1.3 SPECIAL TYPE OF DISTRIBUTION

1.3.1. Introduction :

A random variable is said to have a special type of distribution if its pmf / pdf gets a special form.

Seven such type of distributions are discussed in this chapter. Among these the two discrete distributions Binomial and Poisson are very well known and useful in everyfield of life. The continuous distribution 'Normal distribution' has also a wide range of application in sociology, market research and specially in industry. Definition of each of the distribution is first given and their properties and field of fitness are illustrated in this chapter.

1.3.2. Binomial Distribution.

A discrete random variable X is said to have a binomial distribution with parameters $p(0<p<1)$ and n (a positive integer) if its distribution is given by

X	:	0	1	2	...	...	n
f_i	:	f_0	f_1	f_2	...	...	f_n

where the pmf $f_i = P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$, $i = 0,1,2,\cdots,n$

For instant, one probability mass

$$f_1 = \binom{n}{1} p(1-p)^{n-1} = np(1-p)^{n-1} \text{ etc.}$$

Note : (1) The pmf f_i satisfy the two fundamental properties $f_i \geq 0$ and $\sum_{i=0}^{n} f_i = 1$, which can be easily verified.

(2) When the random variable X has a binomial distribution with parameters n, p we write $X \sim b(n,p)$ and we say X is a binomial variate.

(3) The significance of the parameters n and p would be given in subsequent theorem.

Cases where Binomial Distribution fits.

Let A be an event of a random experiment E. We call "the event A" = 'success' and "the event Not A" = 'failure'. Let p = probability of 'success' in a single trial of E. E be repeated, independently, n times. Let X = number of succss in n trials. Then X may assume the values 0, 1, 2, ... n. For example, the event $(X=3)$ means " 3 success in n trials". It can be shown that $P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$

Thus the distribution of X becomes

X	:	0	1	2	3	...	...	n
f_i	:	f_0	f_1	f_2	f_3	...	...	f_n

where $f_i = P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$ which is the pmf of Binomial distribution.

Thus 'No. of Success' in n trials is a Binomial variate with parameter p = probability of success in a single trial and n = number of trial.

Illustration. The efficiency of a fighter-plane is such that the probability of a bomb hitting a target is $2/5$. The fighter is assigned to completely destroy a camp of enemy-side. The plane carries 6 bombs, i.e., 6 bombs can be aimed at the camp. Here throwing a bomb is the experiment. It can be repeated 6 times; ' A bomb hits the camp' = Success and X = number of success in 6 trials. Then X has the Binomial distribution,

X	:	0	1	2	3	4	5	6
f_i	:	f_0	f_1	f_2	f_3	f_4	f_5	f_6

where $f_i = \binom{6}{i}\left(\frac{2}{5}\right)^i \left(1-\frac{2}{5}\right)^{6-i} = \binom{6}{i}\left(\frac{2}{5}\right)^i \left(\frac{3}{5}\right)^{6-i}$

If it is known that at least four direct hits are necessary to destroy the camp then the probability of complete destruction of the camp .

$$= P(X \geq 4) = P(X=4) + P(X=5) + P(X=6) = f_4 + f_5 + f_6$$

$$= \binom{6}{4}\left(\frac{2}{5}\right)^4\left(\frac{3}{5}\right)^2 + \binom{6}{5}\left(\frac{2}{5}\right)^5\left(\frac{3}{5}\right) + \binom{6}{6}\left(\frac{2}{5}\right)^6 = \frac{112}{625}$$

Theorem. If X has Binomial Distribution with parameter n and p then (i) its mean is np (ii) its variance is npq where $q = 1-p$. [W.B.U.Tech 2005]

Proof. Here $X : 0\ 1\ 2 \cdots n$

and its pmf is $f_i = P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$

(i) Mean $= E(X) = \sum_{i=1}^{n} i f_i \quad = \sum_{i=1}^{n} i\binom{n}{i} p^i (1-p)^{n-i}$

$$= np\sum_{i=1}^{n}\binom{n-1}{i-1} p^{i-1}(1-p)^{n-i}$$

$$= np\sum_{r=0}^{n-1}\binom{n-1}{r} p^r (1-p)^{n-1-r}\text{, replacing } i-1 \text{ by } r$$

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

(ii) Now

$$E\{X(X-1)\} = \sum_{i=0}^{n} i(i-1)\binom{n}{i} p^i (1-p)^{n-i}$$

$$= \sum_{i=0}^{n} i(i-1)\frac{n(n-1)}{i(i-1)}\binom{n-2}{i-2} p^i (1-p)^{n-i}$$

$$= n(n-1)p^2\sum_{i=2}^{n}\binom{n-2}{i-2} p^{i-2}(1-p)^{n-i}$$

$$= n(n-1)p^2\sum_{r=0}^{n-2}\binom{n-2}{r} p^r (1-p)^{n-2-r}\text{, replacing } i-2 \text{ by } r$$

$$= n(n-1)p^2(p+1-p)^{n-2} = n(n-1)p^2.$$

$$\therefore Var(X) = E\{X(X-1)\} - m(m-1)$$

$$= n(n-1)p^2 - np(np-1) = np(1-p) = npq \text{ where } q = 1-p$$

$\therefore$ standard deviation $\sigma = \sqrt{npq}$.

Illustration. An unbiased die is tossed four times. Let 'multiple of three' be success; otherwise it is failure. Here p = probability of sucess in a single trial $= \frac{2}{6} = \frac{1}{3}$. Let X = number of 'multiple of three' appeared among these four trial. Then, as we discussed before, X has Binomial distribution with parameter $n = 4$ and $p = \frac{1}{3}$. The expected number of 'multiple of 3' = mean of $X = 4\times\frac{1}{3} = \frac{4}{3}$. The standard deviation of

$$X = \sqrt{4\times\frac{1}{3}\left(1-\frac{1}{3}\right)} = \sqrt{4\times\frac{1}{3}\times\frac{2}{3}} = \frac{2\sqrt{2}}{3}.$$

Illustrative Examples

Ex. 1. *The mean and s.d of a binomial distribution are respectively 4 and* $\sqrt{\frac{8}{3}}$. *Find the values of n and p. Hence evaluate* $P(X=0)$. [W.B.U.Tech 2006]

We know the mean and s.d of a binomial variate are respectively np and $\sqrt{np(1-p)}$.

$\therefore np = 4$ or, $np(1-p) = \frac{8}{3}$

$\therefore 4(1-p) = \frac{8}{3} \Rightarrow p = \frac{1}{3}$

$\therefore n = 4\times 3 = 12$

$$P(X=0) = f_0 = {}^{12}C_0\left(\frac{1}{3}\right)^0\left(1-\frac{1}{3}\right)^{12-0} = \left(\frac{2}{3}\right)^{12}.$$

Ex. 2. *Comment on the statement "a binomial variate has mean 4 and s.d 3".*

Here, $np = 4$ and $\sqrt{np(1-p)} = 3$ i.e. $np(1-p) = 9$

$\therefore\ 4(1-p)=9 \quad \therefore\ 1-p=\frac{9}{4}\ \therefore\ p=-\frac{5}{4}$ which is not possible since $0<p<1$. So the statement is false.

Ex. 3. *If the mean of a binomial distribution is 3 and the variance is $\frac{3}{2}$, find the probability of obtaining at most 3 success.* [W.B.U.Tech 2007]

Let X be the r.v corresponding to the number of success. Then the *pmf* of X is

$$f_i=P(X=i)={}^nC_i p^i(1-p)^{n-i},\ i=0,1,2,\cdots,n$$

$\therefore$ mean $=np=3$, Variance $=np(1-p)=\frac{3}{2}$

or, $\frac{np(1-p)}{np}=\frac{3}{2}\times\frac{1}{3} \quad \therefore\ p=\frac{1}{2}$ and so $n=6$

Now probability of at most 3 success $=P(X\le 3)$

$=P(X=0)+P(X=1)+P(X=2)+P(X=3)$

$$={}^6C_0\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^6+{}^6C_1\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^5+{}^6C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^4+{}^6C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^3=\frac{21}{32}$$

Ex. 4. *The probability that a pen manufactured by a company will be defective is $\frac{1}{10}$. If 12 such pens are manufactured, find the prob. that (i) exactly two will be defecive (ii) none will be defective (iii) at least two will be defective*

Let 'defective pen' = success

$\therefore$ Prob of success in a single trial $p=\frac{1}{10}$.

The experiment is repeated 12 times.

Let the random variable X = number of defective pens. Then X is a binomial variate where the parameters $n=12$ and $p=\frac{1}{10}$.

$\therefore$ Its pmf is $f_i=P(X=i)=\binom{12}{i}\left(\frac{1}{10}\right)^i\left(1-\frac{1}{10}\right)^{12-i}$

$\therefore$ (i) The required probability $=P(X=2)=\binom{12}{2}\left(\frac{1}{10}\right)^2\left(1-\frac{1}{10}\right)^{12-2}$

$$=\frac{66\times 9^{10}}{10^{12}}=0.2301.$$

(ii) required probality $P(X=0)=\binom{12}{0}\left(\frac{1}{10}\right)^0\left(1-\frac{1}{10}\right)^{12-0}=\left(\frac{9}{10}\right)^{12}$

$=0.2833$.

(iii) $P(X=1)=\binom{12}{1}\left(\frac{1}{10}\right)^1\left(1-\frac{1}{10}\right)^{12-1}=12\times\frac{9^{11}}{10^{12}}=.3755$.

$\therefore\ P(X\ge 2)=1-P(X<2)$

$=1-[P(X=0)+P(X=1)]=0.3412$.

Ex. 5. *A defective die is thrown ten times independently. The probability that an even number will appear 5 times is twice the probability that an even number will appear 4 times. What is the probability that odd face appear in each of the ten throws.*

Let "even face" = "success" and X = number of even face among 10 trials.

$\therefore X$ has Binomial distribution with parameter $n=10$ and p. So the pmf is given by $P(X=i)=f_i={}^{10}C_i p^i(1-p)^{10-i}$.

By problem, $P(X=5)=2P(X=4)$ or, $f_5=2\times f_4$

or, ${}^{10}C_5\times p^5(1-p)^{10-5}=2\times{}^{10}C_4 p^4(1-p)^{10-4}$

or, ${}^{10}C_5 p=2\times{}^{10}C_4(1-p)$

or, $252p=2\times 210(1-p)$ or, $3p=5(1-p)$ or $p=\frac{5}{8}$

Required probability $= P(X=0) = f_0 = {}^{10}C_0 p^0 (1-p)^{10-0}$

$$= \left(1-\frac{5}{8}\right)^{10} = \left(\frac{3}{8}\right)^{10}.$$

Ex. 6. *The overall percentage of failures in a certain examination is 40. What is the probability that out of a group of 6 candidates at least 4 passed the examination?*

Now the overall percentage of success in a certain exam is 60.

Consider the experiment that one student is drawn and seen whether he is passed. Probability that he is a passed student is $p = \frac{60}{100} = 0.6$. The experiment be repeated 6 times. Let X = number of passed-student in 6 such trials. So X is a Binomial variate with parameter $n = 6$, $p = 0.6$.

So the distribution of X is

X	:	0	1	2	3	4	5	6
$P(X=i)=f_i$	:	f_0	f_1	f_2	f_3	f_4	f_5	f_6,

where $f_i = {}^6C_i p^i (1-p)^{6-i} = {}^6C_i (0.6)^i (1-0.6)^{6-i} = {}^6C_i (0.6)^i (0.4)^{6-i}$

Required probability $= P(X \ge 4) = f_4 + f_5 + f_6$

$$= {}^6C_4(0.6)^4(0.4)^{6-4} + {}^6C_5(0.6)^5(0.4)^{6-5} + {}^6C_6(0.6)^6(0.4)^{6-6} = 0.54432.$$

Ex. 7. *A family has 6 children. Find the probability that (i) 3 boys and 3 girls (ii) fewer boys than girls.*

Probability of any particular child being a boy is $\frac{1}{2}$.

Let 'boy' = success. The experiment of noticeing whether the child is boy or girl is repeated 6 times. X = No. of boy

i.e. No. of success

$\therefore X \sim b(n, p)$ where $n = 6$, $p = \frac{1}{2}$.

So the pmf of X is

$$f_i = P(X=i) = {}^6C_i \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{6-i} = {}^6C_i \left(\frac{1}{2}\right)^6.$$

(i) Now Probability of " 3 boys and 3 girls"

$$= P(X=3) = {}^6C_3\left(\frac{1}{2}\right)^6 = 5/16$$

(ii) Probability of "fewer boys than girls'

$$= P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$$

$$= {}^6C_0\left(\frac{1}{2}\right)^6 + {}^6C_1\left(\frac{1}{2}\right)^6 + {}^6C_2\left(\frac{1}{2}\right)^6 = \frac{11}{32}.$$

Ex. 8. *A die is tossed thrice. A success is "getting 1 or 6" on a toss. Find the mean and variance of the number of success.*

Let X denote the number of successs. Clearly X can take the values 0, 1, 2 or 3

and

X follows binomial distribution with $n = 3$

$$p = \text{Probability of success} = \frac{2}{6} = \frac{1}{3}$$

and $$q = 1 - \frac{1}{3} = \frac{2}{3}$$

$$\therefore \text{Mean} = E(X) = np = 3 \times \frac{1}{3} = 1$$

and

$$\text{variance} = npq = 3 \times \frac{1}{3} \times \frac{2}{3} = \frac{2}{3}$$

Ex. 9. *Find the probability distribution of the number of boys in a family with 3 children, assuming equal probabilities for boys and girls. Graph the distributions. Also find the distribution function $F(x)$ for the random variable X.*

Let E be the experiment of picking a child in the family.

The event 'boy' = success.

We have p = probability of boy

$= \frac{1}{2}$ by hypothesis. E is repeated 3 times.

Let X denote the number of boys.

Then X can assume the values 0, 1, 2, 3. X has the Binomial distribution.

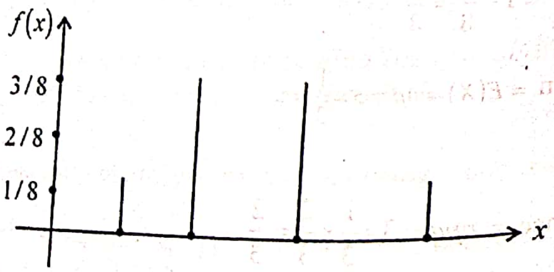
X :	0	1	2	3
f_i :	f_0	f_1	f_2	f_3

$$\text{where } f_i = \binom{3}{i}\left(\frac{1}{2}\right)^i\left(\frac{1}{2}\right)^{3-i} = \binom{3}{i}\left(\frac{1}{2}\right)^3 = \binom{3}{i}\frac{1}{8}$$

Therefore, the probability distribution of boys is

X :	0	1	2	3
f_i :	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

The graph of the distribution is given below :



Using the above table we obtain the distribution function $F(x)$ as

$$F(x) = 0, \quad -\infty < x < 0$$
$$= \frac{1}{8}, \quad 0 \le x < 1$$
$$= \frac{1}{8} + \frac{3}{8}, \quad 1 \le x < 2$$
$$= \frac{1}{2} + \frac{3}{8}, \quad 2 \le x < 3$$
$$= \frac{7}{8} + \frac{1}{8}, \quad 3 \le x < \infty$$

i.e.,
$$F(x) = 0, \quad -\infty < x < 0$$
$$= \frac{1}{8}, \quad 0 \le x < 1$$
$$= \frac{1}{2}, \quad 1 \le x < 2$$
$$= \frac{7}{8}, \quad 2 \le x < 3$$
$$= 1, \quad 3 \le x < \infty.$$

Ex. 10. *Suppose that half the population of a town are consumers of rice. 100 investigators are appointed to find out its truth. Each investigator interviews 10 individuals. How many investigator do you expect to report that three or less of the people interviewed are consumers of rice ?*

Consider the experiment "One investigator is drawn and seen whether he reports that three or less of the people interviewed are consumers of rice" ≡ whether it is success. Let probability of success = p

Now the investigator draws one individual and sees whether he is consumer of rice. Let q be the probability of this event = $\frac{1}{2}$. The investigator repeats this experiment 10 times.

Let Y = Number of such individual among ten. So Y has binomial distribution with parameter $n = 10,\ q = \frac{1}{2}$

Therefore $p = P(Y \le 3)$

$$= P(Y=0)+P(Y=1)+P(Y=2)+P(Y=3)$$

$$= {}^{10}C_0\left(\frac{1}{2}\right)^0\left(1-\frac{1}{2}\right)^{10-0} + {}^{10}C_1\left(\frac{1}{2}\right)^1\left(1-\frac{1}{2}\right)^{10-1}$$

$$+ {}^{10}C_2\left(\frac{1}{2}\right)^2\left(1-\frac{1}{2}\right)^{10-2} + {}^{10}C_3\left(\frac{1}{2}\right)^3\left(1-\frac{1}{2}\right)^{10-3}$$

$$= \left(\frac{1}{2}\right)^{10} + 10\times\frac{1}{2}\left(\frac{1}{2}\right)^9 + 45\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^8 + 120\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)$$

$$= \left(\frac{1}{2}\right)^{10}\{1+10+45+120\} = 176\times\left(\frac{1}{2}\right)^{10}.$$

Now, X has bonomial distribution with parameter $n=100,\ p=176\times\left(\frac{1}{2}\right)^{10}$. Therefore required expectation $= E(X)$

$$= 100\times176\times\left(\frac{1}{2}\right)^{10} = \frac{17600}{2^{10}} = \frac{17600}{1024} \simeq 17$$

$\therefore$ 17 investigators are expected to report so.

Ex. 11. *If X be a binomially distributed with* $E(X) = 2$ *and* $\text{var}(X) = \frac{4}{3}$, *find the distribution of X.*

We have $E(X) = 2, \text{var}(X) = \frac{4}{3}$.

$\therefore\ np = 2$ ...(1)

$np(1-p) = \frac{4}{3}$ ...(2)

Solving (1), (2), we get

$$p = \frac{1}{3}, n = 6,\ q = 1-p = 1-\frac{1}{3} = \frac{2}{3}.$$

$$\therefore\ f_0 = P(X=0) = {}^6C_0\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^6 = \frac{64}{729}$$

$$f_1 = P(X=1) = {}^6C_1\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^5 = \frac{64}{249}$$

$$f_2 = P(X=2) = {}^6C_2\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^4 = \frac{80}{243}.$$

Similarly $f_3 = P(X=3) = \frac{160}{729}, f_4 = P(X=4) = \frac{20}{243}$

$$f_5 = P(X=5) = \frac{4}{243}, f_6 = P(X=6) = \frac{1}{729}.$$

Thus the required distribution of X is

X	:	0	1	2	3	4	5	6
f_i	:	$\frac{64}{729}$	$\frac{64}{243}$	$\frac{80}{243}$	$\frac{160}{729}$	$\frac{20}{243}$	$\frac{4}{729}$	$\frac{1}{729}$.

Ex. 12. *The probability of a man hitting a target is* $\frac{1}{3}$

(a) If he fires 5 times, what is the probability of his hitting the target at least twice?

(b) How many times must he fire so that the probability of his hitting the target at least once is more than 90%.

Here p = probability of hitting $= \frac{1}{3}$.

$\therefore\ q$ = probability of no hit $= \frac{2}{3}$.

(a) Let X be the number of hits. Here $n = 5$

$$\therefore\ P(X \ge 2) = 1 - P(X < 2)$$

$$= 1 - P(X=0) - P(X=1)$$

$$= 1 - {}^5C_0\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^5 - {}^5C_1\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^4$$

$$= \frac{131}{243}.$$

(b) Let n be the smallest number of fires so that the probability of hitting the target of least once is more than 90%.

$\therefore$ By the condition

$$P(X \geq 1) > \frac{90}{100}$$

or, $1 - P(X = 0) > 0.9$

or, $1 - \left(\frac{2}{3}\right)^n > 0.9$

or, $\left(\frac{2}{3}\right)^n < 0.1$

or, $n \log\frac{2}{3} < \log 0.1$

or, $n > \dfrac{\log 0.1}{\log\frac{2}{3}}$

$n > 5 \cdot 679$

$\therefore$ $n = 6$.

Thus he must fire 6 times.

1.3.3. Poisson Distribution.

A discrete random variable X is said to have a poisson distribution with parameter $\mu(> 0)$ if its distribution is given by

X	:	0	1	2	3	...	...
f_i	:	f_0	f_1	f_2	f_3	...	...

where the pmf, $f_i = F(X = i) = \dfrac{e^{-\mu}\mu^i}{i!}$

e.g one probability mass, $f_2 = \dfrac{e^{-\mu}\mu^2}{2!} = \dfrac{e^{-\mu}\mu^2}{2}$ etc.

[W.B.U.T. 2013]

Note : (1) Since the parameter $\mu > 0$, $f_i = \dfrac{e^{-\mu}\mu^i}{i!} \geq 0$.

Again $\displaystyle\sum_{i=0}^{\infty} f_i = \sum_{i=0}^{\infty} \frac{e^{-\mu}\mu^i}{i!} = e^{-\mu}\sum_{\mu=0}^{\infty}\frac{\mu^i}{i!} = e^{-\mu}e^{\mu} = 1$

Thus the pmf f_i satisfies the two fundamental properties of pmf.

(2) When the random variable X has a poisson distribution with parameter $\mu(> 0)$, we write $X \sim P(\mu)$ and we say X is a poisson variate.

(3) Actually poisson distribution is a limiting case of Binomial distribution when n is very large and p is very small so that $\mu = np$ is of finite magnitude.

(4) The significance of the parameter μ is given in next theorem.

Cases where Poisson Distribution fits.

Let us consider a sequence of changes. If the random variable $X(t)$ denotes the number of changes during the interval $(0, t)$, then $X(t)$ assumes the values 0, 1, 2, 3, ...

It can be shown that $P(X(t) = i) = f_i = e^{-\lambda t}\dfrac{(\lambda t)^i}{i!}$, $i = 0, 1, 2, \cdots \cdots$

where λ = number of changes per unit time.

Thus the distribution of $X(t)$ becomes

$X(t)$	:	0	1	2	3	...	...
f_i	:	f_0	f_1	f_2	f_3	...	...

where $f_i = e^{-\lambda t}\dfrac{(\lambda t)^i}{i!}$ which is the pmf of poisson distribution.

Thus 'changes in an interval' is a poisson variate with parameter $\mu = \lambda t$ = average changes in the interval.

Note. (1) The interval $(0, t)$ may not be a time-interval. See the following illustration

(2) We use the notation $X(t)$ in place of X because it depends on t.

Illustration. (i) Let a huge metal sheet be produced by a machine in a factory. Defects are noticed in the sheet. The machine is such that the average number of defects per unit area is 3. A piece of '10 unit area of the sheet is purchased by a company. Let X = Number of defects in this piece. Then X may assume values 0, 1, 2, 3 ... up to ∞. (Here 'change' means 'defect' and the interval $(0,t)$ stands for '10 unit area of the sheet' or $(0,10)$.) Then X (or, $X(10)$) has the poisson distribution

X	:	0	1	2	3	...
f_i	:	f_0	f_1	f_2	f_3	...

where $f_i = e^{-\mu}\dfrac{\mu^i}{i!}$ with the parameter $\mu = 3\times 10 = 30$ = average number of defects per 10 unit area of the sheet.

For example, $f_0 = P(X=0) = e^{-30}\dfrac{30^0}{0!} = e^{-30}$

i.e., probability no defects in the piece $= e^{-30}$

In other words $100e^{-30}\%$ such pieces will be free from defects.

(ii) The number of deaths in a state in one year is a poisson variate

(iii) The number of radio active atoms decaying in time t follows the poisson distribution with parameter $\mu = tλ$ average number of decayed radioactive atoms per unit time.

Theorem. If X is a poisson variate with parameter μ then

(i) Mean of X is μ

(ii) Variance of X is μ [W.B.U.T. 2006,2007,2012,2013]

Proof. The values assumed by X are 0, 1, 2, ..., ... with probability $P(X=i) = f_i = \dfrac{e^{-\mu}\mu^i}{i!}$.

(i) Mean $= E(X) = \sum_{i=0}^{\infty} i\, f_i = \sum_{i=0}^{\infty} i \dfrac{e^{-\mu}\mu^i}{i!}$

$$= e^{-\mu}\sum_{i=0}^{\infty}\frac{\mu^i}{(i-1)!} = e^{-\mu}\left(\frac{\mu}{0!}+\frac{\mu^2}{1!}+\frac{\mu^3}{2!}+\frac{\mu^4}{3!}+\cdots \text{ up to }\infty\right)$$

$$= \mu e^{-\mu}\left(1+\mu+\frac{\mu^2}{2!}+\frac{\mu^3}{3!}+\cdots \text{ up to }\infty\right) = \mu e^{-\mu}\cdot e^{\mu} = \mu$$

(ii) Now, $E\{X(X-1)\} = \sum_{i=0}^{\infty} i(i-1)e^{-\mu}\dfrac{\mu^i}{i!} = e^{-\mu}\mu^2\sum_{i=2}^{\infty}\dfrac{\mu^{i-2}}{(i-2)!}$

$$= e^{-\mu}\mu^2\left(1+\frac{\mu}{1!}+\frac{\mu^2}{2!}+\frac{\mu^3}{3!}+\cdots \cdots \text{ up to }\infty\right)$$

$$= e^{-\mu}\mu^2 e^{\mu} = \mu^2$$

So, by an earlier result, (See the Theorem of Art 2.6) the variance,

$Var(X) = E\{X(X-1)\} - m(m-1)$, where m = mean

$$= \mu^2 - \mu(\mu-1) = \mu$$

Note. In light of the above theorem we considered the parameter $\mu = \lambda t$ in our previous illustration.

1.3.4. Binomial Approximation to Poisson Distribution.

The range of applications of Poisson Distribution becomes more wider as it is used as an approximation for a Binomial distribution. In case of a Binomial distribution when n becomes large, p is small enough so that np is a moderate fixed value, the Binomial variate becomes approximately equal to a Poisson variate. This is given by the following theorem.

Theorem. Let the random variable X follows Binomial distribution with pmf, $f_i = \binom{n}{i}p^i(1-p)^{n-i}$, $i = 0,1,2,\cdots,n$ where n and p are parameters. If $n\to\infty$ and $p\to 0$ such that $np = \mu$, a fixed quantity then

$\lim_{n\to\infty}\binom{n}{i}p^i(1-p)^{n-i} = \dfrac{e^{-\mu}\mu^i}{i!}$ for a fixed i, i.e. f_i of Binomial distribution $\approx f_i$ of Poisson distribution.

Proof. Beyond the scope of the book.

Illustration. Let a box contains 200 fuses. Experience tells that 2% of such fuses are defective. Let us consider the experiment of drawing a fuse and testing whether this is defective or not. Let X = number of defective fuse. $\therefore$ X has Binomial distribution with parameter

$n = 200$, $p = \dfrac{2}{100} = .02$. The pmf of X is

$$f_i = \binom{n}{i} p^i (1-p)^{n-i} = \binom{200}{i} (.02)^i (1-.02)^{200-i}.$$

Here we see n is so large and p is small so we can write

$$\binom{200}{i}(.02)^i(1-.02)^{200-i} \simeq \frac{e^{-200\times.02}}{i!} = \frac{e^{-4}}{i!}$$

Now $P(X \le 3) = f_0 + f_1 + f_2 + f_3 \simeq \frac{e^{-4}}{0!} + \frac{e^{-4}}{1!} + \frac{e^{-4}}{2!} + \frac{e^{-4}}{3!}$.

Illustrative Examples.

Ex. 1. *For a poisson variate if $P(X=2) = P(X=1)$, find $P(X = 1 \text{ or } 0)$. Find also mean of X.*

Let m be the parameter of the poisson variate.

$\therefore P(X=i) = f_i = \frac{e^{-m} m^i}{i!}$.

Now, $f_2 = f_1$, or, $\frac{e^{-m}\cdot m^2}{2!} = \frac{e^{-m}\cdot m^1}{1!}$

$\therefore m = 2$

$\therefore P(X = 1 \text{ or } 0) = f_1 + f_0 = e^{-m}(1+m) = e^{-2}(1+2) = 3e^{-2}$

Mean of $X = m = 2$

Ex. 2. *A car-hire firm has two cars which it hires out day by day. The number of demands for a car on each day is distributed as a poisson distribution with average number of demand per day 1.5. Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused $(e^{-1.5} = 0.2231)$.*

[W.B.U.Tech,2003, 2006,2007]

Let X be the random variable denoting the number of demands for a car on any day. Then X is poisson distributed with parameter $\mu = 1.5$. So its pmf $P(X=i) = f_i = \frac{e^{-\mu}\mu^i}{i!}$ where $\mu = 1.5$.

$\therefore$ Proportion of days on which neither car is used
= Prob. of there being no demand for the car

$$= P(X=0) = \frac{\mu^0 e^{-\mu}}{0!} = e^{-1.5} = 0.2231.$$

Proportion of days on which some demand is refused
= Prob. for the number of demands to be more than two

$$= P(X>2) = 1 - P(X \le 2)$$

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

$$= 1 - \left\{e^{-\mu} + \frac{\mu e^{-\mu}}{1!} + \frac{\mu^2 e^{-\mu}}{2!}\right\} = 1 - e^{-1.5}\left(1 + 1.5 + \frac{(1.5)^2}{2}\right) = 0.19126$$

Ex. 3. *A radio active source emits on the average 2.5 particles per second. Calculate the prob. that 2 or more particles will be emitted in an interval of 4 seconds.*

Here λ = number of changes (which is particle emitted) per unit time on an average $= 2.5$.

Let X be the random variable denoting the number of particles emitted in the given interval. Then X is poisson distributed with parameter μ = average number of particle in 4 seconds $= 2.5 \times 4 = 10$.

So the p.m.f, $f_i = P(X=i) = \mu^i e^{-\mu}/i! = 10^i e^{-10}/i!$

So the required Prob. $= P(X \ge 2) = 1 - P(X < 2)$

$$= 1 - \{P(X=0) + P(X=1)\}$$

$$= 1 - \{e^{-10} + 10e^{-10}\} = 1 - 11e^{-10}.$$

Ex. 4. *In a certain factory turning razor blades, there is a small chance, 1/500 for any blade to be defective. The blades are in packets of 10. Use poisson distribution to calculate the approximate number of packets containing (i) no defective (ii) one defective (iii) two defective blades respectively in one consignment of 10, 000 packets. (Given $e^{-.02} = .9802$).*

[W.B.U.Tech 2004]

On an average there are 1 defective blade per 500 blades.

So the average number of defective blades in a packet of 10 is $10 \times \frac{1}{500} = \frac{1}{50} = 0.02$.

Let X = number of defective blades in a packet. X follows poisson distribution with parameter $\mu = 0.02$. So the pmf is

$$P(X=i) = f_i = \frac{e^{-\mu}\mu^i}{i!}.$$

(i) Now probability that one packet contains no defective blade

$$= P(X=0) = f_0 = \frac{e^{-\mu}\mu^0}{0!} = e^{-\mu} = e^{-0.02} = 0.9802.$$

$\therefore$ Number of packets in the consignment containing no defective blades $= 0.9802 \times 10{,}000 = 9802$

(ii) Probability that one packet contains one defective blade

$$= P(X=1) = f_1 = \frac{e^{-\mu}\mu}{1!} = e^{-0.02} \times 0.02$$

$$= 0.9802 \times 0.02 = 0.019604$$

$\therefore$ Number of blades in the consignment

$$= 0.019604 \times 10{,}000 = 196.04 \simeq 196$$

(iii) $P(X=2) = \frac{e^{-\mu}\mu^2}{2!} = e^{-.02} \times \frac{(.02)^2}{2}$

$$= 0.9802 \times .0002 = 0.00019604$$

$\therefore$ Required number $= 0.00019604 \times 10{,}000 = 1.9604 \simeq 2$.

Ex. 5. *If a random variable has a poisson distribution such that $P(1) = P(2)$, find (i) mean of the distribution (ii) standard derivation (iii) $P(X=4)$* [W.B.U.Tech 2007]

Let X be a poisson variate. Then the p.m.f of X is

$$f_i = P(X=i) = e^{-\mu}\frac{\mu^i}{i!}, \quad i = 0, 1, 2, \cdots$$

As $P(1) = P(2)$, so $e^{-\mu}\mu = e^{-\mu}\frac{\mu^2}{2!}$ $\quad \therefore \mu = 2$

(i) So the mean of the distribution is 2

(ii) Now $Var(X) = \mu = 2$ $\quad \therefore$ standard derivation $= \sqrt{2}$.

(iii) $P(X=4) = e^{-2}\frac{2^4}{4!} = \frac{2}{3}e^2$.

Ex. 6. *Find the probability that at most 5 defective fuses will be found in a box of 200 fuses if experience shows that 2 percent of such fuses are defective.*

Let X denote the number of defective fuses in the box. Then clearly X has a binomial distribution with parameters

$$n = 200, \; p = \frac{2}{100} = \frac{1}{50}$$

$$\therefore \mu = np = 200 \times \frac{1}{50} = 4$$

Using an approximation, by the poisson distribution, we have

$$P(X \le 5) = \sum_{i=0}^{5} \frac{e^{-4}4^i}{i!} = e^{-4}\left(1 + \frac{4}{1!} + \frac{4^2}{2!} + \cdots + \frac{4^5}{5!}\right) = 0.785$$

Ex. 7. *Six coins are tossed 6400 times. Using the poisson distribution find the approximate probability of getting six heads 8 times.*

Let X denote the number of six heads in the toss of six coin. Then X is a binomial variate with parameter $n = 6400$, $p = \left(\frac{1}{2}\right)^6 = \frac{1}{64}$. Here n is so large and p is small, but $\mu = np = 6400 \times \frac{1}{64} = 100$

So using the poisson approximation, we have,

$$P(X=8) = {}^nC_8 p^8 (1-p)^{n-8} \simeq e^{-100}\frac{(100)^8}{8!}.$$

Ex. 8. *2% of the items made by a machine are defective. Find the probability that 3 or more items are defective in a sample of 100 items. (Given $e^{-1} = 0.368$, $e^{-2} = 0.135$, $e^{-3} = 0.498$)*

Consider the experiment – one item is drawn and found whether it is defective (success). Let this experiment be repeated 100 times. X = number of defective items. Then X follows Binomial distribution with parameter $n = 100$, $p = 2\% = \frac{2}{100} = .02$. $\therefore$ the pmf, $f_i = {}^nC_i p^i (1-p)^{n-i}$.

Now since n is large and p is small and $np = 100 \times .02 = 2$ so Binomial pmf is approximately equal to Poisson pmf with parameter np.

$$\therefore f_i \simeq \frac{e^{-2} \cdot 2^i}{i!}$$

Required probability $= P(X \geq 3) = 1 - P(X < 3)$

$$= 1 - \{f_0 + f_1 + f_2\} = 1 - e^{-2}\left(1 + 2 + \frac{2^2}{2!}\right) = 1 - 0.135 \times 5 = 0.325.$$

1.3.5. Uniform or Rectangular Distribution.

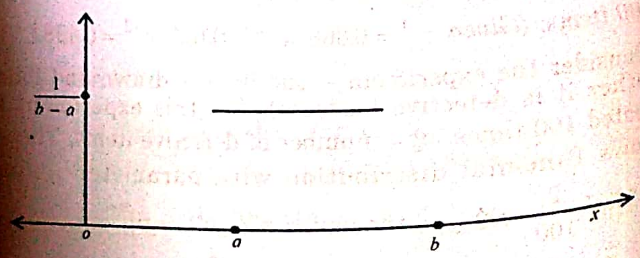
A random variable X is said to have a uniform distribution on the interval $[a, b], -\infty < a < b < \infty$, if its probability density function is given by

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$$

$$= 0, \quad \text{otherwise}$$

where a, b are two parameters of the distribution. Clearly $f(x) \geq 0$, for $a < x < b$ and $\int_{-\infty}^{\infty} f(x)dx = \int_a^b f(x)dx = \int_a^b \frac{dx}{b-a} = 1$. So the fundamental properties of the density function is satisfied. If X has a uniform distribution on $[a, b]$, then we write $X \sim U[a, b]$.

The density curve is shown in the following fig.



The distribution function $F(x)$ may be easily calculated by

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

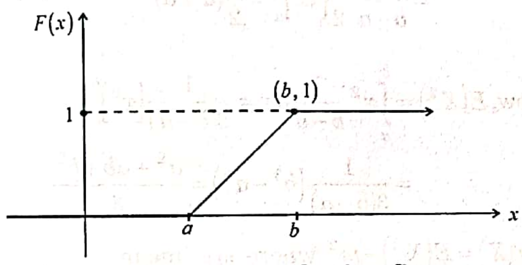
and is given by

$$F(x) = 0, \quad -\infty < x < a$$

$$= \frac{x-a}{b-a}, \quad a \leq x < b$$

$$= 1, \quad b \leq x < \infty$$

The graph of the distribution function is given below.



A Case where Uniform Distribution fits.

Let $[a, b]$ be an interval. A point P is taken at random in the interval $[a, b]$. Let $OP = x$.

Let X be a random variable which assumes the values x. Then X would have uniform distribution with pdf

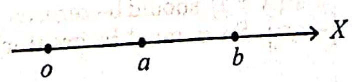
$$f(x) = \frac{1}{b-a}, \quad a < x < b$$

$$= 0, \quad \text{elsewhere}$$

Theorem. If a continuous random variable X has uniform distribution with parameter a and b then

(i) the mean is $\frac{1}{2}(a+b)$ [W.B.U.Tech 2014]

(ii) the variance is $\frac{(a-b)^2}{12}$



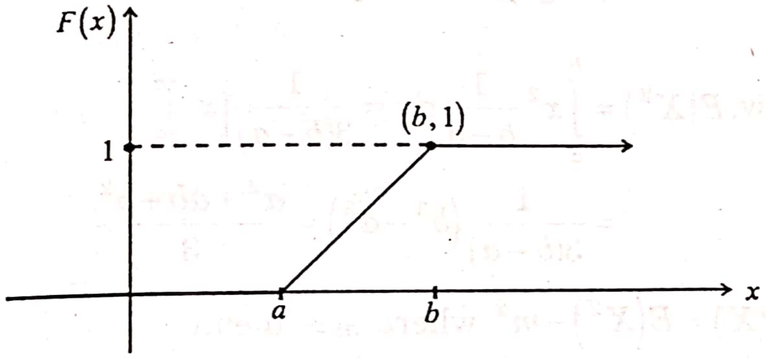
The distribution function $F(x)$ may be easily calculated by

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

and is given by

$$F(x)=0, \qquad -\infty < x < a$$
$$=\frac{x-a}{b-a}, \qquad a \le x < b$$
$$=1, \qquad b \le x < \infty$$

The graph of the distribution function is given below.



A Case where Uniform Distribution fits.

Let [a. b] be an interval. A point P is taken at random in the interval $[a, b]$. Let $OP = x$.

Let X be a random variable which assumes the values x. Then X would have uniform distribution with pdf

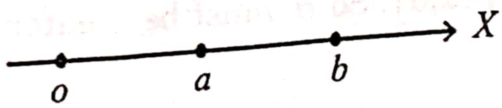
$$f(x)=\frac{1}{b-a}, \quad a < x < b$$

$$=0, \qquad \text{elsewhere}$$

Theorem. If a continuous random variable X has uniform distribution with parameter a and b then

(i) the mean is $\frac{1}{2}(a+b)$

(ii) the variance is $\frac{(a-b)^2}{12}$



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

$$f(x)=\frac{1}{b-a}, \quad a\le x\le b$$
$$=0, \quad \text{elsewhere}$$

(i) the mean $=E(X)=\int_{-\infty}^{\infty} xf(x)dx$

$$=\int_{-\infty}^{a} x\cdot 0dx+\int_{a}^{b} x\cdot\frac{1}{b-a}dx+\int_{b}^{\infty} x\cdot 0dx$$

$$=\frac{1}{b-a}\frac{1}{2}\left[x^2\right]_a^b=\frac{1}{2}(a+b)$$

(ii) Now, $E\left(X^2\right)=\int_a^b x^2\frac{1}{b-a}dx=\frac{1}{3(b-a)}\left[x^3\right]_a^b$

$$=\frac{1}{3(b-a)}\left(b^3-a^3\right)=\frac{a^2+ab+b^2}{3}$$

$\therefore$ $Var(X)=E\left(X^2\right)-m^2$ where $m=$ mean

$$=\frac{a^2+ab+b^2}{3}-\left(\frac{a+b}{2}\right)^2=\frac{(a-b)^2}{12}.$$

Illustrative Examples.

Ex. 1. *If X is uniformly distributed in $[-\alpha,\alpha]$ with $\alpha>0$ then determine α such that* $P(X>1)=\frac{1}{3}$.

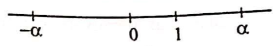
Here p.d.f of X is $f(x)=\frac{1}{2\alpha}$ is $-\alpha<x<\alpha$

$=0$ elsewhere.

when $\alpha<1$,

$P(X>1)$ should be zero, as X lies outside the given interval $[-\alpha,\alpha]$. So α must be greater than 1.

$\therefore$ $P(X>1)$



$$=\int_1^{\alpha} f(x)dx=\frac{\alpha-1}{2\alpha}$$

$\therefore$ $\frac{\alpha-1}{2\alpha}=\frac{1}{3}$.

$\therefore$ $\alpha=3$.

Ex. 2. *If X is uniformly distributed in $-2\le x\le 2$, find* $P\left(|X-1|\ge\frac{1}{2}\right)$.

Here p.d.f of X is

$$f(x)=\frac{1}{4} \text{ is } -2\le x\le 2$$
$$=0, \quad \text{elsewhere.}$$

$\therefore$ $P\left(|X-1|\ge\frac{1}{2}\right)$

$$=P\left(-2\le X\le\frac{1}{2}\right)+P\left(\frac{3}{2}\le X\le 2\right)$$

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

$$=\frac{\frac{1}{2}+2}{4}+\frac{2-\frac{3}{2}}{4}$$

$$=\frac{3}{4}.$$

Ex. 3. *If X is uniformly distributed over* [1, 2] *find U so that* $P\left(X>U+\overline{X}\right)=\frac{1}{6}$. [W.B.U.Tech 2006,2008]

Since X is uniformly distributed over [1, 2] therefore its pdf,

$$f(x) = \frac{1}{2-1}, \quad 1 < x < 2$$

$$= 0, \quad \text{elsewhere}$$

i.e., $f(x) = 1, \quad 1 < x < 2$

$$= 0, \quad \text{elsewhere}$$

$$\therefore\ \bar{X} = E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_{1}^{2} x \cdot 1\, dx = \left[\frac{x^2}{2}\right]_1^2 = \frac{3}{2}$$

Now, $P(X > U + \bar{X}) = \frac{1}{6}$ or, $P\left(U + \frac{3}{2} < X < \infty\right) = \frac{1}{6}$

or, $\int_{U+\frac{3}{2}}^{2} dx = \frac{1}{6}$ or, $2 - U - \frac{3}{2} = \frac{1}{6}$ $\therefore U = \frac{1}{3}$

1.3.6. Exponential Distribution.

Two parameters exponential distribution. A random variable X is said to have a two-parameter-exponential distribution if its probability density function is given by

$$f(x) = \frac{1}{b} e^{-\frac{x-a}{b}}, \quad x \geq a$$

$$= 0, \quad \text{elsewhere}$$

where $a, b (b > a)$ are two parameters of the distribution.

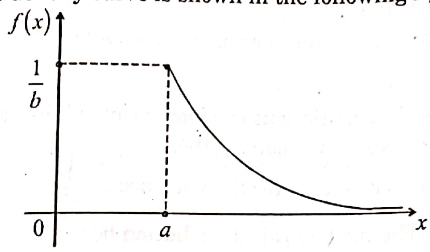
Note. (1) Clearly $f(x) \geq 0$ for all x and

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

So the two fundamental properties of pdf are satisfied

(2) If X has an exponential distribution with parameters a and b we write $X \sim E[a, b]$.

(3) The density curve is shown in the following figure

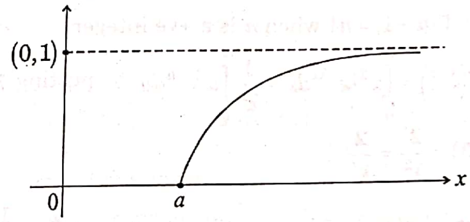


(4) The distribution function $F(x)$ is given by

$$F(x) = 0, \quad -\infty < x < a$$

$$= 1 - e^{-\frac{x-a}{b}}, \quad a \leq x < \infty$$

(5) The graph of the distribution function $F(x)$ is given in the following figure



One Parameter Exponential Distribution

We say that X has one parameter exponential distribution if its pdf is $f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$

$$= 0, \quad \text{elsewhere}$$

where λ is the only parameter.

Note. (1) By the term X has exponential distribution' we mean X has one-parameter exponential distribution.

(2) Clearly the pdf $f(x)$ satisfies the two fundamental properties of pdf.

(3) If X has exponential distribution with parameter λ we write $X \sim E(0,\lambda)$.

(4) This distribution is obtained from the previous by putting $a=0,\ b=\frac{1}{\lambda}$.

Theorem. If a continuous random variable X has exponential distribution with parameter λ then

(i) the mean is $\frac{1}{\lambda}$ (ii) the variance is $\frac{1}{\lambda^2}$

Proof. The p.d.f. of this distribution be

$$f(x)=\lambda e^{-\lambda x}, \quad x\ge 0$$
$$=0, \quad \text{elsewhere}$$

$\therefore$ mean, $m=E(x)=\int_0^\infty x\lambda e^{-\lambda x}dx=\frac{1}{\lambda}\int_0^\infty ue^{-u}du$, by putting $\lambda x=u$

$=\frac{1}{\lambda}\Gamma(2)$ $\left[\because \Gamma(n)=\int_0^\infty e^{-x}x^{n-1}dx \text{ from definition of Gamma function}\right]$

$=\frac{1}{\lambda}$ $[\because \Gamma(n+1)=n! \text{ when } n \text{ is a +ve integer}]$

Now, $E(X^2)=\int_0^\infty x^2\lambda e^{-\lambda x}dx=\frac{1}{\lambda^2}\int_0^\infty u^2e^{-u}du$ by putting $\lambda x=u$

$=\frac{1}{\lambda^2}\Gamma(3)=\frac{2!}{\lambda^2}=\frac{2}{\lambda^2}$

$\therefore Var(X)=E(X^2)-m^2$ [where m is mean] $=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}$.

Illustrative Example.

Ex. 1. *Suppose that during rainy season, on a tropical island, the length of shower has an exponential distribution with average length of shower $\frac{1}{2}$ mins. What is the probability that a shower will last more than three minutes ? If a shower has already lasted for 2 minutes, what is the probability that it will last for at least one more minute ?*

Let X = length of shower in minute.

By problem X has exponential distribution with parameter λ where $\frac{1}{\lambda}=\frac{1}{2}$, i.e., $\lambda=2$

$\therefore$ its p.d.f, $f(x)=2e^{-2x},\ x>0$

Now, Probability that a shower lasts more than three minutes

$$=P(X>3)=\int_3^\infty 2e^{-2x}dx=2\lim_{X\to\infty}\int_3^X e^{-2x}dx$$

$$=2\lim_{X\to\infty}\left[\frac{e^{-2x}}{-2}\right]_3^X=-\lim_{X\to\infty}\left\{\frac{1}{e^{2X}}-\frac{1}{e^6}\right\}=\frac{1}{e^6}$$

Probability that a shower lasts more than two minutes

$$=P(X>2)=\int_2^\infty 2e^{-2x}dx=\frac{1}{e^4}.$$

Now the required probability $=P(X\ge 3/X\ge 2)$

$$=\frac{P((X\ge 3)\cap(X\ge 2))}{P(X\ge 2)}=\frac{P(X\ge 3)}{P(X\ge 2)}=\frac{\frac{1}{e^6}}{\frac{1}{e^4}}=\frac{1}{e^2}.$$

1.3.7. Normal Distribution.

A continuous random variable X is said to have a normal distribution if its probability density function is given by

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

where μ and $\sigma>0$ are the two parameter of the distribution.

Note. (1) In this case we say X is a normal variate with parameters μ and σ and we denote it by $X\sim N(\mu,\sigma)$

(2) Clearly $f(x)\ge 0$ for all x.

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

$$=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-z^2}dz \text{, by putting } z=\frac{x-\mu}{\sigma\sqrt{2}}$$

$$=\frac{2}{\sqrt{\pi}}\int_{0}^{\infty} e^{-z^2}dz=\frac{1}{\sqrt{\pi}}\int_{0}^{\infty} e^{-u}u^{-\frac{1}{2}}du \text{, by putting } z^2=u$$

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

So the two necessary conditions for the probability density function are satisfied.

(3) The significance of the parameters μ and σ are given in the next theorem.

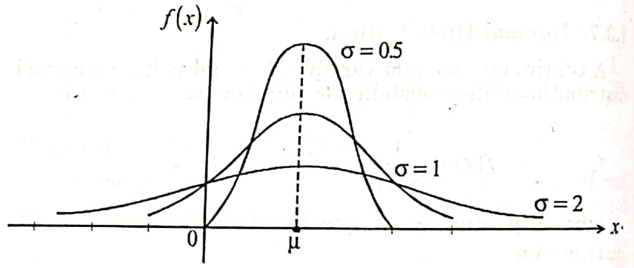
(4) The distribution function $F(x)$ is given by

$$F(x)=\int_{-\infty}^{x} f(u)du=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(u-\mu)^2}{2\sigma^2}}du$$

Normal Curve

The graph of the pdf of a normal variate is called normal curve.

The curve is shown in the following figure for different values of σ.



The normal curve is bell shaped and symmetric about the ordinate $x=\mu$. For small values of σ, the curve has a small peak and as σ increases, the normal curve tends to be flatter.

Theorem 1. If X has normal distribution with parameter μ and σ then (i) the mean of X is μ (ii) the s.d of X is σ.

Proof. Here p.d.f of X is given by

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<\infty$$

(i) Mean $=E(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} xe^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$

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

$$=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}ue^{-u^2}du+\frac{\mu}{\sqrt{\pi}}\int_{-\infty}^{\infty}e^{-u^2}du \text{, by putting } u=\frac{x-\mu}{\sigma\sqrt{2}}$$

$=0+\frac{\mu}{\sqrt{\pi}}2\cdot\int_{0}^{\infty}e^{-u^2}du$ [$\because$ the integrand of the 1st integral is odd and that of the 2nd integral is even]

$$=\frac{\mu}{\sqrt{\pi}}\int_{0}^{\infty}z^{\frac{1}{2}-1}e^{-z}dz \text{, by putting } u^2=z$$

$$=\frac{\mu}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}\right)=\frac{\mu}{\sqrt{\pi}}\sqrt{\pi}=\mu \qquad \left[\because \Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}\right]$$

(ii) Now, $Var(X)=E\{(X-\mu)^2\}$

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

$$=\frac{2\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{\infty}z^2e^{-z^2}dz \text{ by putting } \frac{x-\mu}{\sigma\sqrt{2}}=z$$

$$=\frac{4\sigma^2}{\sqrt{\pi}}\int_{0}^{\infty}z^2e^{-z^2}dz=\frac{2\sigma^2}{\sqrt{\pi}}\int_{0}^{\infty}u^{\frac{1}{2}}e^{-u}du \text{ by putting } z^2=u$$

$$= \frac{2\sigma^2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right) = \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{1}{2}\Gamma\left(\frac{1}{2}\right) \qquad [\because \Gamma(n+1) = n\Gamma(n)]$$

$$= \frac{\sigma^2}{\sqrt{\pi}} \cdot \sqrt{\pi} = \sigma^2$$

$\therefore$ the s.d of X is σ.

A Case where Normal Distribution Fits.

Let a thermal plant supplies electric power in a certain city. Let $X =$ amount of electric power (in watt) supplied by the plant in a day. Clearly X varies from day to day. It can be assumed that X is a continuous variate. It can be shown that X has normal distribution with parameter $\mu =$ average power supply per day and $\sigma =$ standard deviation of all the values assumed by X.

Standard Normal Distribution.

The normal distribution with mean 0 and standard deviation 1 is called standard normal distribution. The random variable having standard normal distribution is called standard normal variate or $X \sim N(0,1)$.

Thus the p.d.f of the standard normal distribution is given by

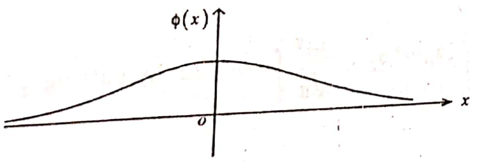
$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty.$$

and the corresponding distribution function is given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du$$

Standard Normal Curve.

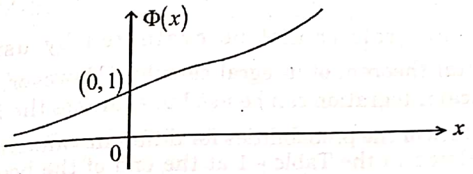
The graph of the p.d.f of a standard normal distribution is called standard normal curve. This is shown in the following figure.



It is symmetric about y axis, bell shaped. It has maximum value at $x = 0$. Since $\int_{-\infty}^{\infty} \phi(x)dx = 1$ so area under this curve is 1. X axis is its asymptote.

Standard Normal Distribution Curve.

The graph of the distribution function $\Phi(x)$ is shown in the following figure.



Theorem 2. (An Important Result)

If the continuous random variable X has normal distribution with parameter μ and σ then $Z = \frac{X-\mu}{\sigma}$ has standard normal distribution.

Proof. That Z has normal distribution is shown in the next chapter. Yet it is observed that the mean of Z,

$$\bar{Z} = E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\{E(X) - E(\mu)\}$$

$$= \frac{1}{\sigma}\{\mu - \mu\} = 0$$

$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}Var(X-\mu)$, from the property of variance

$= \frac{1}{\sigma^2}\{Var(X) - Var(\mu)\}$, from the property of variance

$$= \frac{1}{\sigma^2}\{\sigma^2 - 0\} = 1.$$

Tabulation of the Standard Normal Distribution.

Let the random variable X has standard normal distribution. Then its distribution function

$$\Phi(x) = P(-\infty < X \le x) = \int_{-\infty}^{x} \phi(z)dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz \therefore P(X \le a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{z^2}{2}} dz \text{ and } P(a \le X \le b) = \frac{1}{\sqrt{2\pi}} \int_{a}^{b} e^{-\frac{z^2}{2}} dz$$

These integrals cannot be evaluated by using the fundamental theorem of integral calculus. However, method of numerical integration can be used to evaluate the integral.

The values of the probabilities for different values of a have been tabulated in the Table - 1 at the end of the book. From this we can find $\Phi(x)$ for different values of x.

Since $P(a \le x \le b) = \Phi(b) - \Phi(a)$, we have if X has distribution $N(0,1)$ then from the tabulated value of Φ we can evaluate the probabilities associated with X.

An Illustrative Examples

Let a thermal plant supplies electric power in a certain city. Let $X =$ amount of electric power (in watt) supplied by the plant in a day. Clearly X varies from day to day.

It can be assumed that this X has normal distribution with parameter $\mu =$ average power supply per day and $\sigma =$ standard deviation of all the values assumed by X. Now if it is known that these $m = 300$ and $\sigma = 10$ then by the previous theorem $Z = \dfrac{X - 300}{10}$ is a standard normal variate.

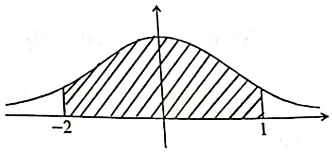
Suppose we are asked to find the number of days on which power supplied will lie between 280 to 310 MW.

Then we find $P(280 < X < 310)$

$$= P\left(\frac{280-300}{10} < \frac{X-300}{10} < \frac{310-300}{10}\right) = P(-2 < Z < 1)$$

$= \int_{-2}^{1} \phi(z)dz$ [where ϕ is the pdf of standard normal distribution]

$=$ area enclosed by standard normal curve, X axis, the ordinates $z = -2$ and $z = 1$ (shown by shade in the figure)



$= 0.4772 + 0.3413$ (obtained from the tabulated value)

$= 0.8185$. So the probability of having power supply between 280 and 310 MW in a day is 0.8185. Thus 81.85% day will receive a power supply between 280 and 310 MW.

1.3.8. Binomial Approximation to Normal Distribution.

Theorem. Let the random variable X follows Binomial distribution with pmf

$$f_i = \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, 1, 2, \cdots, n$$

where n and p are parameters. If $n \to \infty$ and p is not very small then the distribution of the r.v $Z = \dfrac{X - np}{\sqrt{np(1-p)}}$ approches to the standard normal distribution.

Proof. Beyond the scope.

Note. **(1)** In light of above theorem we understand X is approximately a normal variate with mean np and s.d. $\sqrt{npq}$

(2) The variable $Z = \dfrac{X - np}{\sqrt{np(1-p)}}$ is called standardized Binomial variate.

Illustration. Let an unbiased coin is tossed 12 times. Let $X =$ number of heads appeared. Then X has binomial distribution with parameter $n = 12$ and $p = \frac{1}{2} = 0.5$. So th[e] probability masses are

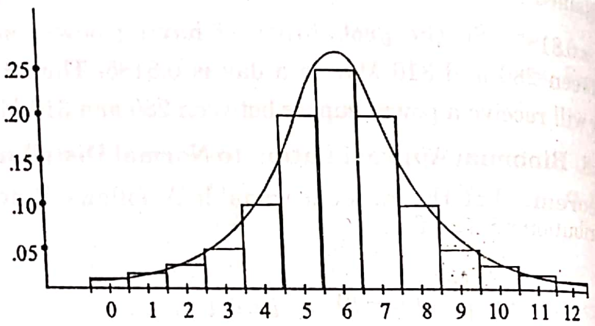
$$f_0 = {}^{12}C_0 (0.5)^0 (0.5)^{12-0} = 0.000244$$

$$f_1 = {}^{12}C_1 (0.5)^1 (0.5)^{12-1} = 0.00292$$

$$f_2 = {}^{12}C_2 (0.5)^2 (0.5)^{12-2} = 0.01611$$

and so on.

In the following diagram the heights of the rectangles a[re] these.



We think n as large. So X is approximately normal variat[e] with parameter $\mu = np = 12 \times 5 = 6$ and

$$\sigma = \sqrt{np(1-p)} = \sqrt{12 \times 5 \times 5} = 1.73$$

i.e., $Z = \frac{X-6}{1.73}$ is approximately a standard normal variat[e]

In the above figure the curve is $(6, 1.73)$ normal curve. Thi[s] approximately fits the height of the rectangle.

(i) Now, suppose we seek $P(4 \le X \le 7)$. But if we find thi[s] from the normal distribution then we must fin[d] $P(3.5 \le X \le 7.5)$ as indicated in the above diagram. Now whe[n]

$X = 3.5$, $Z = \frac{3.5-6}{1.73} = -1.46$, when $X = 7.5, Z = \frac{7.5-6}{1.73} = 0.87$

Then $P(4 \le X \le 7) \simeq P(-1.46 \le X \le 0.87)$

$$= \int_{-1.46}^{0.87} \phi(t)dt$$

$$= \int_{-1.46}^{0} \phi(t)dt + \int_{0}^{0.87} \phi(t)dt$$

$$= 0.4279 + 0.3078$$

(obtained from the statistical table-I)

$$= 0.7357.$$

(ii) Now, suppose we seek $P(X = 8)$.

We calculate this with the help of approximate normal distribution then we find $P(7.5 \le X \le 8.5) = P(0.87 \le Z \le 1.44)$ which would be obtained from statistical table.

Illustrative Examples.

Ex. 1. *If X is normally distributed with zero mean and unit variance find the expectation of* X^2. [*W.B.U.Tech, 2002, 2007*]

By problem $E(X) = 0$, $Var(X) = 1$ or, $E(X^2) - \{E(X)\}^2 = 1$

or, $E(X^2) - 0^2 = 1 \quad \therefore E(X^2) = 1$

Ex. 2. *If X is normally distributed with mean 3 and s.d. 2, find c such that* $P(X > c) = 2\ P(X \le c)$.

Given $\int_{-\infty}^{.43} \phi(t)dt = \frac{1}{3}$. [*W.B.U.Tech 2007*]

Let Z be the standard normal variate. Then $Z = \frac{X-3}{2}$

$$\therefore P(X > c) = P\left(\frac{X-3}{2} > \frac{c-3}{2}\right) = P\left(Z > \frac{c-3}{2}\right)$$

$$= 1 - P\left(Z \le \frac{c-3}{2}\right) = 1 - \int_{-\infty}^{\frac{c-3}{2}} \phi(t)dt$$

Also $P(X \le c) = P\left(\frac{X-3}{2} \le \frac{c-3}{2}\right) = P\left(Z \le \frac{c-3}{2}\right) = \int_{-\infty}^{\frac{c-3}{2}} \phi(t)dt$

As $P(X > c) = 2P(X \le c)$, we have

$$1 - \int_{-\infty}^{\frac{c-3}{2}} \phi(t)dt = 2 \times \int_{-\infty}^{\frac{c-3}{2}} \phi(t)dt \quad \text{or,} \quad \int_{-\infty}^{\frac{c-3}{2}} \phi(t)dt = \frac{1}{3}$$

or, $\frac{c-3}{2} = .43$ (from the given data)

$\therefore\ c = 3.86$.

Ex. 3. *The length of bolts produced by a machine is normally distributed with mean 4 and and s.d. 0.5. A bolt is defective if its length does not lie in the interval* $(3.8, 4.3)$. *Find the percentage of defective bolts produced by the machine.*

$$\left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0.6} e^{-\frac{t^2}{2}} dt = 0.7257,\ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0.4} e^{-\frac{t^2}{2}} dt = 0.6554\right]$$ [W.B.U.Tech 2004]

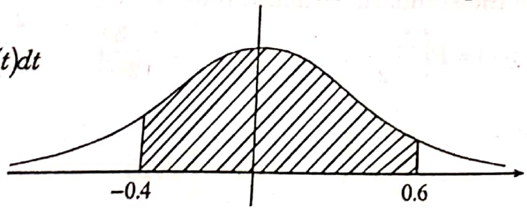
Let $X =$ length of bolt. By problem X has normal distribution with mean $\mu = 4$ and s.d $\sigma = 0.5$. First we shall find the probability $P(3.8 < X < 4.3)$.

Now, $Z = \frac{X-4}{0.5}$ has standard normal distribution. Now when $X = 3.8$, $Z = \frac{3.8-4}{0.5} = -0.4$; when $X = 4.3$, $Z = \frac{4.3-4}{0.5} = 0.6$

$\therefore\ P(3.8 < X < 4.3) = P(-0.4 < Z < 0.6)$

= Area under standard normal curve enclosed between the two ordinate $Z = -.04$ and $Z = 0.6$ (shaded part in figure)

$$= \int_{-0.4}^{0.6} \phi(t)dt$$



$$= \int_{-\infty}^{0.6} \phi(t)dt - \int_{-\infty}^{-0.4} \phi(t)dt = 0.7257 - \int_{0.4}^{\infty} \phi(t)dt \ [\because \phi \text{ curve is symmetric}]$$

$$= 0.7257 - \left(1 - \int_{-\infty}^{0.4} \phi(t)dt\right) = 0.7257 - (1 - 0.6554) = 0.3811$$

$\therefore$ Probability that the length of the bolt lies between 3.8 and 4.3 is 0.3811.

$\therefore$ Probability that the length of the bolt does not lie between 3.8 and $4.3 = 1 - 0.3811 = 0.6189$.

$\therefore$ Probability that the bolt is defective $= 0.6189$

$\therefore$ Percentage of defective bolts produced $= 0.6189 \times 100 = 61.89 \simeq 62$.

Ex. 4. *If the weekly wage of 10,000 workers in a factory follows normal distribution with mean and s.d Rs. 70 and Rs. 5 respectively, find the expected number of workers whose weekly wages are (i) between Rs. 66 and Rs. 72 (ii) less than Rs. 66 and (iii) more than Rs. 72* [W.B.U.Tech 2006]

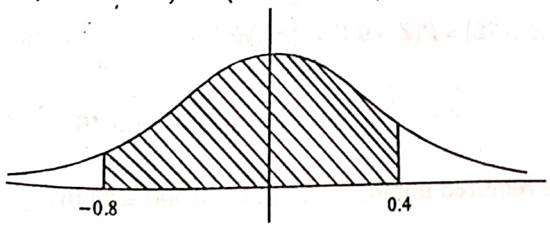
[*Given that* $\frac{1}{\sqrt{2\pi}} \int_0^z e^{-t^2/2} dt = 0.1554$ *and* 0.2881 *according as* $z = 0.4$ *and* $z = 0.8$]

Let $X =$ wage of a worker. X has normal distribution with $\mu = 70$, $\sigma = 5$.

$\therefore\ Z = \frac{X-70}{5}$ has standard normal distribution.

(i) When $X = 66$, $Z = \frac{66-70}{5} = -0.8$; when $X = 72$, $Z = \frac{72-70}{5} = 0.4$

$\therefore\ P(66 < X < 72) = P(-0.8 < Z < 0.4)$



= area under st. normal curve enclosed between the two ordinates $Z=-0.8$ and $Z=0.4$ (shown by shade in the figure)

$$= \int_{-0.8}^{0.4}\phi(t)dt = \int_{-0.8}^{0}\phi(t)dt + \int_{0}^{0.4}\phi(t)dt$$

$$= \int_{0}^{0.8}\phi(t)dt + \int_{0}^{0.4}\phi(t)dt \text{, since } \phi \text{ curve is symmetric about } Y \text{ axis.}$$

$= 0.2881 + 0.1554 = 0.4435$

$\therefore$ Probability that the wage of a worker lies between Rs 66 and Rs. 72 is 0.4435.

$\therefore$ the number of worker whose wage lie between Rs. 66 and Rs. 72 $= 0.4435 \times 10{,}000 = 4435$.

(ii) When $X = 66$, $Z = -0.8$

So, $P(X < 66) = P(Z < -0.8)$

= area under the st. normal curve enclosed on left side of the ordinate $Z = -0.8$

$$= \int_{-\infty}^{-0.8}\phi(t)dt = \int_{0.8}^{\infty}\phi(t)dt \text{ (by symmetry)}$$

$$= 0.5 - \int_{0}^{0.8}\phi(t)dt = 0.5 - 0.2881 = 0.2119$$

$\therefore$ the expected number of workers $= 0.2119 \times 10{,}000 = 2119$

(iii) When $X = 72$, $Z = 0.4$

$$\therefore P(X > 72) = P(Z > 0.4) = \int_{0.4}^{\infty}\phi(t)dt$$

$$= 0.5 - \int_{0}^{0.4}\phi(t)dt = 0.5 - 0.1554 = 0.3446$$

$\therefore$ the required number $= 0.3446 \times 10{,}000 = 3446$.

Ex. 5. *The mean of a normal distribution is 50 and 5% of the values are greater than 60. Find the standard deviation of the distribution (Given that the area under standard normal curve between $z = 0$ and $z = 1.64$ is 0.45)*

Let X be the normal variate. Let its s.d. be σ $\therefore$ $Z = \frac{X-50}{\sigma}$ is standard normal variate. By problem

$$P(X > 60) = \frac{5}{100} = 0.05.$$

When $X = 60$, $Z = \frac{60-50}{\sigma} = \frac{10}{\sigma}$

. from above, $P\left(Z > \frac{10}{\sigma}\right) = 0.05$ ... (1)

From the supplied data we have $P(0 < Z < 1.64) = .45$

or, $P(Z > 1.64) = 0.5 - 0.45 = 0.05$ ... (2)

Comparing (1) and (2) we get $\frac{10}{\sigma} = 1.64$

or, $\sigma = \frac{10}{1.64} = 6.097 \approx 6.1$ $\therefore$ the s.d is 6.1.

Ex. 6. *In a normal distribution, 31% of the items are under 45 and 8% are above 64. Find the mean and standard deviation. [Given*

$$P(0 < Z < 1.405) = 0.42,\ P(-0.496 < Z < 0) = 0.19$$

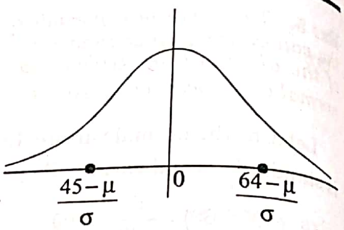
[W.B.U.Tech 2003]

Let X be the normal variate with mean μ and s.d σ. So, $Z = \frac{X-\mu}{\sigma}$ is standard normal variate. Now, by problem, $P(X<45) = \frac{31}{100}$ and $P(X>64) = \frac{8}{100}$. When $X = 45, Z = \frac{45-\mu}{\sigma}$; when $X = 64$, $Z = \frac{64-\mu}{\sigma}$. So from above $P\left(Z < \frac{45-\mu}{\sigma}\right) = 0.31$ and $P\left(Z > \frac{64-\mu}{\sigma}\right) = 0.08$.

Scanned by CamScanner

Since $0.31 < 0.5$ so

$\frac{45-\mu}{\sigma}$ is on negative side.



From above

$$P\left(\frac{45-\mu}{\sigma} < z < 0\right) = 0.5 - 0.31$$

$= 0.19$

Comparing this with the second given data we have

$$\frac{45-\mu}{\sigma} = -1.496$$

or, $45 - \mu = -1.496\,\sigma$ or, $\mu - 0.496\,\sigma = 45$ ... (1)

Since $0.08 < 0.5$ so $\frac{64-\mu}{\sigma}$ is lies on +ve side.

$$\therefore\ P\left(0 < z < \frac{64-\mu}{\sigma}\right) = 0.5 - 0.08 = 0.42$$

Comparing this with the first given data we get

$$\frac{64-\mu}{\sigma} = 1.405$$

or, $64 - \mu = 1.405\,\sigma$

or, $\mu + 1.405\sigma = 64$ ... (2)

Solving (1) and (2) $\sigma = 9.995$ and $\mu = 49.958$.

$\therefore$ Mean $= 49.958$ and s.d. $= 9.995$.

Ex. 7. *If X is normally distributed with mean 12 and s.d. 4, find the probability of (i)* $X \geq 20$ *(ii)* $0 \leq X \leq 12$ *and (iii) also find a such that* $P(X > a) = 0.24$. [Use table]

Let Z be the standard normal variate. Then

$$Z = \frac{X-\mu}{\sigma} = \frac{X-12}{4} \qquad \therefore\ X = 4Z + 12$$

$\therefore$ (i) $P(X \geq 20) = P(4Z + 12 \geq 20) = P(Z \geq 2) = 1 - P(Z < 2) = 1 - \Phi(2)$

$= 1 - 0.9772 = 0.0228$.

(ii) $P(0 \leq X \leq 12) = P(0 \leq 4Z + 12 \leq 12) = P(-3 \leq Z \leq 0)$

$= P(0 \leq Z \leq 3)$ (due to symmetry)

$= \Phi(3) - \Phi(0) = 0.9986 - 0.5 = 0.4986$ (from table)

(iii) $P(X > a) = 0.24$ or, $P(4Z + 12 > a) = 0.24$

or, $P\left(Z > \frac{a-12}{4}\right) = 0.24$ or, $1 - P\left(Z \leq \frac{a-12}{4}\right) = 0.24$

or, $1 - \Phi\left(\frac{a-12}{4}\right) = 0.24$, where Φ is c.d.f of st. normal variate

or, $\Phi\left(\frac{a-12}{4}\right) = 0.76$

or, $\frac{a-12}{4} = 0.71$, from table-I

$\therefore\ a = 14.84$

Ex. 8. *If a random variable X follows a normal distribution such that* $P(9.6 \leq X \leq 13.8) = 0.7008$, *and* $P(X \geq 9.6) = 0.8159$ *where* $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0.9} e^{-t^2/2}dt = 0.8159$, $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1.2} e^{-t^2/2}dt = 0.8849$, *find mean and variance of X.*

Let $E(X) = \mu$, $Var(X) = \sigma^2$. Now, if Φ is c.d.f of st normal variate

$$\Phi(x) = P\left(\frac{X-\mu}{\sigma} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}du$$

$\therefore\ \Phi(0.9) = .8159,\ \Phi(1.2) = 0.8849$

$\therefore\ \Phi(-0.9) = 1 - \Phi(0.9) = 0.1841$

Scanned by CamScanner

Now, $P\left(-0.9 \le \frac{X-\mu}{\sigma} \le 1.2\right) = \Phi(1.2) - \Phi(-0.9)$

$= 0.8849 - 0.1841 = 0.7008$

$= P(9.6 \le X \le 13.8)$

$= P\left(\frac{9.6-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{13.8-\mu}{\sigma}\right)$

$\therefore \frac{13.8-\mu}{\sigma} = 1.2, \ \frac{9.6-\mu}{\sigma} = -0.9.$

Solving we get $\mu = 11.4, \ \sigma = 2$

$\therefore$ mean $= 11.4, \ \text{var}(X) = 4.$

Ex. 9. *A fair cain is tassed 400 times. Using normal approximation to binomial distribution find the probability of obtaining (i) exactly 200 heads (ii) between 190 and 210 heads, both inclusive. Given that the area under standard normal curve between Z=0 and Z=0.05 is 0.0199 and between Z=0 and Z=1.05 is 0.3531.* [W.B.U.Tech 2007]

Let the random variable X denotes the number of heads in 400 tosses. Then clearly X has a binomial distribution with parameter $n = 400, p = \frac{1}{2}$. Since n is large we suppose X is an approximately normal variate with parameter $\mu = np = 400 \times \frac{1}{2} = 200$

and $\sigma = \sqrt{np(1-p)} = \sqrt{400 \times \frac{1}{2} \times \frac{1}{2}} = 10$

i.e. $Z = \frac{X-200}{10}$ is approximate standard normal variate.

(i) Now, $X = 199{\cdot}5 \Rightarrow X = \frac{199{\cdot}5-200}{10} = -0{\cdot}05$ and $X = 200{\cdot}5 \Rightarrow Z = 0{\cdot}05$.

Then using the normal approximation, we have the required probability, $P(X=200) \approx P(199{\cdot}5 \le X \le 200{\cdot}5) = P(-0{\cdot}05 \le Z \le 0{\cdot}05)$

$= 2P(0 \le Z \le 0{\cdot}05) = 2 \times 0{\cdot}0199 = 0{\cdot}0398$

(ii) Now, $X = 189.5 \Rightarrow Z = -1.05$ and $X = 210{\cdot}5 \Rightarrow Z = 1.05$.

Then using the normal approximation, we have the required probability,

$P(190 \le X \le 210)$

$\approx P(189.5 \le X \le 210.5)$ (in terms of normal approximation)

$= P(-1.05 \le Z \le 1.05) = 2P(0 \le Z \le 1.05) = 2 \times 0.3531 = 0.7062$

Ex. 10. *Among 10,000 random digits, find the probability that the digit 3 appears at most 950 times.*

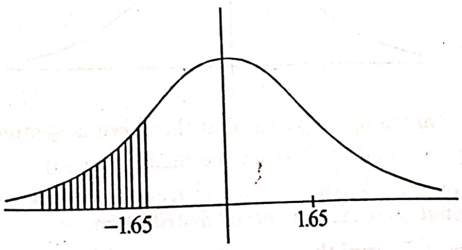
Let X denotes the number of times the digit 3 appears. Then

X is a $b(n, p)$ variate where $n = 10{,}000, \ p = \frac{1}{10}$

$\therefore np = 1000, \ np(1-p) = 10000 \times \frac{1}{10} \times \frac{9}{10} = 900$

$\because n$ is large we find $P(X \le 950)$ by approximating with normal distribution. X is approximately $\left(np, \sqrt{np(1-p)}\right) = (1000, 30)$ normal variate i.e., $Z = \frac{X-1000}{30}$ is approximately standard normal variate. Since discrete variate is approximated by continuous variate so we find $P(X \le 950.5)$.

Now $X = 950.5 \Rightarrow Z = -1.65$



$\therefore$ the required probability $\simeq P(X \le 950.5) = P(Z \le -1.65)$

$= 0.5 - P(-1.65 \le Z < 0) = 0.5 - P(0 \le Z \le 1.65)$

$= 0.5 - \int_0^{1.65} \phi(t)dt = 0.5 - 0.4505$ (From statistical table)

$= 0.0495$.

1.3.9. t - Distribution or Student's Distribution.

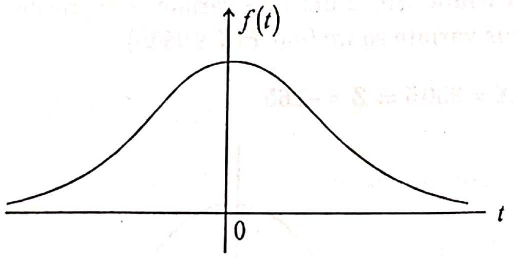
The t-distribution is a continuous distribution whose probability density function is given by

$$f(t) = \frac{1}{\sqrt{n}B\left(\frac{1}{2}, \frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < t < \infty$$

where n is a positive integer called the number of degres of freedom of the distribution and is the only parameter of the distribution. $B\left(\frac{1}{2}, \frac{n}{2}\right)$ is the welknown Beta-Function.

t-Probability Density Curve or *t-Curve*.

The prob. density curve is shown in the following fig. for a fixed n, say for $n = 4$.



From the fig. it is clear that the curve is symmetric about the y-axes and has maximum ordinate at $t = 0$.

Theorem. As the degrees of freedom tends to ∞ the t-distribution tends to a normal distribution.

Proof. Beyond the scope of the text.

Note. Because of having the above result t distribution has an important role in theory of estimation, test of hypothesis. These are discussed in the later chapters.

1.3.10. χ^2-Distribution (Chi-Square Distribution).

χ^2 distribution is a continuous distribution whose probability density function is given by

$$f(\chi^2) = \frac{e^{-\frac{\chi^2}{2}}\left(\frac{1}{2}\chi^2\right)^{\frac{n}{2}-1}}{2\Gamma\left(\frac{1}{2}n\right)}, \quad \chi^2 > 0$$

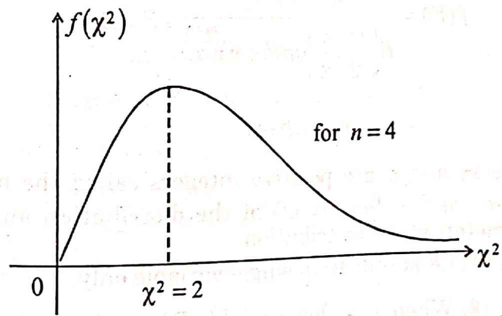
$$= 0, \quad \text{elsewhere}$$

where n is a positive integer called the number of degrees of freedom (d.o.f) of the distribution and is the only parameter of the distribution. $\Gamma\left(\frac{1}{2}n\right)$ is a Gamma-Function

Note. **(1)** χ^2 stands for a single notation-it only tells us that it is always non-negative. But never think it is $\chi \cdot \chi$. **(2)** Here X is a chi-square variate with parameter n. We denote it by $X \sim \chi^2(n)$

The χ^2 Prob. Density Curve or, χ^2-Curve.

It is shown in the following figure for fixed n (say $n = 4$)



From the figure it is clear that $f(\chi^2)$ is maximum at $\chi^2 = 2$. It is positively skew. Starting from 0 it extends to ∞ on the right.

Theorem 1. If X and Y are two independent random variables having χ^2 distribution with degree of freedom n_1 and n_2 then their sum $X+Y$ has χ^2 distribution with d.o.f n_1+n_2.

Proof. Beyond the scope.

Theorem 2. If a χ^2 variate's d.o.f $n \to \infty$ then the variate $\sqrt{2\chi^2} - \sqrt{2n-1}$ tends to a normal variate.

Proof. Beyond the scope of the book.

Theorem 3. If $Z_1, Z_2, \cdots, Z_n$ are n independent standard normal variates then $Z_1^2 + Z_2^2 + \cdots + Z_n^2$ has χ^2 distribution with d.o.f n.

Proof. Beyond the scope.

Note. Because of the above important results χ^2 distribution has a significant role in theory of estimation and test of hypothesis which are discussed in the later chpaters.

1.3.11. F-Distribution (Snedecesor's Distribution).

F distribution is a continuous distribution whose probability density function is given by

$$f(F) = \frac{m^{\frac{m}{2}} n^{\frac{n}{2}} F^{\frac{m}{2}-1}}{B\left(\frac{m}{2}, \frac{n}{2}\right)(mF+n)^{\frac{m+n}{2}}}, F > 0$$

$$= 0, \quad \text{elsewhere}$$

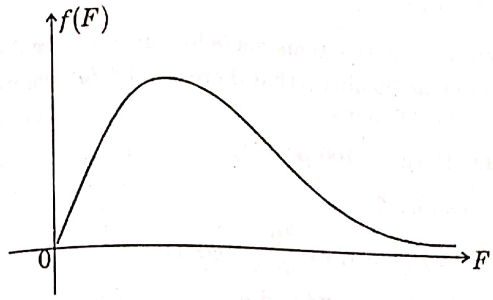
where m and n are positive integers called the number of degrees of freedom (d.o.f) of the distribution and are the parameters of the distribution.

Note : (1) F stands for a single variable only.

(2) When a random variable F has a F distribution with parameter m and n we write $F \sim F(m,n)$ and we say is a F variate with (m,n) degrees of freedom.

The F Prob. Density curve or F-curve

The curve of the pdf of F variate is highly positively skew.



Theorem 1. If χ_1^2 and χ_2^2 are two independent χ^2 variates with m an n degrees of freedom then $F = \frac{n\,\chi_1^2}{m\,\chi_2^2}$ is F variate with m and n degrees of freedom.

Proof. Beyond the scope of the book.

Theorem 2. If F is a F variate with (m,n) d.o.f then $\frac{1}{F}$ is a F variate with (n,m) d.o.f.

Proof. Beyond the scope of the book.

EXERCISES

[I] SHORT ANSWER QUESTIONS

1. If a conference room cannot be reserved for more than 4 hours, find the probability that a given conference lasts more than three hours.
   [Hints : Here $f(x) = \frac{1}{4}$ $P(X \ge 3) = \int_3^4 f(x)dx = \frac{1}{4}$]

2. If X is uniformly distributed in $[-\alpha, \alpha]$ with $\alpha > 0$. Then find α such that $P\left(X < \frac{1}{2}\right) = 0.7$.