POISSON DISTRIBUTION

15. POISSON DISTRIBUTION

Poisson distribution is related to the probability of events which are extremely rare but which have a large number of opportunities for occurrance. Poisson distribution deals with cases like: The number of persons killed in a railway accident during a particular year, the number of children born blind per year in a large city.

16. POISSON DISTRIBUTION AS A LIMITING CASE OF BINOMIAL DISTRIBUTION

(P.T.U., May 2005, 2007, Jan. 2009, May 2011)

Poisson distribution is a limiting case of Binomial Distribution. In B.D when parameters n and p are reasonably known, we can easily find distribution, but in situations where n is very large and p is very small then application of B.D becomes very laborious, then Poisson distribution is applied. So we apply Poisson distribution under following conditions:

(i) n, the number of trials is indefinitely large *i.e.*, $n \to \infty$

(ii) p, the constant probability of success for each trial is indefinitely small *i.e.*, $p \to 0$

(iii) $np = \lambda$ (say) is finite.

Now we will derive Poisson Distribution from Binomial Distribution.

For a Binomial Distribution: $P(X = r) = {}^nC_r q^{n-r} p^r$

$$= \frac{n(n-1)(n-2).....(n-r+1)}{r!} \times (1-p)^{n-r} \times p^r$$

$$= \frac{n(n-1)(n-2).....(n-r+1)}{r!} \times \left(1-\frac{\lambda}{n}\right)^{n-r} \times \left(\frac{\lambda}{n}\right)^r \qquad \left(\text{Since } np = \lambda \therefore p = \frac{\lambda}{n}\right)$$

$$= \frac{\lambda^r}{r!} \times \frac{n(n-1)(n-2).....(n-r+1)}{n^r} \times \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^r}$$

$$= \frac{\lambda^r}{r!}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right).....\left(\frac{n-r+1}{n}\right) \times \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^r}$$

$$= \frac{\lambda^r}{r!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right).....\left(1-\frac{r-1}{n}\right) \times \frac{\left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda}}{\left(1-\frac{\lambda}{n}\right)^r}$$

As $n \to \infty$, each of the $(r-1)$ factors

$$\left(1-\frac{1}{n}\right), \left(1-\frac{2}{n}\right), ..., \left(1-\frac{r-1}{n}\right) \quad \text{tends to 1. Also } \left(1-\frac{\lambda}{n}\right)^r \text{ tends to 1.}$$

Since $\underset{x \to \pm\infty}{\text{Lt}} \left(1+\frac{1}{x}\right)^x = e$, the Naperian base. $\therefore \left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda} \longrightarrow e^{-\lambda}$ as $n \to \infty$

Hence in the limiting case when $n \to \infty$, we have

$$\boxed{P(X=r) = \frac{\lambda^r e^{-\lambda}}{r!}} \qquad (r=0, 1, 2, 3, \ldots\ldots) \qquad \ldots(1)$$

where λ is a finite number $= np$.

(1) represents a Probability Distribution which is called the Poisson Probability Distribution.

Note 1. λ is called the parameter of the distribution.

Note 2. $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots\ldots + \frac{x^n}{n!} + \ldots\ldots$ to ∞.

Note 3. The sum of the probabilities $P(r)$ for $r = 0, 1, 2, 3, \ldots\ldots$ is 1, since

$$P(0) + P(1) + P(2) + P(3) + \ldots\ldots$$

$$= e^{-\lambda} + \frac{\lambda e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!} + \frac{\lambda^3 e^{-\lambda}}{3!} + \ldots\ldots = e^{-\lambda}\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots\ldots\right) = e^{-\lambda} . e^{\lambda} = 1.$$

Poisson distribution was derived in 1837 by a French Mathematician Simeon D Poisson (1781–1840).

7.17. RECURRENCE FORMULA FOR THE POISSON DISTRIBUTION

For Poisson Distribution, $P(r) = \frac{\lambda^r e^{-\lambda}}{r!}$ and $\qquad P(r+1) = \frac{\lambda^{r+1} e^{-\lambda}}{(r+1)!}$

$$\frac{P(r+1)}{P(r)} = \frac{\lambda r!}{(r+1)!} = \frac{\lambda}{r+1} \quad \text{or} \quad P(r+1) = \frac{\lambda}{r+1} P(r), r = 0, 1, 2, 3, \ldots\ldots$$

This is called the recurrence formula for the Poisson Distribution.

7.18. MEAN AND VARIANCE OF THE POISSON DISTRIBUTION (P.T.U., Jan. 2010, May 2014)

For the Poisson Distribution, $P(r) = \frac{\lambda^r e^{-\lambda}}{r!}$

$$\text{Mean } \mu = \sum_{r=0}^{\infty} rP(r) = \sum_{r=0}^{\infty} r . \frac{\lambda^r e^{-\lambda}}{r!}$$

$$= e^{-\lambda} \sum_{r=0}^{\infty} \frac{\lambda^r}{(r-1)!} = e^{\lambda}\left(\lambda + \frac{\lambda^2}{1!} + \frac{\lambda^3}{2!} + \cdots\right) = \lambda e^{-\lambda}\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \cdots\right) = \lambda e^{-\lambda} . e^{\lambda} = \lambda$$

Thus, the mean of the Poisson Distribution is equal to the parameter λ.

$$\text{Variance } \sigma^2 = \sum_{r=0}^{\infty} r^2 P(r) - \mu^2 = \sum_{r=0}^{\infty} r^2 . \frac{\lambda^r e^{-\lambda}}{r!} - \lambda^2 = e^{-\lambda} \sum_{r=1}^{\infty} \frac{r^2 \lambda^r}{r!} - \lambda^2$$

$$= e^{-\lambda}\left[\frac{1^2 . \lambda}{1!} + \frac{2^2 . \lambda}{2!} + \frac{3^2 \lambda^3}{3!} + \frac{4^2 \lambda^4}{4!} + \cdots\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[1 + \frac{2\lambda^2}{1!} + \frac{3\lambda^2}{2!} + \frac{3\lambda^3}{3!} + \cdots\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[1 + \frac{(1+1)\lambda}{1!} + \frac{(1+2)\lambda^2}{2!} + \frac{(1+3)\lambda^3}{3!} + \cdots\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[\left(1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\cdots\right)+\left(\frac{\lambda}{1!}+\frac{2\lambda^2}{2!}+\frac{2\lambda^3}{3!}+\cdots\right)\right]-\lambda^2$$

$$= \lambda e^{-\lambda}\left[e^{\lambda}+\lambda\left(1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\cdots\right)\right]-\lambda^2$$

$$= \lambda e^{-\lambda}\left[e^{\lambda}+\lambda e^{\lambda}\right]-\lambda^2 = \lambda e^{-\lambda}.e^{\lambda}(1+\lambda)-\lambda^2 = \lambda(1+\lambda)-\lambda^2 = \lambda\,.$$

Hence, **the variance of the Poisson Distribution is also λ.**

Thus, **the mean and the variance of the Poisson Distribution are each equal to the parameter λ.**

Note. The mean and the variance of the Poisson Distribution can also be derived from those of the Binomial distribution in the limiting case when $n \to \infty$, $p \to 0$ and $np = \lambda$.

Mean of Binomial Distribution is np.

$\therefore$ Mean of Poisson Distribution $= \underset{n\to\infty}{\text{Lt}}\ np = \underset{n\to\infty}{\text{Lt}}\ \lambda = \lambda$

Variance of Binomial Distribution is $npq = np\,(1-p)$

$\therefore$ Variance of Poisson Distribution $= \underset{n\to\infty}{\text{Lt}}\ np\,(1-p) = \underset{n\to\infty}{\text{Lt}}\ \lambda\left(1-\frac{\lambda}{n}\right) = \lambda\,.$

.19. PROPERTIES OF POISSON DISTRIBUTION (P.D.)

Poisson Distribution (P.D.) is $P(r) = \dfrac{\lambda^r e^{-\lambda}}{r!}$

Properties:

(*i*) It is a discrete Probability Distribution;

(*ii*) n; the number of trials is indefinitely large *i.e.*, $n \to \infty$

(*iii*) p; the probability of success for each trial is very small *i.e.*, $p \to 0$

(*iv*) $np = \lambda$(say), the parameter is finite

(*v*) Mean and variance of P.D are equal and each $= \lambda$, the parameter

(*vi*) Sum of the probabilities $P(r)$ for $r = 1, 2, 3, \ldots \infty$ is always equal to 1.

.20. FITTING A POISSON DISTRIBUTION

For fitting a Poisson's Distribution determine the following from the given data

(*i*) value of λ = mean of the given distribution

$$= \frac{\Sigma f\,(x)}{\Sigma f}$$

(*ii*) N = Total frequency = sum of all the frequencies.

(*iii*) Find the value of $e^{-\lambda}$ *i.e.*, calculate the frequency of zero *i.e.*, $P(X=0)$ where $P(X=r) = \dfrac{\lambda^r e^{-\lambda}}{r!}$

(*iv*) Compute the succession N P (0), N P (1), N P (2) etc.,

(*v*) Prepare the table showing variant $r = 0, 1, 2 \ldots, n$ in the 1st column, expected or theoretical frequencies represented by NP (r) in the 2nd column and the expected frequencies in round figures in the 3rd column.

Note. The sum or total of expected frequencies should also be equal to N.

ILLUSTRATIVE EXAMPLES

Example 1. *If the variance of the Poisson distribution is 2, find the probabilities for r = 1, 2, 3, 4 from the recurrence relation of the Poisson distribution.*

Solution. λ, the parameter of Poisson distribution = Variance = 2

Recurrence relation for the Poisson distribution is

$$P(r+1) = \frac{\lambda}{r+1}P(r) = \frac{2}{r+1}P(r) \quad \ldots(1)$$

Now, $$P(r) = \frac{\lambda^r e^{-\lambda}}{r!} \Rightarrow P(0) = \frac{e^{-2}}{0!} = e^{-2} = 0.1353$$

Putting $r = 0, 1, 2, 3$ in (1), we get

$$P(1) = 2P(0) = 2 \times 0.1353 = 0.2706; \quad P(2) = \frac{2}{2}P(1) = 1.2706$$

$$P(3) = \frac{2}{3}P(2) = \frac{2}{3} \times 0.2706 = 0.1804; \quad P(4) = \frac{2}{4}P(3) = \frac{1}{2} \times 0.1804 = 0.0902.$$

Example 2. *Assume that the probability of an individual coalminer being killed in a mine accident during a year is $\frac{1}{2400}$. Use Poisson's distribution to calculate the probability that in a mine employing 200 miners there will be at least one fatal accident in a year.*

Solution. Here $p = \frac{1}{2400}$, $n = 200$; $\therefore \lambda = np = \frac{200}{2400} = \frac{1}{12} = \cdot 083$

$$\therefore \quad P(r) = \frac{\lambda^r e^{-\lambda}}{r!} = \frac{(\cdot 083)^r e^{-\cdot 083}}{r!}$$

P(at least one fatal accident) = 1 – P(no fatal accident)

$$= 1 - P(0) = 1 - \frac{(\cdot 083)^0 e^{-\cdot 083}}{0!} = 1 \cdot 92 = \cdot 08.$$

Example 3. *Fit a Poisson distribution to the following data and calculate theoretical frequencies.*

x	:	0	1	2	3	4
f	:	122	60	15	2	1

(given $e^{-0.5} = 0.61$) **(P.T.U., May 2007, 2011)**

Solution. Mean of given distribution, $\lambda = \frac{\Sigma fx}{\Sigma f}$

$$\therefore \quad \lambda = \frac{0(122) + 1(60) + 2(15) + 3(2) + 4(1)}{122 + 60 + 15 + 2 + 1} = \frac{100}{200} = \frac{1}{2} = 0.5$$

$\therefore$ Required Poisson distribution $= N\frac{\lambda^r e^{-\lambda}}{r!}$, where $N = \Sigma f = 200$

$$\therefore \quad NP(r) = 200\frac{(0.5)^r e^{-0.5}}{r!} = \frac{200(0.61)(0.5)^r}{r!} = \frac{122(0.5)^r}{r!} \quad (\text{Given } e^{-0.5} = 0.61)$$

∴ Poisson distribution is :

r	$NP(r)$	Theoretical frequencies
0	$P(0)=122$	122
1	$P(1)=\dfrac{122(0.5)}{1!}=61$	61
2	$P(2)=\dfrac{122(0.5)^2}{2!}=15.25$	15
3	$P(3)=\dfrac{122(0.5)^3}{3!}=2.54$	2
4	$P(4)=\dfrac{122(0.5)^4}{4!}=0.32$	0
		Total = 200

Example 4. *If the probability that an individual suffers a bad reaction from a certain injection is 0.001. ind the probability that out of 2000 individuals*

(i) *exactly 3 individuals will suffer a bad reaction*

(ii) *none will suffer a bad reaction*

(iii) *more than one individual will suffer* **(P.T.U., May 2005)**

(iv) *more than two individual will suffer.* **(P.T.U., Jan. 2009)**

(P.T.U., Dec. 2011)

Solution. Here $p=0.001,\ n=2000$

$$\lambda=np=2000\times 0.001=2$$

$$\therefore \quad P(r)=\frac{\lambda^r e^{-\lambda}}{r!}=\frac{2^r e^{-2}}{r!}=\frac{1}{e^2}\cdot\frac{2^r}{r!}$$

(i). P (exactly 3 individual will suffer a bad reaction)

$$\Rightarrow \quad P(3)=\frac{1}{e^2}\frac{2^3}{3!}=\frac{8}{e^2.6}=\frac{4}{3e^2}=\frac{4}{3(2.718)^2}=0.18$$

(ii) $P(\text{none will suffer})=P(0)=\dfrac{1}{e^2}\dfrac{2^0}{0!}=\dfrac{1}{e^2}=\dfrac{1}{(2.718)^2}=0.135$

(iii) $P(\text{more than one})=P(2)+P(3)+P(4)+\ldots\ldots+P(2000)$

$$=1-[P(0)+P(1)]=1-\left[\frac{2^0}{e^2 0!}+\frac{1}{e^2}\frac{2^1}{1!}\right]=1-\frac{1}{e^2}[1+2]=1-3(0.135)=0.595.$$

(iv) $P(\text{more than two})=1-[P(0)+P(1)+P(2)]=1-\left[\dfrac{2^0}{e^2 0!}+\dfrac{2^1}{e^2 1!}+\dfrac{2^2}{e^2 2!}\right]$

$$=1-\frac{1}{e^2}[1+2+2]$$

$$=1-5(0.135)$$

$$=0.325$$

Example 5. *Six coins are tossed 1600 times. Using the Poisson distribution, determine the approximate probability of getting six heads x times.*

Solution. Probability of getting one head with one coin $= \frac{1}{2}$.

$\therefore$ The probability of getting six heads with six coins $= \left(\frac{1}{2}\right)^6 = \frac{1}{64}$

$\therefore$ Average number of six heads with six coins in 6400 throws $= np = 1600 \times \frac{1}{64} = 25$

$\therefore$ The mean of the Poisson distribution $= \lambda = 25$.

Approximate probability of getting six heads x times when the distribution is Poisson

$$= \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(25)^x e^{-25}}{x!}.$$

Example 6. *The probability that a man aged 50 years will die within a year is 0.01125. What is the probability that of 12 such men at least 11 will reach their fifty first birthday?*

Solution. $p =$ the probability that a man aged 50 years will die within a year $= 0.01125$

$n = 12$

$\therefore$ Mean $= \lambda = np = 12 \times 0.01125 = 0.135$

Let X denote the number of men aged fifty years who will die within a year.

$$\therefore \quad P(X = r) = \frac{\lambda^r e^{-\lambda}}{r!}$$

P (at least 11 will reach their 51^{st} birthday)

$= P$ (at the most one man dies within a year)

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

$$= \frac{\lambda^0 e^{-\lambda}}{0!} + \frac{\lambda^1 e^{-\lambda}}{1!} = e^{-\lambda} + \lambda e^{-\lambda} = e^{-\lambda}(1+\lambda)$$

$$= e^{-0.135}(1 + 0.135) = e^{-0.135}(1.135)$$

$$= (0.8731)(1.135) = 0.9916.$$

Example 7. *A product is 0.5% defective and is packed in cartoons of 100. What percentage contains not more than 3 defective?* **(P.T.U., May 2004)**

Solution. $p =$ The probability that a product is defective $= 0.5\%$

$$\therefore \quad p = \frac{0.5}{100} = 0.005$$

$n = 100$

$\lambda = np = 100\,(0.005) = 0.5$

Let X denotes the number of defective products

$$\therefore \quad P(X = r) = \frac{\lambda^r e^{-\lambda}}{r!} = \frac{(0.5)^r e^{-0.5}}{r!}$$

P(not more than 3 defective) $= P(X \le 3) = P(0) + P(1) + P(2) + P(3)$

$$= \frac{(0.5)^0 e^{-0.5}}{0!} + \frac{(0.5)^1 e^{-0.5}}{1!} + \frac{(0.5)^2 e^{-0.5}}{2!} + \frac{(0.5)^3 e^{-0.5}}{3!}$$

$$= e^{-0.5}\left[1 + (0.5) + \frac{0.25}{2} + \frac{0.125}{6}\right]$$

$$= (0.6065)\,[1 + 0.5 + 0.125 + 0.021]$$
$$= (0.6065)\,(1.646) = 0.9983 \text{ (approximately)}$$

Hence, required percentage $= (0.9983) \times 100 = 99.83$

Example 8. *Show that in a Poisson's Distribution with unit mean, mean deviation about mean is $\left(\frac{2}{e}\right)$ times the S.D.*

Solution. In Poisson's distribution, we have

$$P(r) = \frac{\lambda^r e^{-\lambda}}{!r}$$

Given mean $= 1 \quad \therefore \lambda = 1$

$$\text{S.D.} = \sqrt{\lambda} = 1$$

$$\therefore \quad P(r) = \frac{e^{-1}}{!r} = \frac{1}{e\,!r}$$

Now, mean deviation about mean in Poisson's Distribution

$$= \sum_{r=0}^{\infty} P(r)\,|\,r-1\,| \quad \because \text{ mean deviation} = \sum_{i=1} \frac{f_i\,|\,x_i - \bar{x}\,|}{N}$$

Here $\bar{x} = 1$
$x_i = r$ and $NP(r) = f_i$

r	$P(r)$	$\lvert r-1\rvert$	$P(r)\lvert r-1\rvert$
0	$\frac{1}{e}$	1	$\frac{1}{e}$
1	$\frac{1}{e\,1!}$	0	0
2	$\frac{1}{e\,2!}$	1	$\frac{1}{e\,2!}$
3	$\frac{1}{e\,3!}$	2	$\frac{2}{e\,3!}$
4	$\frac{1}{e\,4!}$	3	$\frac{3}{e\,4!}$

and so on.

$$\text{Mean deviation about mean} = \frac{1}{e} + 0 + \frac{1}{e\,2!} + \frac{2}{e\,3!} + \frac{3}{e\,4!} + \ldots \infty$$

$$= \frac{1}{e}\left[1 + 0 + \frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \ldots \infty\right]$$

$$= \frac{1}{e}\left[1 + \left(\frac{1}{1!} - \frac{1}{1!}\right) + \left(\frac{2}{2!} - \frac{1}{2!}\right) + \left(\frac{3}{3!} - \frac{1}{3!}\right) + \left(\frac{4}{4!} - \frac{1}{4!}\right) + \ldots\ldots \infty\right]$$

$$= \frac{1}{e}\left\{1 + \left(\frac{1}{1!} + \frac{2}{2!} + \frac{3}{3!} + \frac{4}{4!} - \ldots\ldots \infty\right) - \left(\frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} - \ldots\ldots \infty\right)\right\}$$

$$= \frac{1}{e}\left\{1 + \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots\ldots \infty\right) - (e-1)\right\}$$

$$= \frac{1}{e}[1 + e - e + 1] = \frac{2}{e} = \frac{2}{e}.1 = \frac{2}{e} \text{ S.D.}$$

TEST YOUR KNOWLEDGE

1. Fit a Poisson distribution to the following:

x :	0	1	2	3	4
f :	192	100	24	3	1

2. If the probability of a bad reaction from a certain injection is 0.001, determine the chance that out of 2000 individuals more than two will get a bad reaction. **(P.T.U., Jan. 2009)**

3. If X is a Poisson variate such that P(X = 2) = 9P(X = 4) + 90P(X = 6), find the standard deviation.

4. If a random variable has a Poisson distribution such that P(1) = P(2), find
(*i*) mean of the distribution (*ii*) P(4).

5. Suppose that X has a Poisson distribution. If $P(X = 2) = \frac{2}{3} P(X = 1)$, find (*i*) P(X = 0) (*ii*) P(X = 3).

6. A certain screw making machine produces on average 2 defective screws out of 100, and packs them in boxes of 500. Find the probability that a box contains 15 defective screws.
[**Hint:** $p = \frac{2}{100}, n = 500 \therefore \lambda = np = 10, e^{-10} = .000045$ Required probability P(15)]

7. The incidence of occupational disease in an industry is such that the workmen have a 10% chance of suffering from it. What is the probability that in a group of 7, five or more will suffer from it?
[**Hint:** $p = \frac{10}{100} = \frac{1}{10}, n = 7, \lambda = \frac{7}{10}; P(r) = \frac{e^{-\lambda}\lambda^r}{r!}$ Required probability = P(5) + P(6) + P(7)]

8. Data was collected over a period of 10 years, showing number of deaths from horse kicks in each of the 200 army corps. The distribution of deaths was as follows:

No. of deaths:	0	1	2	3	4	Total
Frequencies:	109	65	22	3	1	200

Fit a Poisson distribution to the data and calculate the theoretical frequencies. **(P.T.U., 2005)**

9. 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 mean 1.5. Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused. $(e^{-1.5} = 0.2231)$

10. Fit a Poisson distribution to the following data given the number of yeast cells per square for 400 squares:

No. of cells per sq. :	0	1	2	3	4	5	6	7	8	9	10
No. of squares :	103	143	98	42	8	4	2	0	0	0	0

11. The probability that a man aged 35 years will die before reaching the age of 40 years may be taken as 0.018. Out of a group of 400 men, now aged 35 years, what is probability that 2 men will die within the next 5 years?
[**Hint:** $p = 0.018, n = 400, \lambda = np = 7.2, e^{-\lambda} = e^{-7.2} = 0.0007466$]

12. Suppose a book of 585 pages contains 43 typographical errors. If these are randomly distributed throughout the book, what is the probability that 10 pages, selected at random, will be free from errors?
[**Hint:** $p = \frac{43}{585} = 0.0735, n = 10; \lambda = np = 0.735, e^{-\lambda} = 0.4795$ Required probability = P(0)]

13. Six coins are tossed 6400 times. Use Poisson's Distribution to find an expression for probability of getting six heads x times.

ANSWERS

1. $320 \times \frac{e^{0.503}(9.503)^r}{r!}$
2. 0.32
3. 1
4. (i) 2 (ii) $\frac{2}{3e^2}$
5. (i) e^{-4} (ii) $4e^{-4}$
6. $\frac{(10)^{15} e^{-10}}{(15)!} = 0.035$
7. 0.0008
8. $(108.7) \times \frac{(0.61)^r}{r!}$, where $r = 0, 1, 2, 3, 4$

   Theoretical frequencies are 109, 66, 20, 4, 1 respectively
9. 0.1912625
10. Theoretical frequencies are 109, 142, 92, 40, 13, 3, 1, 0, 0, 0, 0
11. 0.01936
12. 0.4795
13. $P(X = x) = \frac{(100)^x e^{-100}}{x!}$.

7.21. So Far We have Dealt with Discrete Probability Distribution where the Variates Take only Integral Values but Now We are Going to Discuss Continuous Probability Distribution

For a continuous function $f(x)$ in the interval $(-\infty, \infty)$

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

$$\text{Variance} = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x)\, dx$$

7.21(a). PROBABILITY DENSITY FUNCTION

As discussed in art 7.5 a continuous function $f(x)$ is called a probability density function in the interval $a \le x \le b$, if it satisfies the following properties:

(i) $f(x) \ge 0$ for all values of x in $a \le x \le b$ (the interval of continuity)

(ii) $\int_a^b f(x)\, dx = P(a \le x \le b)$.

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

NORMAL DISTRIBUTION

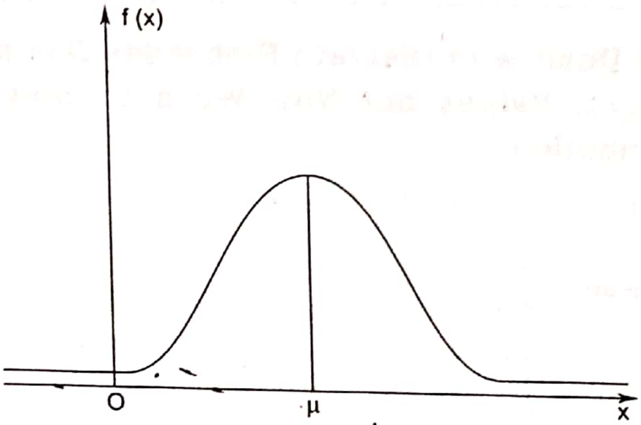
7.22. NORMAL DISTRIBUTION (P.T.U., May 2005, May 2007)

The normal distribution is a continuous distribution. Any quantity whose variation depends on random causes is distributed according to the normal law. Its importance lies in the fact that a large number of distributions approximate to the normal distribution. It can be derived from the Binomial Distribution in the limiting case when n, the number of trials is very large and p, the probability of a success, is close to $\frac{1}{2}$. The general equation of the normal distribution is given by,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$, where the variable x can assume all values from $-\infty$ to $+\infty$. μ and σ, called the parameters of the distribution, are respectively the **mean** and the **standard deviation** of the distribution and $-\infty < \mu < \infty, \sigma > 0$. x is called the normal variate and $f(x)$ is called **probability density function** of the normal distribution.

If a variable x has the normal distribution with mean μ and standard deviation σ, we briefly write $x : N(\mu, \sigma^2)$.

The graph of the normal distribution is called the *normal curve.* It is bell-shapped and symmetrical about the mean μ. The two tails of the curve extend to $+\infty$ and $-\infty$ towards the positive and negative directions of the x-axis respectively and gradually approach the x-axis without ever meeting it. The curve is unimodal and



the mode of the normal distribution coincides with its mean μ. The line $x = \mu$ divides the area under the normal curve above x-axis into two equal parts. Thus, the median of the distribution also coincides with its mean and mode. The area under the normal curve between any two given ordinates $x = x_1$ and $x = x_2$ represents the probability of values falling into the given interval. The total area under the normal curve above the x-axis is 1.

7.23. STANDARD FORM OF THE NORMAL DISTRIBUTION

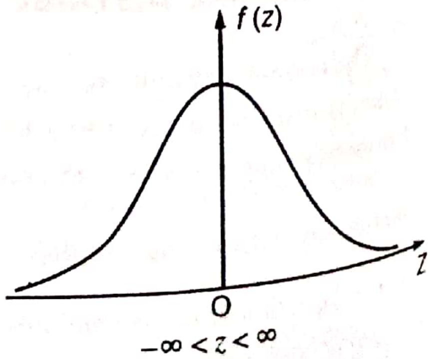
If x is a normal random variable with mean μ and standard deviation σ, then the random variable $z = \frac{x - E(x)}{\sigma} = \frac{x - \mu}{\sigma}$ has the normal distribution with mean 0 and standard deviation 1 written as N(0, 1). The random variable z is called the **standardized (or standard) normal random variable.**

The probability density function for the Normal Distribution in standard form is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad ; -\infty < z < \infty$$

It is free from any parameter.

Probability of x lying between x_1 and x_2 is given by the area under the normal curve from x_1 to x_2 and above x-axis



i.e.,
$$P(x_1 \le x \le x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

In standard form area under the normal probability curve from z_1 to z_2, above z-axis is given by

$$P(z_1 \le z \le z_2) = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} dz, \text{where } z = \frac{x-\mu}{\sigma} \quad \therefore \quad dz = \frac{1}{\sigma} dx. \text{ Also } z_1 = \frac{x_1 - \mu}{\sigma},\ z_2 = \frac{x_2 - \mu}{\sigma}$$

$$= \frac{1}{\sqrt{2\pi}} \left\{ \int_{z_1}^{0} e^{-\frac{z^2}{2}} dz + \int_{0}^{z_2} e^{-\frac{z^2}{2}} dz \right\} = \frac{1}{\sqrt{2\pi}} \left\{ -\int_{0}^{z_1} e^{-\frac{z^2}{2}} dz + \int_{0}^{z_2} e^{-\frac{z^2}{2}} dz \right\}$$

$$= \{-P_1(z) + P_2(z)\}, \text{ where } P(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{z^2}{2}} dz$$

$$= P_2(z) - P_1(z)$$

values of each of the integrals can not be integrated by ordinary means but can be obtained with the help of the table given at the end of the book.

Note. As total area under the normal curve above x-axis is 1

i.e., $P(-\infty < x < \infty) = 1 \ \therefore\ P(0 < x < \infty) = P(-\infty < x < 0) = 0.5.$

7.24. AREA UNDER THE NORMAL CURVE

In the normal curve when we take $z = \dfrac{x-\mu}{\sigma}$, the standard normal curve is formed. The total area under this curve above z-axis is '1' and is divided into two equal parts by $z = 0$.

$\therefore$ The area enclosed by the curve on left hand side $z = 0$ is equal to the area enclosed by the curve on the right hand side of $z = 0$ and each area $= 0.5$. *i.e.*, $P(-\infty \le z \le \infty) = 1$ and $P(-\infty \le z \le 0) = P(0 \le z \le \infty) = 0.5$. The area between the ordinate $z = 0$ and any other ordinate can be noted from the table of normal curves given at the end of the book.

Note 1. The probabilities $P(z_1 \le Z < z_2)$, $P(z_1 < Z \le z_2)$, $P(z_1 \le Z \le z_2)$ and $P(z_1 < Z < z_2)$ are all regarded to be the same.

Note 2. $P(-a \le z \le 0) = P(0 \le z \le a)$

$P(-a \le z \le c) = P(-a \le z \le 0) + P(0 \le z \le c) = P(0 \le z \le a) + P(0 \le z \le c)$

$P(-c \le z \le -d) = P(c \le z \le d) = P(0 \le z \le d) - P(0 \le z \le c)$.

7.25. BASIC PROPERTIES OF THE NORMAL DISTRIBUTION

(P.T.U., Dec. 2003, May 2005, May 2008, Jan. 2009)

(*i*) The Normal Distribution is a continuous Probability Distribution. It can be used to approximate Binomial and Poisson Distributing.

(*ii*) In the probability density function of the Normal Distribution

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

$f(x) \ge 0$ and $\int_{-\infty}^{\infty} f(x)\, dx = 1,$

i.e., the total area under the normal curve above the x-axis is 1. It has bell shaped graph.

(*iii*) As the normal curve has only one maximum point at mean $\therefore$ it has only one mode. So it is unimodal.

(*iv*) The mean, mode and median of this distribution coincide.

(*v*) The normal distribution is symmetrical about its mean. The ordinate at $x = \mu$ divides the whole area into two equal parts. The area to the right as well as to the left of ordinate is 0.5.

(*vi*) The mean deviation in normal distribution $= \frac{4}{5}$ of standard deviation.

(*vii*) All the odd moments about the mean are zero.

(*viii*) The points of inflexion of the normal curve are given by $\mu \pm \sigma$.

7.26. APPLICATIONS OF NORMAL DISTRIBUTION

Normal distribution is applied to problems concerning:

(*i*) Calculation of errors made by chance is experimental measurements

(*ii*) Computation of hit probability of a shot

(*iii*) Statistical inference is almost every branch of science.

7.27. MEAN FOR NORMAL DISTRIBUTION

We know that mean $= \int_{-\infty}^{\infty} x \cdot f(x)\, dx = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$

Put $\quad \frac{x-\mu}{\sigma} = t \quad \therefore\; x = \mu + \sigma t$

$$dx = \sigma\, dt$$

$$\therefore \quad \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{\infty} \frac{\mu + \sigma t}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\, \sigma\, dt$$

$$= \int_{-\infty}^{\infty} \frac{\mu}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\, dt + \int_{-\infty}^{\infty} \frac{\sigma}{\sqrt{2\pi}} t e^{-\frac{1}{2}t^2}\, dt$$

$$= \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2}\, dt + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-\frac{1}{2}t^2}\, dt$$

As $e^{-\frac{1}{2}t^2}$ is an even function $\therefore \quad \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2}\, dt = 2\int_{0}^{\infty} e^{-\frac{1}{2}t^2}\, dt$

As $te^{-\frac{1}{2}t^2}$ is an odd function $\therefore \quad \int_{-\infty}^{\infty} te^{-\frac{1}{2}t^2}\, dt = 0$

$$\int_{-\infty}^{\infty} x f(x)dx = \frac{u}{\sqrt{2\pi}} 2\int_0^{\infty} e^{-\frac{1}{2}t^2} dt$$

Put $\frac{1}{2}t^2 = u \quad \therefore \quad t\,dt = du \quad \text{or} \quad dt = \frac{du}{\sqrt{2u}}$

$$\therefore \int_{-\infty}^{\infty} x f(x)dx = \frac{2u}{\sqrt{2\pi}} \int_0^{\infty} e^{-u} \frac{du}{\sqrt{2u}} = \frac{u}{\sqrt{\pi}} \int_0^{\infty} u^{-\frac{1}{2}} e^{-\mu} du$$

$$= \frac{u}{\sqrt{\pi}} \int_0^{\infty} u^{\frac{1}{2}-1} e^{-u} du$$

$$= \frac{u}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) \qquad \left[\because \int_0^{\infty} e^{n-1} e^{-x} dx = \Gamma(n)\right]$$

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

$$= u$$

28. STANDARD DEVIATION FOR NORMAL DISTRIBUTION

We know that standard deviation is the positive square root of variance

$\therefore$ We will first find out variance of Normal Distribution

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

Put $\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 = t$

$$\therefore \quad \frac{1}{2}\cdot 2\left(\frac{x-\mu}{\sigma}\right)\frac{dx}{\sigma} = dt \quad \text{or} \quad \sqrt{2t}\,\frac{dx}{\sigma} = dt \quad \text{or} \quad dx = \frac{\sigma\,dt}{\sqrt{2t}}$$

$\therefore$ Variance

$$= 2\int_0^{\infty} \frac{2\sigma^2 t}{\sigma\sqrt{2\pi}} e^{-t} \frac{\sigma\,dt}{\sqrt{2t}} = \frac{2\sigma^2}{\sqrt{\pi}} \int_0^{\infty} t^{\frac{1}{2}} e^{-t} dt = \frac{2\sigma^2}{\sqrt{\pi}} \int_0^{\infty} t^{\frac{3}{2}-1} e^{-t} dt$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \Gamma\frac{3}{2} \qquad \left[\because \int_0^{\infty} x^{n-1} e^{-x} dx = \Gamma(n)\right]$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \frac{1}{2} \Gamma\frac{1}{2} \qquad |\because \Gamma(n) = (n-1)\Gamma(n-1)| \text{ [where } n \text{ is a fraction]}$$

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

$\therefore$ Variance $= \sigma^2$ and Standard Deviation $= \sigma$

Note. The parameters μ and σ in the normal equation are the mean and standard deviation of the normal distribution.

ILLUSTRATIVE EXAMPLES

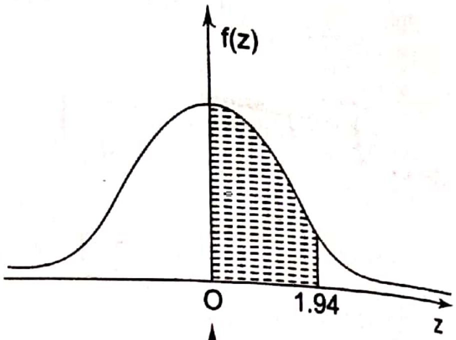
Example 1. *Find the area under the normal curve in the following cases :*

(*i*) $z = 0$ and $z = 1.94$

(*ii*) $z = -0.46$ and $z = 0$

(*iii*) $z = -0.68$ and $z = 2.21$

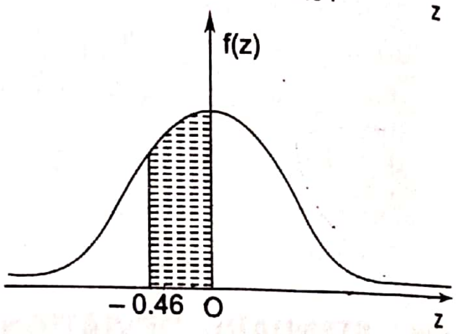
(*iv*) *To the left of* $z = -0.6$.



Solution. (*i*) Required area $= P(0 \le z \le 1.94)$

$= 0.4738$ (From the table of Normal curve given at the end of the book)

(*ii*) Required area

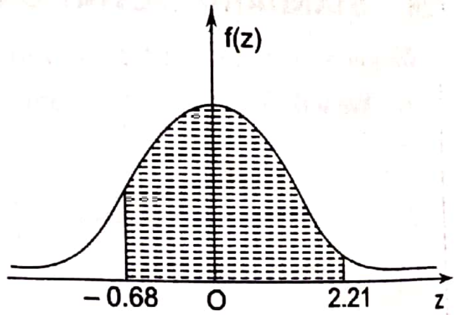


$$= P(-0.46 \le z \le 0)$$
$$= P(0 \le z \le 0.46)$$
$$= 0.1772.$$

$$= 2\int_0^{\infty} \frac{(x-\mu)^2}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

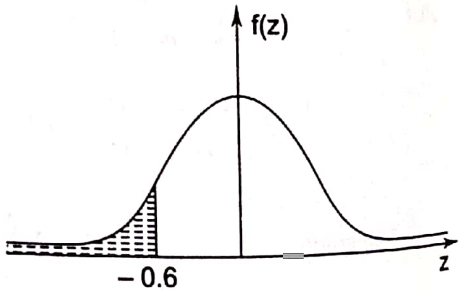
($\because$ Integrand is even)

(*iii*) Required area



$$= P(-0.68 \le z \le 2.21)$$
$$= P(-0.68 \le z \le 0) + P(0 \le z \le 2.21)$$
$$= P(0 \le z \le 0.68) + P(0 \le z \le 2.21)$$
$$= 0.2518 + 0.4865$$
$$= 0.7383.$$

(*iv*) Required area



$$= P(-\infty \le z \le -0.6)$$
$$= 0.5 - P(-0.6 \le z \le 0)$$
$$= 0.5 - P(0 \le z \le 0.6)$$
$$= 0.5 - 0.2257$$
$$= 0.2743.$$

Example 2. *Find the value of z in each of the following cases:*

(*i*) *Area between 0 and z is 0.3621.* (*ii*) *Area to the left of z is 0.7642.*

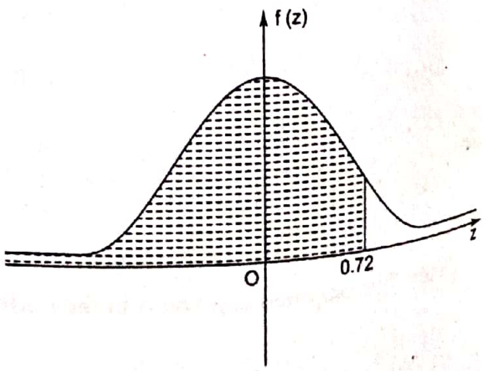
Solution. (*i*) As 0.3621 is < 0.5.

$\therefore$ Area is either on the right or on the left hand side of $z = 0$.

From table, we see that this area 0.3621 corresponds to $z = 1.0 + 0.09$

$\therefore$ $z = \pm(1.0 + 0.09) = \pm 1.09$

$\therefore$ $-1.09 \le z \le 0$ or $0 \le z \le 1.09$



(*ii*) As $0.7642 > 0.5$

$\therefore$ Area between 0 and z is

$$= 0.7642 - 0.5 = 0.2642$$

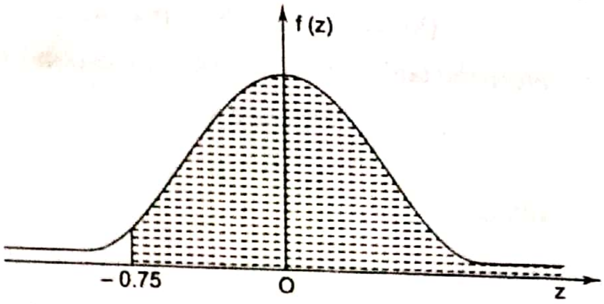
From the table 0.2642 corresponds to $z = 0.7 + 0.02 = 0.72$

$\therefore$ z lies between $-\infty$ and 0.72

i.e., $-\infty < z \le 0.72$ or simply $z \le 0.72$

Example 3. *Let x be normal random variable with mean 10 and S.D. 4. Determine the probability $P(x \geq 7)$.*

Solution. When $x = 7$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{7-10}{4} = -0.75$



$$\therefore \; P(x \geq 7) = P(z \geq -0.75) = P(-0.75 \leq z \leq 0) + P(0 \leq z < \infty)$$
$$= P(0 \leq z \leq 0.75) + 0.5$$
$$= 0.2734 + 0.5$$
$$= 0.7734.$$

Example 4. *A sample of 100 dry battery cells tested to find the length of life produced the following results:*

$$\bar{x} = 12 \text{ hours}, \sigma = 3 \text{ hours}.$$

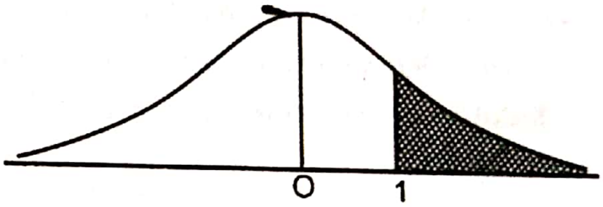
Assuming the data to be normally distributed, what percentage of battery cells are expected to have life.

(i) *more then 15 hours* (ii) *less than 6 hours*

(iii) *between 10 and 14 hours?*

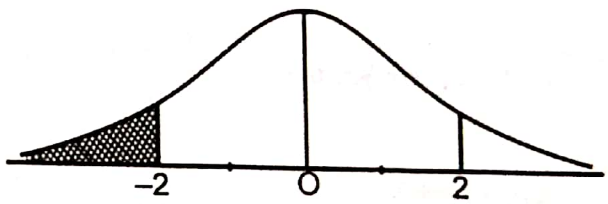
Solution. Here, x denoted the length of life of dry battery cells.

Also $$z = \frac{x-\bar{x}}{\sigma} = \frac{x-12}{3}.$$



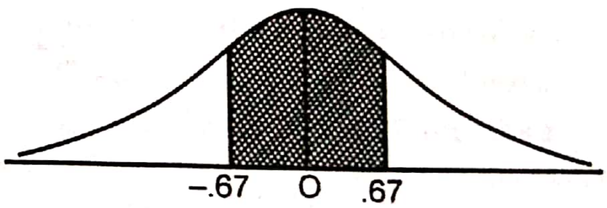
(i) When $x = 15$, $z = 1$

$$\therefore \quad P(x > 15) = P(z > 1)$$
$$= P(0 < z < \infty) - P(0 < z < 1)$$
$$= .5 - 0.3413 = 0.1587 = 15.87\%.$$



(ii) When $x = 6$, $z = -2$

$$\therefore \quad P(x < 6) = P(z < -2)$$
$$= P(z > 2) = P(0 < z < \infty) - P(0 < z < 2)$$
$$= 5 - 0.4772 = 0.0228 = 2.28\%.$$



(iii) When $x = 10$, $z = -\dfrac{2}{3} = -0.67$

When $x = 14$, $z = \dfrac{2}{3} = 0.67$

$$P(10 < x < 14) = P(-0.67 < z < 0.67)$$
$$= 2P(0 < z < 0.67) = 2 \times 0.2487 = 0.4974 = 49.74\%.$$

Example 5. *In a normal distribution, 31% of the items are under 45 and 8% are over 64. Find the mean and standard deviation of the distribution.* **(P.T.U., May 2006, Dec. 2013)**

Solution. Let $\bar{x}$ and σ be the mean and S.D respectively. 31% of the items are under 45.

$\Rightarrow$ Area to the left of the ordinate $x = 45$ is 0.31 when $x = 45$, let $z = z_1$

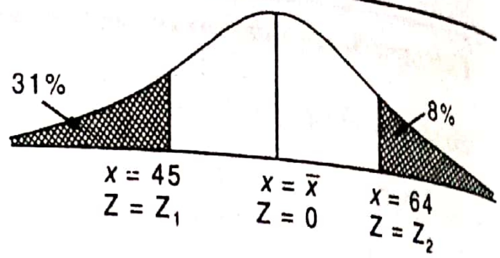
$$P(z_1 < z < 0) = .5 - .31 = .19$$

From the tables, the value of z corresponding to this area is 0.5

$$\therefore \quad z_1 = -0.5 \quad [z_1 < 0]$$

When $x = 64$, let $z = z_2$

$$P(0 < z < z_2) = .5 - .08 = .42$$



From the tables, the value of z corresponding to this area is 1.4.

$$z_2 = 1.4$$

Since, $$z = \frac{x - \bar{x}}{\sigma}$$

$$-0.5 = \frac{45 - \bar{x}}{\sigma} \quad \text{and} \quad 1.4 = \frac{64 - \bar{x}}{\sigma}$$

$\Rightarrow$ $$45 - \bar{x} = -0.5\sigma \quad ...(1)$$

and $$64 - \bar{x} = 1.4\sigma \quad ...(2)$$

Subtracting $-19 = -1.9\sigma$ $\qquad \therefore \sigma = 10$

From (1), $45 - \bar{x} = -0.5 \times 10 = -5$ $\qquad \therefore \bar{x} = 50.$

Example 6. *On a statistics examination the mean score was 78 and S.D was 10.*

(i) Determine standard score of two boys whose score was 93 and 62 respectively.

(ii) Determine the score of two students whose standard score was – 0.6 and 1.4 respectively.

Solution. Mean score = $\mu = 78$ and S.D = $\sigma = 10$

$\therefore$ $$z = \frac{x - 78}{10} \quad ...(1)$$ where x stands for score and z stands for standard score.

(*i*) When $x = 93$, then $z = \dfrac{93 - 78}{10} = 1.5$

when $x = 62$, $\quad z = \dfrac{62 - 78}{10} = -1.6$

3 = -0.6
103 = x+78
10(0.6) = 78 = x

(*ii*) When $z = -0.6$, $x = 10z + 78 = -6 + 78 = 72$

when $z = 1.4$, $\quad x = (1.4)\,10 + 78 = 14 + 78 = 92.$

Example 7. *In an examination taken by 500 candidates, the average and the standard deviation of marks obtained (normally distributed) are 40% and 10%. Find approximately*

(i) how many will pass, if 50% is fixed as a minimum?

(ii) what should be the minimum if 350 candidates are to pass?

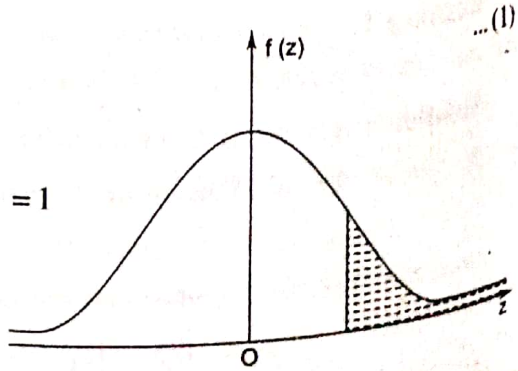
(iii) how many have scored marks above 60%? **(P.T.U., May 2004, Dec. 2011)**

Solution. Average deviation *i.e.*, mean $\mu = 40\% = 0.4$

Standard deviation $\quad$ *i.e.*, $\sigma = 10\% = 0.1$

$\therefore$ $$z = \frac{x - \mu}{\sigma} = \frac{x - 0.4}{0.1} \quad ...(1)$$

(*i*) When 50% is fixed as minimum marks



then $\quad x \geq 50\% = 0.5 \quad \therefore z \geq \dfrac{0.5 - 0.4}{0.1} = \dfrac{0.1}{0.1} = 1$

Probability of pass students = $P(x \geq 0.5) = P(z \geq 1)$

$$= 0.5 - P(0 \leq z \leq 1)$$

$$= 0.5 - 0.3413 = 0.1587$$

Total number of pass students

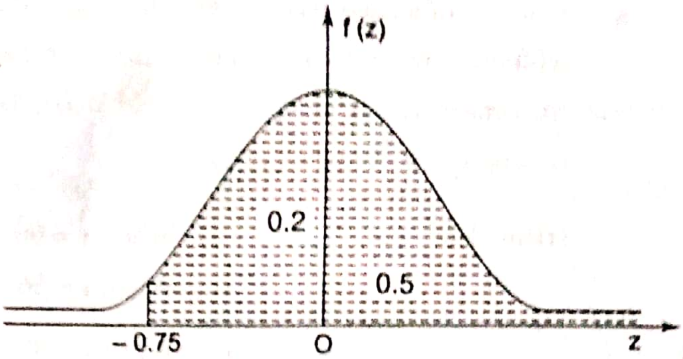
$= 500 \times (0.1587) = 79.35 = 79$ students.

(ii) Number of pass students 350

Probability of pass students $= \dfrac{350}{500} = 0.7.$

As $0.7 > 0.5$

$\therefore$ Area 0.7 = area 0.2 + area 0.5



Area 0.2 will be on the left hand side of O

$\therefore$ From tables we see that 0.2 on the LHS of 'O' corresponds to the value $z = -0.53$

From (1), $$-0.53 = \frac{x - 0.4}{0.1} \quad \text{or } x = 0.347$$

$= 34.7\%.$

$\therefore$ If 350 candidates are to pass then minimum % of marks required is 34.7 *i.e.*, 35%.

(iii) When $x \geq 60\% = 0.6$

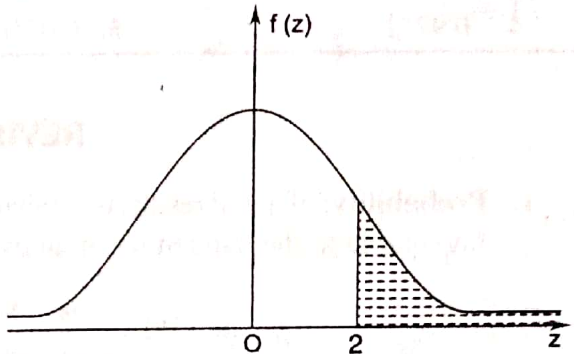
then $$z = \frac{0.6 - 0.4}{0.1} = 2$$

$\therefore$ Probability of students getting more than 60%

$= P(x \geq 0.6) = P(z \geq 2)$

$= 0.5 - P(0 \leq z \leq 2)$

$= 0.5 - 0.4772 = 0.228.$



Total number of students getting more than 60% $= 500 \times 0.0228 = 11.4$

i.e., required number of students = 11.

TEST YOUR KNOWLEDGE

1. The mean height of 500 male students in a certain college is151 cm and the standard deviation is 15 cm. Assuming the heights are normally distributed, find how many students have heights between 120 and 155 cm?
2. An aptitude test from selecting officers in a bank is conducted on 1000 candidates. The average score is 42 and the standard deviation of score is 24. Assuming normal distribution for the scores, find
   (*i*) the number of candidates whose scores exceed 60
   (*ii*) the number of candidates whose scores lie between 30 and 60. **[Hint:** Consult example 4]
3. (*a*) In a normal distribution, 7% of the items are under 35 and 89% are under 63. What are the mean and standard deviation of the distribution? **[Hint:** Consult solved example 5] **(P.T.U., May 2014)**
   (*b*) In a normal distribution 12% of the items are under 30 and 55% are under 60. Find the mean and S.D of the distribution. **[Hint:** $z_1 = -1.175$; $z_2 = 1.0365$] **(P.T.U., Jan. 2008)**
4. Let X denote the number of scores on a test. If X is normally distributed with mean 100 and standard deviation 15, find the probability that X does not exceed 130.
5. It is known from the past experience that the number of telephone calls made daily in a certain community between 3 p.m. and 4 p.m. have a mean of 352 and a standard deviation of 31. What percentage of the time will there be more than 400 telephone calls made in this community between 3 p.m. and 4 p.m.?
   [Hint: $z = \dfrac{400 - 352}{31} = 1.55$, find $P(z \geq 1.55)$]