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 $\boldsymbol{R}$ be the set of real numbers. If $X: S \longrightarrow \boldsymbol{R}$, $i.e.$, 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 $a, b \in \boldsymbol{R}$. 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 consits of $2^3 = 8$ points as given below.

$$S = \{H,T\} \times \{H,T\} \times \{H,T\}$$

$$= \{HH,HT,TH,TT\} \times \{H,T\}$$

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

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,

$$X(HHH) = 3\ ,$$

$$X(HTH) = X(THH) = X(HHT) = 2$$

$$X(TTH) = X(HTT) = X(THT) = 1 \text{ and}$$

$$X(TTT) = 0$$

This shows that X is a random variable.

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

$\{TTT\}, \{TTH,HTT,THT\}, \{HTH,THH,HHT\}$ 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 $p : \boldsymbol{R} \rightarrow \boldsymbol{R}$ defined as

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

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

Probability Distribution: The set of all possible ordered pairs $\left(x_i, p(x_i)\right), i = 1,2, \dots, n, \dots$ 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	…	x_n
$p(x)$	p_1	p_2	p_3	…	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:

$$S = \{H,T\} \times \{H,T\} \times \{H,T\} = \{HH,HT,TH,TT\} \times \{H,T\}$$

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

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 $\boldsymbol{a}$	$\boldsymbol{X = a}$ $\{\boldsymbol{\omega \in S \mid X(\omega) = a}\}$	**No. of favourable cases**	$\boldsymbol{p(x) = P(X = a)}$
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 $\boldsymbol{R}$ 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) $f(x) \geq 0 \quad \forall x \in \boldsymbol{R}$ 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

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

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

Properties of c.d.f.

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

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)

$$P(1 < x < 2) = \int_1^2 f(x)dx$$

$$= \int_1^2 \frac{1}{2}x^2 dx = \frac{1}{9}\left[\frac{x^3}{3}\right]_1^2 = \frac{7}{27}$$

(iii) We have,

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

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

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

If $x \ge 3$, 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

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

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

$$= F(2) - F(1)$$

$$= \frac{2^3}{27} - \frac{1^3}{27} = \frac{8}{27} - \frac{1}{27} = \frac{7}{12}$$

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 are1,3,5, 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 dentoes the number of successes in two throws of a die, then X is a random variable which takes the values 0,1,2.

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

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

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

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

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

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

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

$$F(x) = \begin{cases} 0 & if \quad x < 0 \\ \frac{1}{4} & if \quad 0 \le x < 1 \\ \frac{3}{4} & if \quad 1 \le x < 2 \\ 1 & if \quad x \ge 2 \end{cases}$$

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 ${}^{52}C_2$.

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

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

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

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

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

Hence, the probability distribution of X 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**

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

(i) **Determine k.**

(ii) **Compute $P(x \le 1.5)$**

Solution:

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

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

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

$$\Rightarrow \int_{0}^{1} kx\,dx + \int_{1}^{2} k\,dx + \int_{2}^{3} -k(x-3)\,dx = 1$$

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

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

$$\Rightarrow k\left[\frac{1}{2} + 2 - 1 - \frac{9}{2} + 9 + 2 - 6\right] = 1$$

$$\Rightarrow 2k = 1 \Rightarrow k = \frac{1}{2}$$

(ii)

$$P(x \le 1.5) = \int_{-\infty}^{1.5} f(x)\,dx = \int_{0}^{1} f(x)\,dx + \int_{-\infty}^{1.5} f(x)\,dx$$

$$= k\int_{0}^{1} x\,dx + \int_{1}^{1.5} k\,dx = k\left[\frac{x^2}{2}\right]_0^1 + k[x]_1^{1.5} = k\left[\frac{1}{2} + \frac{1}{2}\right] = k = \frac{1}{2}$$

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 $2, 3, 4, \dots, 12$ 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 ${}^{24}C_2$.

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

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

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

Hence, the probability distribution of X is 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,

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

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

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

$$\Rightarrow \left[3b^2 - 2b^3\right] = \left[1 - 3b^2 + 2b^3\right]$$

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

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

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

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.

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

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 $P(\overline{A}\cap\overline{B})=\frac{1}{3}$**

Solution:

i.

$$P(A)=\int_0^2 f(x)dx=\int_0^1\frac{1}{9}(x+1)dx+\int_1^{\frac{3}{2}}\frac{4}{9}\left(x-\frac{1}{2}\right)dx+\int_{\frac{3}{2}}^2\frac{4}{9}\left(\frac{5}{2}-x\right)dx$$

$$=\frac{1}{9}\left[\frac{x^2}{2}+x\right]_0^1+\frac{4}{9}\left[\frac{x^2}{2}-\frac{x}{2}\right]_1^{\frac{3}{2}}+\frac{4}{9}\left[\frac{5x}{2}-\frac{x^2}{2}\right]_{\frac{3}{2}}^2$$

$$=\frac{1}{9}\left[\left(\frac{1}{2}+1\right)\right]+\frac{4}{9}\left[\left(\frac{9}{8}-\frac{3}{4}\right)\right]+\frac{4}{9}\left[(5-2)-\left(\frac{15}{4}-\frac{9}{8}\right)\right]=\frac{1}{2}$$

$$P(A\cap B)=P(1\le x\le 2)=\int_1^2 f(x)\,dx$$

$$=\int_1^{\frac{3}{2}}\frac{4}{9}\left(x-\frac{1}{2}\right)dx+\int_{\frac{3}{2}}^{2}\frac{4}{9}\left(\frac{5}{2}-x\right)dx$$

$$=\frac{4}{9}\left[\frac{x^2}{2}-\frac{x}{2}\right]_1^{\frac{3}{2}}+\frac{4}{9}\left[\frac{5x}{2}-\frac{x^2}{2}\right]_{\frac{3}{2}}^{2}$$

$$=\frac{4}{9}\left[\left(\frac{9}{8}-\frac{3}{4}\right)-\left(\frac{1}{2}-\frac{1}{2}\right)\right]+\frac{4}{9}\left[(5-2)-\left(\frac{15}{4}-\frac{9}{8}\right)\right]=\frac{1}{3}$$

$$P(B)=\int_1^3 f(x)\,dx$$

$$=\int_1^{\frac{3}{2}}\frac{4}{9}\left(x-\frac{1}{2}\right)dx+\int_{\frac{3}{2}}^{2}\frac{4}{9}\left(\frac{5}{2}-x\right)dx+\int_2^3\frac{1}{9}(4-x)\,dx$$

$$=\frac{4}{9}\left[\frac{x^2}{2}-\frac{x}{2}\right]_1^{\frac{3}{2}}+\frac{4}{9}\left[\frac{5x}{2}-\frac{x^2}{2}\right]_{\frac{3}{2}}^{2}+\frac{1}{9}\left[4x-\frac{x^2}{2}\right]_2^3$$

$$=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{3}{6}=\frac{1}{2}$$

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

ii. $P(\overline{A}\cap\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)$

$$=1-[P(A)+P(B)-P(A\cap B)]=1-\frac{2}{3}=\frac{1}{3}$$

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

$$P(\overline{\mathrm{A}}\cap\overline{\mathrm{B}})=P(x>3)=\int_3^{\infty}f(x)\,dx=\int_3^6 f(x)\,dx+\int_3^{\infty}f(x)\,dx$$

$$=\frac{1}{9}\int_3^6 dx+0=\frac{1}{9}[6-3]=\frac{3}{9}=\frac{1}{3}$$

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, form 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 accidently 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

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

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

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

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

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

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

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

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

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

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:

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

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

$\therefore$ Variance $= \sigma^2 = E(X^2) - \left(E(X)\right)^2 = p - p^2 = p(1 - p) = pq$

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,…,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

$$P(SSFSFFFS \ldots FSF) = P(S)\cdot P(S)\cdot P(F)\cdot P(S)\cdot P(F)\cdot P(F)\cdot P(F)\cdot P(S)\cdot \ldots P(F)\cdot P(S)\cdot P(F)$$

$$= p\ p\ q\ p\ q\ q\ q\ p \ldots q\ p\ q$$

$$= \underbrace{p\cdot p\cdot \ldots \cdot p}_{(x\ times)}\cdot \underbrace{q\cdot q\cdot \ldots \cdot q}_{(n-x\ times)} = p^x q^{n-x}$$

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

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

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 $p = P(head) = \frac{1}{2}$.

Thus $p(x) = P(X = x) = \binom{4}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{4-x}$

$$= \binom{4}{x} \left(\frac{1}{2}\right)^4 = \binom{4}{x} \left(\frac{1}{16}\right) \text{ for } x = 0,1,2,3,4.$$

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: $\mathbf{A}$ and $\mathbf{B}$ play a game in which their chances of winning are in the ratio $\mathbf{3 : 2}$. 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 $p = P(A\ winning) = \frac{3}{5}$ 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:

$$P(A \text{ winning at least 3 out of 5 games}) = P(X \geq 3)$$

$$= P(X = 3) + P(X = 4) + P(X = 5)$$

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

$$= \frac{3^3}{5^5}[10 \times 4 + 5 \times 3 \times 2 + 1 \times 9]$$

$$= \frac{27 \times (40+30+9)}{3125} = 0.68$$

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

- **(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 $\mathbf{\frac{2}{3}}$?**

Solution: Let X be the no. of times a man hitting the target in 7 fires. Here $p = P(man\ hitting\ the\ target) = \frac{1}{4}$ and $n = 7$. Then $X \sim B\left(7, \frac{1}{4}\right)$ and $p(x) = \binom{7}{x}\left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{7-x}$ for $x = 0,1,2, \dots ,7$.

(i) $P(at\ least\ two\ hits) = P(X \geq 2) = 1 - [P(X=0) + P(X=1)]$

$$= 1 - [p(0) + p(1)]$$

$$= 1 - \left\{\binom{7}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^7 + \binom{7}{1}\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^6\right\} = \frac{4547}{8192} = 0.55$$

(ii) Find n such that $P(X \geq 1) > \frac{2}{3}$

$$\Rightarrow 1 - P(X=0) > \frac{2}{3}$$

$$\Rightarrow -1 + P(X=0) < -\frac{2}{3}$$

$$\Rightarrow P(X=0) < \frac{1}{3}$$

$$\Rightarrow \binom{n}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^n < \frac{1}{3}$$

$$\Rightarrow \left(\frac{3}{4}\right)^n < \frac{1}{3}$$

$$\Rightarrow n[\log 3 - \log 4] < \log 1 - \log 3$$

$$\Rightarrow n[\log 4 - \log 3] > \log 3$$

$\Rightarrow n > \frac{\log 3}{\log 4 - \log 3} = 3.8$,since n cannot be fractional, the required number of shots is 4.

Mean of Binomial Distribution:

$$\mu = E(X) = \sum_{x=0}^{n} x\, p(x) = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=1}^{n} x \left(\frac{n}{x}\right)\binom{n-1}{x-1} p^x q^{n-x} = np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} q^{n-x}$$

$$= np(q+p)^{n-1} = np$$

$$\Rightarrow \mu = np$$

Variance of Binomial Distribution:

$\sigma^2 = V(X) = npq$ **(See $\boldsymbol{P_1}$ 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 $\mathbf{4}$ and $\frac{\mathbf{4}}{\mathbf{3}}$ respectively. Find $\boldsymbol{P(X \geq 1)}$.

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

Hence $4(1-p) = \frac{4}{3} \Longrightarrow 1 - p = \frac{1}{3} \Longrightarrow p = \frac{2}{3}$ 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 $n \to \infty$ and $p \to 0$ such that $\lambda = np$ fixed, then $\binom{n}{x} p^x (1-p)^{n-x} \longrightarrow \frac{e^{-\lambda}\lambda^x}{x!}$ which is the p.m.f. of Poisson distribution (**See $\boldsymbol{P_2}$**). 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

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

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

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

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:

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

Let $X=$ Number of errors per page

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

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

P(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:

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

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

$$= e^{-\lambda}\lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \lambda = e^{-\lambda}\lambda^2 e^{\lambda} + \lambda$$

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

The variance of Poisson distribution is given by

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

$\Longrightarrow V(X) = \lambda$

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

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

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

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

$\Longrightarrow \lambda(\lambda - 2) = 0 \Longrightarrow \lambda = 0, 2 \Longrightarrow \lambda = 2$ ($\because$ $\lambda = 0$ is not admissible)

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

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

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

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

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

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

Note:

1. $\binom{x+r-1}{r-1} = \binom{x+r-1}{x} \quad \left(\because \binom{n}{r} = \binom{n}{n-r}\right)$

$$= \frac{(x+r-1)(x+r-2)\ldots(r+1)r}{x!}$$

$$= (-1)^r \frac{(-r)(-r-1)\ldots(-r-x+2)(-r-x+1)}{x!} = (-1)^x \binom{-r}{x}$$

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

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

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

$$\mu = E(x) = \sum_{x=0}^{\infty} x\, p(x) = \sum_{x=1}^{\infty} x\, p(x) = \sum_{x=1}^{\infty} x \binom{-r}{x} p^r (-q)^x$$

$$= p^r \sum_{x=1}^{\infty} x \left(\frac{-r}{x}\right) \binom{-r+1}{x-1} (-q)^x = (-r)(p^r)(-q) \sum_{x=1}^{\infty} \binom{-r+1}{x-1} (-q)^{x-1}$$

$$= (-r)(-q)p^r(1-q)^{-(r+1)}$$

$$= rqp^r p^{-(r+1)} = \frac{rq}{p}$$

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 $y = r, r+1, \ldots$ 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

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

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 $p =$(defective)$= \frac{5}{100} = 0.05.$

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

$$P(Y \geq 4) = 1 - \sum_{y=2}^{3} P(Y = y) = 1 - \{P(Y = 2) + P(Y = 3)\}$$

$$= 1 - \{(0.05)^2 + 2(0.05)^2(0.95)\} = 0.9928$$

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

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

Notation: $X \sim GD(r)$

Mean and variance of GD

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

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

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

Further, $E(Y) = E(X) + 1 = \frac{1}{p}$ 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.

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

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

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

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

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 $\boldsymbol{P_4}$)**.

P1:

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

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

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

$$= \sum_{x=2}^{n} x(x-1)\, p(x) + \sum_{x=0}^{n} x\, p(x)$$

$$= \sum_{x=2}^{n} x(x-1) \binom{n}{x} p^x q^{n-x} + np$$

$$= \sum_{x=2}^{n} x(x-1) \left(\frac{n}{x}\right)\left(\frac{n-1}{x-1}\right)\binom{n-2}{x-2} p^x q^{n-x} + np$$

$$= n(n-1)\, p^2 \sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2} q^{n-x} + np$$

$$= n(n-1)\, p^2 \left(q + p\right)^{n-2} + np$$

$$\Rightarrow E\left(X^2\right) = n(n-1)\, p^2 + np \;\; and \;\; hence$$

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

P2:

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

Proof:

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

$$= \binom{n}{x} \left(\frac{p}{1-p}\right)^x (1-p)^n$$

$$= \frac{n(n-1)(n-2)\ldots(n-x+1)}{x!} \frac{\left(\frac{\lambda}{n}\right)^x}{\left(1-\frac{\lambda}{n}\right)^x} \left(1-\frac{\lambda}{n}\right)^n$$

$$= \frac{n^x\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{x-1}{n}\right)}{x!\left(1-\frac{\lambda}{n}\right)^x} \frac{\lambda^x}{n^x} \left(1-\frac{\lambda}{n}\right)^n$$

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

P3:

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

Proof:

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

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

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

$$= \sum_{x=2}^{\infty} x(x-1)\, p(x) + \sum_{x=0}^{\infty} x\, p(x)$$

$$= \sum_{x=2}^{\infty} x(x-1) \binom{-r}{x} p^r (-q)^x + \frac{rq}{p}$$

$$= \sum_{x=2}^{\infty} x(x-1) \left(\frac{-r}{x}\right)\left(\frac{-r-1}{x-1}\right)\binom{-r+2}{x-2} p^r (-q)^x + \frac{nq}{p}$$

$$= r(r+1)\, p^r (-q)^2 \sum_{x=2}^{\infty} \binom{-r+2}{x-2} (-q)^{x-2} + \frac{nq}{p}$$

$$= r(r+1)\, p^r q^2 (1-q)^{-(r+2)} + \frac{nq}{p} \quad \left(\because \sum_{x=2}^{\infty} \binom{-r+2}{x-2} (-q)^{x-2} = (1-q)^{-(r+2)} \right)$$

$$= r(r+1)\, p^r q^2 p^{-(r+2)} + \frac{nq}{p}$$

$$\Rightarrow E(X^2) = r(r+1)\frac{q^2}{p^2} + \frac{nq}{p}$$

$$Thus,\ V(X) = E(X^2) - (E(X))^2 = r(r+1)\frac{q^2}{p^2} + \frac{nq}{p} - \frac{r^2q^2}{p^2}$$

$$= \frac{r^2q^2}{p^2} + \frac{rq^2}{p^2} + \frac{nq}{p} - \frac{r^2q^2}{p^2}$$

$$= \frac{rq}{p^2}(q+p)$$

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

P4:

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

$$\boldsymbol{\mu = E(X) = \frac{nM}{n}} \quad \textbf{and}$$

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

when $\boldsymbol{\min(n, M) = n.}$

Proof:

We have $p(x) = \dfrac{\dbinom{M}{x}\dbinom{N-M}{n-x}}{\dbinom{N}{n}}$

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

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

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

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

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

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

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

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

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

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

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

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

$$\text{Thus, } V(X) = E\left(X^2\right) - \left(E(X)\right)^2$$

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

$$= \frac{NM(N-M)(N-n)}{N^2(N-1)} \text{ (on simplification)}$$

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

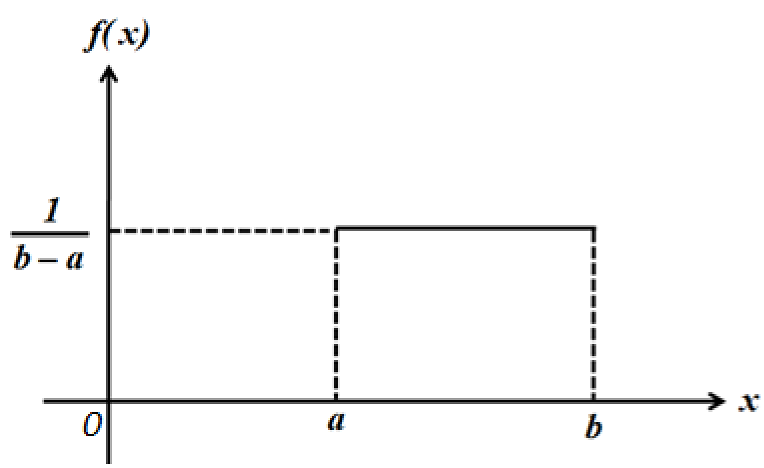
The **continuous probability distributions** are used in a number of applications in *engineering*. For example in *error analysis*, given a set of data or probability distribution, it is possible to estimate the probability that a measurement (temperature, pressure, flow rate) will fall within a desire 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 of 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

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

Notation: $\boldsymbol{X \sim U}(a, b)$. 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

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

The mean of X is given by

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_a^b x\, f(x)\, dx = \int_a^b \frac{x}{b-a}\, dx$$

$$= \left[\frac{x^2}{2(b-a)}\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2} \implies \mu = \frac{b+a}{2}$$

Now, $E(X^2) = \displaystyle\int_{-\infty}^{\infty} x^2\, f(x)\, dx = \int_a^b x^2\, f(x)\, dx = \int_a^b \frac{x^2}{b-a}\, dx$

$$= \left[\frac{x^3}{3(b-a)}\right]_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{b^2+ab+a^2}{3}$$

$$\implies E(X^2) = \frac{b^2+ab+a^2}{3}$$

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

$$= \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4}$$

$$= \frac{b^2-2ab+a^2}{12}$$

$$\implies \sigma^2 = \frac{(b-a)^2}{12}$$

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 $X \sim U(5, 10)$.

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

Normal Distribution

The normal distribution was first discovered by **De – Movire** 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

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

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

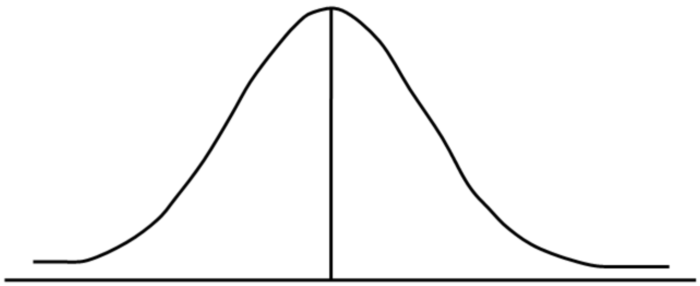
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} exp\left\{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right\}dt$$

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.



Normal probability curve

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

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

and its c. d. f. is given by

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

Area Property of Normal Distribution

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

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

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

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

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

$$= 2\int_{0}^{1} g(z)\,dz \qquad \text{(by symmetry)}$$

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

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

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

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

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

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

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

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

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

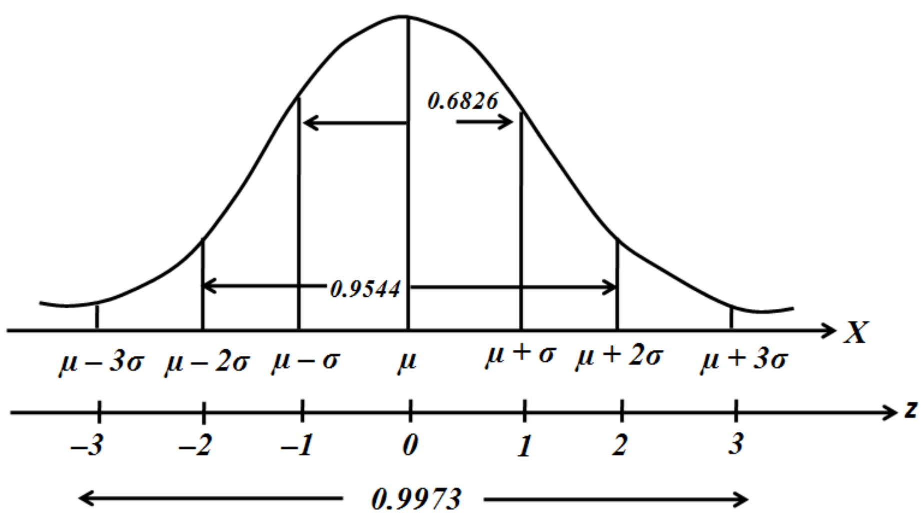
Thus, the probability that a normal variate X lies outside the range $\mu \pm 3\sigma$ is given by

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

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

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

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

$$= \mu . \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz = \mu \times 1 + 0$$

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

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

Variance of Normal distribution

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

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

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

$$V(X) = \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} 2t.e^{-t} \frac{dt}{\sqrt{2t}} = \frac{2\sigma^2}{\sqrt{\pi}} \int_{0}^{\infty} e^{-t}.t^{\frac{3}{2}-1} dt$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) \qquad \text{(Gamma function)}$$

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

$$\implies V(X) = \sigma^2$$

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

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:

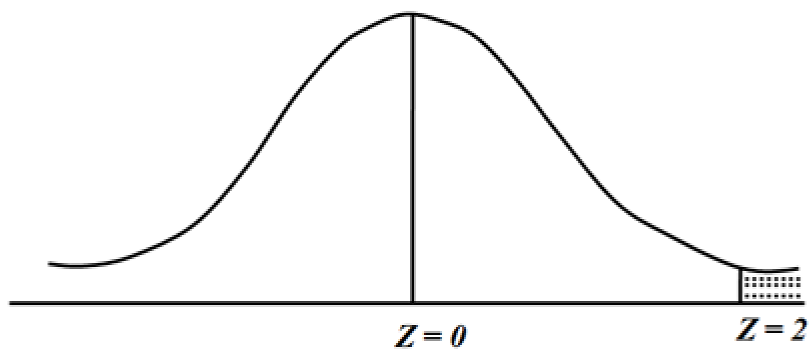
(a) it is given that $\mu = 12$ and $\sigma = 4$ $i.e.$, $X \sim N(12, 16)$

(i) Let $Z = \frac{X-12}{4}$. then $P(X \geq 20) = P\left(\frac{X-12}{4} \geq \frac{20-12}{4}\right)$

$= P(Z \geq 2) = 0.5 - P(0 \leq Z \leq 2)$

$= 0.5 - 0.4772$ (from table)

$= 0.0228$



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

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

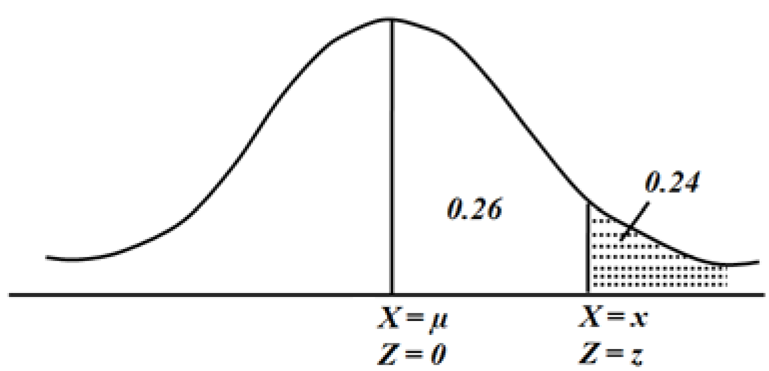
$= P(-3 \leq Z \leq 0)$

$= P(0 \leq Z \leq 3)$ (by symmetry)

$= 0.4986$ (from table)

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

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



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

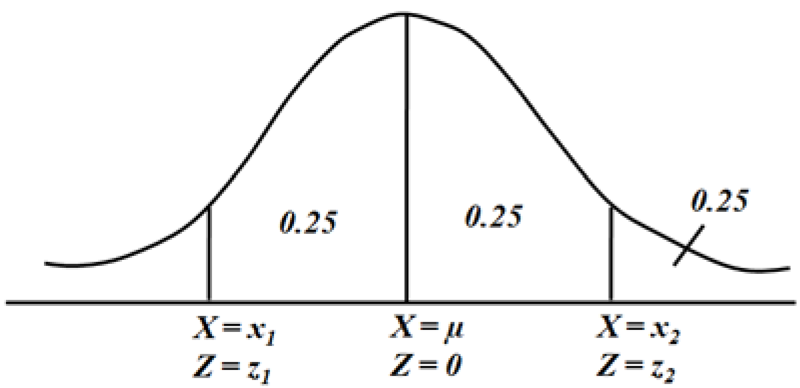
$\therefore$ From normal tables, corresponding to probability 0.26, value of
$z = 0.71$(approximately)

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

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

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

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

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

Similarly, $z_1 = -z_2 \Longrightarrow \frac{x_2-12}{4} = -0.67 \Longrightarrow x_1 = 12 - 4 \times 0.67 = 9.32$

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 $X \sim N(1000, 40000)$

(i) Find $P(X < 800) = P\left(\frac{X-1000}{200} < \frac{800-1000}{200}\right)$

$= P(Z < -1)$, where $Z = \frac{X-1000}{200} \sim N(0,1)$

$= P(Z > 1) = 0.5 - P(0 < Z < 1)$

$= 0.5 - 0.3413 = 0.1587$

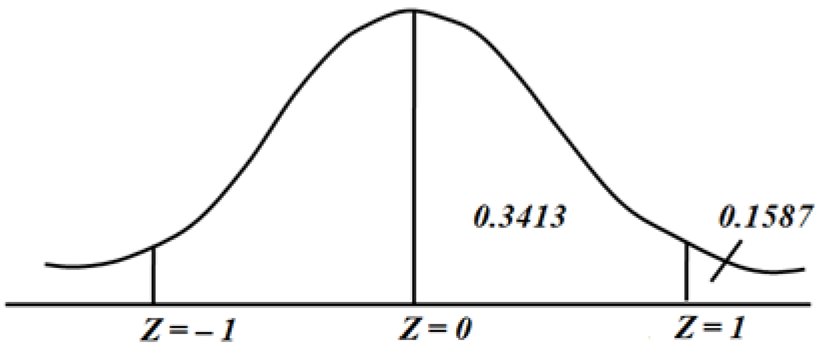
$\therefore$ Out of 10,000 bulbs, number of bulbs which fail in the first 800 hours is

$10{,}000 \times 0.1587 = 1{,}587.$

(ii) Find $P(800 < X < 1200) = P\left(\frac{800-1000}{200} < \frac{X-1000}{200} < \frac{1200-1000}{200}\right)$

$= P(-1 < Z < 1) = 2.P(0 < Z < 1)$

$= 2 \times 0.3413 = 0.6826$



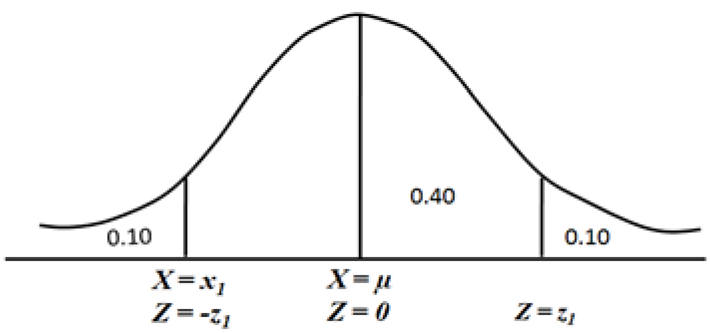
Hence, the expected number of bulbs with life between 800 and 1200 hours of burning life is $10{,}000 \times 0.6826 = 6{,}826$.

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

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

From table corresponding to probability 0.40, we have

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



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

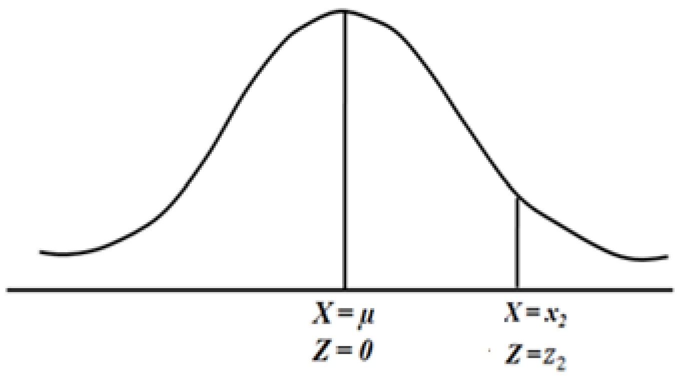
(b) Let 10% of the bulbs be still burning after x_2 hours of burning life. Then we have

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

$\Rightarrow P(Z > z_2) = 0.10$,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, $P(k_1 \le X \le k_2) = \sum_{x=k_1}^{k_2} \binom{n}{x} p^x q^{n-k}$ for some non-negative integers k_1 and k_2 such that $k_1 < k_2$. Since the binomial coefficient $\binom{n}{x}$ grows quite rapidly with n, it is very difficult to compute $P(k_1 \le X \le k_2)$ 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,

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

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

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

$$P(4 \le X \le 7) = \sum_{x=4}^{7} p(x) = \sum_{x=4}^{7} \binom{n}{x} \left(\frac{1}{2}\right)^{x} \left(\frac{1}{2}\right)^{10-x} = \sum_{x=4}^{7} \binom{n}{x} \left(\frac{1}{2}\right)^{10}$$

$$= \frac{\binom{10}{4} + \binom{10}{5} + \binom{10}{6} + \binom{10}{7}}{1024} = \frac{792}{1024} = 0.7734$$

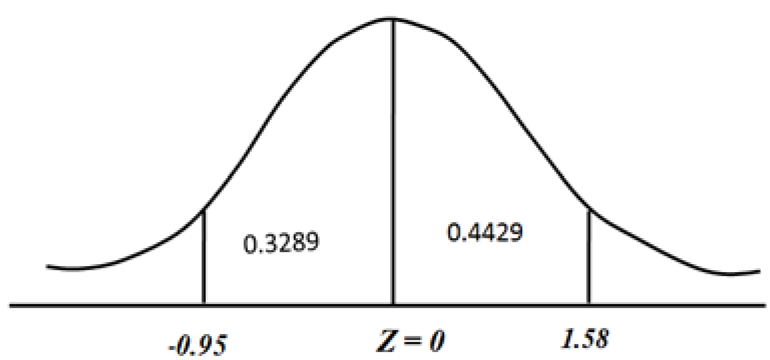
(b) The discrete binomial probability distribution is approximated to continuous normal probability distribution. The integers 4,5,6,7 lie in the interval (3.5 to 7.5). Thus,

$$P(4 \le X \le 7) = P(3.5 \le X \le 7.5) = P\left(\frac{3.5-5}{\sqrt{25}} \le Z \le \frac{7.5-5}{\sqrt{2.5}}\right)$$

$$= P(-0.95 \le Z \le 1.58) = P(-0.95 \le Z \le 0) + P(0 \le Z \le 1.58)$$

$$= P(0 \le Z \le 0.95) + P(0 \le Z \le 1.58)$$

$$= 0.3289 + 0.4429 = 0.7718$$



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

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:

a. The probability that you will wait more than 5 minutes before Ramu is done with the call.
b. 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

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

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

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

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 $X \le t + s$ given that $X > t$ for some nonnegative additional time s.

Thus,

$$P(X \le s + t | X > t) = \frac{P(X \le s+t, X>t)}{P(X>t)} = \frac{P(t<X\le s+t)}{P(X>t)} = \frac{F(s+t)-F(t)}{1-F(t)}$$

$$= \frac{\left(1-e^{-\lambda(s+t)}\right)-\left(1-e^{-\lambda t}\right)}{e^{-\lambda t}} = \frac{e^{-\lambda t}-e^{-\lambda(s+t)}}{e^{-\lambda t}} = 1 - e^{-\lambda s} = F(s) = P(X \le s)$$

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

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.

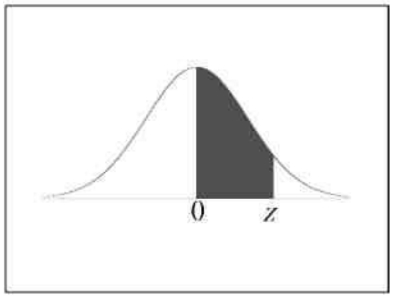
Let X denote the number of arrivals during an interval of t seconds.

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



z	.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	.4 332	.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

2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998

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 $\sigma^2 = 16$.

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

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

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

$$p(x) = P(X > 70) = P\left(Z > \frac{70-65}{5}\right) = P(Z > 1)$$

$$= 0.5 - P(0 < Z < 1)$$

$$= 0.5 - 0.3413 \qquad \text{(Use table)}$$

$$= 0.1587$$

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*.

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

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

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

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

$Y \sim B(6, P)$ 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

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

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 deviatioin 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 uniformaly 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 $500\ hrs$.
   (i) Find the probability that it will give additional $600\ hrs$ life given that the component has been working for the last $300 hrs$.
   (ii) 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.

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

   Find the probability of one such tube lasting for
   (i) at least $800 hrs$.
   (ii) at most $1200 hrs$.
   (iii) 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) $e^{-\frac{1}{4}} - e^{-1}$