = 1 - e^{-6}\{1 + 6 + 18\} = 1 - 25e^{-6}$$

Example 7: In a book of 520 pages, 390 typo-graphical errors occur. Assuming Poisson law for the number of errors per page, find the probability that a random sample of 5 pages will contain no error.

Solution:

$$\lambda = \text{Average number of typo-graphical errors/page} = \frac{390}{520} = 0.75$$

Let X = Number of errors per page

$$\text{Then } X \sim P(0.75) \text{ and } p(x) = P(X = x) = \frac{e^{-0.75}(0.75)^x}{x!}$$

$$P(\text{No error}) = P(X = 0) = p(0) = e^{-0.75}$$

$$P(\text{A random sample of 5 pages contain no error}) = [p(0)]^5 = [e^{-0.75}]^5 = e^{-3.75}$$

Mean of Poisson Distribution : The mean of poisson distribution is given by

$$\mu = E(X) = \sum_{x=0}^{\infty} x p(x) = \sum_{x=1}^{\infty} \frac{x e^{-\lambda} \lambda^x}{x!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

$$\Rightarrow \mu = \lambda$$

Variance of Poisson Distribution:

$$E(X^2) = \sum_{x=0}^{\infty} x^2 P(x) = \sum_{x=0}^{\infty} [x(x-1) + x] p(x)$$

$$\begin{aligned}
&= \sum_{x=2}^{\infty} x(x-1)p(x) + \sum_{x=1}^{\infty} xp(x) = \sum_{x=2}^{\infty} x(x-1) \frac{e^{-\lambda} \lambda^x}{x!} + \lambda \\
&= e^{-\lambda} \lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \lambda = e^{-\lambda} \lambda^2 e^{\lambda} + \lambda
\end{aligned}$$

$$\Rightarrow E(X^2) = \lambda^2 + \lambda$$

The variance of Poisson distribution is given by

$$V(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

$$\Rightarrow V(X) = \lambda$$

Note that *for Poisson distribution, mean and variance are equal.*

Example 8: If X and Y are independent Poisson variates such that

$P(X = 1) = P(X = 2)$ and $P(Y = 2) = P(Y = 3)$, find the variance of $X - 2Y$.

Solution: Let $X \sim P(\lambda)$ and $Y \sim P(\mu)$. Then

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} \text{ for } x = 0, 1, 2, \dots; \lambda > 0 \text{ and}$$

$$P(Y = y) = \frac{e^{-\mu} \mu^y}{y!} \text{ for } y = 0, 1, 2, \dots; \mu > 0.$$

$$\text{Since } P(X = 1) = P(X = 2); \lambda e^{-\lambda} = \frac{e^{-\lambda} \lambda^2}{2} \Rightarrow \lambda^2 - 2\lambda = 0$$

$$\Rightarrow \lambda(\lambda - 2) = 0 \Rightarrow \lambda = 0, 2 \Rightarrow \lambda = 2 (\because \lambda = 0 \text{ is not admissible})$$

$$\text{Since } P(Y = 2) = P(Y = 3), \text{ then } \frac{e^{-\mu} \mu^2}{2} = \frac{e^{-\mu} \mu^3}{6} \Rightarrow \mu^3 - 3\mu^2 = 0$$

$$\Rightarrow \mu^2(\mu - 3) = 0 \Rightarrow \mu = 0, 3 \Rightarrow \mu = 3.$$

$$V(X - 2Y) = 1^2 V(X) + (-2)^2 V(Y) = \lambda + 4\mu = 2 + 4 \times 3 = 2 + 12 = 14$$

Negative Binomial (or Pascal) Distribution:

Let X denote the number of failures before the r^{th} success in a sequence of Bernoulli trials. Then the number of trials required is $X + r$.

Derivation of $P(X = x)$:

In $x + r$ trials, the last trial must be a success whose probability is p . In the remaining $(x + r - 1)$ trials, we must have $(r - 1)$ successes whose probability is $\binom{x + r - 1}{r - 1} p^{r-1} q^x$ (Using binomial distribution).

Thus, by multiplication theorem, we have

$$p(x) = P(X = x) = \binom{x + r - 1}{r - 1} p^{r-1} q^x \cdot p = \binom{x + r - 1}{r - 1} p^r q^x$$

Definition: A random variable X is said to follow a **Negative binomial distribution**(NBD) with parameters r and p if its p.m.f is given by

$$p(x) = P(X = x) = \begin{cases} \binom{x + r - 1}{r - 1} p^r q^x & , \quad x = 0, 1, 2, \dots \\ 0 & , \quad \text{otherwise} \end{cases}$$

Notation: $X \sim NB(r, p)$.

Note:

$$\begin{aligned} 1. \quad \binom{x + r - 1}{r - 1} &= \binom{x + r - 1}{x} \quad \left(\because \binom{n}{r} = \binom{n}{n - r} \right) \\ &= \frac{(x + r - 1)(x + r - 2) \dots (r + 1)r}{x!} \\ &= (-1)^r \frac{(-r)(-r - 1) \dots (-r - x + 2)(-r - x + 1)}{x!} = (-1)^x \binom{-r}{x} \end{aligned}$$

Thus, the p.m.f. of NBD can be written as

$$p(x) = \begin{cases} \binom{-r}{x} p^r (-q)^x & , \quad x = 0, 1, 2, \dots \\ 0 & , \quad \text{otherwise} \end{cases}$$

Further, it is the $(x + 1)^{th}$ term in the expansion of $p^r(1 - q)^{-r}$, a binomial expansion with negative index. Therefore, the distribution is known as negative binomial distribution.

$$2. \sum_{x=0}^{\infty} p(x) = p^r \sum_{x=0}^{\infty} \binom{-r}{x} (-q)^x = p^r (1 - q)^{-r} = p^r p^{-r} = 1$$

Mean of NBD: The mean of NBD is given by

$$\begin{aligned} \mu = E(x) &= \sum_{x=0}^{\infty} x p(x) = \sum_{x=1}^{\infty} x p(x) = \sum_{x=1}^{\infty} x \binom{-r}{x} p^r (-q)^x \\ &= p^r \sum_{x=1}^{\infty} x \binom{-r}{x} \binom{-r+1}{x-1} (-q)^x = (-r)(p^r)(-q) \sum_{x=1}^{\infty} \binom{-r+1}{x-1} (-q)^{x-1} \\ &= (-r)(-q)p^r (1 - q)^{-(r+1)} \\ &= rq p^r p^{-(r+1)} = \frac{rq}{p} \end{aligned}$$

Variance of NBD: $\sigma^2 = \frac{rq}{p^2}$ (See P_3 for proof)

3. Notice that $\frac{\mu}{\sigma^2} = p < 1$ and this implies that mean is smaller than variance in NBD.

4. If Y = Number of trials required to get r^{th} success, then $Y = X + r$ and

$$P(Y = y) = P(X + r = y) = P(X = y - r) = \binom{y-1}{r-1} p^r q^{y-r}$$

for $y = r, r + 1, \dots$ and $E(y) = E(X) + r = \frac{rq}{p} + 1$ and $V(Y) = V(X) = \frac{rq}{p^2}$

Real life examples

- 1) Number of tails before the third head.
- 2) Number of girls before the second son.
- 3) Number of non-defectives before the third defective.

Example 9: Find the probability that there are two daughters before the second son in a family when probability of a son is 0.5.

Solution: Let X = the number of daughters before second son

$$\text{Then } P(X = x) = \binom{x+2-1}{2-1} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^x = \binom{x+1}{1} \left(\frac{1}{2}\right)^{x+2}$$

$$\text{and } P(X = 2) = \binom{3}{1} \left(\frac{1}{2}\right)^4 = \frac{3}{16}$$

Example 10: An item is produced in large numbers. The machine is known to produce 5% defectives. A quality control inspector is examining the items by taking them at random. What is the probability that at least 4 items are to be examined in order to get 2 defectives?

Solution: Let Y = No. of items to be examined in order to get 2 defectives. Here $p = (\text{defective}) = \frac{5}{100} = 0.05$.

$$\text{Then } P(Y = y) = \binom{y-1}{2-1} (0.05)^2 (0.95)^{y-2}$$

$$\Rightarrow P(Y = y) = (y-1)(0.05)^2 (0.95)^{y-2}$$

We want to find

$$\begin{aligned} P(Y \geq 4) &= 1 - \sum_{y=2}^3 P(Y = y) = 1 - \{P(Y = 2) + P(Y = 3)\} \\ &= 1 - \{(0.05)^2 + 2(0.05)^2(0.95)\} = 0.9928 \end{aligned}$$

Geometric distribution:

Let X denotes the number of failures before the first success in a sequence of Bernoulli trials. Then the required number of trials is $X + 1$.

Geometric distribution is a particular case of negative binomial distribution with $r = 1$.

Definition: A random variable X is said to follow a **Geometric distribution** (GD) with parameter p if its p.m.f. is given by

$$p(x) = P(X = x) = \begin{cases} p \cdot q^x & \text{for } x = 0, 1, 2, \dots \\ 0 & , \text{ otherwise} \end{cases}$$

Notation: $X \sim GD(r)$

Mean and variance of GD

$$\mu = \frac{q}{p} \text{ and } \sigma^2 = \frac{q}{p^2} \text{ (take } r = 1 \text{ in } \mu \text{ and } \sigma^2 \text{ of NBD)}$$

Note: Let Y = Number of trials required to get first success, then $Y = X + 1$ and

$$P(Y = y) = P(X + 1 = y) = P(X = y - 1) = \begin{cases} p q^{y-1} & \text{for } y = 1, 2, 3, \dots \\ 0 & , \text{ otherwise} \end{cases}$$

$$\text{Further, } E(Y) = E(X) + 1 = \frac{1}{p} \text{ and } V(Y) = V(X) = \frac{q}{p^2}.$$

Real life examples

- 1) Number of tails before the third head
- 2) Number of girls before the second son
- 3) Number of non-defectives before the first defective

Example 11: find the probability that there are two daughters before the first son in a family where probability of a son is 0.5 .

Solution: Let X = Number of daughters before the first son.

$$\text{Then } P(X = x) = \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right)^x = \left(\frac{1}{2}\right)^{x+1} \text{ for } x = 0, 1, 2, \dots$$

$$\text{and } P(X = 2) = \left(\frac{1}{2}\right)^{2+1} = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

Hyper geometric Distribution:

Consider an urn with N balls, M of which are white and $N - M$ are red. Suppose we draw a sample of n balls at random with replacement. Let X denote the

number of white balls in the sample. Then $X \sim B(n, p)$ where $p = \frac{M}{N}$ which remains same for all trials and outcomes of different trials are independent. The p.m.f of X is given by $P(X = x) = \binom{n}{x} \left(\frac{M}{N}\right)^x \left(1 - \frac{M}{N}\right)^{n-x}$ for $i = 1, 2, \dots, n$.

If the sample is selected without replacement, p is not same for all trials and outcomes of different trials are not independent and hence binomial distribution cannot be applied.

Derivation of $P(X = x)$: The number of all possible samples without replacement $= \binom{N}{x}$.

The number of samples in which there are x white balls and $(n - x)$ red balls $= \binom{M}{x} \binom{N - M}{n - x}$.

Thus $p(x) = P(X = x) = \frac{\binom{M}{x} \binom{N - M}{n - x}}{\binom{N}{n}}$.

Definition: A random variable X is said to follow the **hyper geometric distribution** if its p.m.f is given by

$$p(x) = P(X = x) = \begin{cases} \frac{\binom{M}{x} \binom{N - M}{n - x}}{\binom{N}{n}} & , x = 0, 1, \dots, \min(n, M) \\ 0 & , \text{otherwise} \end{cases}$$

Example 12: A bag contains 4 white balls and 3 green balls. Three balls are drawn. What is the probability that 2 are white.

Solution: $N = 4 + 3 = 7$, $M = 4$, $n = 3$

X = Number of white balls and

$$P(X = 2) = \frac{\binom{4}{2} \binom{3}{1}}{\binom{7}{3}} = \frac{4 \times 3 \times 3 \times 6}{2 \times 7 \times 6 \times 5} = \frac{18}{35}$$

Note:
$$\sum_{x=0}^{\min(n,M)} p(x) = \begin{cases} \sum_{x=0}^n \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} = \frac{\binom{N}{n}}{\binom{N}{n}} = 1 & \text{if } \min(n, M) = n \\ \sum_{x=0}^n \frac{\binom{M}{x} \binom{N-M}{M-x}}{\binom{N}{M}} = \frac{\binom{N}{M}}{\binom{N}{M}} = 1 & \text{if } \min(n, M) = M \end{cases}$$

Mean and variance of Hyper geometric distribution:

The mean is given by $\mu = \frac{nM}{N}$ and variance is given by $\sigma^2 = \frac{NM(N-M)(N-n)}{N^2(N-1)}$ if $\min(n, M) = n$ (**For proof, see P_4**).

P1:

If $X \sim B(n, p)$, then show that $V(X) = npq$.

Proof: We already derived that $E(X) = \sum_{x=0}^n x p(x) = np$.

$$\begin{aligned} E(X^2) &= \sum_{x=0}^n x^2 p(x) = \sum_{x=0}^n [x(x-1) + x] p(x) \\ &= \sum_{x=2}^n x(x-1) p(x) + \sum_{x=0}^n x p(x) \\ &= \sum_{x=2}^n x(x-1) \binom{n}{x} p^x q^{n-x} + np \\ &= \sum_{x=2}^n x(x-1) \binom{n}{x} \binom{n-1}{x-1} \binom{n-2}{x-2} p^x q^{n-x} + np \\ &= n(n-1) p^2 \sum_{x=2}^n \binom{n-2}{x-2} p^{x-2} q^{n-x} + np \\ &= n(n-1) p^2 (q + p)^{n-2} + np \\ &\Rightarrow E(X^2) = n(n-1) p^2 + np \text{ and hence} \end{aligned}$$

$$V(X) = E(X^2) - (E(X))^2 = n(n-1) p^2 + np - n^2 p^2 = npq$$

P2:

If $\lambda = np$, then $\lim_{n \rightarrow \infty} \lim_{p \rightarrow 0} \binom{n}{x} p^x q^{n-x} = \frac{e^{-\lambda} \lambda^x}{x!}$.

Proof:

$$\begin{aligned} \binom{n}{x} p^x q^{n-x} &= \binom{n}{x} p^x (1-p)^{n-x} \\ &= \binom{n}{x} \left(\frac{p}{1-p} \right)^x (1-p)^n \\ &= \frac{n(n-1)(n-2)\dots(n-x+1)}{x!} \frac{\left(\frac{\lambda}{n} \right)^x}{\left(1 - \frac{\lambda}{n} \right)^x} \left(1 - \frac{\lambda}{n} \right)^n \\ &= \frac{n^x \left(1 - \frac{1}{n} \right) \left(1 - \frac{2}{n} \right) \dots \left(1 - \frac{x-1}{n} \right)}{x! \left(1 - \frac{\lambda}{n} \right)^x} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n} \right)^n \end{aligned}$$

Then $\lim_{n \rightarrow \infty} \lim_{p \rightarrow 0} \binom{n}{x} p^x q^{n-x} = \frac{e^{-\lambda} \lambda^x}{x!} \quad \left(\because \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n} \right)^n = e^{-\lambda} \right)$

P3:

For negative binomial distribution, show that $V(X) = \frac{rq}{p^2}$.

Proof:

We already derived that $p(x) = \binom{-r}{x} p^r (-q)^x$

$$E(X) = \sum_{x=0}^{\infty} x p(x) = \frac{rq}{p}$$

$$\begin{aligned} E(X^2) &= \sum_{x=0}^{\infty} x^2 p(x) = \sum_{x=0}^{\infty} [x(x-1) + x] p(x) \\ &= \sum_{x=2}^{\infty} x(x-1) p(x) + \sum_{x=0}^{\infty} x p(x) \\ &= \sum_{x=2}^{\infty} x(x-1) \binom{-r}{x} p^r (-q)^x + \frac{rq}{p} \\ &= \sum_{x=2}^{\infty} x(x-1) \binom{-r}{x} \binom{-r-1}{x-1} \binom{-r+2}{x-2} p^r (-q)^x + \frac{nq}{p} \\ &= r(r+1) p^r (-q)^2 \sum_{x=2}^{\infty} \binom{-r+2}{x-2} (-q)^{x-2} + \frac{nq}{p} \\ &= r(r+1) p^r q^2 (1-q)^{-(r+2)} + \frac{nq}{p} \left(\because \sum_{x=2}^{\infty} \binom{-r+2}{x-2} (-q)^{x-2} = (1-q)^{-(r+2)} \right) \\ &= r(r+1) p^r q^2 p^{-(r+2)} + \frac{nq}{p} \\ &\Rightarrow E(X^2) = r(r+1) \frac{q^2}{p^2} + \frac{nq}{p} \end{aligned}$$

$$\begin{aligned}
 \text{Thus, } V(X) &= E(X^2) - (E(X))^2 = r(r+1)\frac{q^2}{p^2} + \frac{nq}{p} - \frac{r^2q^2}{p^2} \\
 &= \frac{r^2q^2}{p^2} + \frac{rq^2}{p^2} + \frac{nq}{p} - \frac{r^2q^2}{p^2} \\
 &= \frac{rq}{p^2}(q+p)
 \end{aligned}$$

$$\Rightarrow V(X) = \frac{rq}{p^2}$$

P4:

For hyper geometric distribution, show that mean and variance are given by

$$\mu = E(X) = \frac{nM}{n} \text{ and}$$

$$\sigma^2 = V(X) = \frac{NM(N-M)(N-n)}{N^2(N-1)}$$

when $\min(n, M) = n$.

Proof:

$$\text{We have } p(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

$$E(X) = \sum_{x=0}^n x p(x) = \sum_{x=0}^n \frac{x \binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= \sum_{x=1}^n \frac{x \binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= \frac{M}{\binom{N}{n}} \sum_{x=1}^n \binom{M-1}{x-1} \binom{N-M}{n-x} = \frac{M}{\binom{N}{n}} \binom{N-1}{n-1}$$

$$= \frac{M}{\binom{N}{n} \binom{N-1}{n-1}} \left(\because \sum_{x=1}^n \binom{M-1}{x-1} \binom{N-M}{n-x} = \binom{N-1}{n-1} \right)$$

$$\Rightarrow \mu = \frac{nM}{N}$$

$$\text{Now, } E(X^2) = \sum_{x=0}^n x^2 p(x) = \sum_{x=0}^n [x(x-1) + x] p(x) = \sum_{x=2}^n x(x-1) p(x) + \sum_{x=0}^n x p(x)$$

$$= \sum_{x=2}^n \frac{x(x-1) \binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} + \mu$$

$$= \sum_{x=2}^n \frac{x(x-1) \left(\frac{M}{x}\right) \left(\frac{M-1}{x-1}\right) \binom{M-2}{x-2} \binom{N-M}{n-x}}{\binom{N}{n}} + \mu$$

$$= \frac{M(M-1)}{\binom{N}{n}} \sum_{x=2}^n \binom{M-2}{x-2} \binom{N-M}{n-x} + \mu$$

$$= \frac{M(M-1)}{\binom{N}{n}} \binom{N-2}{n-2} + \mu \left(\because \sum_{x=2}^n \binom{M-2}{x-2} \binom{N-M}{n-x} = \binom{N-2}{n-2} \right)$$

$$= \frac{M(M-1)}{\left(\frac{N}{n}\right) \left(\frac{N-1}{n-1}\right) \binom{N-2}{n-2}} \binom{N-2}{n-2} + \mu$$

$$\Rightarrow E(X^2) = \frac{n(n-1)M(M-1)}{N(N-1)} + \frac{nM}{N}$$

$$\begin{aligned}
 \text{Thus, } V(X) &= E(X^2) - (E(X))^2 \\
 &= \frac{n(n-1)M(M-1)}{N(N-1)} + \frac{nM}{N} - \frac{n^2M^2}{N^2} \\
 &= \frac{NM(N-M)(N-n)}{N^2(N-1)} \text{ (on simplification)}
 \end{aligned}$$

2.4. Discrete Probability Distributions

Exercise

1. Ten coins are tossed simultaneously. Find the probability of getting at least seven heads.
2. A multiple choice test consists of 8 questions with 3 answers to each question (of which only one is correct). A student answers each question by rolling a balanced die and checking the first answer if he gets 1 or 2, the second answer if he gets 3 or 4 and the third answer if he gets 5 or 6. To get a distinction, the student must secure at least 75% correct answers. If there is no negative marking, what is the probability that the student secures a distinction?
3. In a precision bombing attack there is a 50% chance that any one bomb will strike the target. Two direct hits are required to destroy the target completely. How many bombs must be dropped to give a 99% chance or better of completely destroying the target?
4. A manufacturer of pins knows that 5% of his product is defective. If he sells pins in boxes of 100 and guarantees that not more than 10 pins will be defective, what is the probability that a box fails to meet the guaranteed quality?
5. An insurance company insures 4,000 people against loss of both eyes in a car accident. Based on previous data, the rates were computed on the assumption that on the average 10 persons in 1,00,000 will have car accident each year that result in this type of injury. What is the probability that more than 3 of the insured will collect on their policy in a given year?

6. A manufacturer, who produces medicine bottles, finds that 0.1% of the bottles are defective. The bottles are packed in boxes containing 500 bottles. A drug manufacturer buys 100 boxes from the product of bottles. Find how many boxes will contain (i) no defective, and (ii) at least two defectives.
7. Six coins are tossed 6,400 times. Using Poisson distribution, find the approximate probability of getting six heads two times.
8. Find the probability that a person tossing 3 coins will get either all heads or all tails, for the second time on the fifth toss.
9. If the probability is 0.4 that a child exposed to a certain contagious disease will catch it, what is the probability that the tenth child exposed to the disease will be third to catch it.
10. If the probabilities of having a male or female child are both 0.5, find the probability that
 - a) a family's fourth child is their first son
 - b) a family's seventh child is their second daughter
 - c) a family's tenth child is their fourth or fifth son

Answers:

1. $\frac{176}{1024}$

2. $\frac{129}{729 \times 9}$

3. 11

4. $1 - e^{-5} \sum_{x=0}^{10} \frac{5^x}{x!}$

5. 0.0008

6. (i) 61 (ii) 9

7. $\frac{e^{-100}(100)^2}{2}$

8. $\binom{4}{1} \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^3$

9. $\binom{9}{2} (0.4)^3 (0.6)^7$

10. a) $(0.5)^4$ b) $6(0.5)^7$ c) $210(0.5)^{10}$

2.5

Continuous Probability Distributions

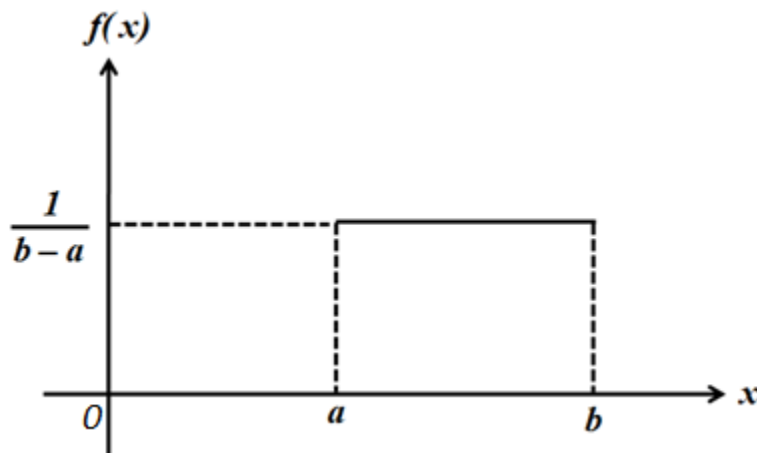
The **continuous probability distributions** are used in a number of applications in *engineering*. For example in *error analysis*, given a set of data or probability distribution, it is possible to estimate the probability that a measurement (temperature, pressure, flow rate) will fall within a desired range, and hence determine how reliable an instrument or piece of equipment is. Also, one can calibrate an instrument (ex. Temperature sensor) from the manufacturer on a regular basis and use a probability distribution to see if the variance in the instruments' measurements increases or decreases over time.

Uniform Distribution

A continuous random variable (c. r. v.) X is said to have a uniform distribution over the interval $[a, b]$ if its p. d. f. is given by

$$f(x) = \begin{cases} \frac{1}{b-a} & , \quad a \leq x \leq b \\ 0 & , \quad \text{otherwise} \end{cases}$$

Notation: $X \sim U(a, b)$. Read as X follows uniform distribution with parameters a and b . It is used to model events that are equally likely to occur at any time within a given time interval. The plot of p. d. f. is given below:



The cumulative distribution function (c. d. f.) of X is given by

$$F(x) = p(X \leq x) = \begin{cases} 0 & , \quad x < a \\ \frac{x-a}{b-a} & , \quad a \leq x \leq b \\ 1 & , \quad x \geq b \end{cases}$$

The mean of X is given by

$$\begin{aligned} \mu = E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_a^b x f(x) dx = \int_a^b \frac{x}{b-a} dx \\ &= \left[\frac{x^2}{2(b-a)} \right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2} \Rightarrow \mu = \frac{b+a}{2} \end{aligned}$$

$$\begin{aligned} \text{Now, } E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_a^b x^2 f(x) dx = \int_a^b \frac{x^2}{b-a} dx \\ &= \left[\frac{x^3}{3(b-a)} \right]_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{b^2+ab+a^2}{3} \\ &\Rightarrow E(X^2) = \frac{b^2+ab+a^2}{3} \end{aligned}$$

Thus, the variance of X is given by $\sigma^2 = E(X^2) - (E(X))^2$

$$\begin{aligned} &= \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4} \\ &= \frac{b^2-2ab+a^2}{12} \\ &\Rightarrow \sigma^2 = \frac{(b-a)^2}{12} \end{aligned}$$

Example 1: The time that a professor takes to grade a paper is uniformly distributed between 5 *minutes* and 10 *minutes*. Find the mean and variance of the time the professor takes to grade a paper.

Solution: Let X denotes the time the professor takes to grade a paper. Then $X \sim U(5, 10)$.

$$\mu = E(X) = \frac{10+5}{2} = 7.5 \text{ and } \sigma^2 = V(X) = \frac{(10-5)^2}{12} = \frac{25}{12} \text{ (minutes)}^2$$

Normal Distribution

The normal distribution was first discovered by **De – Moivre** and **Laplace** as the limiting form of Binomial distribution. Through a historical error it was credited to Gauss who first made reference to it as the distribution of errors in Astromy. Gauss used the normal curve to describe theory of accidental errors of measurements involved in the calculation of orbits of heavenly bodies.

Definition: A c. r. v. X is said to have a **normal distribution** with parameters μ and σ^2 if its p. d. f. is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right\}, -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$$

The c. d. f. of X is given by

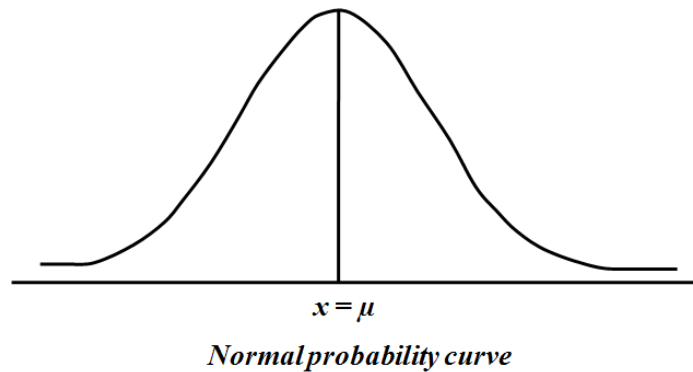
$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp \left\{ -\frac{1}{2} \left(\frac{t-\mu}{\sigma} \right)^2 \right\} dt$$

Notation: $X \sim N(\mu, \sigma^2)$. Read as X follows normal distribution with parameters μ and σ^2 .

Note:

1. The graph of $f(x)$ is famous ***bell – shaped*** curve and is symmetric about the line $X = \mu$. The top of the bell is directly above μ . For large values of σ , the curve tends

to flatten out and for small values of σ , it has a sharp peak. The curve of $f(x)$ is given below.



2. Whenever the random variable is continuous and the probabilities of it are increasing and then decreasing, in such cases we can think of using normal distribution.

Real life examples:

- 1) The heights of students.
- 2) The weights of students.
- 3) The diameters of bolts manufactured.
- 4) The lives of electrical bulbs manufactured.

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

Standard Normal distribution

If $X \sim N(\mu, \sigma^2)$ then $Z = \frac{X - \mu}{\sigma}$ is known as **standard normal distribution** with mean $E(Z) = 0$, with variance $V(Z) = 1$ and we write $Z \sim N(0, 1)$. Its p. d. f. is given by

$$g(z) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}z^2\right), -\infty < z < \infty$$

and its c. d. f. is given by

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^z g(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left\{-\frac{1}{2}t^2\right\}dt$$

Area Property of Normal Distribution

$$\text{If } X \sim N(\mu, \sigma^2), \text{ then } P(\mu < X < x_1) = \int_{\mu}^{x_1} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{x_1} \exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^2\right\}dx$$

Let $Z = \frac{X-\mu}{\sigma}$. Then $X - \mu = \sigma Z$.

If $X = \mu$, then $Z = 0$. If $X = x_1$, then $Z = \frac{x_1-\mu}{\sigma} = z_1$ (say).

$$\therefore P(\mu < X < x_1) = P(0 < Z < z_1) = \int_0^{z_1} g(z)dz = \frac{1}{\sqrt{2\pi}} \int_0^{z_1} \exp\left\{-\frac{1}{2}z^2\right\}dz$$

where $g(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$ is the p. d. f. of standard normal variate. The definite integral

$\int_0^{z_1} g(z)dz$ is known as **normal probability integral** and gives the area under standard normal curve between the ordinates at $z = 0$ and $z = z_1$. These areas have been tabulated for different values of z_1 at intervals of 0.01 in the table given at the **end of the module**.

In particular, the probability that the random variable X lies in the interval $(\mu - \sigma, \mu + \sigma)$ is given by

$$P(\mu - \sigma < X < \mu + \sigma) = P(-1 < Z < 1)$$

$$= \int_{-1}^1 g(z)dz$$

$$= 2 \int_0^1 g(z)dz \quad (\text{by symmetry})$$

$$= 2 \times 0.3413 \quad (\text{from table})$$

$$\Rightarrow P(\mu - \sigma < X < \mu + \sigma) = 0.6826$$

$$\text{Similarly, } P(\mu - 2\sigma < X < \mu + 2\sigma) = P(-2 < Z < 2)$$

$$= 2 \times P(0 < Z < 2) = 2 \times 0.4772 (\text{see table})$$

$$\Rightarrow P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9544$$

$$\text{and } P(\mu - 3\sigma < X < \mu + 3\sigma) = P(-3 < Z < 3)$$

$$= 2 \times P(0 < Z < 3)$$

$$= 2 \times 0.49865 \quad (\text{see table})$$

$$\Rightarrow P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.9973$$

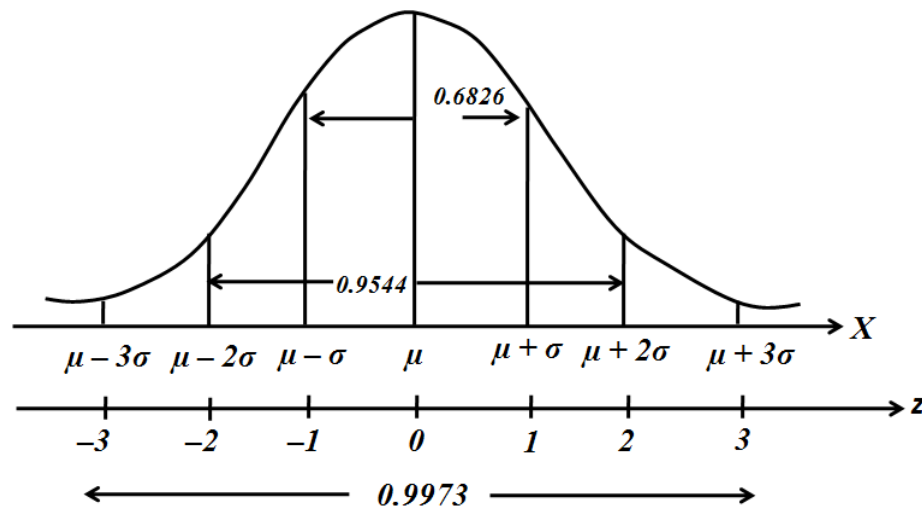
Thus, the probability that a normal variate X lies outside the range $\mu \pm 3\sigma$ is given by

$$P(|x - \mu| > 3\sigma) = P(|Z| > 3) = 1 - P(|Z| \leq 3)$$

$$= 1 - P(-3 \leq Z \leq 3) = 1 - 0.9973 = 0.0027$$

Thus, in all probability, we should expect a normal variate to lie within the range $\mu \pm 3\sigma$, though theoretically, it may range from $-\infty$ to ∞ .

The probabilities computed above are exhibited in the following figure.



Note: The Gamma function defined below is used to evaluate mean and variance of the normal distribution.

$$\Gamma(n) = \int_0^{\infty} e^{-x} x^{n-1} dx \text{ for } n > 0$$

$$\Gamma(n+1) = n \Gamma(n)$$

$$\Gamma(n+1) = n! , \text{ where } n \text{ is a positive integer.}$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Mean of Normal distribution

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Let $z = \frac{x-\mu}{\sigma}$. Then $x = \mu + \sigma z$, $dx = \sigma dz$ and

$$\begin{aligned} E(X) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma z) e^{-\frac{1}{2}z^2} dz \\ &= \mu \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz = \mu \times 1 + 0 \end{aligned}$$

Note that the integral in first term is 1 since total probability is one and the integral in the second term is zero since the integral is an odd function.

Therefore, Mean = $E(X) = \mu$

Variance of Normal distribution

$$V(X) = E(X - E(X))^2 = E(X - \mu)^2 \quad (\because E(X) = \mu)$$

$$= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x - \mu)^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Let $\frac{x-\mu}{\sigma} = z$. Then $x - \mu = \sigma z$, $dx = \sigma dz$ and

$$V(X) = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{1}{2}z^2} dz = \frac{2\sigma^2}{\sqrt{2\pi}} \int_0^{\infty} z^2 e^{-\frac{1}{2}z^2} dz$$

(Since the integrand is an even function)

Let $\frac{1}{2}z^2 = t \Rightarrow z = \sqrt{2t}$ and $dz = \frac{dt}{\sqrt{2t}}$. Then

$$\begin{aligned} V(X) &= \frac{2\sigma^2}{\sqrt{2\pi}} \int_0^{\infty} 2t \cdot e^{-t} \frac{dt}{\sqrt{2t}} = \frac{2\sigma^2}{\sqrt{\pi}} \int_0^{\infty} e^{-t} t^{\frac{3}{2}-1} dt \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) \quad (\text{Gamma function}) \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{1}{2} \cdot \Gamma\left(\frac{1}{2}\right) = \frac{\sigma^2}{\sqrt{\pi}} \cdot \sqrt{\pi} = \sigma^2 \\ &\Rightarrow V(X) = \sigma^2 \end{aligned}$$

Note: Standard deviation $= \sqrt{V(X)} = \sqrt{\sigma^2} = \sigma$

Example 2: If X is normally distributed with mean 12 and standard deviation , then

(a) Find the probabilities of the following :

- (i) $X \geq 20$**
- (ii) $X \leq 20$ and**
- (iii) $0 \leq X \leq 12$**

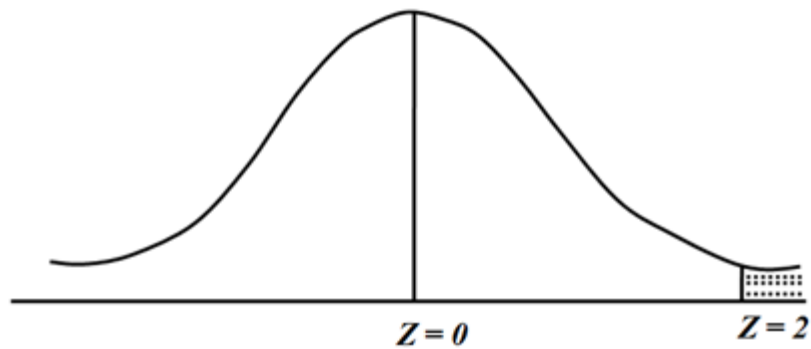
(b) Find x when $P(X > x) = 0.24$

(c) Find x_1 and x_2 when $P(x_1 < X < x_2) = 0.5$ and $P(X > x_2) = 0.25$

Solution:

(a) it is given that $\mu = 12$ and $\sigma = 4$ i. e., $X \sim N(12, 16)$

$$\begin{aligned} \text{(i)} \quad \text{Let } Z &= \frac{X-12}{4} \text{ . then } P(X \geq 20) = P\left(\frac{X-12}{4} \geq \frac{20-12}{4}\right) \\ &= P(Z \geq 2) = 0.5 - P(0 \leq Z \leq 2) \\ &= 0.5 - 0.4772 \quad (\text{from table}) \\ &= 0.0228 \end{aligned}$$



$$(ii) \quad P(X \leq 20) = 1 - P(X \geq 20) = 1 - 0.0228 = 0.9722$$

$$(iii) \quad P(0 \leq X \leq 12) = P\left(\frac{0-12}{4} \leq \frac{X-12}{4} \leq \frac{12-12}{4}\right)$$

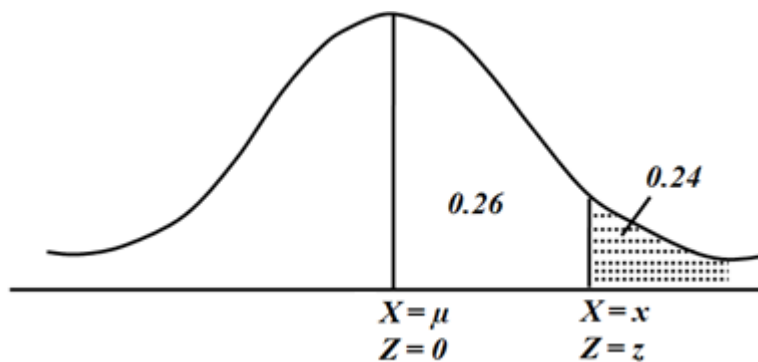
$$= P(-3 \leq Z \leq 0)$$

$$= P(0 \leq Z \leq 3) \quad (\text{by symmetry})$$

$$= 0.4986 \quad (\text{from table})$$

$$(b) \quad P(X > x) = 0.24$$

$$\Rightarrow P\left(\frac{X-12}{4} > \frac{x-12}{4}\right) = 0.24 \Rightarrow P(Z > z) = 0.24, \text{ where } z = \frac{x-12}{4}$$



$$\Rightarrow P(0 < Z < z) = 0.5 - 0.24 - 0.26$$

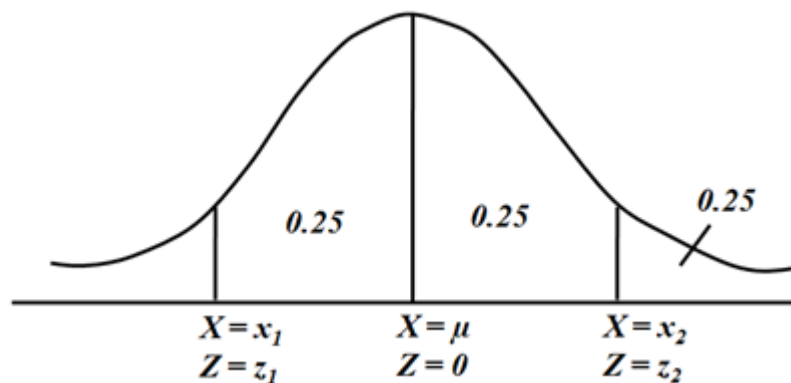
$\therefore$ From normal tables, corresponding to probability 0.26, value of $z = 0.71$ (approximately)

Hence $0.71 = z = \frac{x-12}{4} \Rightarrow x = 0.71 \times 4 + 12 = 14.84$

(c) We are given that $P(x_1 < X < x_2) = 0.5$ and $P(X > x_2) = 0.25$

$$\Rightarrow P\left(\frac{x_1-12}{4} < \frac{X-12}{4} < \frac{x_2-12}{4}\right) = 0.5 \text{ and } P\left(\frac{X-12}{4} > \frac{x_2-12}{4}\right) = 0.25$$

$$\Rightarrow P(z_1 < Z < z_2) = 0.5 \text{ and } P(Z > z_2) = 0.25, \text{ where } z_1 = \frac{x_1-12}{4} \text{ and } z_2 = \frac{x_2-12}{4}$$



By symmetry of normal curve, $z_1 = -z_2$. Find z_2 such that $P(0 < Z < z_2) = 0.25$

Corresponding to probability 0.25 from the normal table, we have $z_2 = 0.67$ approximately. Thus

$$\frac{x_2-12}{4} = 0.67 \Rightarrow x_2 = 12 + 4 \times 0.67 = 14.68$$

$$\text{Similarly, } z_1 = -z_2 \Rightarrow \frac{x_1-12}{4} = -0.67 \Rightarrow x_1 = 12 - 4 \times 0.67 = 9.32$$

Example 3: The local authorities in a certain city install 10,000 electric lamps in the streets of the city. If these lamps have an average life of 1,000 burning hours with a standard deviation of 200 hours, assuming normality, what number of lamps might be expected to fail

- (i) in the first 800 and 1200 burning hours?
- (ii) between 800 and 1200 burning hours?

After what period of burning hours would you expect that

- (a) 10% of the lamps would fail?
- (b) 10% of the lamps would be still burning?

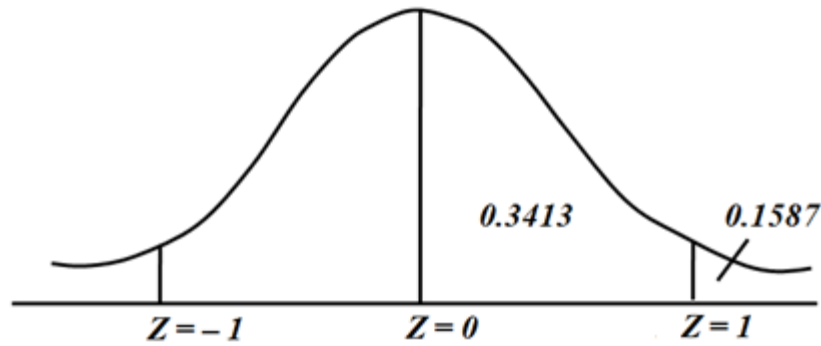
Solution:

Let X denote the life of a bulb in burning hours. Here $\mu = 1000$, $\sigma = 200$ and $X \sim N(1000, 40000)$

$$\begin{aligned} \text{(i)} \quad \text{Find } P(X < 800) &= P\left(\frac{X-1000}{200} < \frac{800-1000}{200}\right) \\ &= P(Z < -1), \text{ where } Z = \frac{X-1000}{200} \sim N(0, 1) \\ &= P(Z > 1) = 0.5 - P(0 < Z < 1) \\ &= 0.5 - 0.3413 = 0.1587 \end{aligned}$$

$\therefore$ Out of 10,000 bulbs, number of bulbs which fail in the first 800 hours is
 $10,000 \times 0.1587 = 1,587.$

$$\begin{aligned} \text{(ii)} \quad \text{Find } P(800 < X < 1200) &= P\left(\frac{800-1000}{200} < \frac{X-1000}{200} < \frac{1200-1000}{200}\right) \\ &= P(-1 < Z < 1) = 2 \cdot P(0 < Z < 1) \\ &= 2 \times 0.3413 = 0.6826 \end{aligned}$$



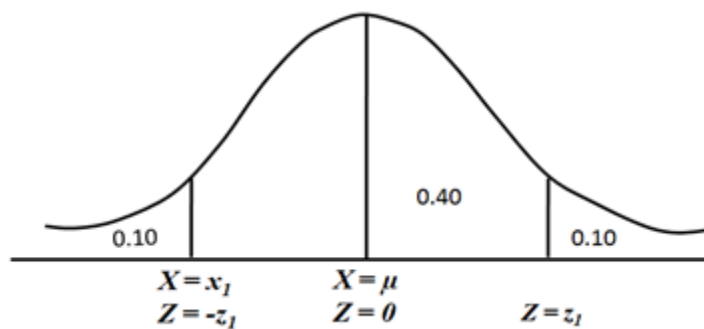
Hence, the expected number of bulbs with life between 800 and 1200 hours of burning life is $10,000 \times 0.6826 = 6,826$.

(a) Let 10% of the bulbs fail after x_1 hours of burning life. Then we have to find x_1 such that

$$\begin{aligned}
 P(X < x_1) &= 0.10 \Rightarrow P\left(\frac{X-1000}{200} < \frac{x_1-1000}{200}\right) = 0.10 \\
 &\Rightarrow P(Z < -z_1) = 0.10, \text{ where } z_1 = -\left(\frac{x_1-1000}{200}\right) \\
 &\Rightarrow P(Z > z_1) = 0.10 \\
 &\Rightarrow P(0 < Z < z_1) = 0.5 - 0.10 = 0.40
 \end{aligned}$$

From table corresponding to probability 0.40, we have

$$\begin{aligned}
 z_1 &= 1.28 \Rightarrow -\left(\frac{x_1-1000}{200}\right) = 1.28 \\
 &\Rightarrow x_1 = 1000 - 1.28 \times 200 = 1000 - 256 = 744.
 \end{aligned}$$



Thus, after 744 hours of burning life, 10% of the bulbs will fail.

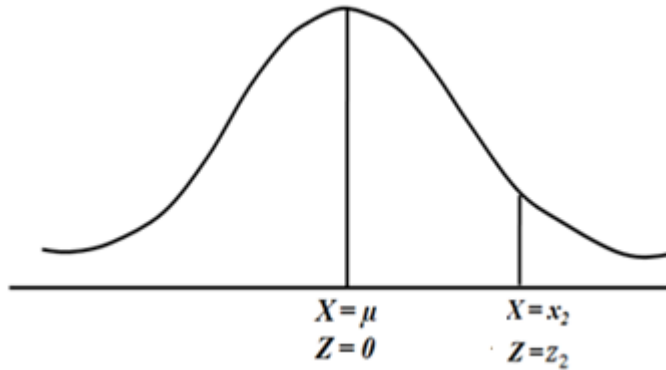
(b) Let 10% of the bulbs be still burning after x_2 hours of burning life. Then we have

$$P(X > x_2) = 0.10 \Rightarrow P\left(\frac{X-1000}{200} > \frac{x_2-1000}{200}\right) = 0.10$$

$$\Rightarrow P(Z > z_2) = 0.10, \text{ where } z_2 = \frac{x_2 - 1000}{200}$$

From normal tables, $z_2 = 1.28$ and hence

$$\frac{x_2 - 1000}{200} = 1.28 \Rightarrow x_2 = 1000 + 1.28 \times 200 = 1000 + 256 = 1256$$



Hence, after 1,256 hours of burning life, 10% of the bulbs will be still burning.

De Moivre-Laplace Theorem (Normal Approximation to Binomial Distribution)

Let $X \sim B(n, p)$. Then its p.m.f. is given by $p(x) = \binom{n}{x} p^x q^{n-x}$ for $x = 0, 1, 2, \dots, n$. The mean and variance of X are given by $\mu = np$ and $\sigma^2 = npq$ respectively. Now,

$$P(k_1 \leq X \leq k_2) = \sum_{x=k_1}^{k_2} \binom{n}{x} p^x q^{n-x} \text{ for some non-negative integers } k_1 \text{ and } k_2 \text{ such that}$$

$k_1 < k_2$. Since the binomial coefficient $\binom{n}{x}$ grows quite rapidly with n , it is very difficult to compute $P(k_1 \leq X \leq k_2)$ for large n . In this context, normal approximation to binomial distribution is extremely useful.

Let $Z = \frac{X - \mu}{\sigma} = \frac{X - np}{\sqrt{npq}}$. If n is large with neither p nor q close to zero, the binomial distribution can be approximated by the standard normal distribution. Thus,

$$\lim_{n \rightarrow \infty} P(k_1 \leq X \leq k_2) = \lim_{n \rightarrow \infty} P\left(\frac{k_1 - np}{\sqrt{npq}} \leq Z \leq \frac{k_2 - np}{\sqrt{npq}}\right) = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{1}{2}z^2} dz$$

$$\text{where } z_1 = \frac{k_1 - np}{\sqrt{npq}} \text{ and } z_2 = \frac{k_2 - np}{\sqrt{npq}}$$

This is a very good approximation when both np and npq are greater than 5.

Example 4: A coin is tossed 10 times. Find the probability of getting between 4 and 7 heads inclusive using the (a) binomial distribution and (b) the normal approximation to the binomial distribution.

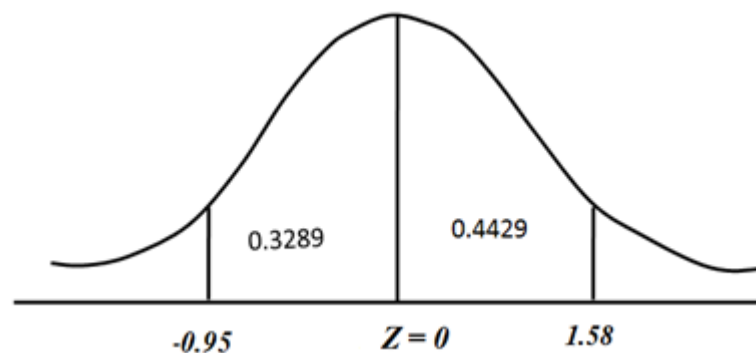
Solution:

(a) Let X denote the number of heads in 10 tosses. Then $X \sim B\left(10, \frac{1}{2}\right)$ and $\mu = np = 5$ and $\sigma^2 = npq = 2.5$ and

$$\begin{aligned} P(4 \leq X \leq 7) &= \sum_{x=4}^7 p(x) = \sum_{x=4}^7 \binom{n}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x} = \sum_{x=4}^7 \binom{n}{x} \left(\frac{1}{2}\right)^{10} \\ &= \frac{\binom{10}{4} + \binom{10}{5} + \binom{10}{6} + \binom{10}{7}}{1024} = \frac{792}{1024} = 0.7734 \end{aligned}$$

(b) The discrete binomial probability distribution is approximated to continuous normal probability distribution. The integers 4,5,6,7 lie in the interval (3.5 to 7.5). Thus,

$$\begin{aligned} P(4 \leq X \leq 7) &= P(3.5 \leq X \leq 7.5) = P\left(\frac{3.5-5}{\sqrt{2.5}} \leq Z \leq \frac{7.5-5}{\sqrt{2.5}}\right) \\ &= P(-0.95 \leq Z \leq 1.58) = P(-0.95 \leq Z \leq 0) + P(0 \leq Z \leq 1.58) \\ &= P(0 \leq Z \leq 0.95) + P(0 \leq Z \leq 1.58) \\ &= 0.3289 + 0.4429 = 0.7718 \end{aligned}$$



Exponential distribution: A c.r.v. X is said to follow **exponential distribution** with parameter λ if its p.m.f. is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases}$$

The c.d.f. is given by

$$F(x) = P(X \leq x) = \int_0^x f(t) dt = \lambda \int_0^x e^{-\lambda t} dt = \lambda \left[\frac{e^{-\lambda t}}{-\lambda} \right]_0^x = 1 - e^{-\lambda x}$$

$$\Rightarrow F(x) = 1 - e^{-\lambda x}$$

Notation: $X \sim E(\lambda)$. Read as X follows exponential distribution with parameter λ .

Real life examples of exponential distribution

1. The time taken to serve a customer at a petrol pump, railway booking counter or any other service facility.
2. The period of time for which an electronic component operates without any breakdown.
3. The time between two successive arrivals at any service facility.

Mean and Variance of exponential distribution

For $r \geq 1$, $E(X^r) = \int_0^\infty x^r f(x) dx = \lambda \int_0^\infty x^r e^{-\lambda x} dx$

Let $\lambda x = t$. Then t varies between 0 to ∞ and $dx = \frac{dt}{\lambda}$. Then

$$E(X^r) = \lambda \int_0^\infty \left(\frac{t}{\lambda} \right)^r e^{-t} \frac{dt}{\lambda} = \frac{1}{\lambda^r} \int_0^\infty e^{-t} t^{(r+1)-1} dt$$

$$\Rightarrow E(X^r) = \frac{\Gamma(r+1)}{\lambda^r} = \frac{r!}{\lambda^r} \text{ (using Gamma function)}$$

Thus, mean $= \mu = E(X) = \frac{1}{\lambda}$ and $E(X^2) = \frac{2!}{\lambda^2} = \frac{2}{\lambda^2}$

Hence $\sigma^2 = V(X) = E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$

Therefore, $\mu = \frac{1}{\lambda}$ and $\sigma^2 = \frac{1}{\lambda^2}$.

Example 5: Assume that the length of phone calls made at a particular telephone booth is exponentially distributed with a mean of 3 minutes. If you arrive at the telephone booth just as Ramu was about to make a call, find the following:

- The probability that you will wait more than 5 minutes before Ramu is done with the call.
- The probability that Ramu's call will last between 2 minutes and 6 minutes.

Solution: Let X be a r.v. that denotes the length of calls made at the telephone booth. Since the mean length of calls $\frac{1}{\lambda} = 3$, the p.d.f. is given by

$$f(x) = \frac{1}{3} e^{-\frac{x}{3}}$$

$$\text{a. } P(X > 5) = \int_5^{\infty} f(x) dx = \frac{1}{3} \int_5^{\infty} e^{-\frac{x}{3}} dx = \left[e^{-\frac{x}{3}} \right]_5^{\infty} = e^{-\frac{5}{3}}$$

$$\text{b. } P[2 \leq X \leq 6] = \int_2^6 f(x) dx = \frac{1}{3} \int_2^6 e^{-\frac{x}{3}} dx = \left[-e^{-\frac{x}{3}} \right]_2^6 = e^{-\frac{2}{3}} - e^{-2}$$

Memory lessness property of exponential distribution

The exponential distribution is used extensively in reliability engineering to model the lifetimes of systems. Suppose the life X of an equipment is exponentially distributed with a mean of $\frac{1}{\lambda}$. Assume that the equipment has not failed by time t . We want to find the probability that $X \leq t + s$ given that $X > t$ for some nonnegative additional time s .

Thus,

$$\begin{aligned} P(X \leq s + t | X > t) &= \frac{P(X \leq s + t, X > t)}{P(X > t)} = \frac{P(t < X \leq s + t)}{P(X > t)} = \frac{F(s + t) - F(t)}{1 - F(t)} \\ &= \frac{(1 - e^{-\lambda(s + t)}) - (1 - e^{-\lambda t})}{e^{-\lambda t}} = \frac{e^{-\lambda t} - e^{-\lambda(s + t)}}{e^{-\lambda t}} = 1 - e^{-\lambda s} = F(s) = P(X \leq s) \end{aligned}$$

$$\Rightarrow P(X \leq s + t | X > t) = P(X \leq s)$$

This indicates that the process only remembers the present and not the past.

Example 6: In example 5, Ramu, who is using the phone at the telephone booth, had already talked for 2 minutes before you arrived. According to the memory lessness property of the exponential distribution, the mean time until Ramu is done with the call is still 3 minutes. The random variable forgets the length of time the call had lasted before you arrived.

Relationship between exponential and Poisson distributions

Let λ denote the mean number of arrivals per unit of time, say per second. Then the mean number of arrivals in t seconds is λt .

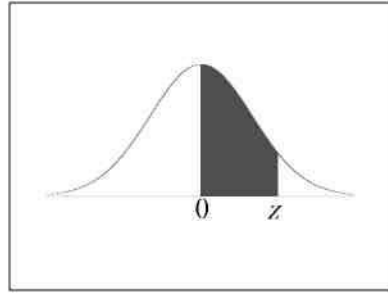
Let X denote the number of arrivals during an interval of t seconds.

Let Y denote the time between two successive arrivals.

If $X \sim P(\lambda t)$ i.e., $p(x) = P(X = x) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}$ for $x = 0, 1, 2, \dots$; $t \geq 0$, then

$Y \sim E(\lambda)$ i.e., $f(x) = \lambda e^{-\lambda t}$.

Standard Normal Distribution Table



<i>z</i>	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952

[illegible]

P1:

The weights in pounds of parcels arriving at a package delivery company's warehouse can be modeled by an $N(5, 16)$ normal random variable X .

- a) What is the probability that a randomly selected parcel weighs between 1 and 10 pounds?
- b) What is the probability that a randomly selected parcel weighs more than 9 pounds?

Solution:

Since $X \sim N(5, 16)$, we have $\mu = 5$ and $\sigma^2 = 16$.

$$\begin{aligned} \text{a) } P(1 < X < 10) &= P\left(\frac{1-5}{4} < Z < \frac{10-5}{4}\right) = P(-1 < Z < 1.25) \\ &= P(-1 < Z < 0) + P(0 < Z < 1.25) \\ &= P(0 < Z < 1) + P(0 < Z < 1.25) \\ &= 0.3413 + 0.3943 = 0.7356 \quad (\text{Use table}) \end{aligned}$$

$$\begin{aligned} \text{b) } P(X > 9) &= P\left(Z > \frac{9-5}{4}\right) = P(Z > 1) \\ &= 0.5 - P(0 < Z < 1) \\ &= 0.5 - 0.3413 \\ &= 0.1587 \quad (\text{Use table}) \end{aligned}$$

P2:

The marks obtained by a number of students for a certain subject are assumed to be approximately normally distributed with mean value 65 and with a standard deviation of 5. If 3 students are selected at random from this set, what is the probability that exactly 2 of them will have marks over 70?

Solution:

Let X denotes marks obtained by a student in a certain subject. Then $X \sim N(65, 25)$ and

$$\begin{aligned} p(x) = P(X > 70) &= P\left(Z > \frac{70-65}{5}\right) = P(Z > 1) \\ &= 0.5 - P(0 < Z < 1) \\ &= 0.5 - 0.3413 \quad \text{(Use table)} \\ &= 0.1587 \end{aligned}$$

Let Y denote the number of students who got more than 70 marks in a sample of 3 students.

Then $Y \sim B(3, p)$ and $P(Y = 2) = \binom{3}{2} (0.1587)^2 (0.8413)$

P3:

The time taken by a person while speaking over a telephone is exponentially distributed with mean 4 *minutes*.

- i) Find the probability that he speaks for more than 6 *minutes* but less than 7 *minutes*.
- ii) Out of 6 calls that he makes, what is the probability that exactly 2 calls take him more than 3 minutes each.
- iii) How many calls out of 100 are expected to take more than 3 minutes each?

Solution:

Let X be the time taken (in minutes) per call. We are given that X is exponentially distributed with mean 4 *minutes*.

$$\therefore f(x) = \begin{cases} \frac{1}{4}e^{-\frac{x}{4}} & , \quad x > 0 \\ 0 & , \quad \text{otherwise} \end{cases}$$

$$(i) \quad P(6 < X < 7) = \frac{1}{4} \int_6^7 e^{-\frac{x}{4}} dx = \frac{1}{4} \left[\frac{e^{-\frac{x}{4}}}{-\frac{1}{4}} \right]_6^7 = e^{-\frac{6}{4}} - e^{-\frac{7}{4}} (= 0.04936)$$

$$(ii) \quad P(X > 3) = \frac{1}{4} \int_3^\infty e^{-\frac{x}{4}} dx = \frac{1}{4} \left[\frac{e^{-\frac{x}{4}}}{-\frac{1}{4}} \right]_3^\infty = e^{-\frac{3}{4}} (= 0.4724)$$

Let Y denote the number of calls each with more than 3 minutes out of 6 calls. Then

$$Y \sim B(6, P) \text{ where } P = 0.4724$$

$$\therefore P(Y = 2) = \binom{6}{2} (0.4764)^2 (0.5276)^4 = 0.2594$$

- (iii) Expected number of calls out of 100 that will be longer than 3 minutes each
= $100 \times P(X > 3)$
= $100 \times 0.4724 = 47.24 = 47$ (approximately)

P4:

The mileage (in thousands of miles) which car owners get with a certain kind of tyres is a random variable having probability density function

$$f(x) = \begin{cases} \frac{1}{10} e^{-\frac{x}{10}} & , \quad x > 0 \\ 0 & , \quad \textit{otherwise} \end{cases}$$

Find the probability that one of these tyres will last

- i) at most 5,000 miles.
- ii) anywhere from 8,000 to 12,000 miles.

Solution:

Let X represents the mileage in thousands of miles

- i) $P(X \leq 5) = F(5) = 1 - e^{-\frac{5}{10}} = 1 - e^{-\frac{1}{2}} = 1 - 0.6065 = 0.3935$
- ii) $P(8 < X < 12) = F(12) - F(8) = e^{-0.8} - e^{-1.2} = 0.148$

2.5. Continuous Probability Distributions

Exercise:

- 1) X is a normal variate with mean 30 and standard deviation 5. Find the probabilities that
(i) $26 \leq X \leq 40$ (ii) $X \geq 45$ (iii) $|X - 30| > 5$
- 2) There are 600 engineering students in the B.Tech. classes of a university and the probability for any student to need a copy of a particular book from the university library on any day is 0.05. How many copies of the book should be kept in the university library so that the probability may be greater than 0.90 that none of the students needing a copy from the library has to come back disappointed? (use normal approximation to the binomial distribution)
- 3) In a distribution exactly normal, 10.03% of the items are under $25kg$ weight and 89.97% of the items are under $70kg$ weight. What are the mean and standard deviation of the distribution?
- 4) In an examination it is laid down that a student passes if he secures 30 percent or more marks. He is placed in the first, second or third division according as he secures 60% or more marks, between 45% and 60% marks and marks between 30% and 45% respectively. He gets distinction in case he secures 80% or more marks. It is noticed from the result that 10% of the students failed in the examination, where as 5% of them obtained distinction. Calculate the percentage of students placed in the examination. (Assume normal distribution to marks)
- 5) A sample of 100 items is taken at random from a batch known to contain 40% defectives. What is the probability that the sample contains (i) at least 44 defectives and (ii) exactly 44 defectives? (use normal approximation to the binomial distribution)

- 6) If X is uniformly distributed with mean 1 and variance $\frac{4}{3}$, find $P(X < 0)$.
- 7) Subway trains on a certain line run every half hour between mid-night and six in the morning. What is the probability that a man entering the station at a random time during this period will have to wait at least twenty minutes?
- 8) Suppose the life time of an electric component has exponential distribution with a mean life of 500 *hrs*.
- Find the probability that it will give additional 600 *hrs* life given that the component has been working for the last 300*hrs*.
 - Find the probability that it will work for more than 600*hrs*.
- 9) The life(in hours) of electronic tubes manufactured by a certain process is known to have p.d.f.

$$f(x) = \begin{cases} \frac{1}{400} e^{-\frac{1}{400}(x-400)} & , \quad x \geq 400 \\ 0 & , \quad \text{otherwise} \end{cases}$$

Find the probability of one such tube lasting for

- at least 800*hrs*.
- at most 1200*hrs*.
- anywhere between 500 and 800*hrs*.

Answers:

1) (i) 0.7653 (ii) 0.00135 (iii) 0.3174

2) 37

3) $\mu = 47.5$ and $\sigma = 17.578$

4) 34%

5) (i) 0.2376 (ii) 0.0584

6) $\frac{1}{4}$

7) $\frac{1}{3}$

8) (i) $e^{-\frac{6}{5}}$ (ii) $e^{-\frac{6}{5}}$

9) (i) e^{-1} (ii) e^{-2} (iii) $e^{-\frac{1}{4}} - e^{-1}$

2.2

Bivariate random variable

In the real life situations more than one variable effects the outcome of a random experiment. For example, consider an electronic system consisting of two components. Suppose the system will fail if both the components fail. The probability distribution of the life of the system depends jointly on the probability distributions of lives of the components. Knowing the probability distributions of lives of the components will not provide us the enough information. What we need is the probability distribution of the simultaneous behavior of lives of the components. A pair of random variables is known as a **bivariate random variable**. The individual random variables in the pair may be related.

Bivariate random variable: let S be the sample space associated with a random experiment. Let R be the real line. If $(X, Y): S \rightarrow R \times R$, i.e., $(X, Y)(\omega) = (X(\omega), Y(\omega)) \quad \forall \omega \in S$, then the pair (X, Y) is known as a bivariate random variable.

Note:

1. If X and Y are both discrete random variables, then (X, Y) is a bivariate discrete random variable.
2. If X and Y are both continuous random variables, then (X, Y) is a bivariate continuous random variable.

Joint probability mass function: Let (X, Y) be a bivariate discrete random variable, which takes the values (x_i, y_j) for $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n$. Let

$$p(x_i, y_j) = P(X = x_i, Y = y_j) \quad \forall i \text{ and } j$$

Then $p(x_i, y_j) \geq 0 \quad \forall i \text{ and } j$ and $\sum_{i=1}^m \sum_{j=1}^n p(x_i, y_j) = 1$. The function $p(x, y)$ is

known as **joint probability mass function** (j.p.m.f) of (X, Y) .

Marginal probability mass functions: Let (X, Y) be a bivariate discrete random variable with joint probability mass function given by $p(x, y)$. The marginal probability mass function of X and Y are given by

$$p_1(x_i) = \sum_{j=1}^n p(x_i, y_j) \text{ for } i = 1, 2, 3, \dots, m \text{ and}$$

$$p_2(y_j) = \sum_{i=1}^m p(x_i, y_j) \text{ for } j = 1, 2, 3, \dots, n$$

respectively.

Note: X and Y are independent if and only if $p(x_i, y_j) = p_1(x_i)p_2(y_j) \forall (i, j)$

Conditional probability mass functions: Let (X, Y) be a bivariate discrete random variable with joint probability mass function given by $p(x, y)$. The conditional probability mass function of X given $y = y_j$ and the conditional probability mass function of Y given $x = x_i$ are given by

$$p_{1|2}(x_i | y_j) = \frac{p(x_i, y_j)}{p_2(y_j)} \text{ for } i = 1, 2, 3, \dots, m \text{ and}$$

$$p_{2|1}(y_j | x_i) = \frac{p(x_i, y_j)}{p_1(x_i)} \text{ } j = 1, 2, 3, \dots, n$$

respectively.

Example 1: A fair coin is tossed three times. Let X be a random variable that takes the value 0 if the first toss is a tail and the value 1 if the first toss is a head and Y be a random variable that defines the total number of heads in the three tosses. Then

- i. Determine the joint, marginal and conditional mass functions of X and Y .
- ii. Are X and Y independent?

Solution:

i. The sample space and values of X and Y are given in the following table:

Out comes in sample space	Value of X	Value of Y
HHH	1	3
HHT	1	2
HTH	1	2
HTT	1	1
THH	0	2
THT	0	1
TTH	0	1
TTT	0	0

Here X takes the values 0 and 1 and Y takes the values 0, 1, 2 and 3. Then the j.p.m.f of (x, y) is computed as below:

$$p(0, 0) = P(X = 0, Y = 0) = P(\{TTT\}) = \frac{1}{8}$$

$$p(0, 1) = P(X = 0, Y = 1) = P(\{THT, TTH\}) = \frac{2}{8} = \frac{1}{4}$$

$$p(0, 2) = P(X = 0, Y = 2) = P(\{THH\}) = \frac{1}{8}$$

$$p(0, 3) = P(X = 0, Y = 3) = 0$$

$$p(1, 0) = P(X = 1, Y = 0) = 0$$

$$p(1, 1) = P(X = 1, Y = 1) = P(\{HTT\}) = \frac{1}{8}$$

$$p(1, 2) = P(X = 1, Y = 2) = P(\{HTH, HHT\}) = \frac{2}{8} = \frac{1}{4}$$

$$p(1, 3) = P(X = 1, Y = 3) = P(\{HHH\}) = \frac{1}{8}$$

The m.p.m.f of X is given by

$$p_1(0) = p(0, 0) + p(0, 1) + p(0, 2) + p(0, 3) = \frac{1}{8} + \frac{2}{8} + \frac{1}{8} + 0 = \frac{1}{2} \text{ and}$$

$$p_1(1) = p(1,0) + p(1,1) + p(1,2) + p(1,3) = 0 + \frac{1}{8} + \frac{2}{8} + \frac{1}{8} = \frac{1}{2}$$

The m.p.m.f of Y is given by

$$p_2(0) = p(0,0) + p(1,0) = \frac{1}{8} + 0 = \frac{1}{8}$$

$$p_2(1) = p(0,1) + p(1,1) = \frac{2}{8} + \frac{1}{8} = \frac{3}{8}$$

$$p_2(2) = p(0,2) + p(1,2) = \frac{1}{8} + \frac{2}{8} = \frac{3}{8}$$

$$p_2(3) = p(0,3) + p(1,3) = 0 + \frac{1}{8} = \frac{1}{8}$$

The conditional p.m.f of X given Y is computed below:

$$p_{1|2}(0|0) = \frac{p(0,0)}{p_2(0)} = \frac{\frac{1}{8}}{\frac{1}{8}} = 1, \quad p_{1|2}(1|0) = \frac{p(1,0)}{p_2(0)} = \frac{0}{\frac{1}{8}} = 0$$

$$p_{1|2}(0|1) = \frac{p(0,1)}{p_2(1)} = \frac{\frac{2}{8}}{\frac{3}{8}} = \frac{2}{3}, \quad p_{1|2}(1|1) = \frac{p(1,1)}{p_2(1)} = \frac{\frac{1}{8}}{\frac{3}{8}} = \frac{1}{3}$$

$$p_{1|2}(0|2) = \frac{p(0,2)}{p_2(2)} = \frac{\frac{1}{8}}{\frac{3}{8}} = \frac{1}{3}, \quad p_{1|2}(1|2) = \frac{p(1,2)}{p_2(2)} = \frac{\frac{2}{8}}{\frac{3}{8}} = \frac{2}{3}$$

$$p_{1|2}(0|3) = \frac{p(0,3)}{p_2(3)} = \frac{0}{\frac{1}{8}} = 0, \quad p_{1|2}(1|3) = \frac{p(1,3)}{p_2(3)} = \frac{\frac{1}{8}}{\frac{1}{8}} = 1$$

The conditional p.m.f. of Y given X is computed as below:

$$p_{2|1}(0|0) = \frac{p(0,0)}{p_1(0)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}, \quad p_{2|1}(1|0) = \frac{p(0,1)}{p_1(0)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

$$p_{2|1}(2|0) = \frac{p(0,2)}{p_1(0)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}, \quad p_{2|1}(3|0) = \frac{p(0,3)}{p_1(0)} = \frac{0}{\frac{1}{2}} = 0$$

$$p_{2|1}(0|1) = \frac{p(1,0)}{p_1(1)} = \frac{0}{\frac{1}{2}} = 0, \quad p_{2|1}(1|1) = \frac{p(1,1)}{p_1(1)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}$$

$$p_{2|1}(2|1) = \frac{p(1,2)}{p_1(1)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}, \quad p_{2|1}(3|1) = \frac{p(1,3)}{p_1(1)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}$$

- ii. Here $p(0,0) = \frac{1}{8}$, $p_1(0) = \frac{1}{2}$ and $p_2(0) = \frac{1}{8}$
 Since $p(0,0) \neq p_1(0)p_2(0)$, X and Y are not independent.

Example 2 : The j.p.m.f. of (X, Y) is given by

$$p(x,y) = \begin{cases} k(2x+y) & \text{for } x=1,2; y=1,2 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant

- Find the value of k .
- Find marginal and conditional p.m.f.s.
- Are X and Y independent.

Solution:

- a. Since $p(x,y)$ is a j.p.m.f, $\sum_{x=1}^2 \sum_{y=1}^2 p(x,y) = 1$

$$\sum_{x=1}^2 \sum_{y=1}^2 p(x,y) = k \sum_{x=1}^2 \sum_{y=1}^2 (2x+y) = k(3+4+5+6) = 18k = 1.$$

$$\text{Thus } k = \frac{1}{18}$$

- b. The m.p.m.f. of X is given by

$$p_1(x) = \sum_{y=1}^2 p(x,y) = \frac{1}{18} \sum_{y=1}^2 (2x+y) = \frac{1}{18} [(2x+1) + (2x+2)] = \frac{4x+3}{18}$$

$$\text{Thus, } p_1(x) = \frac{4x+3}{18} \text{ for } x = 1, 2.$$

The m.p.m.f of Y is given by

$$p_2(y) = \sum_{x=1}^2 p(x,y) = \frac{1}{18} \sum_{x=1}^2 (2x+y) = \frac{1}{18} [(2+y) + (4+y)] = \frac{2y+6}{18} = \frac{y+3}{9}$$

Thus, $p_2(y) = \frac{y+3}{9}$ for $y = 1, 2$

The c.p.m.f. of X given Y is given by

$$p_{1|2}(x|y) = \frac{p(x,y)}{p_2(y)} = \frac{\frac{1}{18}(2x+y)}{\frac{1}{18}(2y+6)} = \frac{2x+y}{2y+6}$$

Thus, $p_{1|2}(x|y) = \frac{2x+y}{2y+6}$ for $x = 1, 2$

The c.p.m.f. of Y given X is given by

$$p_{2|1}(y|x) = \frac{p(x,y)}{p_1(x)} = \frac{\frac{1}{18}(2x+y)}{\frac{1}{18}(4x+3)} = \frac{2x+y}{4x+3}$$

Thus, $p_{2|1}(y|x) = \frac{2x+y}{4x+3}$ for $y = 1, 2$

c. Note that $p_1(x) \cdot p_2(y) = \frac{4x+3}{18} \cdot \frac{y+3}{9} \neq p(x,y)$.

Thus, X and Y are not independent.

Joint probability density function: Let (X, Y) be a bivariate continuous random variable. Let

$$P(a \leq x \leq b, c \leq y \leq d) = \int_a^b \int_c^d f(x,y) dx dy$$

for some real numbers a, b, c, d such that $a < b$ and $c < d$. Then

i. $f(x, y) \geq 0 \forall (x, y)$ and

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

and the function $f(x, y)$ is known as the **joint probability density function** of the bivariate continuous random variable (X, Y) .

Marginal probability density function: Let (X, Y) be a bivariate continuous random variable with j.p.d.f. $f(x, y)$. The marginal probability density functions of X and Y are given by

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad \text{and} \quad f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

respectively.

Note: X and Y are independent if and only if $f(x, y) = f_1(x) \cdot f_2(y)$

Conditional probability density functions: Let (X, Y) be a bivariate continuous random variable with j.p.d.f. $f(x, y)$. Let $f_1(x)$ and $f_2(y)$ be the m.p.d.fs of X and Y respectively. The conditional probability density function of X given Y and the conditional probability density function of Y given X are given by

$$f_{1|2}(x|y) = \frac{f(x, y)}{f_2(y)} \quad \text{and}$$

$$f_{2|1}(y|x) = \frac{f(x, y)}{f_1(x)}$$

respectively.

Cumulative distribution function: The cumulative distribution of a bivariate random variable (X, Y) is defined by

$$F(x, y) = P(X \leq x, Y \leq y) \quad \text{and}$$

$$F(x,y) = \begin{cases} \sum_{t \leq x} \sum_{s \leq y} p(t,s) & \text{if } (X,Y) \text{ is a d.r.v with j.p.m.f. } p(x,y) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t,s) dt ds & \text{if } (X,Y) \text{ is a c.r.v with j.p.d.f. } f(x,y) \end{cases}$$

Properties of cumulative distribution function

1. $0 \leq F(x,y) \leq 1$
2. $F(\infty, \infty) = 1, F(-\infty, -\infty) = 0$
3. $P(a < X \leq b, Y \leq d) = F(b, d) - F(a, d)$ and
 $P(X \leq b, c < Y \leq d) = F(b, d) - F(b, c)$
4. $P(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c)$

Marginal cumulative distribution function: Let (X,Y) be a bivariate random variable with c.d.f. $F(x,y)$. The marginal cumulative distribution functions of X and Y are given by $F_1(x) = F(x, \infty)$ and $F_2(y) = F(\infty, y)$ respectively.

Note:

If (X,Y) is a bivariate continuous random variable with c.d.f. $F(x,y)$, then its j.p.d.f. is given by

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

Example 3: The j.p.d.f. of (X,Y) is given by

$$f(x,y) = \begin{cases} e^{-(x+y)} & \text{for } 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

- a. Find marginal p.d.fs of X and Y .
- b. Are X and Y independent?

Solution:

- a. The m.p.d.f of X is given by

$$f_1(x) = \int_0^{\infty} f(x,y) dy = \int_0^{\infty} e^{-(x+y)} dy = e^{-x} \int_0^{\infty} e^{-y} dy = e^{-x} \cdot 1 = e^{-x}$$

$$= \begin{cases} e^{-x} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

The m.p.d.f. of Y is given by

$$\begin{aligned} f_2(y) &= \int_0^{\infty} f(x,y) dx = \int_0^{\infty} e^{-(x+y)} dx = e^{-y} \int_0^{\infty} e^{-x} dx = e^{-y} \cdot 1 = e^{-y} \\ &= \begin{cases} e^{-y} & \text{for } y \geq 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

b. Since $f_1(x) \cdot f_2(y) = e^{-x} \cdot e^{-y} = e^{-(x+y)} = f(x,y)$, X and Y are independent.

Example 4: The j.p.d.f. of (X, Y) is given by

$$f(x,y) = \begin{cases} xe^{-x(y+1)} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

a. Determine the marginal and conditional p.d.fs

b. Are X and Y independent.

Solution: The m.p.d.f of X is given by

$$f_1(x) = \int_0^{\infty} f(x,y) dy = \int_0^{\infty} xe^{-x(y+1)} dy = xe^{-x} \int_0^{\infty} e^{-xy} dy = xe^{-x} \left[\frac{e^{-xy}}{-x} \right]_{y=0}^{y=\infty} = e^{-x}$$

$$\Rightarrow f_1(x) = e^{-x} \text{ for } 0 < x < \infty$$

The m.p.d.f. of Y is given by

$$f_2(y) = \int_0^{\infty} f(x,y) dx = \int_0^{\infty} xe^{-x(y+1)} dx = \left[\frac{x \cdot e^{-x(y+1)}}{y+1} \right]_0^{\infty} + \frac{1}{y+1} \int_0^{\infty} e^{-x(y+1)} dx$$

(using integration by parts)

$$= 0 - \frac{1}{(y+1)^2} \left[e^{-x(y+1)} \right]_0^{\infty} = \frac{1}{(y+1)^2}$$

$$\Rightarrow f_2(y) = \frac{1}{(y+1)^2} \text{ for } 0 < y < \infty$$

The conditional p.d.f of X given Y is given by

$$f_{1|2}(x|y) = \frac{f(x, y)}{f_2(y)} = \frac{x \cdot e^{-x \cdot (y+1)}}{\frac{1}{(y+1)^2}} = x(y+1)^2 e^{-x \cdot (y+1)}$$

$$\Rightarrow f_{1|2}(x|y) = x(y+1)^2 e^{-x \cdot (y+1)} \text{ for } 0 < x < \infty$$

The conditional p.d.f of Y given X is given by

$$f_{2|1}(x|y) = \frac{f(x, y)}{f_1(x)} = \frac{x \cdot e^{-x \cdot (y+1)}}{e^{-x}} = x \cdot e^{-xy}$$

$$\Rightarrow f_{2|1}(x|y) = x \cdot e^{-xy} \text{ for } 0 < x < \infty$$

Note that $f_1(x) \cdot f_2(y) = e^{-x} \cdot \frac{1}{(y+1)^2} \neq f(x, y)$. Hence, X and Y are not independent.

Example 5: The j.p.d.f of (X, Y) is given by $f(x, y) = kx^3y$ for $0 < x < 2, 0 < y < 1$.

- Find k
- Find the m.p.d.fs of X and Y
- Are X and Y independent.

Solution:

$$\begin{aligned} \text{a. We have } \int_0^2 \int_0^1 f(x, y) dx dy &= \int_0^2 \int_0^1 kx^3y dx dy = k \int_0^2 x^3 \left(\int_0^1 y dy \right) dx = k \int_0^2 x^3 \frac{1}{2} dx \\ &= \frac{k}{2} \left[\frac{x^4}{4} \right]_0^2 = \frac{k}{8} \times 16 = 2k \end{aligned}$$

$$\text{Now, } \int_0^2 \int_0^1 f(x, y) dx dy = 1 \Rightarrow 2k = 1 \Rightarrow k = \frac{1}{2}$$

The j.p.d.f of (X, Y) is given by

$$f(x, y) = \frac{1}{2}x^3y \text{ for } 0 < x < 2, 0 < y < 1$$

The m.p.d.f of X is given by

$$f_1(x) = \int_0^1 f(x, y) dy = \frac{1}{2}x^3 \int_0^1 y dy = \frac{1}{2}x^3 \left[\frac{y^2}{2} \right]_0^1 = \frac{1}{4}x^3$$

$$\Rightarrow f_1(x) = \frac{x^3}{4} \text{ for } 0 < x < 2.$$

The m.p.d.f. of Y is given by

$$f_2(y) = \int_0^2 f(x, y) dx = \frac{1}{2}y \int_0^2 x^3 dx = \frac{1}{2}y \left[\frac{x^4}{4} \right]_0^2 = 2y$$

$$\Rightarrow f_2(y) = 2y \text{ for } 0 < y < 1.$$

b. Note that $f_1(x)f_2(y) = \frac{x^3}{4} \cdot 2y = \frac{x^3y}{2} = f(x, y), \forall(x, y)$

Since $f_1(x)f_2(y) = f(x, y), \forall(x, y)$ X and Y are independent.

Dr Suguna Rao K

P1:

The c.d.f. of bivariate discrete random variable (x, y) is given by

$$F(x, y) = \begin{cases} \frac{1}{8} & \text{for } x=1, y=1 \\ \frac{5}{8} & \text{for } x=1, y=2 \\ \frac{1}{4} & \text{for } x=2, y=1 \\ 1 & \text{for } x=2, y=2 \end{cases}$$

Find a) j.p.m.f. of (x, y)

b) m.p.m.f. of x

c) m.p.m.f. of y

Solution:

The j.p.m.f. is obtained from the relationship $F(x, y) = \sum_{t \leq x} \sum_{s \leq y} p(t, s)$. Thus

$$F(1, 1) = \frac{1}{8} = p(1, 1)$$

$$F(1, 2) = p(1, 1) + p(1, 2) = \frac{5}{8} \Rightarrow p(1, 2) = \frac{5}{8} - \frac{1}{8} = \frac{1}{2}$$

$$F(2, 1) = p(1, 1) + p(2, 1) = \frac{1}{4} \Rightarrow p(2, 1) = \frac{1}{4} - \frac{1}{8} = \frac{1}{8}$$

$$F(2, 2) = p(1, 1) + p(1, 2) + p(2, 1) + p(2, 2) = 1 \Rightarrow p(2, 2) = \frac{1}{4}$$

Thus, the j.p.m.f is given by

$$p(x, y) = \begin{cases} \frac{1}{8} & \text{for } x=1, y=1 \\ \frac{1}{2} & \text{for } x=1, y=2 \\ \frac{1}{8} & \text{for } x=2, y=1 \\ \frac{1}{4} & \text{for } x=2, y=2 \end{cases}$$

The m.p.m.f of X is given by

$$p_1(x) = \begin{cases} p(1,1) + p(1,2) = \frac{5}{8} & \text{for } x=1 \\ p(2,1) + p(2,2) = \frac{3}{8} & \text{for } x=2 \end{cases}$$

The m.p.m.f of Y is given by

$$p_2(y) = \begin{cases} p(1,1) + p(2,1) = \frac{1}{4} & \text{for } y=1 \\ p(1,2) + p(2,2) = \frac{3}{4} & \text{for } y=2 \end{cases}$$

P2:

The j.p.m.f. of (X, Y) is given in the following table:

$X \backslash Y$	1	2	3	4	Total
1	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{10}{36}$
2	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{9}{36}$
3	$\frac{5}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{8}{36}$
4	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{5}{36}$	$\frac{9}{36}$
Total	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	1

Find a) The m.p.m.fs of x and y

b) Conditional p.m.f. of x given $y = 1$

c) Conditional p.m.f. of y given $x = 2$

Solution:

a) The marginal p.m.f of x is given by

$$p_1(x) = \sum_y p(x, y)$$

$$\begin{aligned}\therefore p_1(1) &= \sum_{y=1}^4 p(1, y) = p(1, 1) + p(1, 2) + p(1, 3) + p(1, 4) \\ &= \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{10}{36}\end{aligned}$$

Similarly,

$$p_1(2) = \sum_{y=1}^4 p(2, y) = \frac{9}{36}$$

$$p_1(3) = \sum_{y=1}^4 p(3, y) = \frac{8}{36} \quad \text{and}$$

$$p_1(4) = \sum_{y=1}^4 p(4, y) = \frac{9}{36}$$

Thus, the m.p.m.f. of x is given in the following table:

x	1	2	3	4
$p_1(x)$	$\frac{10}{36}$	$\frac{9}{36}$	$\frac{8}{36}$	$\frac{9}{36}$

Similarly, we can obtain the m.p.d.f. of y as given in the following table:

y	1	2	3	4
$p_2(y)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{9}{36}$

b) The conditional p.m.f. of x given $y = 1$ is given by

$$p_{1|2}(x|1) = \frac{p(x,1)}{p_2(1)} \text{ for } x = 1, 2, 3, 4.$$

$$p_{1|2}(1|1) = \frac{p(1,1)}{p_2(1)} = \frac{\frac{4}{36}}{\frac{11}{36}} = \frac{4}{11}.$$

Similarly, we can find

$$p_{1|2}(2|1) = \frac{1}{11}, p_{1|2}(3|1) = \frac{5}{11} \text{ and } p_{1|2}(4|1) = \frac{1}{11}$$

Hence, the conditional p.m.f. of x given $y = 1$ is given in the following table:

x	1	2	3	4
$p_{1 2}(x 1)$	$\frac{4}{11}$	$\frac{1}{11}$	$\frac{5}{11}$	$\frac{1}{11}$

Similarly, we can obtain the conditional p.m.f. of y given $x = 2$ as given in the following table:

y	1	2	3	4
$p_{2 1}(y 2)$	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{9}$

P3:

If X and Y are two random variables having j.p.d.f.

$$f(x, y) = \begin{cases} \frac{1}{8}(6 - x - y), & 0 < x < 2, \quad 2 < y < 4 \\ 0, & \text{otherwise} \end{cases}$$

Find (i) $P(X < 1 \cap Y < 3)$ (ii) $P(X + Y < 3)$ and (iii) $P(X < 1 | Y < 3)$.

Solution:

$$(i) P(X < 1 \cap Y < 3) = \int_0^1 \int_2^3 f(x, y) dx dy = \frac{1}{8} \int_0^1 \int_2^3 (6 - x - y) dx dy = \frac{3}{8}$$

$$(ii) P(X + Y < 3) = \frac{1}{8} \int_0^1 \int_2^{3-x} (6 - x - y) dx dy = \frac{5}{24}$$

$$(iii) P(X < 1 | Y < 3) = \frac{P(X < 1 \cap Y < 3)}{P(Y < 3)}$$

$$\text{But } P(Y < 3) = \frac{1}{8} \int_0^2 \int_2^3 (6 - x - y) dx dy = \frac{5}{8}$$

$$\text{Hence } P(X < 1 | Y < 3) = \frac{\frac{3}{8}}{\frac{5}{8}} = \frac{3}{5}$$

P4:

The j.p.d.f. of (X, Y) is given by

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

a) Find marginal and conditional p.d.fs of x and y .

b) Check for independence of x and y .

Solution:

a) The m.p.d.f. of X is given by

$$f_1(x) = \int_0^x f(x, y) dy = \int_0^x 2 dy = 2x$$
$$\Rightarrow f_1(x) = 2x, 0 < x < 1$$

The m.p.d.f. of Y is given by

$$f_2(y) = \int_y^1 f(x, y) dx = \int_y^1 2 dx = 2(1 - y)$$
$$\Rightarrow f_2(y) = 2(1 - y), 0 < y < 1$$

The conditional p.d.f of X given Y is given by

$$f_{1|2}(x|y) = \frac{f(x, y)}{f_2(y)} = \frac{2}{2(1 - y)} = \frac{1}{1 - y}, y < x < 1$$

The conditional p.d.f of Y given X is given by

$$f_{2|1}(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{2}{2x} = \frac{1}{x}, 0 < y < x < 1$$

b) Since $f_1(x) \cdot f_2(y) = 2x \cdot 2(1 - y) = 4x(1 - y) \neq f(x, y)$, X and Y are not independent.

2.3

Mathematical Expectation

The term expectation is used for the process of averaging when a random variable is involved. It is the number used to locate the centre of the probability distribution (p.m.f or p.d.f) of a random variable. A probability distribution is described by certain satisfied measures which are computed using mathematical expectation (or expectation)

Let X be a random variable defined on a sample space S . Let $g(\cdot)$ be a function of X such that $g(X)$ is a random variable. Then the **expected value of $g(X)$** is defined by

$$E(g(X)) = \begin{cases} \sum_x g(x) p(x) & \text{if } X \text{ is a d.r.v with p.m.f. } p(x) \\ \int_{-\infty}^{\infty} g(x) f(x) dx & \text{if } X \text{ is a c.r.v with p.d.f. } f(x) \end{cases} \text{-----(1)}$$

provided these values exist.

Mean and moments:

i. Let $g(X) = X$. Then, by formula (1), **expected value of X** is defined by

$$E(X) = \mu = \begin{cases} \sum_x x p(x) & \text{if } X \text{ is a d.r.v with p.m.f. } p(x) \\ \int_{-\infty}^{\infty} x f(x) dx & \text{if } X \text{ is a c.r.v with p.d.f. } f(x) \end{cases}$$

Then $E(X)$ is called the **mean of the random variable X** and it is denoted by μ .

ii. Let $g(X) = (X - A)^r$ where A is an arbitrary constant and r is a non negative integer. Then the formula (1) gives

$$E(X - A)^r = \mu'_r = \begin{cases} \sum_x (x - A)^r p(x) & \text{if } X \text{ is a d.r.v with p.m.f. } p(x) \\ \int_{-\infty}^{\infty} (x - A)^r f(x) dx & \text{if } X \text{ is a c.r.v with p.d.f. } f(x) \end{cases}$$

The quantity $E(X - A)^r$ is called the **r^{th} moment about A** and it is denoted by μ'_r .
If $A = 0$, then μ'_r are known as **Raw Moments**.

iii. Let $g(X) = (X - E(X))^r = (X - \mu)^r$. Then the formula (1) gives

$$E(X - \mu)^r = \mu_r = \begin{cases} \sum_x (x - \mu)^r p(x) & \text{if } X \text{ is a d.r.v with p.m.f. } p(x) \\ \int_{-\infty}^{\infty} (x - \mu)^r f(x) dx & \text{if } X \text{ is a c.r.v with p.d.f. } f(x) \end{cases}$$

The function $E(X - \mu)^r$ is called the **r^{th} central moment of X** and it is denoted by μ_r

iv. If $r = 2$, then $\mu_2 = \sigma^2 = E(X - \mu)^2$ and it is known as the **variance of the random variable X** and it is denoted by $V(X)$ or σ^2 .

v. Mean (μ) and variance (σ^2) are important statistical measures of a probability distribution.

Example 1: Let X be a d.r.v with the p.m.f. given below:

x	-3	6	9
$p(x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

Find $E(X)$ and $E(X^2)$.

Solution:

$$E(X) = \sum_x x p(x) = -3 \times \frac{1}{6} + 6 \times \frac{1}{2} + 9 \times \frac{1}{3} = -\frac{1}{2} + 3 + 3 = \frac{11}{2}$$

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

Example 2: Find the expectation of the number on a die when thrown.

Solution: Let X be the random variable representing the number on a die when thrown. Then X can take any one of the values 1,2,3,4,5,6 each with equal probability $\frac{1}{6}$. Hence

$$\begin{aligned} E(X) &= \sum_x x p(x) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} \\ &= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{1}{6} \cdot \frac{6 \cdot 7}{2} = \frac{7}{2} \\ \Rightarrow E(X) &= \frac{7}{2} \end{aligned}$$

Example 3: Two unbiased dice are thrown. Find the expected values of the sum of numbers of points on them.

Solution: Define X is the sum of the numbers obtained on the two dice and $X = 2, 3, 4, \dots, 12$ and its probability distribution is given by

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

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

Example 4: In four tosses of a coin, let X be the number of heads. Find the mean and variance of X .

Solution: The sample space S consists of $2^4 = 16$ outcomes and the following table gives the outcomes and the value of X for each outcome is

S.No	Out come	X
1	$TTTT$	0
2	$TTTH$	1
3	$TTHT$	1
4	$TTHH$	2
5	$THTT$	1
6	$THTH$	2
7	$THHT$	2
8	$THHH$	3
9	$HTTT$	1
10	$HTTH$	2
11	$HTHT$	2
12	$HTHH$	3
13	$HHTT$	2
14	$HHTH$	3
15	$HHHT$	3
16	$HHHH$	4

$$p(0) = \frac{1}{16}, p(1) = \frac{4}{16}, p(2) = \frac{6}{16}, p(3) = \frac{4}{16}, p(4) = \frac{1}{16}$$

The p.m.f of X is given in the following table:

x	0	1	2	3	4
$p(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

$$E(X) = \sum_x x p(x) = 0 \times \frac{1}{16} + 1 \times \frac{4}{16} + 2 \times \frac{6}{16} + 3 \times \frac{4}{16} + 4 \times \frac{1}{16}$$

$$= \frac{1}{16} (0 + 4 + 12 + 12 + 4) = \frac{32}{16} = 2$$

$$\Rightarrow E(X) = 2$$

$$V(X) = E(X - 2)^2 = \sum (x - 2)^2 p(x)$$

$$\begin{aligned}
 &= (0-2)^2 \times \frac{1}{16} + (1-2)^2 \times \frac{4}{16} + (2-2)^2 \times \frac{6}{16} + (3-2)^2 \times \frac{4}{16} + (4-2)^2 \times \frac{1}{16} \\
 &= 4 \times \frac{1}{16} + 1 \times \frac{4}{16} + 0 \times \frac{6}{16} + 1 \times \frac{4}{16} + 4 \times \frac{1}{16} = \frac{1}{16} (4 + 4 + 4 + 4) = \frac{16}{16} = 1 \\
 &\Rightarrow V(X) = 1
 \end{aligned}$$

Example 5: Find the mean and variance of the random variable X , whose p.d.f is given by

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

Solution:

$$E(X) = \int_0^2 x \cdot f(x) dx = \frac{1}{2} \int_0^2 x \cdot dx = \frac{1}{2} \left[\frac{x^2}{2} \right]_0^2 = 1 - 0 = 1$$

$\Rightarrow$ Mean of the random variable X is 1.

$$\text{Variance} = E(X - 1)^2 =$$

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

$\Rightarrow$ Variance of the random variable X is $\frac{1}{3}$

Example 6: Find the mean of the random variable X whose p.d.f. is given by

$$f(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Solution:

$$E(X) = \int_0^{\infty} x f(x) dx = \frac{1}{2} \int_0^{\infty} x e^{-x} dx = \left[-x e^{-x} \right]_0^{\infty} + \int_0^{\infty} e^{-x} dx = 0 + \left[-e^{-x} \right]_0^{\infty} = 0 + 1 = 1$$

$$\Rightarrow E(X) = 1$$

Theorems on Mathematical Expectation:

The following theorems are proved by assuming that the random variables are continuous. If the random variables are discrete, the proof remains the same except replacing integration by summation.

Theorem 1: If X is a random variable and a and b are constants then

$$E(aX + b) = aE(X) + b.$$

Proof: Let X be a c.r.v with p.d.f. $f(x)$. Then

$$E(aX + b) =$$

$$\int_{-\infty}^{\infty} (ax + b) f(x) dx = a \int_{-\infty}^{\infty} x f(x) dx + b \int_{-\infty}^{\infty} f(x) dx = aE(X) + b \quad \left(\because \int_{-\infty}^{\infty} f(x) dx = 1 \right)$$

Corollary 1: If $b = 0$, then $E(aX) = aE(X)$

Corollary 2: If $X = 1$ and $b = 0$, then $E(a) = a$

Theorem 2: Addition Theorem of mathematical expectation.

If X and Y are random variables, then $E(X + Y) = E(X) + E(Y)$ provided all the expectations exist.

Proof: Let X and Y be continuous random variables with j.p.d.f. $f(x, y)$ and m.p.d.fs be $f_1(x)$ and $f_2(y)$ respectively. Then by definition,

$$E(X) = \int_{-\infty}^{\infty} x f_1(x) dx \quad \text{and} \quad E(Y) = \int_{-\infty}^{\infty} y f_2(y) dy$$

$$\text{Now, } E(X + Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f(x, y) dx dy$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x,y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x,y) dx dy \\
&= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x,y) dy \right] dx + \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f(x,y) dx \right] dy \\
&= \int_{-\infty}^{\infty} x f_1(x) dx + \int_{-\infty}^{\infty} y f_2(y) dy = E(X) + E(Y)
\end{aligned}$$

$$\therefore E(X + Y) = E(X) + E(Y)$$

Generalization: If $X_1, X_2, X_3 \dots X_n$ are random variables, then $E(X_1 + X_2 + X_3 + \dots + X_n) = E(X_1) + E(X_2) + E(X_3) \dots + E(X_n)$ provided all the expectations exist.

Theorem 3: Multiplication Theorem of mathematical Expectations

If X and Y are independent random variables, then $E(XY) = E(X)E(Y)$.

Proof: Let X and Y be continuous random variables with j.p.d.f. $f(x, y)$ and m.p.d.fs be $f_1(x)$ and $f_2(y)$ respectively. Then by definition,

$$E(X) = \int_{-\infty}^{\infty} x f_1(x) dx \quad \text{and} \quad E(Y) = \int_{-\infty}^{\infty} y f_2(y) dy$$

$$\text{Now, } E(X.Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (xy) f(x,y) dx dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x,y) f_1(x) f_2(y) dx dy \quad (\because X \text{ and } Y \text{ are independent})$$

$$= \left(\int_{-\infty}^{\infty} x f_1(x) dx \right) \left(\int_{-\infty}^{\infty} y f_2(y) dy \right) = E(X)E(Y)$$

Generalization: If $X_1, X_2, X_3 \dots X_n$ are independent random variables, then $E(X_1 X_2 X_3 \dots X_n) = E(X_1)E(X_2)E(X_3) \dots E(X_n)$.

Theorem 4: Mathematical expectation of a linear combination of random variables.

Let $X_1, X_2, X_3 \dots X_n$ be any n random variables and $a_1, a_2, a_3, \dots, a_n$ be any n constants. Then

$$E(a_1X_1 + a_2X_2 + a_3X_3 + \dots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + a_3E(X_3) + \dots + a_nE(X_n)$$

provided all the expectations exist.

The proof follows using Theorem 1 and generalization of Theorem 2.

Theorem 5: $V(X) = E(X^2) - (E(X))^2$

Proof: $V(X) = E[X - E(X)]^2$

$$= E[X^2 - 2XE(X) + (E(X))^2]$$

$$= E(X^2) - 2E(X E(X)) + E(E(X))^2$$

$$= E(X^2) - 2E(X)E(X) + (E(X))^2 \because E(X) \text{ is a constant and } E(E(X)) = E(X)$$

$$\Rightarrow V(X) = E(X^2) - 2(E(X))^2 + (E(X))^2$$

$$\Rightarrow V(X) = E(X^2) - (E(X))^2$$

Note: The formula is simple to use instead of $E(X - E(X))^2$.

Theorem 6: If X is a random variable, and a and b are constants, then $V(ax + b) = a^2 V(X)$.

Proof: Let $Y = aX + b$. Then $E(Y) = E(aX + b) = aE(X) + b$ and

$$Y - E(Y) = a(X - E(X))$$

$$\Rightarrow E(Y - E(Y))^2 = a^2 E(X - E(X))^2$$

$$\Rightarrow V(Y) = a^2 V(X) \Rightarrow V(aX + b) = a^2 V(X)$$

Corollary 1: If $a = 0$, then $V(b) = 0$ i. e., variance of a constant is zero.

Corollary 2: If $b = 0$, then $V(aX) = a^2V(X)$

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

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

Note:

1. If X and Y are independent, then $\text{Cov}(X, Y) = 0$
2. $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$ where a and b are constants.
3. $\text{Cov}(X + a, Y + b) = \text{Cov}(X, Y)$.

Theorem 7: Variance of a linear combination of random variables.

Let $X_1, X_2, X_3, \dots, X_n$ be any n random variables and $a_1, a_2, a_3, \dots, a_n$ are n constants, then

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

Proof:

$$\text{Let } U = \sum_{i=1}^n a_i X_i, \text{ then } E(U) = \sum_{i=1}^n a_i E(X_i) \text{ and } U - E(U) = \sum_{i=1}^n a_i (X_i - E(X_i))$$

$$\begin{aligned}
&\Rightarrow (U - E(U))^2 = \\
&\sum_{i=1}^n a_i^2 (X_i - E(X_i))^2 + 2 \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n a_i a_j (X_i - E(X_i))(X_j - E(X_j)) \\
&\Rightarrow E(U - E(U))^2 = \\
&\sum_{i=1}^n a_i^2 E(X_i - E(X_i))^2 + 2 \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n a_i a_j E[(X_i - E(X_i))(X_j - E(X_j))] \\
&\Rightarrow V\left(\sum_{i=1}^n a_i X_i\right) = V(U) = \sum_{i=1}^n a_i^2 V(X_i) + 2 \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n a_i a_j \text{cov}(X_i, X_j)
\end{aligned}$$

Note:

1. If $X_1, X_2, X_3, \dots, X_n$ are independent, then $V\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 V(X_i)$
2. If $a_1 = a_2 = 1$ and $a_3 = \dots = a_n = 0$, then
 $V(X_1 + X_2) = V(X_1) + V(X_2) + 2\text{Cov}(X_1, X_2)$
3. If $a_1 = 1, a_2 = -1$ and $a_3 = \dots = a_n = 0$, then
 $V(X_1 - X_2) = V(X_1) + V(X_2) - 2\text{Cov}(X_1, X_2)$
4. If X_1 and X_2 are independent, then $V(X_1 \pm X_2) = V(X_1) + V(X_2)$

Example 7: The j.p.d.f. of X and Y is given by

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

Find

- i. m.p.d.fs of X and Y
- ii. c.p.d.fs of X and Y
- iii. $V(X)$ and $V(Y)$

iv. Covariance between X and Y

Solutions:

$$i. \quad f_1(x) = \int_0^1 f(x, y) dy = \int_0^1 (2 - x - y) dy = \left[2y - xy - \frac{y^2}{2} \right]_0^1 = 2 - x - \frac{1}{2} = \frac{3}{2} - x$$

$$f_1(x) = \begin{cases} \frac{3}{2} - x & , 0 < x < 1 \\ 0 & , \text{otherwise} \end{cases}$$

$$\text{Similarly } f_2(y) = \begin{cases} \frac{3}{2} - y & , 0 < y < 1 \\ 0 & , \text{otherwise} \end{cases}$$

$$ii. \quad f_{1|2}(x|y) = \frac{f(x, y)}{f_2(y)} = \frac{2 - x - y}{\frac{3}{2} - y}, \quad 0 < x, y < 1$$

$$\text{and } f_{2|1}(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{2 - x - y}{\frac{3}{2} - x}, \quad 0 < x, y < 1$$

$$iii. \quad E(X) = \int_0^1 x f_1(x) dx = \int_0^1 x \left(\frac{3}{2} - x \right) dx = \frac{5}{12} \text{ and}$$

$$E(X^2) = \int_0^1 x^2 f_1(x) dx = \int_0^1 x^2 \left(\frac{3}{2} - x \right) dx = \frac{1}{4}$$

$$\text{Thus } V(X) = E(X^2) - (E(X))^2 = \frac{1}{4} - \frac{25}{144} = \frac{11}{144}$$

$$\text{Similarly } V(Y) = \frac{11}{144}$$

$$iv. \quad E(XY) = \int_0^1 \int_0^1 xy f(x, y) dx dy = \int_0^1 \int_0^1 xy (2 - x - y) dx dy = \frac{1}{6} \text{ (verify!)}$$

$$\therefore \text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{6} - \frac{5}{12} \cdot \frac{5}{12} = -\frac{1}{144}$$

P1:

Let X be a d.r.v. with p.m.f given by

$$p(x) = \begin{cases} \frac{1}{3}, & x=0 \\ \frac{2}{3}, & x=2 \end{cases}$$

Find the mean and variance of X .

Solution:

Here $p(0) = \frac{1}{3}$ and $p(2) = \frac{2}{3}$

Then

$$E(X) = \sum_x x p(x) = 0 \times \frac{1}{3} + 2 \times \frac{2}{3} = \frac{4}{3}$$

and

$$E(X^2) = \sum_x x^2 p(x) = 0^2 \times \frac{1}{3} + 2^2 \times \frac{2}{3} = \frac{8}{3}$$

$$\therefore V(X) = E(X^2) - (E(X))^2 = \frac{8}{3} - \frac{16}{9} = \frac{8}{9}$$

Thus, mean is equal to $\frac{4}{3}$ and variance is equal to $\frac{8}{9}$.

P2:

Find the mean and variance of the random variable, whose p.d.f. is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

Solution:

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x f(x) dx = \lambda \int_0^{\infty} x e^{-\lambda x} dx = \lambda \left[\left(-\frac{x e^{-\lambda x}}{\lambda} \right)_0^{\infty} + \frac{1}{\lambda} \int_0^{\infty} e^{-\lambda x} dx \right]$$

$$= \lambda \left[0 - \left(\frac{e^{-\lambda x}}{\lambda^2} \right)_0^{\infty} \right] = \lambda \left(\frac{1}{\lambda^2} \right) = \frac{1}{\lambda}$$

$$\text{and } E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^{\infty} x^2 f(x) dx = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx$$

$$= \lambda \left[\left(\frac{x^2 \cdot e^{-\lambda x}}{\lambda} \right)_0^{\infty} + 2 \int_0^{\infty} \frac{x \cdot e^{-\lambda x}}{\lambda} dx \right] = \lambda \left[0 + \frac{2}{\lambda} \int_0^{\infty} x \cdot e^{-\lambda x} dx \right] = \lambda \left[\frac{2}{\lambda} \cdot \frac{1}{\lambda^2} \right] = \frac{2}{\lambda^2}$$

$$\text{Thus } V(X) = E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

P3:

A student doing a summer internship in a company was asked to model the life term of certain equipment that the company makes. After a series of tests, the student proposed that the life time of the equipment can be modeled by a random variable X that has p.d.f.

$$f(x) = \begin{cases} \frac{x e^{-\frac{x}{10}}}{100}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

- Show that $f(x)$ is a valid p.d.f.
- What is the probability that the lifetime of the equipment exceeds 20?
- Find the mean?

Solution:

$$\begin{aligned} \text{a. Consider } \int_0^{\infty} f(x) dx &= \frac{1}{100} \int_0^{\infty} x e^{-\frac{x}{10}} dx \\ &= \frac{1}{100} \left\{ \left[-10x e^{-\frac{x}{10}} \right]_0^{\infty} + 10 \int_0^{\infty} e^{-\frac{x}{10}} dx \right\} \\ &= \frac{1}{100} \left\{ 0 - 100 \left[e^{-\frac{x}{10}} \right]_0^{\infty} \right\} = \frac{1}{100} \times 100 = 1 \end{aligned}$$

Thus, $f(x)$ is a valid p.d.f.

$$\text{b. } P(X > 20) = \int_{20}^{\infty} f(x) dx = \frac{1}{100} \left\{ \left[-10x e^{-\frac{x}{10}} \right]_{20}^{\infty} + 10 \int_{20}^{\infty} e^{-\frac{x}{10}} dx \right\}$$

$$= \frac{1}{100} \left\{ 200e^{-2} - \left[100e^{-\frac{x}{10}} \right]_{20}^{\infty} \right\} = 2e^{-2} + e^{-2} = 3e^{-2} = \frac{3}{e^2} = 0.406$$

c. Mean = $E(X) = \int_0^{\infty} x f(x) dx = \frac{1}{100} \int_0^{\infty} x^2 e^{-\frac{x}{10}} dx$

$$= \frac{1}{100} \left[10x^2 e^{-\frac{x}{10}} \right]_0^{\infty} + \frac{20}{100} \int_0^{\infty} x e^{-\frac{x}{10}} dx$$

$$= 0 + 20 = 20$$

P4:

A box contains a white and b black balls. c balls are drawn at random. Find the expected value of the number of white balls drawn.

Solution:

Let a variable X_i associated with i^{th} drawn, be defined as follows:

$$X_i = \begin{cases} 1, & \text{if the } i^{th} \text{ ball drawn is white} \\ 0, & \text{if the } i^{th} \text{ ball drawn is black} \end{cases}$$

Then the number of S of white balls among c balls drawn is given by

$$S = X_1 + X_2 + \cdots + X_c = \sum_{i=1}^c X_i$$

$$\Rightarrow E(S) = \sum_{i=1}^c E(X_i)$$

$$\text{But } P(X_i = 1) = P(\text{ of drawing a white ball }) = \frac{a}{a+b}$$

$$\text{and } P(X_i = 0) = P(\text{ of drawing a black ball }) = \frac{b}{a+b}$$

$$\therefore E(X_i) = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(X_i = 1) = \frac{a}{a+b} \text{ and hence}$$

$$E(S) = \sum_{i=1}^c E(X_i) = \sum_{i=1}^c \frac{a}{a+b} = \frac{ac}{a+b}.$$

2.6

Functions of Random Variables

The previous modules discussed basic properties of events defined in a given sample space and the random variables used to represent those events. The fundamental assumption that was made in those modules is that events can always be defined by **random variables**. However, in many applications, the events are functions of other events. For example, the time until a complex system fails is a function of the time to failure of the individual components that make up the system. This means that the random variable used to represent the time to failure of the complex system is a function of the random variables used to represent the times to failure of the component parts of the system. This module deals with functions of random variables. Because of the complexity involved in computing the c.d.fs and p.d.fs of multiple random variables, the discussion is restricted to functions of at most two random variables.

Functions of One Random Variable: Let X be a r.v. with p.d.f. (or p.m.f.) $f_X(x)$ and c.d.f. $F_X(x)$. Let Y be the new random variable that is a function of X . That is,

$$Y = g(X)$$

Then we are interested in computing p.d.f (or p.m.f.) $f_Y(y)$ and c.d.f. $F_Y(y)$ of Y .

For example, let $Y = X + 5$. Then

$$F_Y(y) = P(Y \leq y) = P[X + 5 \leq y] = P[X \leq y - 5] = F_X(y - 5)$$

Linear Functions: Consider the function $g(X) = aX + b$, where a and b are constants. The c.d.f of Y is given by

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} = P[aX + b \leq y] \\ &= P\left[X \leq \frac{y-b}{a}\right] = F_X\left(\frac{y-b}{a}\right) \end{aligned}$$

where a is positive .The p.d.f. of Y is given by

$$f_Y(y) = \frac{d}{dy} (F_Y(y)) = \frac{d}{dy} \left(F_X \left(\frac{y-b}{a} \right) \right) = \left(\frac{d}{du} (F_X(u)) \right) \left(\frac{du}{dy} \right)$$

where $u = \frac{y-b}{a}$ and $\frac{du}{dy} = \frac{1}{a}$. Thus,

$$f_Y(y) = \left(\frac{1}{a} \right) f_X(u) = \left(\frac{1}{a} \right) f_X \left(\frac{y-b}{a} \right)$$

If $a < 0$, we have,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(aX + b \leq y) = P(aX \leq y - b) \\ &= P \left(X \geq \frac{y-b}{a} \right) = 1 - \left\{ P \left[X \leq \frac{y-b}{a} \right] - P \left[X = \frac{y-b}{a} \right] \right\} \quad (\because a < 0) \end{aligned}$$

The change in sign on the second line arises from the fact that a is negative. If X is continuous, $P \left[X = \frac{y-b}{a} \right] = 0$. Thus, the c.d.f and p.d.f for the case of negative a are given by

$$\begin{aligned} F_Y(y) &= 1 - P \left[X \leq \frac{y-b}{a} \right] \\ &= 1 - F_X \left(\frac{y-b}{a} \right) \end{aligned}$$

$$\text{Therefore, } f_Y(y) = \frac{d}{dy} (F_Y(y)) = - \left(\frac{1}{a} \right) f_X \left(\frac{y-b}{a} \right)$$

Therefore, the general p.d.f. of Y is given by

$$f_Y(y) = \frac{1}{|a|} f_X \left(\frac{y-b}{a} \right)$$

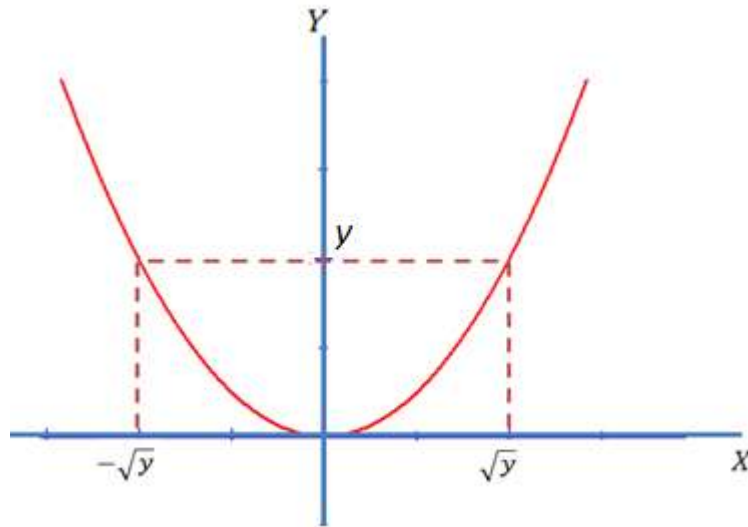
Example 1: Find the p.d.f of Y in terms of the p.d.f of X if $Y = 2X + 7$.

Solution: From the results obtained above,

$$F_Y(y) = F_X \left(\frac{y-7}{2} \right)$$

$$\text{and } f_Y(y) = \left(\frac{1}{2} \right) f_X \left(\frac{y-7}{2} \right)$$

Power Functions: Consider the quadratic function $Y = X^2$. The plot of Y against X is shown in the following figure where we see that for one value of Y there are two values of X , namely $\sqrt{Y}$ and $-\sqrt{Y}$.



Thus, the c.d.f of Y is given by

$$\begin{aligned} F_Y(y) &= P[Y \leq y] = P[X^2 \leq y] \\ &= P[|X| \leq \sqrt{y}], \quad y > 0 \\ &= P[-\sqrt{y} \leq X \leq \sqrt{y}] \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$

The p.d.f of Y is given by

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{d}{dy} [F_X(\sqrt{y}) - F_X(-\sqrt{y})]$$

Let $u = \sqrt{y} = y^{\frac{1}{2}}$. Thus, $\frac{du}{dy} = \frac{1}{2}y^{-\frac{1}{2}}$ and

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} [F_X(\sqrt{y}) - F_X(-\sqrt{y})] \\ &= \frac{d}{du} (F_X(u)) \frac{du}{dy} + \frac{d}{du} (F_X(-u)) \frac{du}{dy} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} y^{-\frac{1}{2}} \left[\frac{d}{du} (F_X(u)) + \frac{d}{du} (F_X(-u)) \right] \\
&= \frac{1}{2} y^{-\frac{1}{2}} [f_X(\sqrt{y}) + f_X(-\sqrt{y})] \\
&= \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}}, y > 0
\end{aligned}$$

If $f_X(x)$ is an even function, then $f_X(x) = f_X(-x)$ and $F_X(-x) = 1 - F_X(x)$. Thus, we have

$$f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}} = \frac{2f_X(\sqrt{y})}{2\sqrt{y}} = \frac{f_X(\sqrt{y})}{\sqrt{y}}$$

Example2: Find the p.d.f of the random variable $Y = X^2$, where X is the standard normal random variable.

Solution: Since the p.d.f. of X is given by $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, which is an even function, we know that

$$\begin{aligned}
F_Y(y) &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \\
&= 2F_X(\sqrt{y}) - 1
\end{aligned}$$

Therefore, if we let $u = \sqrt{y}$, then

$$\begin{aligned}
f_Y(y) &= \frac{dF_Y(y)}{dy} = 2 \frac{dF_X(u)}{du} \frac{1}{2\sqrt{y}} \\
&= \frac{1}{\sqrt{y}} f_X(\sqrt{y}) \\
&= \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, y > 0
\end{aligned}$$

Sum of Two Independent Random variables

Consider two independent continuous random variables X and Y . We are interested in computing the c.d.f and p.d.f of their sum $g(X, Y) = S = X + Y$. The random variable S can be used to model the reliability of systems with stand-by connections, as shown in *fig.1* In such systems, the component A whose time-to-failure is represented by the random variable X is the primary component, and the component B whose time-to-failure is represented by the random variable Y is the backup component that is brought into operation when the primary component fails. Thus, S represents the time until the system fails, which is the sum of the lifetimes of both components.

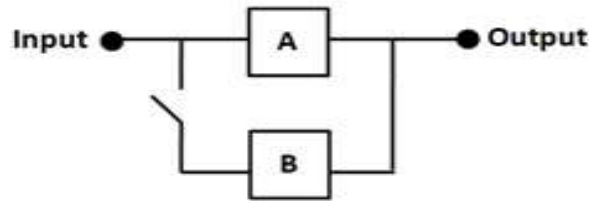


fig. 1

Their c.d.f. can be obtained as follows:

$$F_S(s) = P[S \leq s] = P[X + Y \leq s] = \iint_D f_{XY}(x, y) dx dy$$

where D is the set $D = \{(x, y) | x + y \leq s\}$, which is the area to the left of the line $s = x + y$ as shown in *fig. 2*.

Thus,

$$\begin{aligned} F_S(s) &= \int_{-\infty}^{\infty} \int_{-\infty}^{s-y} f_{XY}(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{s-y} f_X(x) f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{s-y} f_X(x) dx \right\} f_Y(y) dy \\ &= \int_{-\infty}^{\infty} F_X(s - y) f_Y(y) dy \end{aligned}$$

The p.d.f. of S is obtained by differentiating the c.d.f. , as follows:

$$\begin{aligned} f_S(s) &= \frac{d}{ds} F_S(s) = \frac{d}{ds} \int_{-\infty}^{\infty} F_X(s-y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \frac{d}{ds} F_X(s-y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} f_X(s-y) f_Y(y) dy \end{aligned}$$

where we have assumed that we can interchange differentiation and integration. The expression on the right-hand side is a well-known result in signal analysis

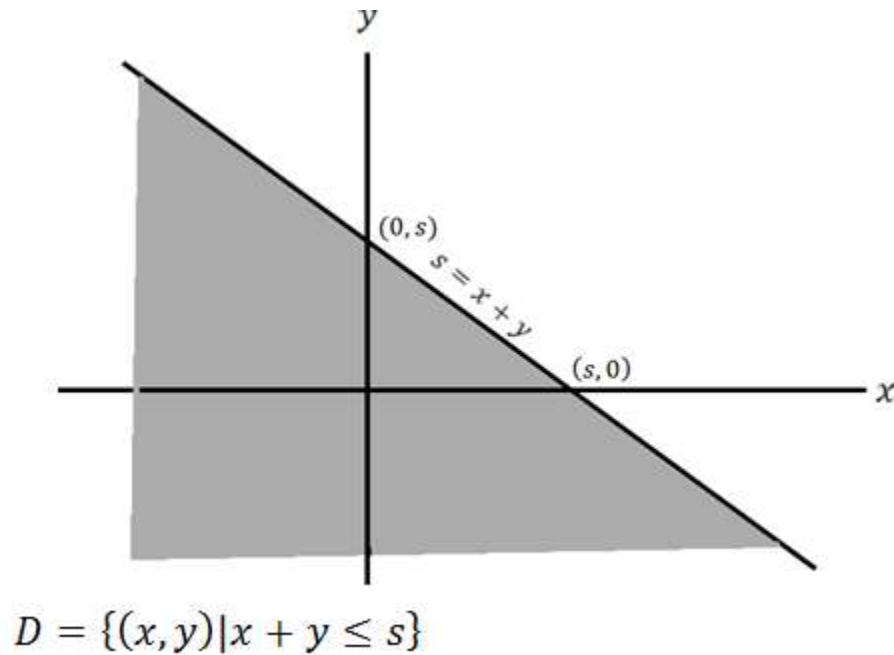


fig. 2

called the **convolution integral**. Thus, we find that the p.d.f of the sum S of two independent random variables X and Y is the convolution of the p.d.fs of the two random variables; that is,

$$f_S(s) = f_X(s) f_Y(s)$$

Example 3: Find the p.d.f. of the sum of X and Y if the two random variables are independent random variables with the common p.d.f.

$$f_X(u) = f_Y(u) = \begin{cases} \frac{1}{4} & 0 < u < 4 \\ 0 & \text{otherwise} \end{cases}$$

Solution: The limits of integration of the p.d.f of $S = X + Y$ can be computed with the aid of *fig. 3* When $0 \leq s \leq 4$ (see *fig. 3 (a)* where $f_Y(s - x)$ is shown in dashed lines),

$$f_S(s) = \int_0^s \frac{1}{16} dy = \frac{s}{16}$$

For $4 < s < 8$ (see *fig. 3 (b)*), we obtain

$$f_S(s) = \int_{s-4}^4 \frac{1}{16} dy = \frac{8-s}{16}$$

Thus ,

$$f_S(s) = \begin{cases} \frac{s}{16} & , 0 \leq s \leq 4 \\ \frac{8-s}{16} & , 4 < s < 8 \\ 0 & , \text{otherwise} \end{cases}$$

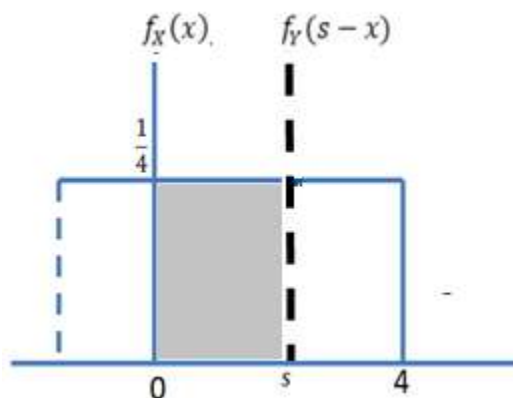


fig 3(a)

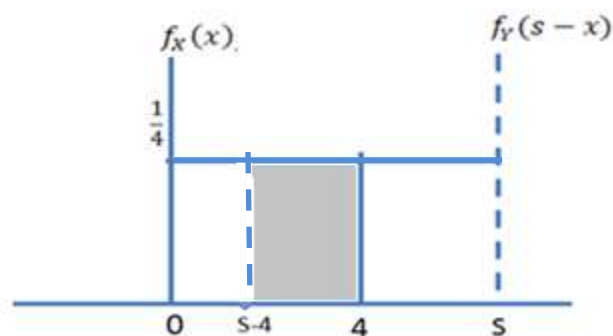
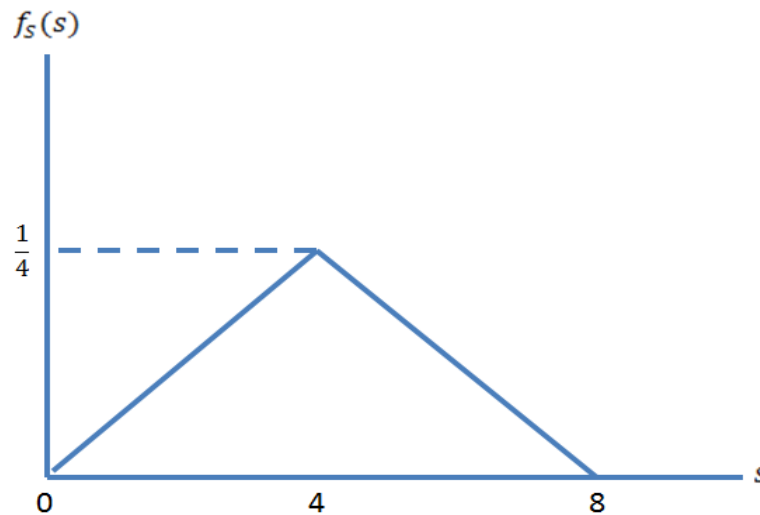


fig 3(b)

Fig. 3: Convolution of p.d.fs (a) $0 \leq s \leq 4$ and (b) $4 \leq s \leq 8$

The p.d.f of $S = X + Y$ is illustrated in the following figure.



Example 4: The time X between consecutive snowstorms in winter is a random variable with the p.d.f.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Assume it has not snowed up until now. What is the p.d.f. of the time U until the second snowstorm?

Solution: Let X be the random variable that denotes the time until the first snowstorm from the reference time, and let Y be the random variable that denotes the time between the first snowstorm and the second snowstorm. If we assume that the times between snowstorms are independent, then X and Y are independent and identically distributed random variables. That is, the p.d.f of Y is given by

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Thus, $U = X + Y$, and the p.d.f. of U is given by

$$f_U(u) = \int_0^{\infty} f_X(x)f_Y(u-x)dx$$

Since $f_X(x) = 0$ when $x < 0$, $f_Y(u - x) = 0$ when $u - x < 0$ (or $x > u$). Thus, the range of interest in the integration is $0 \leq x \leq u$, and we obtain

$$\begin{aligned} f_U(u) &= \int_0^u f_X(x) f_Y(u - x) dx \\ &= \int_0^u \lambda e^{-\lambda x} \lambda e^{-\lambda(u-x)} dx = \lambda^2 e^{-\lambda u} \int_0^u dx \\ &= \lambda^2 u e^{-\lambda u} \quad u \geq 0 \end{aligned}$$

This is the Erlang – 2 distribution.

Note: A random variable X is said to follow Erlang- k distribution if its p.d.f. is given by

$$f(x) = \begin{cases} \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} & , \quad k = 1, 2, \dots; \lambda > 0; x \geq 0 \\ 0 & , \quad x < 0 \end{cases}.$$

Sum of Two Discrete Random Variables

The examples above deal with continuous random variables.

Let $Z = X + Y$, where X and Y are discrete random variables. Then the p.m.f of Z is given by

$$\begin{aligned} p_Z(z) &= P[Z = z] = P[X + Y = z] = \sum_{k \leq z} P[X = k, Y = z - k] \\ &= \sum_{k \leq z} p_{XY}[k, z - k] \end{aligned}$$

If X and Y are independent random variables, then the p.m.f. of Z is the convolution of the p.m.f of X and the p.m.f of Y . That is,

$$p_Z(z) = \sum_{k \leq z} p_{XY}(k, z - k) = \sum_{k \leq z} p_X(k) p_Y(z - k)$$

Sum of Two Independent Binomial Random Variables

Let X and Y be two independent binomial random variables with parameters (n, p) and (m, p) , respectively and their sum be Z ; that is, $Z = X + Y$. Then the p.m.f of Z is given by

$$\begin{aligned} p_Z(z) &= P[X + Y = z] \\ &= \sum_{k=0}^n P[X = k, Y = z - k] = \sum_{k=0}^n P[X = k]P[Y = z - k] \\ &= \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \binom{m}{z-k} p^{z-k} (1-p)^{m-z+k} \\ &= p^z (1-p)^{n+m-z} \sum_{k=0}^n \binom{n}{k} \binom{m}{z-k} \end{aligned}$$

Using the combinatorial identity $\binom{n+m}{z} = \sum_{k=0}^n \binom{n}{k} \binom{m}{z-k}$, we obtain

$$p_Z(z) = \binom{n+m}{z} p^z (1-p)^{n+m-z}$$

This result shows that the sum of two independent binomial random variables with parameters (n, p) and (m, p) is a binomial random variable with parameter $(n + m, p)$.

Minimum of Two Independent Random Variables

Consider two independent continuous random variables X and Y . We are interested in a random variable U that is the minimum of X and Y ; that is, $U = \min(X, Y)$. The random variable U can be used to represent the reliability of systems with series connections, as shown in *fig. 4*. Such systems are operational as long as all components are operational. The first component to fail causes the system to fail. Thus, if in the example shown in *fig. 4*, the times-to-failure are

represented by the random variables X and Y , then S represents the time until the system fails, which is the minimum of the lifetimes of the two components.

The c.d.f. of U can be obtained as follows:

$$F_U(u) = P[U \leq u] = P[\min(X, Y) \leq u] = P[(X \leq u, X \leq Y) \cup (Y \leq u, X > Y)]$$

Since $P[A \cup B] = P[A] + P[B] - P[A \cap B]$, we have that $F_U(u) = F_X(u) + F_Y(u) - F_{XY}(u, u)$. Also, since X and Y are independent, we obtain the c.d.f. and p.d.f. of U as follows:

$$F_U(u) = F_X(u) + F_Y(u) - F_{XY}(u, u) = F_X(u) + F_Y(u) - F_X(u)F_Y(u)$$

$$\begin{aligned} f_U(u) &= \frac{d}{du} F_U(u) = f_X(u) + f_Y(u) - f_X(u)F_Y(u) - F_X(u)f_Y(u) \\ &= f_X(u)\{1 - F_Y(u)\} + f_Y(u)\{1 - F_X(u)\} \end{aligned}$$

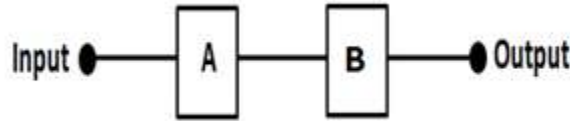


fig. 4

fig. 4: Series connection modeled by random variable U

Example 5: Assume that $U = \min(X, Y)$, where X and Y are independent random variables with the respective p.d.fs

$$f_X(x) = \lambda e^{-\lambda x} \quad x \geq 0$$

$$f_Y(y) = \mu e^{-\mu y} \quad y \geq 0$$

where $\lambda > 0$ and $\mu > 0$. What is the p.d.f. of U ?

Solution: We first obtain the c.d.fs of X and Y , which are as follows:

$$F_X(x) = P[X \leq x] = \int_0^x \lambda e^{-\lambda w} dw = 1 - e^{-\lambda x}$$

$$F_Y(y) = P[Y \leq y] = \int_0^y \mu e^{-\mu w} dw = 1 - e^{-\mu y}$$

Thus, the p.d.f of U is given by

$$\begin{aligned} f_U(u) &= f_X(u)\{1 - F_Y(u)\} + f_Y(u)\{1 - F_X(u)\} \\ &= \lambda e^{-\lambda u} e^{-\mu u} + \mu e^{-\mu u} e^{-\lambda u} \\ &= (\lambda + \mu)e^{-(\lambda+\mu)u}, \quad u \geq 0 \end{aligned}$$

This is exponential distribution with mean $\frac{1}{\lambda+\mu}$.

Maximum of Two Independent Random Variables

Consider two independent continuous random variables X and Y . We are interested in the c.d.f. and p.d.f. of the random variable W that is the maximum of the two random variables; that is, $W = \max(X, Y)$. The random variable W can be used to represent the reliability of systems with parallel connections, as shown in

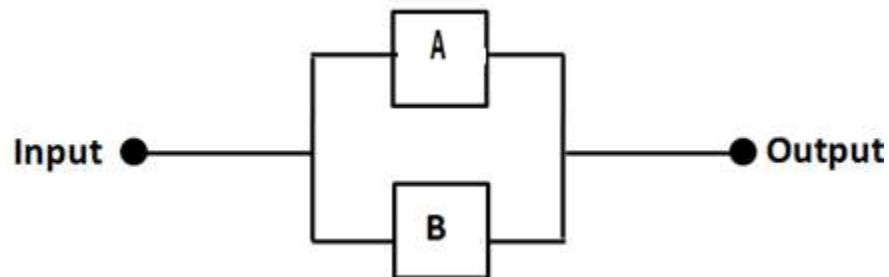


fig. 5

fig. 5: Parallel connection modeled by the random variable W

In such systems, we are interested in passing a signal between the two endpoints through either the component labeled A or the component labeled B . Thus, as long as one or both components are operational, the system is operational. This implies that the system is declared to have failed when both paths become unavailable. That is, the reliability of the system depends on the reliability of the last component to fail.

The c.d.f of W can be obtained by noting that if the greater of the two random variables is less than or equal to w , then the smaller random variable must also be less than or equal to w . Thus,

$$\begin{aligned} F_W(w) &= P[W \leq w] = P[\max(X, Y) \leq w] = P[(X \leq w) \cap (Y \leq w)] \\ &= F_{XY}(w, w) \end{aligned}$$

Since X and Y are independent, we obtain the c.d.f and p.d.f of W as follows:

$$\begin{aligned} F_W(w) &= F_{XY}(w, w) = F_X(w)F_Y(w) \\ f_W(w) &= \frac{d}{dw} F_W(w) = f_X(w)F_Y(w) + F_X(w)f_Y(w) \end{aligned}$$

Example 6: Assume that $W = \max(X, Y)$, where X and Y are independent random variables with the respective p.d.fs:

$$f_X(x) = \lambda e^{-\lambda x} \quad x \geq 0$$

$$f_Y(y) = \mu e^{-\mu y} \quad y \geq 0$$

where $\lambda > 0$ and $\mu > 0$. What is the pdf of W .

Solution: We first obtain the c.d.fs of X and Y , which are as follows:

$$F_X(x) = P[X \leq x] = \int_0^x \lambda e^{-\lambda z} dz = 1 - e^{-\lambda x}$$

$$F_Y(y) = P[Y \leq y] = \int_0^y \mu e^{-\mu z} dz = 1 - e^{-\mu y}$$

Thus, the p.d.f of W is given by

$$\begin{aligned} f_W(w) &= f_X(w)F_Y(w) + F_X(w)f_Y(w) \\ &= \lambda e^{-\lambda w} (1 - e^{-\mu w}) + \mu e^{-\mu w} (1 - e^{-\lambda w}) \\ &= \lambda e^{-\lambda w} + \mu e^{-\mu w} - (\lambda + \mu)e^{-(\lambda + \mu)w} \quad w \geq 0 \end{aligned}$$

Note that the mean of W is $\frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu}$.

Two Functions of Two Random Variables

Let X and Y be two random variables with a given joint p.d.f $f_{XY}(x, y)$. Assume that U and W are two functions of X and Y ; that is, $U = g(X, Y)$ and $W = h(X, Y)$. Sometimes it is necessary to obtain the joint p.d.f of U and W , $f_{UW}(u, w)$, in terms of the p.d.fs of X and Y .

It can be shown that $f_{UW}(u, w)$ is given by

$$f_{UW}(u, w) = \frac{f_{XY}(x_1, y_1)}{|J(x_1, y_1)|} + \frac{f_{XY}(x_2, y_2)}{|J(x_2, y_2)|} + \dots + \frac{f_{XY}(x_n, y_n)}{|J(x_n, y_n)|}$$

where $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ are real solutions of the equations $u = g(x, y)$ and $w = h(x, y)$; and $J(x, y)$ is called the **Jacobian** of the transformation $\{u = g(x, y), w = h(x, y)\}$ and defined by

$$J(x, y) = \begin{vmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{vmatrix} = \frac{\partial g}{\partial x} \cdot \frac{\partial h}{\partial y} - \frac{\partial g}{\partial y} \cdot \frac{\partial h}{\partial x}$$

Example 7: Let $U = g(X, Y) = X + Y$ and $W = h(X, Y) = X - Y$. Find $f_{UW}(u, w)$.

Solution: The unique solution to the equations $u = x + y$ and $w = x - y$ is $x = \frac{u+w}{2}$ and $y = \frac{u-w}{2}$. Thus, there is only one set of solutions. Since

$$J(x, y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2$$

we obtain

$$f_{UW}(u, w) = \frac{f_{XY}(x, y)}{|J(x, y)|} = \frac{1}{|-2|} f_{XY}\left(\frac{u+w}{2}, \frac{u-w}{2}\right) = \frac{1}{2} f_{XY}\left(\frac{u+w}{2}, \frac{u-w}{2}\right)$$

Application of the Transformation Method

Assume that $U = g(X, Y)$, and we are required to find the p.d.f. of U . We can use the above transformation method by defining an auxiliary function $W = X$ or $W = Y$ so we can obtain the joint p.d.f. $f_{UW}(u, w)$ of U and W . Then we obtain the required marginal p.d.f. $f_U(u)$ as follows:

$$f_U(u) = \int_{-\infty}^{\infty} f_{UW}(u, w) dw$$

Example 8: Find the p.d.f. of the random variable $U = X + Y$, where the joint p.d.f. of X and Y , $f_{XY}(x, y)$, is given.

Solution: We define the auxiliary random variable $W = X$. Then the unique solution to the two equations is $x = w$ and $y = u - w$, and the Jacobian of the transformation is

$$J(x, y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

Since there is only one solution to the equations, we have that

$$f_{UW}(u, w) = \frac{f_{XY}(w, u - w)}{|-1|} = f_{XY}(w, u - w)$$

$$f_U(u) = \int_{-\infty}^{\infty} f_{UW}(u, w) dw = \int_{-\infty}^{\infty} f_{XY}(w, u - w) dw$$

This reduces to the convolution integral. We obtained earlier when X and Y are independent.

Example 9: Find the p.d.f. of the random variable $U = X - Y$, where the joint p.d.f. of X and Y is given.

Solution: We define the auxiliary random variable $W = X$. Then the unique solution to the two equations is $x = w$ and $y = w - u$, and the Jacobian of the transformation is

$$J(x, y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} = 1$$

Since there is only one solution to the equations, we have that

$$f_{UW}(u, w) = f_{XY}(w, w - u)$$

$$f_U(u) = \int_{-\infty}^{\infty} f_{UW}(u, w) dw = \int_{-\infty}^{\infty} f_{XY}(w, w - u) dw$$

Example 10: The joint p.d.f of two random variables X and Y is given by $f_{XY}(x, y)$. If we define the random variable $U = XY$, determine the p.d.f of U .

Solution: We define the auxiliary random variable $W = X$. Then the unique solution to the two equations is $x = w$ and $y = \frac{u}{x} = \frac{u}{w}$, and the Jacobian of the transformation is

$$J(x, y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} y & x \\ 1 & 0 \end{vmatrix} = -x = -w$$

Since there is only one solution to the equations, we have that

$$f_{UW}(u, w) = \frac{f_{XY}(x, y)}{|J(x, y)|} = \frac{1}{|w|} f_{XY}\left(w, \frac{u}{w}\right)$$

$$f_U(u) = \int_{-\infty}^{\infty} f_{UW}(u, w) dw = \int_{-\infty}^{\infty} \frac{1}{|w|} f_{XY}\left(w, \frac{u}{w}\right) dw$$

P1:

Obtain the p.d.f. of $Z = X + Y$, where X and Y are two independent random variables with the following p.d.fs:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & ; \quad a < x < b \\ 0 & ; \quad \text{otherwise} \end{cases}$$
$$f_Y(y) = \begin{cases} \frac{1}{d-c} & ; \quad c < y < d ; \quad d - c < b - a \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Solution:

The two p.d.fs are shown in the *fig. 1*. To evaluate the limits of integration of the p.d.f. of Z , we consider the following regions represented by the diagram shown in *fig. 2*.

When $z < a + c$, $f_Z(z) = 0$. When $a + c \leq z \leq a + d$ (see *figure 2(i)*, we obtain)

$$f_Z(z) = \frac{1}{(b-a)(d-c)} \int_a^{z-c} dy = \frac{z-c-a}{(b-a)(d-c)}$$

When $a + d \leq z \leq b + c$ (see *fig. 2(ii)*), we obtain)

$$f_Z(z) = \frac{1}{(b-a)(d-c)} \int_{z-d}^{z-c} dy = \frac{1}{b-a}$$

When $b + c \leq z \leq b + d$ (see *fig. 2(ii)*), we obtain

$$f_Z(z) = \frac{1}{(b-a)(d-c)} \int_{z-d}^b dy = \frac{b+d-z}{(b-a)(d-c)}$$

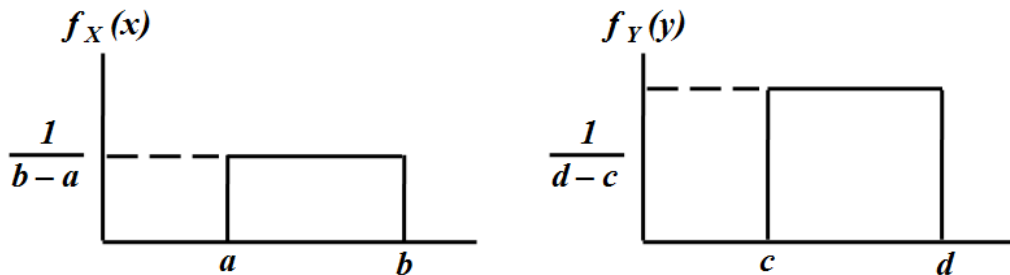


Figure 1: p.d.fs of X and Y

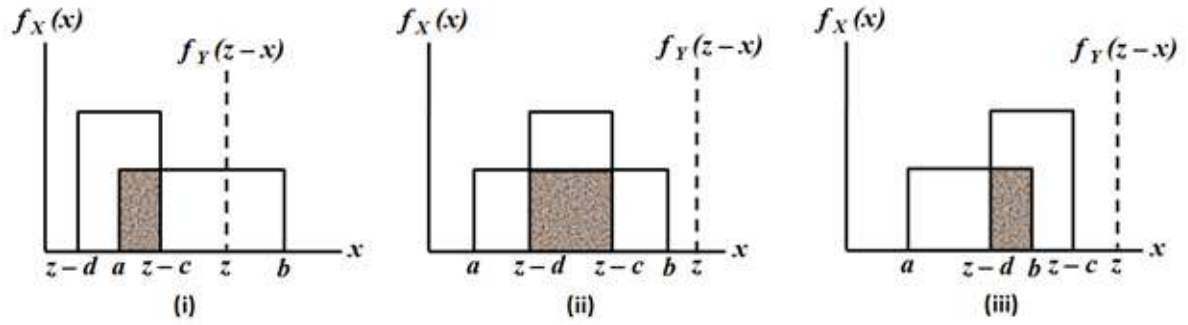


Figure 2: Convolution of the p.d.fs for different values z

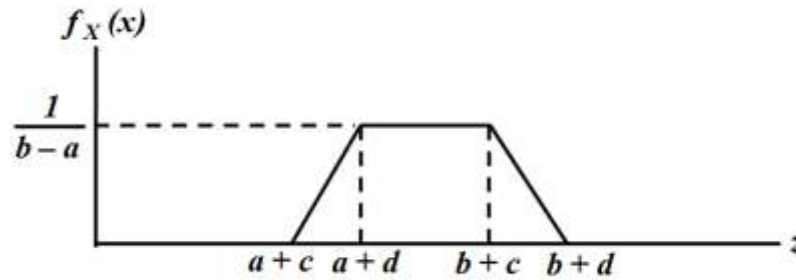
Finally, when $z > b + d$, $f_Z(z) = 0$, thus, the p.d.f of Z is given by

$$f_Z(z) = \begin{cases} 0 & , \quad z < a + c \\ \frac{z - a - c}{(b - a)(d - c)} & , \quad a + c \leq z \leq a + d \\ \frac{1}{b - a} & , \quad a + d \leq z \leq b + c \\ \frac{b + d - z}{(b - a)(d - c)} & , \quad b + c \leq z \leq b + d \\ 0 & , \quad z > b + d \end{cases}$$

The p.d.f is graphically illustrated in the following figure which is a **trapezoid**.

Note that when $b - a = d - c$, the p.d.f reduces to an **isosceles triangle** centered at $z = \frac{a+c+b+d}{2}$. In the special case when $a = c$ and $b = d$, the

isosceles triangle is centered at $z = a + b$.



P2:

Find the p.d.f of U , which is the sum of X and Y that are independent random variables with the following p.d.fs:

$$\begin{aligned}f_X(x) &= \lambda e^{-\lambda x} \quad , \quad x \geq 0 \\f_Y(y) &= \lambda^2 y e^{-\lambda y} \quad , \quad y \geq 0\end{aligned}$$

Solution:

Since X and Y are independent random variables and the p.d.f of U is given by

$$\begin{aligned}f_U(u) &= \int_0^u f_X(x) f_Y(u-x) dx \\&= \int_0^u \lambda e^{-\lambda x} \lambda^2 (u-x) e^{-\lambda(u-x)} dx \\&= \lambda^3 e^{-\lambda u} \int_0^u (u-x) dx \\&= \lambda^3 e^{-\lambda u} \left[ux - \frac{x^2}{2} \right]_0^u = \lambda^3 e^{-\lambda u} \left[u^2 - \frac{u^2}{2} \right] \\&= \frac{\lambda^3 u^2 e^{-\lambda u}}{2} \\&= \frac{\lambda^3 u^2 e^{-\lambda u}}{2!} \quad u \geq 0\end{aligned}$$

This is known as **Erlang – 3 distribution**.

2.6. Functions of Random Variables:

Exercise

- 1) The p.m.f. of X is given by

x	-2	-1	1	2
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Find the probability distribution of $Y = 4X + 3$.

- 2) A r.v. X is exponentially distributed with parameter 1. Find the p.d.f. of $y = \sqrt{x}$

- 3) Let X be a c.r.v with p.d.f.

$$f(x) = \begin{cases} \frac{x}{12} & , \quad 1 < x < 5 \\ 0 & , \quad otherwise \end{cases}$$

find the p.d.f. of $y = 2X + 4$

- 4) If the c.r.v X has the p.d.f.

$$f(x) = \begin{cases} \frac{2}{9}(x+1) & , \quad -1 < x < 2 \\ 0 & , \quad otherwise \end{cases}$$

find the p.d.f. of $y = X^2$

- 5) Find the p.d.f of $W = X + Y$ where X and Y are independent r.v s with the following p.d.f s:

$$f_X(x) = \lambda e^{-\lambda x}, x \geq 0$$

and $f_Y(y) = \mu e^{-\mu y}, y \geq 0$, where $\lambda \neq \mu$.

- 6) The j.p.d.f of two random variables X and Y is given by $f_{XY}(xy)$. Find the p.d.f of $V = \frac{X}{Y}$

ANSWERS

1.

Y	-5	-1	7	11
$p(y)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$2. f(y) = \begin{cases} 2y e^{-y^2}, & y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$3. f(y) = \begin{cases} \frac{y-4}{48}, & 6 < y < 14 \\ 0, & \text{otherwise} \end{cases}$$

$$4. f(y) = \begin{cases} \frac{2}{9\sqrt{y}}, & 0 < y < 1 \\ \frac{1}{9} \left(\frac{\sqrt{y}+1}{\sqrt{y}} \right), & 1 < y < 4 \\ 0, & \text{otherwise} \end{cases}$$

$$5. f_W(w) = \begin{cases} \frac{\lambda\mu}{\lambda-\mu} (e^{-\mu w} - e^{-\lambda w}), & w \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$6. f_V(v) = \int_{-\infty}^{\infty} |w| f_{XY}(vw, w) dw$$

2.7

Correlation coefficient and Bivariate Normal Distribution

Meaning of correlation:

In a bivariate distribution we may be interested to find out if there is any **correlation** or **covariance** between the two variables under study. If the change in one variable affects a change in the other variable, the variables are said to be **correlated**. If the two variables deviate in the same direction, *i. e.*, if the increase (or decrease) in one results in a corresponding increase (or decrease) in the other, **correlation** is said to be **positive**. But, if they constantly deviate in the opposite directions, *i. e.*, if increase (or decrease) in one results in corresponding decrease (or increase) in the other, **correlation** is said to be **negative**. For example, the correlation between (i) the heights and weights of a group of persons, and (ii) the income and expenditure; is positive and the correlation between (i) price and demand of a commodity and (ii) the volume and pressure of a perfect gas; is negative. *Correlation is said to be perfect* if the deviation in one variable is followed by a corresponding and proportional deviation in the other.

Karl Pearson's Coefficient of Correlation:

As a measure of intensity or degree of linear relationship between two variables, **Karl Pearson**, a British Biometrician developed a formula called **correlation coefficient**. Correlation coefficient between two variables X and Y , usually denoted by $\rho(X, Y)$ or ρ_{XY} , is a numerical measure of linear relationship between them and is defined by

$$\rho(X, Y) = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} = \frac{\text{cov}(X, Y)}{\sqrt{V(X)} \sqrt{V(Y)}}$$

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

$$\sigma_X^2 = V(X) = E[(X - E(X))^2] \text{ and } \sigma_Y^2 = V(Y) = E[(Y - E(Y))^2]$$

Note:

1. $\rho(X, Y)$ provides a measure of linear relationship between X and Y . For non – linear relationship, however, it is not suitable.
2. Karl Pearson's correlation coefficient is also called **product – moment correlation coefficient**.

Properties:

1. $-1 \leq \rho(X, Y) \leq 1$. If $\rho = -1$, the **correlation is perfect and negative**. If $\rho = 1$, the **correlation is perfect and positive**.
2. Correlation coefficient is independent of change of origin and scale. That is, if $U = \frac{X-a}{h}$ and $V = \frac{Y-b}{k}$, then $\rho(U, V) = \rho(X, Y)$

Theorem: Two independent variables are uncorrelated.

Proof:

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

$$\Rightarrow \sigma_{XY} = E(X, Y) - E(X).E(Y) \quad \text{..... (1)}$$

If X and Y are independent, then

$$E(XY) = E(X).E(Y) \quad \text{..... (2)}$$

From (1) and (2), if X and Y are independent, then $\rho(X, Y) = 0$

The converse need not be true. That is, uncorrelated variables need not be independent.

Example 1 : Let $X \sim N(0, 1)$ and $Y = X^2$. Then $E(X) = E(X^3) = 0$.

Solution: Consider $\text{cov}(X, Y) = E(XY) - E(X).E(Y) = E(X^3) - E(X).E(X^2)$

$$= 0 - 0 = 0$$

$\Rightarrow \text{cov}(X, Y) = 0$ but X and Y are related by $Y = X^2$.

Thus, uncorrelated variables need not be independent.

Note: The converse is true if the joint distribution of (X, Y) is bivariate normal.

Example 2: The j.p.m.f of (X, Y) is given below:

$Y \backslash X$	-1	1
0	$\frac{1}{8}$	$\frac{3}{8}$
1	$\frac{2}{8}$	$\frac{2}{8}$

Find the correlation coefficient between X and Y

Solution : Computation of marginal p.m.fs

$Y \backslash X$	-1	1	$g(y)$
0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{4}{8}$
1	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{4}{8}$
$p(x)$	$\frac{3}{8}$	$\frac{5}{8}$	1

We have

$$E(X) = \sum x p(x) = (-1) \times \frac{3}{8} + 1 \times \frac{5}{8} = -\frac{3}{8} + \frac{5}{8} = \frac{2}{8} = \frac{1}{4},$$

$$E(X^2) = \sum x^2 P(x) = (-1)^2 \frac{3}{8} + 1^2 \times \frac{5}{8} = \frac{3}{8} + \frac{5}{8} = 1, \text{ then}$$

$$V(X) = E(X^2) - (E(X))^2 = 1 - \left(\frac{1}{4}\right)^2 = 1 - \frac{1}{16} = \frac{15}{16}$$

Similarly, $E(Y) = \sum y g(y) = 0 \times \frac{4}{8} + 1 \times \frac{4}{8} = \frac{4}{8} = \frac{1}{2}$

$E(Y^2) = \sum y^2 g(y) = 0^2 \times \frac{4}{8} + 1^2 \times \frac{4}{8} = \frac{4}{8} = \frac{1}{2}$ and

$V(Y) = E(Y^2) - (E(Y))^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$

Further, $E(XY) = 0 \times (-1) \times \frac{1}{8} + 0 \times 1 \times \frac{3}{8} + 1 \times (-1) \times \frac{2}{8} + 1 \times 1 \times \frac{2}{8} = 0$

Thus, $cov(X, Y) = E(XY) - E(X)E(Y)$

$$0 - \frac{1}{4} \times \frac{1}{2} = -\frac{1}{8}$$

$$\therefore \rho(X, Y) = \frac{cov(X, Y)}{\sqrt{V(X)} \sqrt{V(Y)}} = -\frac{\frac{1}{8}}{\sqrt{\frac{15}{16} \times \frac{1}{4}}} = -\frac{1}{\sqrt{15}} = -0.2582$$

Example 3: Two random variables X and Y have the joint probability density function

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

Find correlation coefficient between X and Y .

Solution: By symmetry in x and y we have $f_1(x) = f_2(y)$, $E(X) = E(Y)$ and $V(X) = V(Y)$

The m.p.d.f X is given by

$$f_1(x) = \int_0^1 f(x, y) dy = \int_0^1 (2 - x - y) dy = \frac{3}{2} - x$$

Thus, $f_1(x) = \begin{cases} \frac{3}{2} - x & , \quad \text{if } 0 < x < 1 \\ 0 & , \quad \text{otherwise} \end{cases}$

Consider.

$$E(X) = \int_0^1 x f_1(x) dx = \int_0^1 x \left(\frac{3}{2} - x \right) dx = \int_0^1 \left(\frac{3}{2}x - x^2 \right) dx = \frac{5}{12}$$

$$E(X^2) = \int_0^1 x f_1(x) dx = \int_0^1 x^2 \left(\frac{3}{2} - x \right) dx = \int_0^1 \left(\frac{3}{2}x^2 - x^3 \right) dx = \frac{1}{4}$$

Further,

$$\begin{aligned} E(XY) &= \int_0^1 \int_0^1 xy f(x, y) dx dy = \int_0^1 \int_0^1 xy (2 - x - y) dx dy \\ &= \int_0^1 y \left(\int_0^1 (2x - x^2 - xy) dx \right) dy = \int_0^1 y \left[2 \cdot \frac{x^2}{2} - \frac{x^3}{3} - \frac{yx^2}{2} \right]_0^1 dy \\ &= \int_0^1 y \left(1 - \frac{1}{3} - \frac{y}{2} \right) dy = \int_0^1 y \left(\frac{2}{3} - \frac{y}{2} \right) dy \\ &= \int_0^1 \left(\frac{2}{3}y - \frac{y^2}{2} \right) dy = \left[\frac{y^3}{3} - \frac{y^3}{6} \right]_0^1 = \frac{1}{3} - \frac{1}{6} = \frac{1}{6} \end{aligned}$$

$$\therefore E(XY) = \frac{1}{6}$$

$$\text{Thus, } V(X) = E(X^2) - (E(X))^2 = \frac{1}{4} - \left(\frac{5}{12} \right)^2 = \frac{1}{4} - \frac{25}{144} = \frac{36-25}{144} = \frac{11}{144}$$

$$\text{and } cov(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{6} - \left(\frac{5}{12} \right)^2 = \frac{1}{6} - \frac{25}{144} = \frac{24-25}{144} = -\frac{1}{144}$$

$\therefore$ The correlation coefficient is given by

$$\rho(X, Y) = \frac{cov(X, Y)}{\sqrt{V(X)} \sqrt{V(Y)}} = -\frac{\frac{1}{144}}{\sqrt{\frac{11}{144}} \sqrt{\frac{11}{144}}} = -\frac{\frac{1}{144}}{\frac{11}{144}} = -\frac{1}{11}$$

Bivariate Normal Distribution:

The bivariate normal distribution is a generalization of a normal distribution for a single value.

Let X and Y be two normally correlated variables with correlation coefficient ρ . Let $E(X) = \mu_1$, $V(X) = \sigma_1^2$, $E(Y) = \mu_2$ and $V(Y) = \sigma_2^2$.

Definition: The bivariate continuous random variable (X, Y) is said to follow bivariate normal distribution with parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and ρ if its p.d.f. is given by

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right\} \right];$$
$$-\infty < x, y, \mu_1, \mu_2 < \infty, \sigma_1 > 0, \sigma_2 > 0 \text{ and } -1 < \rho < 1.$$

Notation: $(X, Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Read as (X, Y) follows **bivariate normal distribution** with parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and ρ .

Note: The curve $z = f(x, y)$ which is the equation of a surface in three dimensions is called the **Normal correlation surface**.

Marginal p.d.fs of X and Y : The m.p.d.f of X is given by

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

Let $v = \frac{y-\mu_2}{\sigma_2}$, then $y = \mu_2 + \sigma_2 v$ and $dy = \sigma_2 dv$

Therefore,

$$f_1(x) = \frac{\sigma_2}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ \left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho v \left(\frac{x-\mu_1}{\sigma_1} \right) + v^2 \right\} \right] dv$$
$$= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2} \left(\frac{x-\mu_1}{\sigma_1} \right)^2 \right] \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ v - \rho \left(\frac{x-\mu_1}{\sigma_1} \right) \right\}^2 \right] dv$$

Let $\frac{1}{\sqrt{1-\rho^2}} \left[v - \rho \left(\frac{x-\mu_1}{\sigma_1} \right) \right] = t$. Then $dv = \sqrt{1-\rho^2} dt$

$$\begin{aligned}
\therefore f_1(x) &= \frac{1}{2\pi\sigma_1} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] \int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt \\
&= \frac{1}{2\pi\sigma_1} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] \cdot \sqrt{2\pi} \\
\Rightarrow f_1(x) &= \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] \text{ for } -\infty < x < \infty
\end{aligned}$$

Similarly, it can be shown that

$$f_2(y) = \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right] \text{ for } -\infty < y < \infty$$

Hence $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$.

Note: If $(X, Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$

Conditional p.d.fs of X and Y

The conditional probability density function (c.p.d.f.) of X for given Y is given by

$$\begin{aligned}
f_{1|2}(x|y) &= \frac{f(x, y)}{f_2(y)} \\
&= \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2(1-(1-\rho^2))\right\}\right] \\
&= \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1)^2 - 2\rho\frac{\sigma_1}{\sigma_2}(x-\mu_1)(y-\mu_2) + \frac{\sigma_1^2}{\sigma_2^2}\rho^2(y-\mu_2)^2\right\}\right] \\
&= \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1) - \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right\}^2\right]
\end{aligned}$$

$$\text{Therefore, } f_{1|2}(x|y) = \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1) - \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right\}^2\right]$$

which is the univariate normal distribution with mean

$$E(X|Y = y) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2) \text{ and}$$

$$V(X|Y = y) = \sigma_1^2(1 - \rho^2)$$

Thus, the c.p.d.f of X for fixed Y is given by

$$(X|Y = y) \sim N \left[\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (y - \mu_2), \sigma_1^2 (1 - \rho^2) \right]$$

Similarly, the c.p.d.f of Y for fixed $X = x$ is given by

$$f_{2|1}(y|x) = \frac{1}{\sigma_2 \sqrt{2\pi} \sqrt{1-\rho^2}} \exp \left[\frac{-1}{2(1-\rho^2)\sigma_2^2} \left\{ (y - \mu_2) - \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1) \right\}^2 \right], -\infty < y < \infty$$

Thus, the c.p.d.f of Y for fixed X is given by

$$(Y|X = x) \sim N \left[\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1), \sigma_2^2 (1 - \rho^2) \right]$$

Example 4: If $(X, Y) \sim BN(5, 10, 1, 25, \rho)$ where $\rho > 0$, find ρ when $P(4 < Y < 16|X = 5) = 0.954$

Solution:

Here $\mu_1 = 5, \mu_2 = 10, \sigma_1^2 = 1, \sigma_2^2 = 25$. We know that $(Y|X = x) \sim N[\mu, \sigma^2]$

where $\mu = \mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1)$ and $\sigma^2 = \sigma_2^2 (1 - \rho^2)$.

Here $\mu = 10 + \rho \times \frac{5}{1} (5 - 5) = 10$ and $\sigma^2 = 25(1 - \rho^2)$

Thus $(Y|X = 5) \sim N[10, 25(1 - \rho^2)]$. We want to find ρ so that $P(4 < Y < 16|X = 5) = 0.954$

Let $Z = \frac{Y - \mu}{\sigma} = \frac{Y - 10}{5\sqrt{1-\rho^2}} \sim N(0, 1) \Rightarrow P\left(\frac{4-10}{\sigma} < Z < \frac{16-10}{\sigma}\right) = 0.954$

$$\Rightarrow P\left(-\frac{6}{\sigma} < Z < \frac{6}{\sigma}\right) = 0.954 \Rightarrow P\left(0 < Z < \frac{6}{\sigma}\right) = 0.477$$

From standard normal table, we have $\frac{6}{\sigma} = 2 \Rightarrow \sigma = 3 \Rightarrow \sigma^2 = 9$

$$\Rightarrow 25(1 - \rho^2) = 9 \Rightarrow 1 - \rho^2 = \frac{9}{25} \Rightarrow \rho^2 = 1 - \frac{9}{25} = \frac{16}{25} \Rightarrow \rho = \frac{4}{5} = 0.8$$

Example 5: Find $cor(X, Y)$ for the jointly normal distribution

$$f(x, y) = \frac{1}{2\pi\sqrt{3}} \exp \left[-\frac{\{(2x - y)^2 + 2xy\}}{6} \right], -\infty < x, y < \infty$$

Solution: Given $(X, Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Then its p.d.f. is given by

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right\} \right] \quad (1)$$

We have

$$f(x, y) = \frac{1}{2\pi\sqrt{3}} \exp \left[-\frac{\{(2x - y)^2 + 2xy\}}{6} \right], -\infty < x, y < \infty$$

$$i.e., f(x, y) = \frac{1}{2\pi\sqrt{3}} \exp \left[-\frac{\{4x^2 + y^2 - 2xy\}}{6} \right] \quad (2)$$

Comparing (1) and (2), we get $\mu_1 = \mu_2 = 0$. Then (1) becomes

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[-\frac{\left\{ \frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} - 2\rho \frac{xy}{\sigma_1\sigma_2} \right\}}{2(1-\rho^2)} \right] \quad (3)$$

Comparing (2) and (3), we find

$$\sigma_1\sigma_2\sqrt{1-\rho^2} = \sqrt{3}, \sigma_1^2(1-\rho^2) = \frac{3}{4}, \sigma_2^2(1-\rho^2) = 3 \text{ and } \frac{\rho}{\sigma_1\sigma_2(1-\rho^2)} = \frac{1}{3}$$

On solving we get $\sigma_1^2 = 1, \sigma_2^2 = 4, \rho^2 = \frac{1}{4}$

$$\text{Thus } cor(X, Y) = \rho = \sqrt{\frac{1}{4}} = \pm \frac{1}{2}$$

Example 6: Determine the parameters of the bivariate normal distribution

$$f(x, y) = c \exp \left[-\frac{\{16(x-2)^2 - 12(x-2)(y+3) + 9(y+3)^2\}}{216} \right]$$

Solution: If $(X, Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[-\frac{\left\{ \left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right\}}{2(1-\rho^2)} \right]$$

Comparing these functions, we get

$$\mu_1 = 2, \mu_2 = -3, c = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}, \frac{16}{216} = \frac{1}{2(1-\rho^2)\sigma_1^2}$$

$$\frac{9}{216} = \frac{1}{2(1-\rho^2)\sigma_2^2}, \frac{12}{216} = \frac{2\rho}{2\sigma_1\sigma_2(1-\rho^2)}$$

$$\therefore (1-\rho^2)\sigma_1^2 = \frac{27}{4}, (1-\rho^2)\sigma_2^2 = 12, \sigma_1\sigma_2(1-\rho^2) = 18\rho$$

$$\Rightarrow (1-\rho^2)^2\sigma_1^2\sigma_2^2 = 81 = (18\rho)^2 \Rightarrow \rho^2 = \frac{1}{4},$$

Further, $\sigma_1 = 3$ and $\sigma_2 = 4$.

$$\text{Thus, } c = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} = \frac{1}{2\pi \times 3 \times 4 \sqrt{1-\frac{1}{4}}} = \frac{1}{12\pi\sqrt{3}}$$

$$\therefore (X, Y) \sim BN\left(2, 3, 9, 16, \frac{1}{2}\right)$$

Example 7: If $X \sim N(\mu, \sigma^2)$ and $(Y|x) \sim N(x, \sigma^2)$, show that

$$(X, Y) \sim BN(\mu, \mu, \sigma^2, 2\sigma^2, \rho) .$$

Solution: We are given that

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right], -\infty < x < \infty$$

$$g(y|x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{1}{2} \left(\frac{y-x}{\sigma} \right)^2 \right], -\infty < y < \infty$$

$$\therefore h(x, y) = g(y|x)f(x) = \frac{1}{2\pi\sigma^2} \exp \left[-\frac{1}{2} \left\{ \left(\frac{x-\mu}{\sigma} \right)^2 + \left(\frac{y-x}{\sigma} \right)^2 \right\} \right]$$

$$\text{Consider } \left(\frac{y-x}{\sigma} \right)^2 = \left(\frac{y-\mu+\mu-x}{\sigma} \right)^2 = \left(\frac{y-\mu}{\sigma} \right)^2 + \left(\frac{x-\mu}{\sigma} \right)^2 - 2 \left(\frac{x-\mu}{\sigma} \right) \left(\frac{y-\mu}{\sigma} \right)$$

Thus, $h(x, y) = \frac{1}{2\pi\sigma^2} \exp \left[-\frac{1}{2} \left\{ 2 \left(\frac{x-\mu}{\sigma} \right)^2 + \left(\frac{y-\mu}{\sigma} \right)^2 - 2 \left(\frac{x-\mu}{\sigma} \right) \left(\frac{y-\mu}{\sigma} \right) \right\} \right]$

The bivariate normal p.d.f. is given by

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ \left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right\} \right]$$

On comparing $h(x, y)$ with $f(x, y)$, we get

$$\sigma_1\sigma_2\sqrt{1-\rho^2} = \sigma^2 \quad , \quad \sigma_1^2(1-\rho^2) = \frac{1}{2}\sigma^2$$

$$\frac{\sigma_1\sigma_2(1-\rho^2)}{\rho} = \sigma^2 \quad , \quad \sigma_2^2(1-\rho^2) = \sigma^2, \quad \mu_1 = \mu_2 = \mu$$

On solving, we get $\rho^2 = \frac{1}{2}$ $\sigma_2^2 = 2\sigma^2$, $\sigma_1^2 = \sigma^2$.

Thus, $(X, Y) \sim BN \left(\mu, \mu, \sigma^2, 2\sigma^2, \frac{1}{\sqrt{2}} \right)$

Example 8: The variables X and Y are connected by the equation $aX + bY + c = 0$. Show that the correlation between them is -1 if signs of a and b are same and $+1$ if they are different signs.

Solution: Given $aX + bY + c = 0 \Rightarrow aE(X) + bE(Y) + c = 0$

$$\therefore a[X - E(X)] + b[Y - E(Y)] = 0 \Rightarrow [X - E(X)] = -\frac{b}{a}[Y - E(Y)]$$

$$\therefore \text{cov}(X, Y) = E[\{X - E(X)\}\{Y - E(Y)\}] = -\frac{b}{a}E(Y - E(Y))^2 = -\frac{b}{a}\sigma_Y^2 \text{ and}$$

$$\sigma_X^2 = E(X - E(X))^2 = \frac{b^2}{a^2}E(Y - E(Y))^2 = \frac{b^2}{a^2}\sigma_Y^2$$

$$\therefore \rho = \frac{\text{cov}(X, Y)}{\sigma_X \cdot \sigma_Y} = \frac{-\frac{b}{a}\sigma_Y^2}{\sqrt{\sigma_Y^2} \sqrt{\frac{b^2}{a^2}\sigma_Y^2}} = \frac{-\frac{b}{a}\sigma_Y^2}{\left| \frac{b}{a} \right| \sigma_Y^2} = \frac{-\frac{b}{a}}{\left| \frac{b}{a} \right|}$$

$$\therefore \rho = \frac{\text{cov}(X, Y)}{\sigma_X \cdot \sigma_Y} = \begin{cases} 1, & \text{if } a \text{ and } b \text{ have opposite signs} \\ -1, & \text{if } a \text{ and } b \text{ have same signs} \end{cases}$$

2.7 .Correlation coefficient and Bivariate Normal Distribution

Exercise:

1. Find the correlation coefficient between X and Y for each of the j.p.d.f. $f(x, y)$ of (X, Y) given below:

$$\begin{aligned} \text{(i)} \quad f(x, y) &= \begin{cases} \frac{3}{2}(x^2 + y^2) & , \quad 0 \leq x \leq 1, \quad 0 \leq y \leq 1 \\ 0 & , \quad \text{otherwise} \end{cases} \\ \text{(ii)} \quad f(x, y) &= \begin{cases} (x + y) & , \quad 0 \leq x, y \leq 1 \\ 0 & , \quad \text{otherwise} \end{cases} \\ \text{(iii)} \quad f(x, y) &= \begin{cases} 2xy & , \quad 0 < x < 1, \quad 0 < y < 1 \\ 0 & , \quad \text{otherwise} \end{cases} \end{aligned}$$

2. If X, Y and Z are uncorrelated r.vs with 0 mean and standard deviations 5, 12 and 9 respectively and $U = X + Y$ and $V = Y + Z$, then find the correlation coefficient between U and V .
3. If X, Y, Z are uncorrelated r.vs having same variance, find the correlation coefficient between $(X + Y)$ and $(Y + Z)$.
4. If the independent r.vs X and Y have variance 36 and 16 respectively, find the correlation coefficient between $(X + Y)$ and $(X - Y)$.

ANSWERS

1. (i) -0.2055 (ii) $-\frac{1}{11}$ (iii) 0.8

2. $\frac{48}{65}$

3. $\frac{1}{2}$

4. $\frac{5}{13}$