CHAPTER - 5

CONTINUOUS PROBABILITY DISTRIBUTIONS

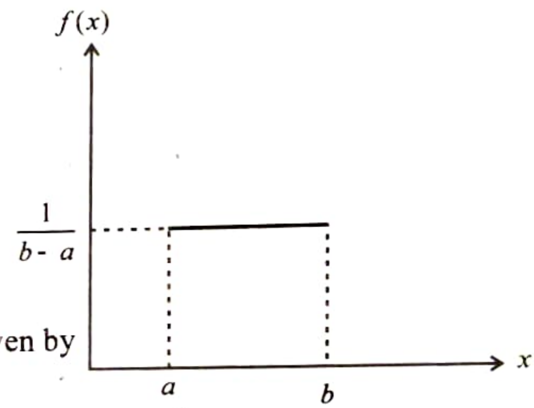
[Uniform, Normal, Exponential and Gamma Distributions]

5.1 INTRODUCTION

In the preceding chapter we have discussed the discrete probability distributions : Binomial and Poisson. In this chapter, we shall study the continuous probability distributions namely, Normal, Exponential and Gamma distributions. Before we discuss these distributions, we first give a brief introduction of Uniform (or Rectangular) distribution.

5.2 UNIFORM (OR RECTANGULAR) DISTRIBUTION

Uniform distribution is a continuous distribution. A random variable X is said to be uniformly distributed over the interval $-\infty < a < b < \infty$, if its probability density function is constant over the entire range



i.e. $f(x) = \begin{cases} k, a < x < b \\ 0, \text{elsewhere} \end{cases}$

where k is a constant given by

$$\int_a^b f(x)\, dx = 1 \Rightarrow k(b-a) = 1 \Rightarrow k = \frac{1}{b-a}$$

Hence the probability density function is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a}, \text{ for } a \le x \le b \\ 0, \text{ for } x < a \text{ or } x > b \text{ (or) for all other values of } x. \end{cases}$$

Graph of a uniform Distribution :

The graph of $f(x)$ is a flat line. The area under it between any two points a and b will be the area of a rectangle with height $\dfrac{1}{b-a}$ and width $(b-a)$

Since the probability density curve $y = f(x)$ resembles a rectangle on the x-axis and between the ordinates at $x = a$ and $x = b$. The Uniform Distribution is also known as Rectangular Distribution.

For continuous Uniform distribution, the mean μ is the expected value of a random variable having the probability density $f(x)$.

$$\therefore \text{ Mean, } \mu = E(X) = \int_{-\infty}^{\infty} x\, f(x) dx = \int_a^b x.f(x) dx = \int_a^b x.\frac{1}{b-a} dx$$

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

Similarly, Variance, $\sigma^2 = E(X^2) - [E(X)]^2 = \frac{(b-a)^2}{12}$

Properties of Uniform Distribution :

1. $P(x_1 \le X \le x_2) = \frac{x_2 - x_1}{b - a}$

2. Its mean $= \frac{a+b}{2}$

3. Its variance $= \frac{(b-a)^2}{12}$

Illustrations :

1. If a shuttle bus has a cycle time of 50 minutes the waiting time would be uniformly distributed between 0 and 50 minutes.
2. Journey in Metro Rail in Hyderabad from Nagole to Hitech city. If a user comes to a stop at random time and wait till the facility arrives, the waiting time will be uniformly distributed between a minimum of zero and maximum equal to cycle time (say 10 minutes).

NORMAL DISTRIBUTION

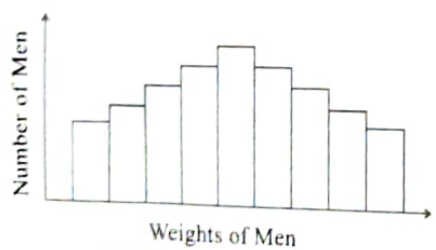
The normal law of error stands out in the mankind as one of the broadest generalisations of natural philosophy. It serves as the guiding instrument in researches in the physical and social sciences and in engineering, medicine and agriculture. It is an indispensable tool for the analysis and interpretation of the basic data obtained by Observation and experience.

- W.J. Youden

Normal distribution is applicable in the following situations:

1. Life of items subjected to wear and tear like tyres, batteries, bulbs, currency notes, etc.
2. Length and diameter of certain products like pipes, screws and discs.
3. Height and weight of baby at birth.
4. Aggregate marks obtained by students in an examination.
5. Weekly sales of an item in a store.

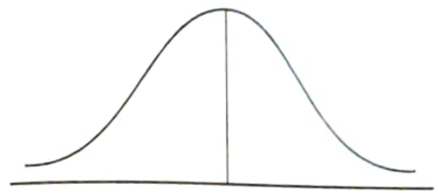
The most common pattern of distribution of a continuous variable, found in nature, is of the following type:



Distribution of Weights of Men

Maximum number of men will be around the average weight, and as one goes to the left or right, the number of persons will keep on reducing. In fact, even other human characteristics like age, intelligence, etc. follow this pattern or distribution.

This type of distribution was discovered by Karl Gauss. For this reason, it is called Gaussian Distribution. But, it is found so often in real life that it is called Normal Distribution. The name which is commonly used. If the curve is drawn in a continuous manner, it would look as follows.



Shape of Normal Distribution

5.3 NORMAL DISTRIBUTION

The distributions discussed so far are discrete distributions *i.e.*, distributions in which the variate can take only integral values. Examples of such distributions are the numbers of defective items manufactured in a factory, the numbers of persons suffering from a disease, etc. Now we shall consider a continuous distribution, namely the *Normal Distribution*. A continuous distribution is a distribution in which the variate can take all values within a given range. Examples of continuous distributions are the heights of persons, the speed of a vehicle, etc.

The Normal distribution was first discovered by English Mathematician De-Moivre (1667-1745) in 1733 and further refined by French Mathematician Laplace (1749-1827) in 1774 and independently by Karl Friedrich Gauss (1777-1855). Normal distribution is also known as Gaussian Distribution. It is another limiting form of the Binomial Distribution for large values of n when neither p nor q is very small. The Normal Distribution is, therefore, derived from the Binomial Distribution by increasing the number of trials indefinitely.

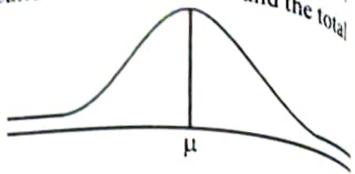
Definition : A random variable X is said to have a Normal Distribution, if its density function or probability distribution is given by **[JNTU (H) Dec. 09 (Set No. 2), (H) Dec. 2014]**

$$f(x;\, \mu, \sigma) = \frac{1}{\sigma.\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}};\ -\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0 \quad \ldots (1)$$

where μ is the Mean and σ is the Standard deviation of x. As can be seen, the function, called probability density function of the normal distribution, depends on two values μ and σ. These are referred as the two parameters of the normal distribution.

The curve has maximum value at $x = \mu$ and tapers off on either side but never touches the horizontal line. The curve on the left side goes upto $-\infty$, and on the right side it goes upto $+\infty$. However, as much as 99.73% of the area under the curve lies between $(\mu - 3\sigma)$ and $(\mu + 3\sigma)$ and only 0.277% of the area lies beyond these points.

The random vairable X is then said to be a normal random variable or normal variate. The curve representing the Normal Distribution (1) is called the normal curve and the total area bounded by the curve and the x - axis is one.



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

The area under the curve between the ordinates $x = a$ and $x = b$, where $a < b$, represents the probability that x lies between a and b. *i.e.,* $P(a < x < b)$.

Thus $P(a < x < b) =$ area under the normal curve between the vertical lines $x = a$ and $x = b$, which is $\int_a^b f(x)\,dx$.

Note : A random variable X with mean μ and variance σ^2 and following the probability law (1) is expressed by $X \sim N(\mu, \sigma^2)$.

5.4 NORMAL DISTRIBUTION AS A LIMITING FORM OF BINOMIAL DISTRIBUTION

[JNTU (A) Dec. 2009, (K) Nov. 2011 (Set No. 1)]

Normal distribution is a limiting case of the Binomial distribution under the following conditions :

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

(*ii*) neither p nor q is very small.

The probability of x successes in a series of n independent trials is

$$p(x) = {}^nC_x\, p^x q^{n-x};\; x = 0, 1, 2, \ldots, n \qquad \ldots (1)$$

Let us now consider the variate

$$Z = \frac{X - np}{\sqrt{npq}}, X = 0, 1, 2, \ldots\ldots, n \qquad \ldots (2)$$

When $X = 0, Z = \frac{-np}{\sqrt{npq}} = -\sqrt{\frac{np}{q}}$ and

When $X = n,\ Z = \frac{n - np}{\sqrt{npq}} = \frac{n(1-p)}{\sqrt{npq}}$

$$= \frac{nq}{\sqrt{npq}} = \sqrt{\frac{nq}{p}}$$

Thus in the limit as $n \to \infty$, Z takes the values from $-\infty$ to ∞. Hence the distribution of X will be a continuous distribution over the range $-\infty$ to ∞.

We want the limiting form of (1) under the above conditions.

Hence $\lim_{n\to\infty} p(x) = \lim_{n\to\infty} \frac{n!}{x!(n-x)!} p^x q^{n-x}$

Using Stirling's approximation for $n!$ as $n \to \infty$, $\lim_{n\to\infty} n! \approx \sqrt{2\pi}\, e^{-n} n^{n+\frac{1}{2}}$, we get

$$\lim_{n\to\infty} p(x) = \lim_{n\to\infty} \left[\frac{\sqrt{2\pi}\, e^{-n} n^{n+\frac{1}{2}} . p^x q^{n-x}}{\left(\sqrt{2\pi}\, e^{-x} x^{x+\left(\frac{1}{2}\right)}\right) . \left(\sqrt{2\pi}\, e^{-(n-x)} (n-x)^{n-x+\left(\frac{1}{2}\right)}\right)} \right]$$

$$= \lim_{n\to\infty} \left[\frac{1}{\sqrt{2\pi}} . \frac{n^{n+\left(\frac{1}{2}\right)}}{x^{x+\frac{1}{2}} (n-x)^{n-x+\left(\frac{1}{2}\right)}} . p^x q^{n-x} \right]$$

$$= \lim_{n\to\infty} \left[\frac{1}{\sqrt{2\pi}} . \frac{1}{\sqrt{npq}} . \frac{(np)^{x+\frac{1}{2}} (nq)^{n-x+\frac{1}{2}}}{x^{x+\frac{1}{2}} (n-x)^{n-x+\frac{1}{2}}} \right]$$

$$= \lim_{n\to\infty} \left[\frac{1}{\sqrt{2\pi}} . \frac{1}{\sqrt{npq}} . \left(\frac{np}{x}\right)^{x+\frac{1}{2}} . \left(\frac{nq}{n-x}\right)^{n-x+\frac{1}{2}} \right] \qquad \ldots (3)$$

Now, substitute $N = \left(\frac{x}{np}\right)^{x+\frac{1}{2}} . \left(\frac{n-x}{nq}\right)^{n-x+\frac{1}{2}}$

Then (3) becomes $\lim_{n\to\infty} p(x) = \lim_{n\to\infty} \frac{1}{\sqrt{2\pi}} . \frac{1}{\sqrt{npq}} . \frac{1}{N} \qquad \ldots (4)$

From (2), we have

$$X = np + Z . \sqrt{npq} \qquad \ldots (4.1)$$

$$\Rightarrow \frac{X}{np} = 1 + Z . \sqrt{\frac{q}{np}} \qquad \ldots (4.2)$$

Also $n - X = n - np - Z . \sqrt{npq} = n(1-p) - Z . \sqrt{npq}$

$$= nq - Z . \sqrt{npq} \qquad \ldots (4.3)$$

$$\therefore \frac{n - X}{nq} = 1 - Z . \sqrt{\frac{p}{nq}} \qquad \ldots (4.4)$$

Now $\log N = \left(x + \frac{1}{2}\right) \log\left(\frac{x}{np}\right) + \left(n - x + \frac{1}{2}\right) . \log\left(\frac{n-x}{nq}\right)$

$$= \left(np + Z.\sqrt{npq} + \frac{1}{2}\right) . \log\left(1 + Z . \sqrt{\frac{q}{np}}\right) + \left(nq - Z\sqrt{npq} + \frac{1}{2}\right) . \log\left(1 - Z . \sqrt{\frac{p}{nq}}\right)$$

using (4.1), (4.2), (4.3), (4.4)

$$= \left(np + Z . \sqrt{npq} + \frac{1}{2}\right) \left[Z . \sqrt{\frac{q}{np}} - \frac{1}{2} Z^2 \left(\frac{q}{np}\right) + \frac{1}{3} Z^3 \left(\frac{q}{np}\right)^{3/2} - \ldots\ldots \right]$$

$$+ \left(nq - Z\sqrt{npq} + \frac{1}{2}\right) \left[-Z\sqrt{\frac{p}{nq}} - \frac{1}{2} Z^2 \left(\frac{p}{nq}\right) - \frac{1}{3} Z^3 \left(\frac{p}{nq}\right)^{3/2} - \ldots\ldots \right]$$

$$= \left\{ \left(Z\sqrt{npq} - \frac{1}{2}qZ^2 + \frac{1}{3}Z^3\frac{q^{3/2}}{\sqrt{np}} + Z^2 q - \frac{1}{2}Z^3\frac{q^{3/2}}{\sqrt{np}} + \frac{1}{2}Z\sqrt{\frac{q}{np}} - \frac{1}{4}Z^2\frac{q}{np} + \cdots \right) \right.$$

$$\left. + \left(-Z\sqrt{npq} - \frac{1}{2}Z^2 p - \frac{1}{3}Z^3\frac{p^{3/2}}{\sqrt{nq}} + Z^2 p + \frac{1}{2}Z^3\frac{p^{3/2}}{\sqrt{nq}} - \frac{1}{2}Z\sqrt{\frac{p}{nq}} - \frac{1}{4}Z^2\frac{p}{nq} + \cdots \right) \right\}$$

i.e., $\log N = \left[-\frac{1}{2}Z^2(p+q) + Z^2(p+q) + \frac{Z}{2\sqrt{n}}\left\{ \sqrt{\frac{q}{p}} + \sqrt{\frac{p}{q}} \right\} + O\{n^{-1/2}\} \right]$

$$= \frac{Z^2}{2} + O(n^{-1/2}) \to \frac{Z^2}{2} \text{ as } n \to \infty$$

$$\lim_{n\to\infty} \log N = \frac{Z^2}{2} \Rightarrow \lim_{n\to\infty} N = e^{Z^2/2}$$

Now, from (4) we have

$$\lim_{n\to\infty} p(x) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} e^{-Z^2/2}$$

or $P(z \le Z \le z + dz) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{npq}} e^{-Z^2/2}$ .... (5)

But $Z = \frac{X - np}{\sqrt{npq}} = \frac{x-\mu}{\sigma}$

To obtain the distribution of X, we put $Z = \frac{x-\mu}{\sigma}$ in (5). We get

$$P(x \le X \le x + dx) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$$

5.5 CONSTANTS OF NORMAL DISTRIBUTION

1. **Mean of Normal Distribution** **[JNTU 06S, (H) Dec. 09 (Set No.2), (H) Dec. 2014]**

Consider the Normal Distribution with b, σ as the parameters.

Then $f(x; b, \sigma) = \frac{e^{-\frac{(x-b)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$

The mean $\mu = E(X)$ is given by

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

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z + b) e^{-\frac{z^2}{2}} dz \quad \left[\text{Putting } z = \frac{x-b}{\sigma} \text{ so that } dz = \frac{dx}{\sigma}\right]$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{z^2}{2}} dz + \frac{b}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz$$

$$= 0 + \frac{b}{\sqrt{2\pi}} \cdot 2\int_0^{\infty} e^{-\frac{z^2}{2}} dz \quad \left[\because z e^{-\frac{z^2}{2}} \text{ is odd function and } e^{-\frac{z^2}{2}} \text{ is even function}\right]$$

$$= \frac{2b}{\sqrt{2\pi}} \cdot \frac{\sqrt{\pi}}{\sqrt{2}} \quad \left[\because \int_0^{\infty} e^{-\frac{z^2}{2}} dz = \sqrt{\frac{\pi}{2}}\right]$$

$$= b$$

$\therefore$ **Mean, $\mu = b$**

2. **Variance of Normal Distribution**

By definition,

$$\text{Variance} = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$$

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

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z)^2 e^{-\frac{z^2}{2}} dz \quad \left[\text{Putting } z = \frac{x-b}{\sigma} \text{ so that } dz = \frac{dx}{\sigma}\right]$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} dz$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \cdot 2\int_0^{\infty} z^2 e^{-\frac{z^2}{2}} dz \quad [\because \text{Integrand is even function}]$$

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_0^{\infty} 2t\, e^{-t} \cdot \frac{dt}{\sqrt{2t}} \quad \left[\text{Putting } \frac{z^2}{2} = t \text{ so that } dz = \frac{dt}{\sqrt{2t}}\right]$$

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

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

Hence **variance = σ^2**

Thus the Standard deviation of the Normal Distribution is σ.

3. **Mode of Normal Distribution** **[JNTU 2004 S, (H) Dec. 2014]**

Mode is the value of x for which $f(x)$ is maximum. That is, mode is the solution of $f'(x) = 0$ and $f''(x) < 0$.

By definition, we have

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

Differentiating w.r.t. x, we get

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

Now $f'(x) = 0 \Rightarrow x-\mu = 0$ *i.e.,* $x = \mu$

$$f''(x) = -\frac{1}{\sigma^2}\left[(x-\mu).f'(x)+f(x)\right]$$

$$= -\frac{1}{\sigma^2}\left[(x-\mu).-\frac{(x-\mu)}{\sigma^2}.f(x)+f(x)\right]$$

$$= \frac{-f(x)}{\sigma^2}\left[1-\frac{(x-\mu)^2}{\sigma^2}\right]$$

At the point $x = \mu$, we have

$$f''(x) = -\left[\frac{f(x)}{\sigma^2}\right]_{x=\mu} = -\frac{1}{\sigma^2.\sqrt{2\pi}.\sigma} < 0$$

Hence $x = \mu$ is the mode of the Normal Distribution.

4. **Median of Normal distribution** **[JNTU 2004 S, (H) Dec. 2014]**

If M is the median of the normal distribution, we have

$$\int_{-\infty}^{M} f(x)\,dx = \frac{1}{2}$$

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

i.e. $$\frac{1}{\sigma.\sqrt{2\pi}}\int_{-\infty}^{\mu} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx + \frac{1}{\sigma.\sqrt{2\pi}}\int_{\mu}^{M} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx = \frac{1}{2} \quad \ldots (1)$$

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

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

$$\therefore \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\mu} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx = \frac{1}{\sigma.\sqrt{2\pi}}\int_{-\infty}^{0} e^{-\frac{z^2}{2}}.\sigma dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{-\frac{z^2}{2}}dz$$

$$= \frac{1}{\sqrt{2\pi}}\int_{0}^{\infty} e^{-\frac{z^2}{2}}dz \text{ (by symmetry)} = \frac{1}{\sqrt{2\pi}}.\sqrt{\frac{\pi}{2}} = \frac{1}{2} \quad \ldots (2)$$

$\therefore$ From (1) and (2), we have

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

$$\Rightarrow \frac{1}{\sigma.\sqrt{2\pi}}\int_{\mu}^{M} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx = 0 \Rightarrow \int_{\mu}^{M} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx = 0$$

$$\Rightarrow \mu = M \left[\because \text{ if } \int_a^b f(x)\,dx = 0 \text{ then } a = b, \text{ where } f(x) > 0\right]$$

Hence for the Normal Distribution, Mean = Median

Note : From above, we notice that for the Normal distribution mean, median and mode coincide.

i.e., **Mean = Median = Mode**

Hence the distribution is symmetrical. **[JNTU 2004S, 2008S (Set No.1,2)]**

5. **Mean deviation from the Mean for Normal Distribution**

[JNTU 2001S, (A) Dec. 2009 (Set No. 4)]

By definition,

Mean deviation (about mean) $= \int_{-\infty}^{\infty} |x-\mu|\,f(x)\,dx$

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

$$= \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|z|\,e^{-z^2/2}dz \quad \left[\text{Putting } \frac{x-\mu}{\sigma} = z\right]$$

$$= \frac{\sigma}{\sqrt{2\pi}}.2\int_{0}^{\infty}|z|.e^{-\frac{z^2}{2}}dz$$

[$\because$ Integrand is an even function]

$$= \sqrt{\frac{2}{\pi}}.\sigma\int_0^{\infty} z\,e^{-\frac{z^2}{2}}dz \quad (\because |z| = z \text{ in } [0, \infty])$$

$$= \sqrt{\frac{2}{\pi}}.\sigma\int_0^{\infty} e^{-t}dt \quad \left[\text{Putting } \frac{z^2}{2} = t\right]$$

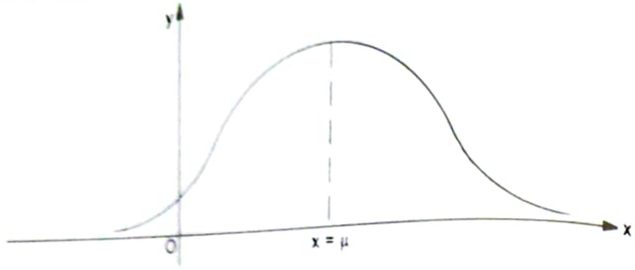
$$= \sqrt{\frac{2}{\pi}}.\sigma\left(-e^{-t}\right)_0^{\infty} = \sqrt{\frac{2}{\pi}}.\sigma(0+1)$$

$$= \sigma.\sqrt{\frac{2}{\pi}} = \frac{4\sigma}{5} \text{ (approximately)}$$

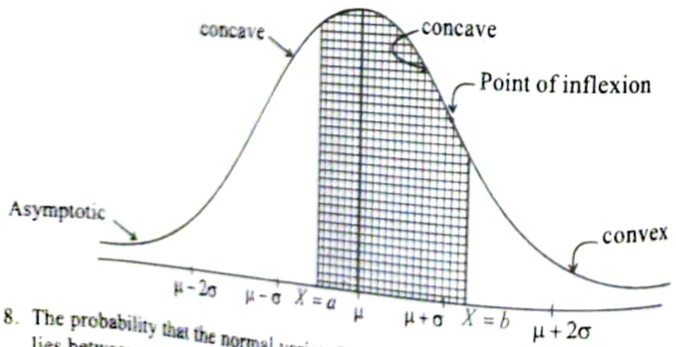
Hence the mean deviation from the Mean for Normal distribution is equal to $\frac{4}{5}$ times of Standard deviation approximately.

5.6 CHIEF CHARACTERISTICS OF THE NORMAL DISTRIBUTION [JNTU 2004]

1. The graph of the Normal distribution $y = f(x)$ in the xy-plane is known as the normal curve.



2. The curve is a bell shaped curve and symmetrical with respect to mean *i.e.*, about the line $x = \mu$ and the two tails on the right and the left sides of the mean (μ) extends to infinity. The top of the bell is directly above the mean μ.
3. Area under the normal curve represents the total population.
4. Mean, median and mode of the distribution coincide at $x = \mu$ as the distribution is symmetrical. So normal curve is unimodal (has only one maximum point).
5. x-axis is an asymptote to the curve.
6. Linear combination of independent normal variates is also a normal variate.
7. The points of inflexion of the curve are at $x = \mu \pm \sigma$ and the curve changes from concave to convex at $x = \mu + \sigma$ to $x = \mu - \sigma$.

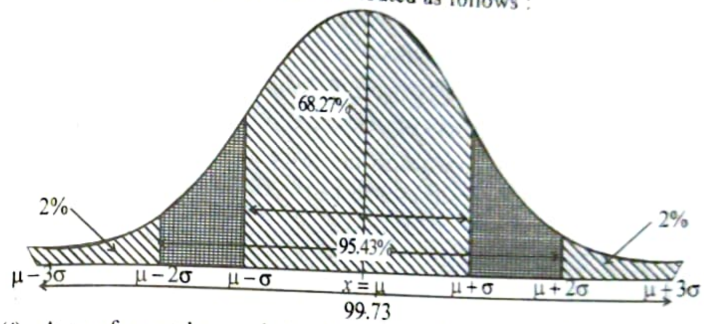


8. The probability that the normal variate X with mean μ and standard deviation σ lies between x_1 and x_2 is given by

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

Since (1) depends on the two parameters μ and σ, we get different normal curves for different values of μ and σ and it is an impracticable task to plot all such normal curves. Instead, by putting $z = \frac{x-\mu}{\sigma}$, the R.H.S. of equation (1) becomes independent of the two parameters μ and σ. Here z is known as the standard variable.

9. Area under the normal curve is distributed as follows :



(*i*) Area of normal curve between $\mu - \sigma$ and $\mu + \sigma$ is 68.27%
i.e., $P(\mu - \sigma < X < \mu + \sigma) = 0.6826$.

(*ii*) Area of normal curve between $\mu - 2\sigma$ and $\mu + 2\sigma$ is 95.43%.

(*iii*) Area of normal curve between $\mu - 3\sigma$ and $\mu + 3\sigma$ is 99.73%.

Standard Normal Distribution :

Definition : The Normal distribution with mean $(\mu) = 0$ and S. D.$(\sigma) = 1$, is known as *Standard Normal Distribution*.

The random variable that follows this distribution is denoted by z. If a variable x follows normal distribution with mean, μ and *s.d.*, σ, the variable z defined as

$$z = \frac{x-\mu}{\sigma}$$

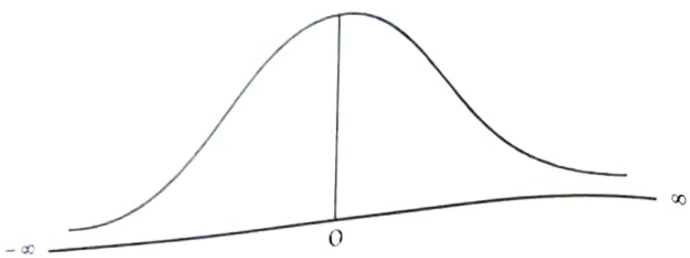
has standard normal distribution with mean 0 and *s.d.* as 1. This is also referred as z-score.

Standardized Variable :

If a random variable X has mean μ and variance σ^2, then the corresponding variable $z = \frac{X-\mu}{\sigma}$ has the mean 0 and variable 1.

The variable z is called the standardized variable, corresponding to X.

The shape of the Normal distribution is as follows:



It may be noted that the curve has the same shape as the normal distribution but, with the special properties that $\mu = 0$ and $\sigma = 1$ and has maximum at $z = 0$, and extends from $-\infty$ to $+\infty$ like normal distribution.

This is not the case with Binomial or Poisson distributions. They have different skewness (asymmetry) for different parameters.

We can transform any normal distribution to the standarised form because the normal distribution has the same shape whatever its parameter (μ and σ) may be.

If we take $\frac{x-\mu}{\sigma} = z$ in $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, we obtain the Standard Normal Distribution with mean zero and S.D. $(\sigma) = 1$ defined by the density function

$$p(z) = \frac{1}{\sqrt{2\pi}} e^{-(z^2/2)}, \quad -\infty < z < \infty$$

Here z is called the standard normal variate with mean 0 and standard deviation 1 and in short, we say that $p(z)$ is $N(0,1)$.

Area under any normal curve is found from the table of standard normal probability distribution showing the area between the mean and any value of the normally distributed random variable. For given values of μ and σ, and a specific value, x, of the random variable, the standardized variate z is derived from the formula $z = \frac{x-\mu}{\sigma}$. The purpose of standardisation of the normal distribution is that we can make use of the tables of the area of the standard curve $f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ (representing probability) for the various points along the x - axis.

(The standard normal tables are given at the end of the book).

The standard normal distribution is also known as Unit Normal Distribution or Z-Distribution.

The standard normal curve helps us to find the areas within two assigned limits under the curve. The areas between the standard normal curve drawn at two assigned limits a and b will give the proportion of cases for which the values of z lie between a and b. Thus the area between two assigned limits a and b under the standard normal curve will represent the probability that z will lie between a and b. It is denoted by $P(a \le z \le b)$.

Uses of Normal Distribution:

1. The normal distribution can be used to approximate Binomial and Poisson distributions.
2. It has extensive use in sampling theory. It helps us to estimate parameter from stastistic and to find confidence limits of the parameter.
3. It has a wide use in testing Statistical Hypothesis and Tests of significance in which it is always assumed that the population from which the samples have been drawn should have normal distribution.
4. It serves as a guiding instrument in the analysis and interpretation of statistical data.

Some other uses and situations where normal distribution is applicable are mentioned in the section 5.8.

5.7 AREA UNDER THE NORMAL CURVE (NORMAL PROBABILITY INTEGRAL) CONVERSION INTO STANDARD NORMAL FORM

By taking $z = \frac{x-\mu}{\sigma}$, the standard normal curve is formed. The total area under this curve is 1. The area under the curve is divided into two equal parts by $z = 0$. Left hand side area and right hand side area to z = 0 is 0.5. The area between the ordinate $z = 0$ and any other ordinate can be noted from the table given in the subsequent pages.

The probability that random variable X will lie between $X = \mu$ and $X = x_1$ is given by

$$P(\mu < X < x_1) = \int_{\mu}^{x_1} f(x)dx = \frac{1}{\sigma.\sqrt{2\pi}} \int_{\mu}^{x_1} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sigma.\sqrt{2\pi}} \int_{0}^{z_1} e^{-\frac{z^2}{2}} .\sigma\, dz \quad \left[\text{Putting } \frac{x-\mu}{\sigma} = z \text{ and } \frac{x_1-\mu}{\sigma} = z_1\right]$$

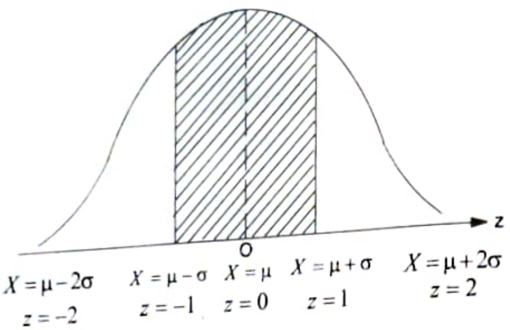
i.e., $P(\mu < X < x_1) = P(0 < z < z_1)$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{z_1} e^{-\frac{z^2}{2}} dz = A(z_1) \text{ say.}$$

i.e., $$A(z_1) = \frac{1}{\sqrt{2\pi}} \int_{0}^{z_1} e^{-\frac{z^2}{2}} dz.$$

The function $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ is the probability function of standard normal variate.

The definite integral $\int_{0}^{z_1} e^{-\frac{z^2}{2}} dz$ is known as the *normal probability integral* and gives the area under standard normal curve between $z = 0$ and $z = z_1$.



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

$$P(\mu - \sigma < X < \mu + \sigma) = \int_{\mu-\sigma}^{\mu+\sigma} f(x)\, dx$$

$$\Rightarrow P(-1 < z < 1) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-\frac{z^2}{2}}\, dz \quad \left(z = \frac{x-\mu}{\sigma}\right)$$

$$= \frac{2}{\sqrt{2\pi}} \int_{0}^{1} e^{-\frac{z^2}{2}}\, dz = 2\int_{0}^{1} \phi(z)\, dz \text{ where } \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

$$= 2\,(0.3413) \quad \text{(from tables)}$$

$$= 0.6826$$

Note : Since the integrand is an even function, it follows that $A(-z_1) = A(z_1)$.

Similarly, $P(\mu - 2\sigma < X < \mu + 2\sigma) = P(-2 < z < 2) = \int_{-2}^{2} \phi(z)\, dz$

$$= 2\int_{0}^{2} \phi(z)\, dz = 2\,(0.4772) = 0.9544$$

How to Find Probability Density of Normal Curve

The probability that the normal variate X with mean μ and standard deviation σ lies between two specific values x_1 and x_2 with $x_1 \le x_2$ can be obtained using area under the standard normal curve as follows :

Step 1 : Perform the change of scale $z = \frac{x-\mu}{\sigma}$ and find z_1 and z_2 corresponding to the values of x_1 and x_2 respectively.

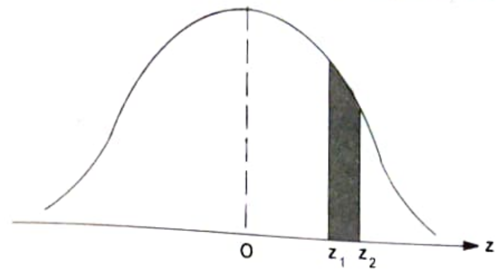
Step 2 (a) : To find $P(x_1 \le X \le x_2) = P(z_1 \le z \le z_2)$

Case 1 : If both z_1 and z_2 are positive (or both negative), then

$$P(x_1 \le X \le x_2) = |A(z_2) - A(z_1)|$$

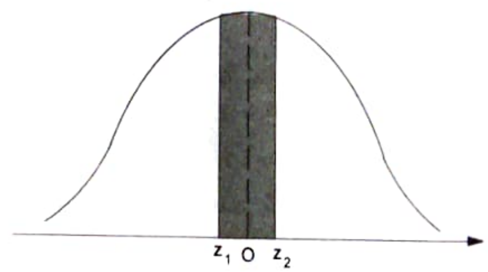
= (Area under the normal curve from 0 to z_2)

− (Area under the normal curve from 0 to z_1)



Case 2 : If $z_1 < 0$ and $z_2 > 0$, then

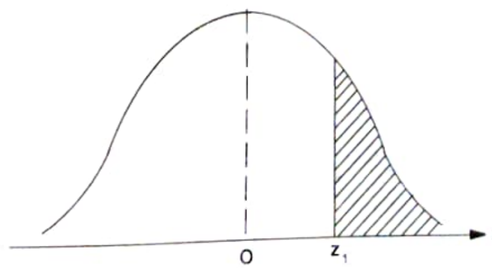
$$P(x_1 \le X \le x_2) = A(z_2) + A(z_1)$$



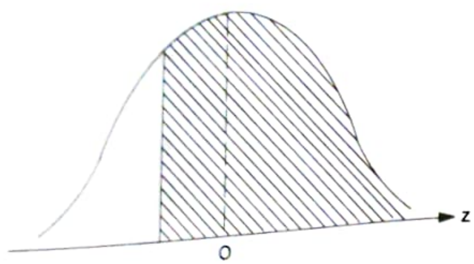
Step 2 (b) : To find $P(z > z_1)$

Case 1 : If $z_1 > 0$ then

$$P(z > z_1) = 0.5 - A(z_1) \quad [\because P(z < 0) = P(z > 0) = \tfrac{1}{2}]$$



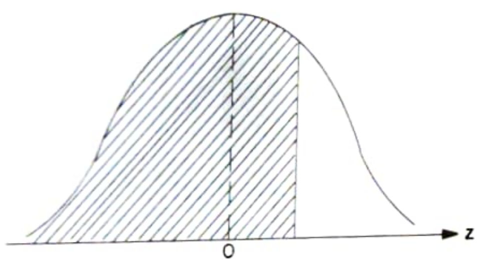
Case 2 : If $z_1 < 0$, then $P(z > z_1) = 0.5 + A(z_1)$



Step 2 (c) : To find $P(z < z_1) = 1 - P(z > z_1)$

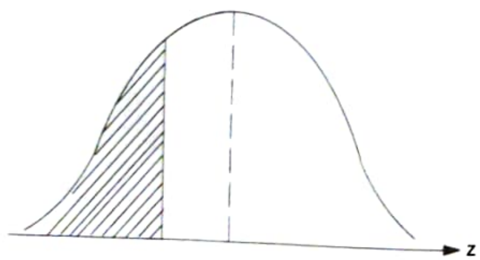
Case 1 : If $z_1 > 0$, then

$$P(z < z_1) = 1 - P(z > z_1) = 1 - [0.5 - A(z_1)] = 0.5 + A(z_1)$$



Case 2 : If $z_1 \leq 0$, then

$$P(z < z_1) = 1 - P(z > z_1) = 1 - [0.5 + A(z_1)]$$
$$= 0.5 - A(z_1)$$



5.8 THE IMPORTANCE AND APPLICATIONS OF THE NORMAL DISTRIBUTION

[JNTU 2000, 200S, 20005 (Set No. 4)]

Normal distribution plays a very important role in statistical theory because of the following reasons :

1. Data obtained from Psychological, Physical and Biological measurements approximately follow Normal distribution. I.Q. scores, heights and weights of individuals etc., are examples of measurements which are normally distributed or nearly so.

2. Most of the distributions that are encountered in practice, for example, Binomial, Poisson, Hypergeometric, etc. can be approximated to Normal distribution. If the number of trials n is indefinitely large and neither p nor q is very small, then Binomial distribution tends to Normal distribution. If the parameter $\lambda \to \infty$, then Poisson distribution tends to Normal distribution.

3. Since the Normal distribution is a limiting case of the Binomial distribution for exceptionally large numbers, it is applicable to many applied problems in kinetic theory of gases and fluctuations in the magnitude of an electric current.

4. Even if a variable is normally distributed, it can sometimes be brought to normal form by simple transformation of the variable.

5. For large samples, any statistic (i.e., sample mean, sample S.D., etc.) approximately follows Normal distribution and as such it can be studied with the help of normal curve.

6. Normal curve is used to find confidence limits of the population parameters.

7. The proofs of all the tests of significance in sampling are based upon the fundamental assumption that the population from which the samples have been drawn is normal.

8. Normal distribution finds large applications in Statistical Quality Control in industry for finding control limits.

Table : Area under standard normal curve from 0 to $\frac{x-\mu}{\sigma}$

$\frac{x-\mu}{\sigma}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0159	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1084	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1916	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2611	.2642	.2671	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4430	.4441
1.6	.4452	.4463	.4474	.4485	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4654	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4762	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4865	.4868	.4871	.4875	.5878	.4881	.4884	.4887	.4992
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946.	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4980	.4980	.4981
2.9	.4981	.4982	.4983	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993

SOLVED EXAMPLES

Example 1 : For a normally distributed variate with mean 1 and standard deviation 3, find the probabilities that (*i*) $3.43 \le x \le 6.19$ (*ii*) $-1.43 \le x \le 6.19$

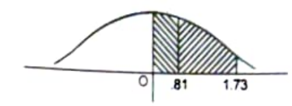
Solution : Given $\mu = 1$ and $\sigma = 3$

(*i*) When $x = 3.43$,

$$z = \frac{x-\mu}{\sigma} = \frac{3.43-1}{3} = \frac{2.43}{3} = 0.81 = z_1 \text{ (say)}$$

When $x = 6.19$,

$$z = \frac{x-\mu}{\sigma} = \frac{6.19-1}{3} = \frac{5.19}{3} = 1.73 = z_2 \text{ (say)}$$



$\therefore$ $P(3.43 \le x \le 6.19) = P(0.81 \le z \le 1.73)$

$= |A(z_2) - A(z_1)|$

$= |A(1.73) - A(0.81)|$

$= 0.4582^* - 0.2910$ (from Normal tables)

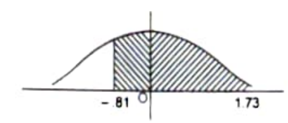
$= 0.1672$ (Cross hatched area in the figure)

(*ii*) When $x = -1.43$,

$$z = \frac{x-\mu}{\sigma} = \frac{-1.43-1}{3} = -0.81 = z_1 \text{ (say)}$$

When $x = 6.19$,

$$z = \frac{x-\mu}{\sigma} = \frac{6.19-1}{3} = 1.73 = z_2 \text{ (say)}$$



$\therefore$ $P(-1.43 \le x \le 6.19) = P(-0.81 \le z \le 1.73)$

$= A(z_2) + A(z_1)$

$= A(1.73) + A(-0.81)$

$= A(1.73) + A(0.81)$ **[$\because$ A(–z) = A(z)]**

$= 0.4582 + 0.2910 = 0.7492$ = shaded area in the figure.

Example 2 : If X is a normal variate with mean 30 and standard deviation 5. Find (*i*) $P(26 \le X \le 40)$ (*ii*) $P(X \ge 45)$ **[JNTU 2000, 2000S, (K) Nov. 2009 (Set No. 1)]**

Solution : Given mean, $\mu = 30$ and S.D., $\sigma = 5$

(*i*) When $x = 26$, $z = \frac{x-\mu}{\sigma} = \frac{26-30}{5} = -0.8 = z_1$ (say)

When $x = 40$, $z = \frac{x-\mu}{\sigma} = \frac{40-30}{5} = 2 = z_2$ (say)

$\therefore$ $P(26 \le X \le 40) = P(-0.8 \le z \le 2)$

$= A(z_2) + A(z_1) = A(2) + A(-0.8)$

$= 0.4772 + 0.2881 = 0.7653$ (from normal distribution tables)

* See Normal Distribution Table given, read the data in row 1.7 and column 0.03.

(ii) When $x = 45$, $z = \frac{x-\mu}{\sigma} = \frac{45-30}{5} = 3 = z_1$ (say)

$\therefore P(X \geq 45) = P(z_1 \geq 3)$

$= 0.5 - A(z_1) = 0.5 - A(3)$

$= 0.5 - 0.49865 = 0.00135$

Example 3 : In a Normal distribution, 7% of the items are under 35 and 89% are under 63. Determine the mean and variance of the distribution.

[JNTU Jan. 2007, (H) Dec. 2011, (K) Nov. 2012 (Set No. 2), (H) III yr. Nov. 2015]

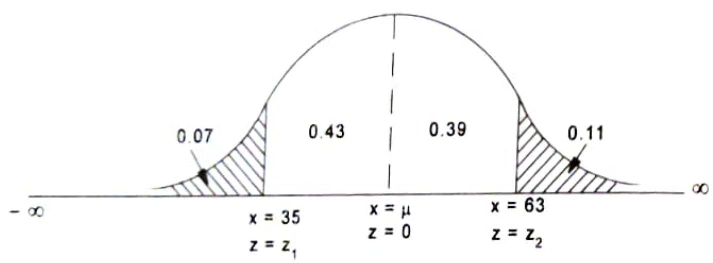
(OR) Find the mean and standard deviation of a normal distribution in which 7% of items are under 35 and 89 % are under 63. **[JNTU (H) Nov., 2009, (A) Nov. 2010, (H) May 2017]**

Solution : Let μ be the mean (at $z = 0$) and σ the standard deviation of the normal curve. 7% of the items are under 35 means the area to the left of the ordinate $x = 35$.

Given $P(X < 35) = 0.07$ and $P(X < 63) = 0.89$

$\therefore \quad P(X > 63) = 1 - P(X < 63) = 1 - 0.89 = 0.11$

The points $X = 35$ and $X = 63$ are shown in the following Figure.



When $x = 35$, $z = \frac{x-\mu}{\sigma} = \frac{35-\mu}{\sigma} = -z_1$ (say) ... (1)

When $x = 63$, $z = \frac{x-\mu}{\sigma} = \frac{63-\mu}{\sigma} = z_2$ (say) ... (2)

From the figure, we have

$P(0 < z < z_2) = 0.39 \Rightarrow z_2 = 1.23$ (from Normal tables)

and $P(0 < z < z_1) = 0.43 \Rightarrow z_1 = 1.48$

From (1), we have $\frac{35-\mu}{\sigma} = -1.48$ ... (3)

From (2), we have $\frac{63-\mu}{\sigma} = 1.23$ ... (4)

(4) – (3) gives

$\frac{28}{\sigma} = 2.71 \Rightarrow \sigma = \frac{28}{2.71} = 10.332$

From (3), $35 - \mu = -1.48\ (\sigma) = (-1.48)\ (10.332) = -15.3$

$\therefore \mu = 35 + 15.3 = 50.3$

and variance $= \sigma^2 = 106.75$

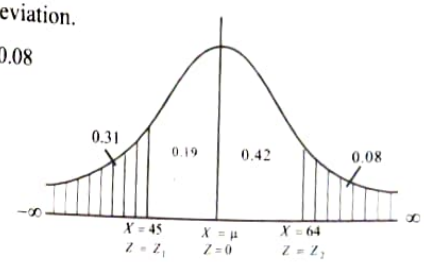
Note : Standard deviation, $\sigma = \sqrt{106.75} = 10.332$

Example 4 : In a normal distribution 31% of the items are under 45 and 8% are over 64. Find the mean and variance of the distribution.

[JNTU (A) Dec. 2009, (K) Nov. 2012, Mar. 2014 (Set No. 2)]

Solution : Let X be the continuous random variable.

Let μ be the mean and σ the standard deviation.



Given $P(X < 45) = 0.31$ and $P(X > 64) = 0.08$

Standard variable, $z = \frac{X-\mu}{\sigma}$

When X = 45, let $z = z_1$

so that $z_1 = \frac{45-\mu}{\sigma}$ ... (1)

$\therefore \int_{-\infty}^{z_1} \phi(z)dz = 0.31$ or $\int_{-\infty}^{0} \phi(z)dz - \int_{z_1}^{0} \phi(z)dz = 0.31$

$\therefore \int_{z_1}^{0} \phi(z)dz = \int_{-\infty}^{0} \phi(z)dz - 0.31 = 0.5 - 0.31 = 0.19$

Hence $P(0 < z < z_1) = 0.19 \Rightarrow z_1 = -0.5$ (from table) ... (2)

When X = 64, $z = \frac{64-\mu}{\sigma} = z_2$ (say) ... (3)

$\therefore \int_{z_2}^{\infty} \phi(z)dz = 0.08$ or $\int_{0}^{\infty} \phi(z)dz - \int_{0}^{z_2} \phi(z)dz = 0.08$

Hence $\int_{0}^{z_2} \phi(z)dz = \int_{0}^{\infty} \phi(z)dz - 0.08 = 0.5 - 0.08 = 0.42$

Thus $P(0 < z < z_2) = 0.42 \Rightarrow z_2 = 1.4$ (from tables) ... (4)

From (1) and (2), we have $\frac{45-\mu}{\sigma} = -0.5 \Rightarrow 45 - \mu = -0.5\sigma$ ... (5)

From (3) and (4), we have $\frac{64-\mu}{\sigma} = 1.4 \Rightarrow 64 - \mu = 1.4\sigma$ ... (6)

(5) – (6) gives $(45-\mu) - (64-\mu) = -0.5\sigma - 1.4\sigma$

$\Rightarrow -19 = -1.9\sigma \quad \therefore \sigma = \frac{19}{1.9} = 10$

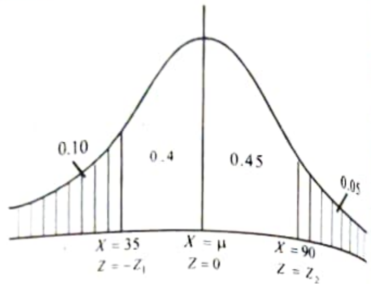
From (5), $\mu = 45 + 0.5\sigma = 45 + 0.5(10) = 50$

Hence mean $= 50$ and standard deviation $= 10$

Example 5 : Suppose 10 percent of the probability for a normal distribution $N(\mu, \sigma^2)$ is below 35 and 5 percent above 90. What are the values of μ and σ ? **[JNTU (K) 2009(Set No.2)]**

Solution : If $X \sim N(\mu, \sigma^2)$ then we are given $P(X < 35) = 0.10$ and $P(X > 90) = 0.05$.

The points $X = 35$ and $X = 90$ are located as shown in Figure.



Standard variable, $z = \dfrac{X - \mu}{\sigma}$

When $X = 35$, $z = \dfrac{35 - \mu}{\sigma} = -z_1$ (say)

and when $X = 90$,

$z = \dfrac{90 - \mu}{\sigma} = z_2$ (say)

Thus we have $P(0 < z < z_1) = 0.4$ and $P(0 < z < z_2) = 0.45$.

Hence from Normal tables, we have $z_1 = 1.3$ and $z_2 = 1.65$

$\therefore \dfrac{35 - \mu}{\sigma} = -1.3$ and $\dfrac{90 - \mu}{\sigma} = 1.65$

Subtracting, we get $\dfrac{55}{\sigma} = 2.95 \Rightarrow \sigma = \dfrac{55}{2.95} = 18.64$

$\therefore \mu = 35 + 1.3\sigma = 35 + 1.3(18.64) = 35 + 24.24 = 59.24$

Hence $\mu = 59.24$ and $\sigma = 18.64$

Example 6 : The mean and standard deviation of the marks obtained by 1000 students in an examination are respectively 34.5 and 16.5. Assuming the normality of the distribution, find the approximate number of students expected to obtain marks between 30 and 60.

[JNTU (A) Nov. 2010, (H) Sept. 2017]

Solution : Given mean, $\mu = 34.5$ and S.D., $\sigma = 16.5$

When $x = 30$, $z = \dfrac{x - \mu}{\sigma} = \dfrac{30 - 34.5}{16.5} = -0.27 = z_1$ (say)

When $x = 60$, $z = \dfrac{x - \mu}{\sigma} = \dfrac{60 - 34.5}{16.5} = 1.54 = z_2$ (say)

$\therefore P(30 \le x \le 60) = P(z_1 \le z \le z_2)$

$= A(z_2) + A(z_1)$ [$\because z_1 < 0$ and $z_2 > 0$]

$= A(1.54) + A(-0.27) = A(1.54) + A(0.27)$

$= 0.4382 + 0.1084$ [From normal distribution tables]

$= 0.5916$

$\therefore$ The number of students who get marks between 30 and 60

$= 0.5916 \times 1000 = 591.6$

Hence 592 students get marks between 30 and 60.

Example 7 : The marks obtained in mathematics by 1000 students is normally distributed with mean 78% and standard deviation 11%. Determine

(i) How many students got marks above 90%

(ii) What was the highest mark obtained by the lowest 10% of the students

(iii) Within what limits did the middle of 90% of the students lie.

[JNTU 2006S, 2007, (A) Nov. 2010, Apr. 2012, (K) Nov. 2011 (Set No. 2), (H) Nov. 2012]

Solution : Given mean, $\mu = 0.78$ and S.D., $\sigma = 0.11$

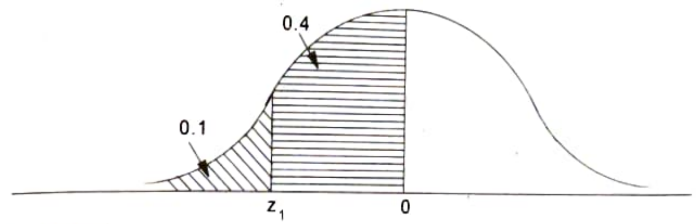
(i) When $x = 0.9$, $z = \dfrac{x - \mu}{\sigma} = \dfrac{0.9 - 0.78}{0.11} = 1.09 = z_1$ (say)

Hence the number of students with marks more than 90%

$= 0.1379 \times 1000$

$= 137.9 \simeq 138$

(ii) The 0.1 area to the left of z corresponds to the lowest 10% of the students



From Figure,

$0.4 = 0.5 - 0.1 = 0.5 -$ Area from 0 to z_1

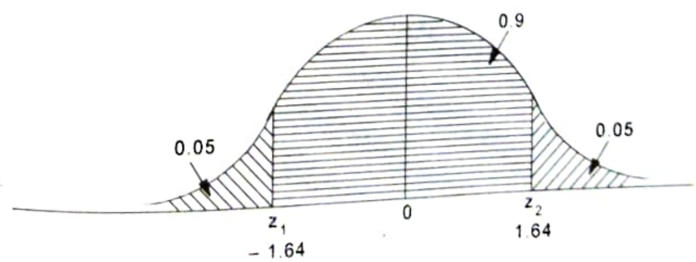
$\therefore z_1 = -1.28$ (from tables)

Thus $-1.28 = \dfrac{x - \mu}{\sigma} = \dfrac{x - 0.78}{0.11} \Rightarrow x = 0.78 - 1.28\,(0.11) = 0.6392$

Hence the highest mark obtained by the lowest 10% of students

$= 0.6392 \times 1000 \simeq 64\%$

(iii)



Middle 90% correspond to 0.9 area, leaving 0.05 area on both sides. Then the corresponding z's are $\pm$ 1.64 [since if the area from 0 to z is 0.45, then $z = 1.64$]

$$\therefore -1.64 = z_1 = \frac{x_1 - \mu}{\sigma} = \frac{x_1 - 0.78}{0.11}$$

$$\Rightarrow x_1 = 0.78 - 1.64\,(0.11) = 0.5996 \text{ or } 59.96\%$$

and $1.64 = z_2 = \dfrac{x_2 - \mu}{\sigma} = \dfrac{x_2 - 0.78}{0.11}$

$$\Rightarrow \quad x_2 = 1.64(0.11) + 0.78 = 0.9604 \text{ or } 96.04\%$$

Thus the middle 90% have marks in between 60 to 96.

Example 8 : Suppose the weights of 800 male students are normally distributed with mean $\mu = 140$ pounds and standard deviation 10 pounds. Find the number of students whose weights are (*i*) between 138 and 148 pounds (*ii*) more than 152 pounds

[JNTU 2006 S (Set No.2)]

Solution : Let μ be the mean and σ be the standard deviation. Then $\mu = 140$ pounds and $\sigma = 10$ pounds.

(*i*) When $x = 138$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{138-140}{10} = -0.2 = z_1$ (say)

When $x = 148$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{148-140}{10} = 0.8 = z_2$ (say)

$$\begin{aligned}\therefore P(138 \le x \le 148) &= P(-0.2 \le z \le 0.8)\\ &= A(z_2) + A(z_1) = A(0.8) + A(-0.2)\\ &= A(0.8) + A(0.2)\\ &= 0.2881 + 0.0793 = 0.3674\end{aligned}$$

Hence the number of students whose weights are between 138 pounds and 148 pounds $= 0.3674 \times 800 \approx 294$

(*ii*) When $x = 152$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{152-140}{10} = 1.2 = z_1$ (say)

$$\begin{aligned}\therefore P(x > 152) &= P(z > z_1)\\ &= 0.5 - A(z_1) = 0.5 - A(1.2)\\ &= 0.5 - 0.3849 = 0.1151\end{aligned}$$

$\therefore$ Number of students whose weights are more than 152 pounds

$$= 800 \times 0.1151 = 92$$

Example 9 : The marks obtained in Statistics in a certain examination found to be normally distributed. If 15% of the students $\ge$ 60 marks, 40% < 30 marks, find the mean and standard deviation.

[JNTU 2006S (Set No.3)]

Solution : Let μ be the mean (at $z = 0$) and σ the standard deviation of the normal distribution.

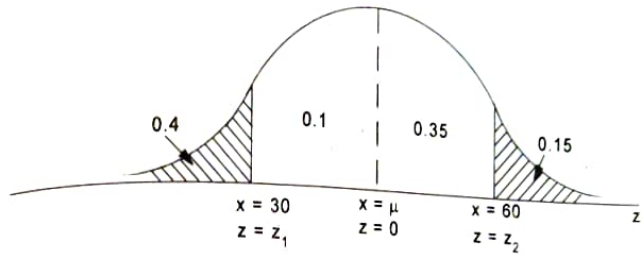
Let the variable X denote the marks in statistics.

Then we are given

$$P(X < 30) = 0.4 \text{ and } P(X \ge 60) = 0.15$$

When $X = 30$, $z = \dfrac{X-\mu}{\sigma} = \dfrac{30-\mu}{\sigma} = -z_1$ (say) ... (1)

When $X = 60$, $z = \dfrac{X-\mu}{\sigma} = \dfrac{60-\mu}{\sigma} = z_2$ (say) ... (2)



$\therefore P(0 < z < z_2) = 0.5 - 0.15 = 0.35$

and $P(0 < z < z_1) = P(-z_1 < z < 0)$ (By symmetry)

$$= 0.5 - 0.4 = 0.1$$

$\therefore$ From normal tables, we get

$$z_1 = 0.25 \text{ and } z_2 = 1.04$$

Hence $\dfrac{30-\mu}{\sigma} = -0.25$ or $\dfrac{\mu-30}{\sigma} = 0.25$, using (1) ... (3)

and $\dfrac{60-\mu}{\sigma} = 1.04$, using (2) ... (4)

(3) + (4) gives

$\dfrac{30}{\sigma} = 1.29 \quad \therefore \sigma = \dfrac{30}{1.29} = 23.26$

$\therefore \mu = 0.25\,\sigma + 30$ [From (3)]

$$= 0.25\,(23.26) + 30 = 35.81$$

Hence mean = 35.81 and standard deviation = 23.26

Example 10 : If X is a normal variate, find the area A

(*i*) to the left of $z = -1.78$

(*ii*) to the right of $z = -1.45$

(*iii*) corresponding to $-0.8 \le z \le 1.53$

(iv) to the left of $z = -2.52$ and to the right of $z = 1.83$

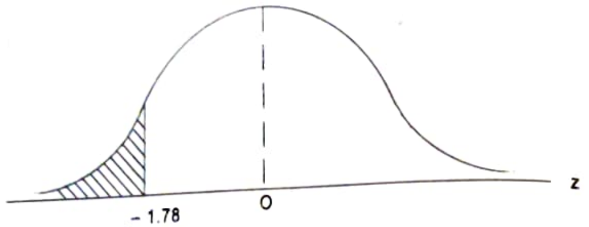
[JNTU 2004S, 2005S, 2007(Set No.3)]

Solution :

(*i*) Required area, A

$= 0.5 - \text{Area } (0 \text{ to } -1.78) = 0.5 - \text{Area } (0 \text{ to } 1.78)$ (By symmetry)

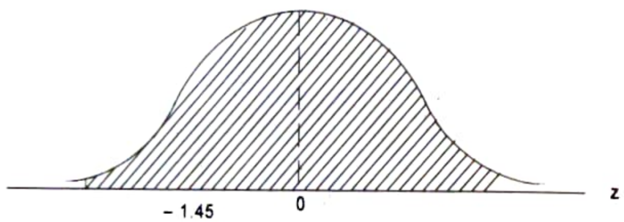
$= 0.5 - 0.4625$ (from tables) $= 0.0375$



(*ii*) Required area, $A = 0.5 + \text{Area from } 0 \text{ to } -1.45$

$= 0.5 + \text{Area from } 0 \text{ to } 1.45$ (By symmetry)

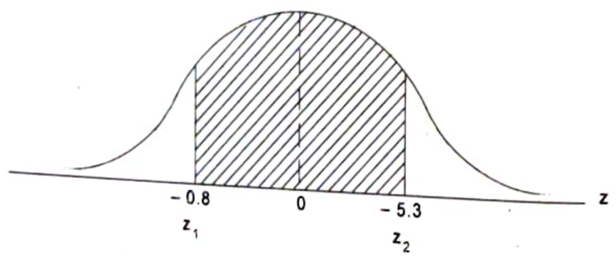
$= 0.5 + 0.4265 = 0.9265$



(*iii*) Required area $= A(z_2) + A(z_1)$

$= A(1.53) + A(-0.8) = A(1.53) + A(0.8)$

$= 0.437 + 0.2881 = 0.7251$



Aliter :

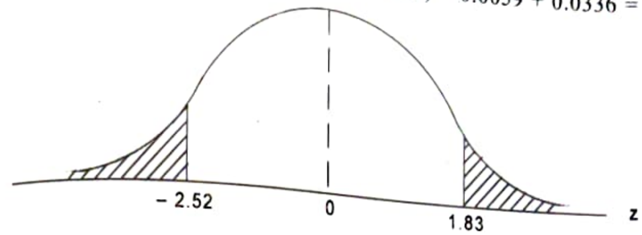
Required area = Area from (0 to – 0.8) + Area from (0 to 1.53)

= Area from (0 to 0.8) + Area from (0 to 1.53)

$= 0.4370 + 0.2881 = 0.7251$

(*iv*) Required area = [0.5 – Area from 0 to 2.52] + [0.5 – Area from 0 to 1.83]

$= (0.5 - 0.4941) + (0.5 - 0.4664) = 0.0059 + 0.0336 = 0.0395$



Example 11 : In a sample of 1000 cases, the mean of a certain test is 14 and standard deviation is 2.5. Assuming the distribution to be normal, find

(*i*) how many students score between 12 and 15 ?

(*ii*) how many score above 18 ?

(*iii*) how many score below 18 ? **[JNTU 2004 S, 2007S (Set No. 2,3), (H) Nov. 2015]**

Solution : Let μ be the mean and σ the sandard deviation of the normal distribution. Then we are given

$\mu = 14$ and $\sigma = 2.5$

Let the variable X denote the score in a test.

(*i*) When $X = 12$, $z = \dfrac{X-\mu}{\sigma} = \dfrac{12-14}{2.5} = -0.8 = z_1$ (say).

When $X = 15$, $z = \dfrac{X-\mu}{\sigma} = \dfrac{15-14}{2.5} = 0.4 = z_2$ (say)

$\therefore \quad P(12 < X < 15) = P(-0.8 < z < 0.4)$

$= A(z_2) + A(z_1) = A(0.4) + A(-0.8)$

$= A(0.4) + A(0.8)$ (due to symmetry)

$= 0.1554 + 0.2881$

$= 0.4435$

Hence number of students score between 12 and 15

$= 1000 \times 0.4435 = 443$ (approximately)

(*ii*) When $X = 18$, $z = \dfrac{X-\mu}{\sigma} = \dfrac{18-14}{2.5} = 1.6$

$\therefore \quad P(X > 18) = P(z > 1.6)$

$= 0.5 - A(1.6) = 0.5 - 0.4452$

$= 0.0548$

Hence number of students score above 18 $= 1000 \times 0.0548$

$= 54.8 = 55$ (approximately)

(iii) $P(X < 18) = P(z < 1.6)$

$= 0.5 + A(1.6) = 0.5 + 0.4452 = 0.9452$

Aliter : $P(X < 18) = 1 - P(X > 18) = 1 - 0.0548 = 0.9452$

$\therefore$ Number of students score below 18 = 1000 × 0.9452 = 945

Example 12 : A sales tax officer has reported that the average sales of the 500 business that he has to deal with during a year is Rs. 36,000 with a standard deviation of 10,000. Assuming that the sales in these business are normally distributed, find

(i) the number of business as the sales of which are Rs. 40,000.

(ii) the percentage of business the sales of which are likely to range between Rs. 30,000 and Rs. 40,000. **[JNTU 2005 (Set No. 1, 2, 3)]**

Solution : Let μ be the mean and σ the standard deviation of the sales. Then we are given that μ = 36000 and σ = 10000

Let the variable X denote the sales in the business

When $X = 40000,\ z = \dfrac{X-\mu}{\sigma} = \dfrac{40000-36000}{10000} = 0.4$

When $X = 30000,\ z = \dfrac{X-\mu}{\sigma} = \dfrac{30000-36000}{10000} = -0.6$

(i) $P(X > 40000) = P(z > 0.4)$

$= 0.5 - A(0.4) = 0.5 - 0.1554 = 0.3446$

Number of business as the sales of which are Rs. 40,000

$= 500 \times 0.3446 = 172$ (approximately)

(ii) $P(30000 < X < 40000) = P(-0.6 < z < 0.4)$

$= A(0.4) + A(-0.6) = A(0.4) + A(0.6)$

$= 0.1554 + 0.2257 = 0.3811$

$\therefore$ The required percentage of business = 38.11%

Example 13 : If the masses of 300 students are normally distributed with mean 68 kgs and standard deviation 3 kgs, how many students have masses

(i) Greater than 72 kgs (ii) Less than or equal to 64 kgs

(iii) Between 65 and 71 kgs inclusive.

[JNTU 2005, 2008, (A) Nov. 2010, (H) Dec. 2011, (K) May 2013 (Set No. 2)]

Solution : Let μ be the mean and σ the standard deviation of the distribution. Then

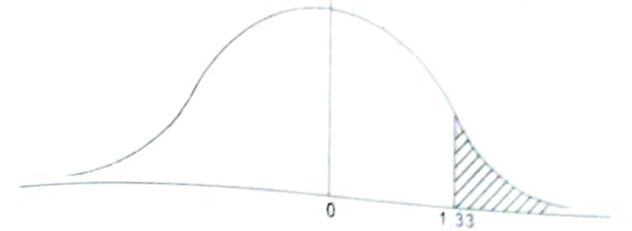
μ = 68 kgs and σ = 3 kgs

Let the variable X denote the masses of students

(i) When $X = 72,\ z = \dfrac{X-\mu}{\sigma} = \dfrac{72-68}{3} = 1.33$

$\therefore\ P(X > 72) = P(z > 1.33)$

$= 0.5 - A(1.33) = 0.5 - 0.4082 = 0.0918$

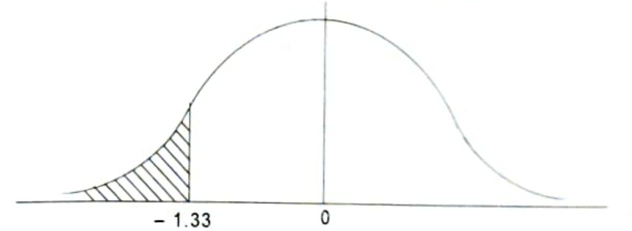


Number of students with more than 72 kgs = 300 × 0.0918 = 28 (approximately)

(ii) When $X = 64,\ z = \dfrac{X-\mu}{\sigma} = \dfrac{64-68}{3} = -1.33$

$\therefore\ P(X \le 64) = P(z \le -1.33)$

$= 0.5 - A(1.33) = 0.5 - 0.4082 = 0.0918$



Number of students have masses less than or equal to 64 Kg

$= 300 \times 0.0918 = 28$ (approximately)

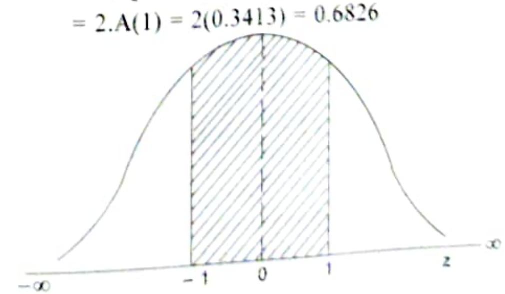
(iii) When $X = 65,\ z = \dfrac{X-\mu}{\sigma} = \dfrac{65-68}{3} = -1 = z_1$ (say)

When $x = 71,\ z = \dfrac{X-\mu}{\sigma} = \dfrac{71-68}{3} = 1 = z_2$ (say)

$\therefore\ P(65 \le X \le 71) = P(-1 \le z \le 1)$

$= A(z_2) + A(z_1) = A(1) + A(-1) = A(1) + A(1)$

$= 2.A(1) = 2(0.3413) = 0.6826$



$\therefore$ Required number of students = 300 × 0.6826 = 205 (approximately)

Example 14 : In an examination it is laid down that a student passes if he secures 40 percent or more. He is placed in the first, second and third division according as he secures 60% or more marks, between 50% and 60% marks and marks between 40% and 50% respectively. He gets a distinction in case he secures 75% or more. It is noticed from the results that 10% of the students failed in the examination, whereas 5% of them obtained distinction. Calculate the percentage of students placed in the second division. (Assume normal distribution of marks). **[JNTU (K) May 2013 (Set No. 1)]**

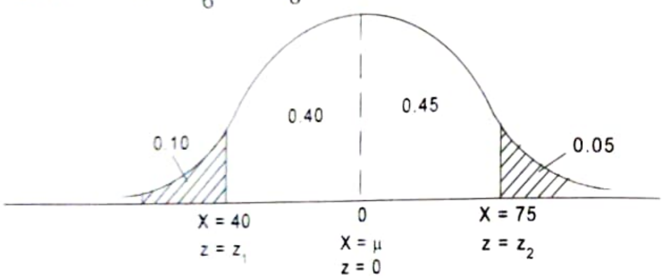
Solution : Let μ be the mean and σ the standard deviation of the normal distribution of marks.

Let the variable X denote the marks in the examination. Then we are given

$P(X < 40) = 0.10$ and $P(X \geq 75) = 0.05$

When $X = 40$, $z = \dfrac{X-\mu}{\sigma} = \dfrac{40-\mu}{\sigma} = -z_1$ (say)

When $X = 75$, $z = \dfrac{X-\mu}{\sigma} = \dfrac{75-\mu}{\sigma} = z_2$ (say)



$\therefore \quad P(0 < z < z_2) = 0.5 - 0.05 = 0.45$

and $\quad P(0 < z < z_1) = P(-z_1 < z < 0) = 0.5 - 0.10 = 0.40$

$\therefore \quad$ From normal tables, we have $z_1 = 1.28$ and $z_2 = 1.64$

Hence $\dfrac{40-\mu}{\sigma} = -1.28 \Rightarrow \dfrac{\mu-40}{\sigma} = 1.28$ .... (1) and $\dfrac{75-\mu}{\sigma} = 1.64$ .... (2)

(1) + (2) gives $\dfrac{35}{\sigma} = 2.92 \Rightarrow \sigma = \dfrac{35}{2.92} = 11.99 \simeq 12$

From (1), $\mu = 40 + \sigma\,(1.28) = 40 + 12\,(1.28) \approx 55$

The probability 'p' that a candidate is placed in the second division is equal to the probability that his score lies between 50 and 60. That is,

$$p = P(50 < X < 60) = P(-0.42 < Z < 0.42) \quad \left[\because Z = \frac{X-55}{12}\right]$$
$$= A(0.42) + A(-0.42)$$
$$= A(0.42) + A(0.42) \quad \text{(due to symmetry)}$$
$$= 2A(0.42) = 2(0.1628) = 0.3256 = 0.32 \text{ (approximately)}$$

Hence 32% candidates get second division in the examination.

Example 15 : A manufacturer knows from experience that the resistance of resistors he produces is normal with mean 100 ohms and standard deviation 2 ohms. What percentage of resistors will have resistance between 98 ohms and 102 ohms ?

Solution : Let μ be the mean and σ the standard deviation of the normal distribution. Then we are given that

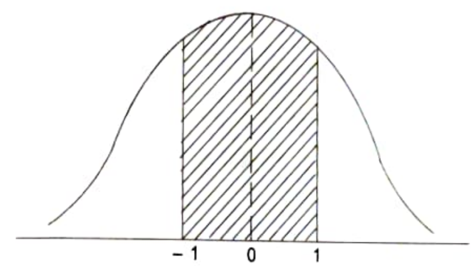
$\mu = 100$ ohms and $\sigma = 2$ ohms

Let the variable X represents the resistance of resistors

When $X = 98$ ohms, $z = \dfrac{X-\mu}{\sigma} = \dfrac{98-100}{2} = -1$

and when $X = 102$ ohms, $z = \dfrac{X-\mu}{\sigma} = \dfrac{102-100}{2} = 1$

$$\therefore P(98 < X < 102) = P(-1 < z < 1)$$
$$= A(1) + A(-1) = A(1) + A(1) = 2A(1)$$
$$= 2(0.3413) = 0.6826$$



$\therefore$ Percentage of resistors having resistance between 98 ohms and 102 ohms $= 68.26 = 68$ (approximately).

Example 16 : Given that the mean height of students in a class is 158 cms with standard deviation of 20 cms. Find how many students heights lie between 150 cms and 170 cms, if there are 100 students in the class. **[JNTU 2007, 2007S, (H) Dec. 2011 (Set No. 4)]**

Solution : We have

Mean, $\mu = 158$ cm and standard deviation, $\sigma = 20$ cms

$\therefore \quad z = \dfrac{X-\mu}{\sigma} = \dfrac{X-158}{20}$

When $X = 150$, $z = \dfrac{150-158}{20} = \dfrac{-2}{5} = -0.4$

When $X = 170$, $z = \dfrac{170-158}{20} = \dfrac{3}{5} = 0.6$

$$\therefore P(150 \leq X \leq 170) = P(-0.4 \leq z \leq 0.6)$$
$$= P(-0.4 \leq z \leq 0) + P(0 \leq z \leq 0.6)$$
$$= P(0 \leq z \leq 0.4) + P(0 \leq z \leq 0.6), \text{ due to symmetry}$$
$$= 0.1554 + 0.2257 = 0.3811$$

Number of students whose height lie between 150 cms and 170 cms

= Probability × total no. of students

= 0.3811 × 100 = 38 (∵ Number of students should be integer.)

Example 17 : The mean and standard deviation of a normal variable are 8 and 4 respectively.

[JNTU (H) 2009 (Set No. 1)]

Find (*i*) $p(5 \le X \le 10)$ (*ii*) $p(x \ge 5)$

Solution : Given mean, $\mu = 8$ and standard deviation, $\sigma = 4$.

(*i*) When $x = 5$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{5-8}{4} = \dfrac{-3}{4} = -0.75 = z_1$ (say)

When $x = 10$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{10-8}{4} = \dfrac{2}{4} = \dfrac{1}{2} = 0.5 = z_2$ (say)

$\therefore P(5 \le X \le 10) = P(-0.75 \le z \le 0.5)$

$= A(z_2) + A(z_1) = A(-0.75) + A(0.5)$

$= A(0.75) + A(0.5)$ [∵ $A(-z) = A(z)$]

$= 0.2734 + 0.1916 = 0.465$

(*ii*) When $x = 5$, $z_1 = -0.75$, from above

$\therefore P(X \ge 5) = P(z_1 \ge -0.75) = 0.5 - A(z_1)$

$= 0.5 - A(-0.75) = 0.5 - 0.2734 = 0.2266$

Example 18 : In a test on 2000 electrical bulbs, it was found that the life of a particular make, was normally distributed with an average life of 2040 hours and S.D. of 40 hrs. Estimate the number of bulbs likely to burn for (*i*) more than 2140 hrs. (*ii*) between 1920 and 2080 hrs. (*iii*) less than 1960 hrs.

[JNTU (K) June 2015 (Set No.1)]

Solution : Given mean, $\mu = 2040$ hrs and S.D., $\sigma = 40$ hrs

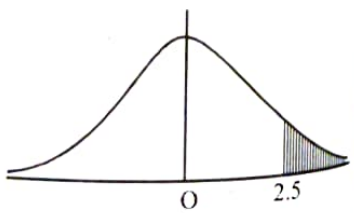
(*i*) When $x = 2140$,

$z = \dfrac{x-\mu}{\sigma} = \dfrac{2140-2040}{40} = \dfrac{100}{40} = 2.5$

$\therefore P(x > 2140) = P(z > 2.5)$

$= 0.5 - P(0 \le z \le 2.5)$

$= 0.5 - 0.4938 = 0.0062$

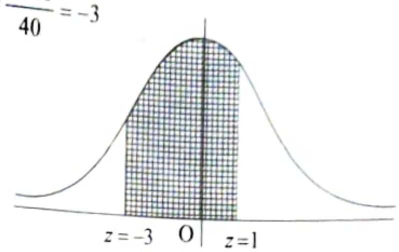


Hence the number of bulbs likely to burn for more than 2140 hours $= 0.0062 \times 2000 = 12.4 \simeq 12$.

(*ii*) When $x_1 = 1920$, $z = \dfrac{x_1-\mu}{\sigma} = \dfrac{1920-2040}{40} = \dfrac{-120}{40} = -3$

When $x_2 = 2080$, $z = \dfrac{x_2-\mu}{\sigma} = \dfrac{2080-2040}{40}$

$= \dfrac{40}{40} = 1$



$\therefore P(1920 \le x \le 2080) = P(-3 \le z \le 1)$

$= P(-3 \le z \le 0) + P(0 \le z \le 1)$

$= P(0 \le z \le 3) + P(0 \le z \le 1)$ (by symmetry)

$= 0.4986 + 0.3413 = 0.8399$

Hence number of bulbs which are likely to burn between 1920 hrs. and 2080 hrs.

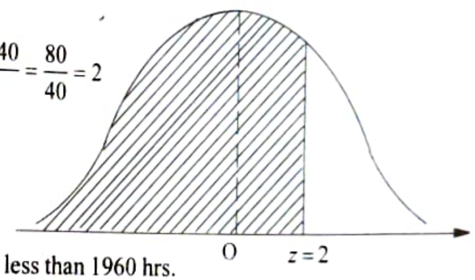
$= 0.8399 \times 2000 = 1679.8 \approx 1680$

(*iii*) When $x = 1960$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{1960-2040}{40} = \dfrac{80}{40} = 2$

$\therefore P(x \le 1960) = P(z \le 2)$

$= 0.5 + P(0 \le z \le 2)$

$= 0.5 + 0.4772 = 0.9772$



Hence number of bulbs likely to burn less than 1960 hrs.

$= 0.9772 \times 2000 = 1954$

Example 19 : If the marks obtained by a number of students in a certain subject are approximately normally distributed with mean 65 and standard deviation 5. If three students are selected at random from this group what is the probability that at least one of them would have scored above 75?

[JNTU (K) May 2010 (Set No. 2)]

Solution : Here μ = mean = 65 and σ = S.D. = 5

We have $z = \dfrac{x-\mu}{\sigma}$ when $x = 75$, $z = \dfrac{75-65}{5} = \dfrac{10}{5} = 2$

$\therefore p = P(x > 75) = P(z > 2) = 0.5 - P(0 \le z \le 2) = 0.5 - 0.4772 = 0.0228$

Hence the required probability that out of 3 students, at least one of them will have marks over 75 is given by

${}^3C_1 pq^2 + {}^3C_2 p^2 q^{3-2} + {}^3C_3 p^3 q^{3-3} = 3p(1-p)^2 + 3p^2(1-p) + p^3$ [∵ $q = 1 - p$]

$= 3(0.0228)(1-0.0228)^2 + 3(0.0228)^2(1-0.0228) + (0.0228)^3$

$= 0.0668$.

Example 20 : If X is normally distributed with mean 2 and variance 0.1, then find $P(|X-2| \ge 0.01)$?

[JNTU (K) May 2010 (Set No. 3)]

Solution : Here $\mu = 2, \sigma = 0.1$

when $X = 1.99, z = \frac{x-\mu}{\sigma} = \frac{1.99-2}{0.1} = -0.1$

when $X = 2.01, z = \frac{2.01-2}{0.1} = \frac{0.01}{0.1} = 0.1$

$\therefore P(|X-2| < 0.01) = P(1.99 < X < 2.01)$

$= P(-0.1 < z < 0.1) = P(-0.1 < z < 0) + P(0 < z < 0.1)$

$= 2P(0 < z < 0.1)$ [by symmetry]

$= 2 \times 0.0398 = 0.0796$

$\therefore P(|X-2| \geq 0.01) = 1 - P(|X-2| < 0.01) = 1 - 0.0796 = 0.9204$

Example 21 : The mean height of students in a college is 155 cms and standard deviation is 15. What is the probability that the mean height of 36 students is less than 157 cms.

[JNTU (H) Dec. 2011 (Set No. 4), Sept. 2017]

Solution : Let X be the continuous random variable denoting the height of students. Then mean, $\mu = 155$ and S. D., $\sigma = 15$.

We know that the standardized value of X is $z = \frac{X-\mu}{\sigma}$

when $X = 157, z = \frac{157-155}{15} = \frac{2}{15} = 0.13$

Probability that the height of student is less than 157 cm. $= P(X < 157)$

$= P(z < 0.13)$

= area under the normal curve to the right of $z = 0.13$

$= 0.0517$

$\therefore$ Required probability $= 0.0517 \times 36 = 1.86$

Example 22 : The mean inside diameter of a sample of 200 washers produced by a Machine is 500 cms with standard deviation 0.005 cms. The purpose of which these washers are intended a maximum tolerance in the diameter 0.495 to 0.505 cms. otherwise the washers are considered defective. Determine the precentage of defective washers produced by the machine, assuming the diameters are normally distributed. **[JNTU (H) Apr. 2012 (Set No. 3)]**

Solution : We are given

μ = Mean of the inside diameters = 0.500 cm

σ = Standard deviation = 0.005 cm

The tolerance limits of non - defective tubes are 0.495 cm and 0.505 cm.

When $x = 0.495$ (standardized value of 0.495),

$$z = \frac{x-\mu}{\sigma} = \frac{0.495-0.500}{0.005} = -1$$

When $x = 0.505$ (*i.e.* standardized value of 0.505),

$$z = \frac{x-\mu}{\sigma} = \frac{0.505-0.500}{0.005} = 1$$

$\therefore$ Proportion of non-defective washers

= Area under the standard normal curve between $z = -1$ and $z = +1$

$= 2$ (Area between $z = 0$ and $z = 1$)

$= 2 \times 0.3413 = 0.6826 = 68.26\,\%$

Hence the required percentage of defective washers

$= 100\% - 68.26\% = 31.74\%$

[(or) Probability of defective washers $= 1 -$ probability of non-defective washers

$= 1 - p(0.495 \leq x \leq 0.505)$

$= 1 - p(-1 \leq z \leq 1)$

$= 1 - 2 \times p(0 \leq z \leq 1)$

$= 1 - 2(0.3413) = 0.3174$

$\therefore$ Percentage of defective washers $= 31.74 \approx 32$]

Example 23 : 1000 students have written an examination the mean of test is 35 and standard deviation is 5. Assuming the distribution to be normal, find:

(*i*) How many students marks lie between 25 and 40?

(*ii*) How many students get more than 40?

(*iii*) How many students get below 20?

(*iv*) How many students get more than 50? **[JNTU(K) Dec. 2013 (Set No. 1)]**

Solution : Here $\mu = 35$ and $\sigma = 5$

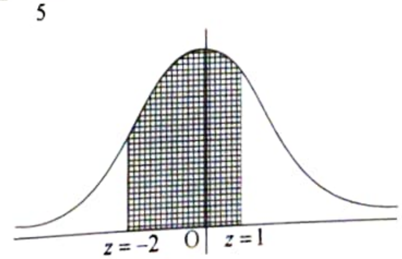
We know, standarised variate, $z = \frac{x-\mu}{\sigma} = \frac{x-35}{5}$

(*i*) When $x_1 = 25$, then

$z = \frac{25-35}{5} = \frac{-10}{5} = -2$

When $x_2 = 40$, then

$z = \frac{40-35}{5} = \frac{5}{5} = 1$



$\therefore$ Probability of students whose marks lie between 25 and 40 $= P(25 \leq X \leq 40)$

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

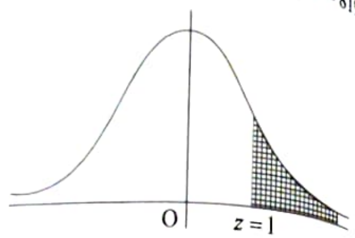
$= P(0 \le z \le 2) + P(0 \le z \le 1)$, by symmetry

$= 0.4772 + 0.3415$ [From Normal Tables]

$= 0.8185$

Number of students whose marks lie between 25 and 40 $= 1000 \times 0.8185 = 819$ (approximately)

(ii) When $x = 40$, then $z = \frac{40-35}{5} = \frac{5}{5} = 1$



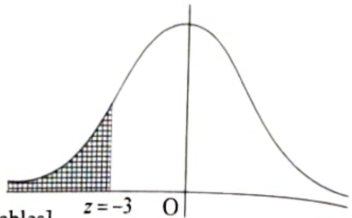
$P(X > 40) = P(z > 1) = 0.5 - P(z \le 1)$

$= 0.5 - 0.3415 = 0.1585$

[From Normal Tables]

$\therefore$ Number of students whose marks are greater than 40 $= 1000 \times 0.1585 = 159$ (approximately)

(iii) When $x = 20$, then $z = \frac{20-35}{5} = \frac{-15}{5} = -3$



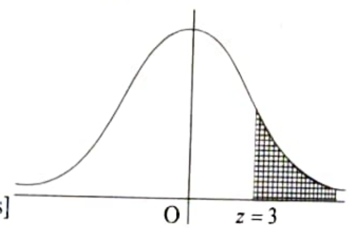
$P(X < 20) = P(z \le -3) = 0.5 - P(-3 \le z \le 0)$

$= 0.5 - P(0 \le z \le 3)$, by symmetry

$= 0.5 - 0.499 = 0.001$ [From Normal Tables]

$\therefore$ Number of students whose marks are less than 20 $= 1000 \times 0.001 = 1$

(iv) When $x = 50$, $z = \frac{50-35}{5} = \frac{15}{5} = 3$



$\therefore P(X > 50) = P(z > 3) = 0.5 - P(0 \le z \le 3)$

$= 0.5 - 0.499$

$= 0.001$ [From Normal Tables]

Hence number of students whose marks are greater than 50 $= 1000 \times 0.001 = 1$.

Example 24 : If X has the binomial distribution with mean 25 and probability of success 1/5, find $P(X < \mu - 2\sigma)$, where μ and σ^2 are the mean and variance of the distribution?

[JNTU (K) Nov. 2009 (Set No.3)]

Solution : Given mean, $\mu = 25 = np$ and $p = \frac{1}{5}$

$\Rightarrow n\left(\frac{1}{5}\right) = 25 \quad \therefore n = 125$

Now $q = 1 - p = 1 - \frac{1}{5} = \frac{4}{5}$

Variance, $\sigma^2 = npq = 125 \times \frac{1}{5} \times \frac{4}{5} = 20$

$\therefore \sigma = \sqrt{20} = 4.47$

Now $\mu - 2\sigma = 25 - 8.94 = 16.06$

$\therefore P(X < \mu - 2\sigma) = P(X < 16.06) = P\left(z < \frac{x-\mu}{\sigma}\right) = P\left(z < \frac{16.06-25}{4.47}\right)$

$= P(z < -2) = 0.5 - A(2) = 0.5 - 0.4772$ (from normal tables)

$= 0.0228$

5.9 NORMAL APPROXIMATION TO THE BINOMIAL DISTRIBUTION

The Normal distribution can be used to approximate the Binomial distribution. Suppose the number of successes x ranges from x_1 to x_2. Then the probability of getting x_1 to x_2 successes is given by

$$\sum_{r=x_1}^{x_2} n_{c_r} p^r q^{n-r}$$

For large n, the calculation of binomial probabilities is very difficult. In such cases the binomial curve can be replaced by the normal curve and the required probability is computed. We consider two cases.

Case 1 : When $p = q = \frac{1}{2}$

Even when n is not large, Binomial distribution (B.D) can be approximated by Normal distribution (N.D).

Mean of the B.D. = np and S.D. = $\sqrt{npq}$

Hence for the corresponding Normal distribution, μ and σ are known.

We know that $z = \frac{x-\mu}{\sigma}$

Let z_1 and z_2 be the values of z, corresponding to x_1 and x_2 of x respectively. Then

$$P(x_1 < x < x_2) = P(z_1 < z < z_2) = \int_{z_1}^{z_2} \phi(z)\, dz$$

and this can be determined by using the normal tables.

Case 2 : When $p \ne q$

For large n, we can approximate the binomial curve by the normal curve and calculate the probability.

For any success x, real class interval is $\left(x - \frac{1}{2}, x + \frac{1}{2}\right)$. Hence z_1 must correspond to the lower limit of the x_1 class and z_2, to the upper limit of the x_2 class.

Hence $z_1 = \frac{\left(x_1 - \frac{1}{2}\right) - \mu}{\sigma} = \frac{x_1 - \frac{1}{2} - np}{\sqrt{npq}}$

and $z_2 = \frac{\left(x_2 + \frac{1}{2}\right) - \mu}{\sigma} = \frac{x_2 + \frac{1}{2} - np}{\sqrt{npq}}$

$\therefore$ The required probability $= \int_{z_1}^{z_2} \phi(z)\,dz$ and this can be evaluated using normal tables.

SOLVED EXAMPLES

Example 1 : Find the probability that out of 100 patients between 84 and 95 inclusive will survive a heart-operation given that the chances of survival is 0.9.

Solution : Here $p = 0.9$ and $n = 100$ $\quad \therefore q = 1 - p = 1 - 0.9 = 0.1$

Standard deviation $= \sqrt{npq} = \sqrt{90(0.1)} = \sqrt{9} = 3$

The required probability $= \sum_{r=84}^{95} {}^{100}C_r (0.9)^r (0.1)^{100-r}$

We have to sum up a large number of terms of the binomial. To avoid it, we can replace the binomial distribution by a normal distribution.

Here $x_1 = 84$ and $x_2 = 95$. Hence while calculating z_1, z_2 corresponding to 84 and 95, we take $P(84 \le x \le 95) = P(83.5 < x < 95.8)$ and calculate z_1, z_2 correspondingly,

Hence $z_1 = \frac{84-(1/2)-90}{3} = \frac{-13}{6}$ and $z_2 = \frac{95+(1/2)-90}{3} = \frac{11}{6}$

Hence the required probability $= \int_{-13/6}^{11/6} \phi(z)\,dz$

$$= P\left(-\frac{13}{6} \le z \le \frac{11}{6}\right)$$

$$= A\left(\frac{11}{6}\right) + A\left(\frac{-13}{6}\right)$$

$$= A\left(\frac{11}{6}\right) + A\left(\frac{13}{6}\right)$$

$$= A(1.83) + A(2.17)$$

$$= 0.4664 + 0.4850 = 0.9514$$

Example 2 : Eight coins are tossed together. Find the probability of getting 1 to 4 heads in a single toss.

Solution : Here $p = q = \frac{1}{2}$

Mean, $\mu = np = 8 \times \frac{1}{2} = 4$ and standard deviation, $\sigma = \sqrt{npq} = \sqrt{4 \times \frac{1}{2}} = \sqrt{2}$

Even though $n = 8$ is not sufficiently large, we can approximate the symmetric binomial by a normal distribution.

Here $x_1 = 1$ and $x_2 = 4$

Thus $z_1 = \frac{x_1 - \frac{1}{2} - \mu}{\sigma} = \frac{1 - \frac{1}{2} - 4}{\sqrt{2}} = \frac{-7}{2\sqrt{2}} = -2.47$

and $z_2 = \frac{x_2 + \frac{1}{2} - \mu}{\sigma} = \frac{4 + \frac{1}{2} - 4}{\sqrt{2}} = \frac{1}{2\sqrt{2}} = 0.35$

The required probability $= \int_{-2.47}^{0.35} \phi(z)\,dz$

$$= P(-2.47 \le z \le 0.35) = A(0.35) + A(-2.47)$$

$$= A(0.35) + A(2.47)$$

$$= 0.4932 + 0.1368 = 0.63$$

Example 3 : Find the probability that by guess-work a student can correctly answer 25 to 30 questions in a multiple-choice quiz consisting of 80 questions. Assume that in each question with four choices, only one choice is correct and student has no knowledge of the subject. **[JNTU (K) II Sem June 2015 (Set No. 1)]**

Solution : Here $p = \frac{1}{4}$. So $q = 1 - p = 1 - \frac{1}{4} = \frac{3}{4}$. Mean, $\mu = np = 80\left(\frac{1}{4}\right) = 20$

and $\sigma = \sqrt{npq} = \sqrt{20 \times \frac{3}{4}} = \sqrt{15}$, $x_1 = 25$ and $x_2 = 30$

Thus $z_1 = \frac{\left(x_1 - \frac{1}{2}\right) - \mu}{\sigma} = \frac{\left(25 - \frac{1}{2}\right) - 20}{3.87} = \frac{4.5}{3.87} = 1.16$

and $z_2 = \frac{\left(x_2 + \frac{1}{2}\right) - \mu}{\sigma} = \frac{\left(30 + \frac{1}{2}\right) - 20}{3.87} = \frac{10.5}{3.87} = 2.71$

Hence the required probability $= P(25 \le X \le 30)$

$$= P(1.16 \le z \le 2.71)$$

$$= |A(2.71) - A(1.16)|$$

$$= 0.4966 - 0.3770 = 0.1196$$

Example 4 : Find the probability of getting an even number on face 3 to 5 times in throwing 10 dice together. **[JNTU (H) May 2012 (Set No. 2)]**

Solution : p = Probability of getting an even number on face

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

$\therefore q = 1 - p = \frac{1}{2}$

Here $n = 10$

Mean, $\mu = np = 10\left(\frac{1}{2}\right) = 5$ and $\sigma = \sqrt{npq} = \sqrt{\frac{5}{2}} = 1.58$

Also $x_1 = 3$ and $x_2 = 5$

$$\text{Thus } z_1 = \frac{\left(x_1 - \frac{1}{2}\right) - \mu}{\sigma} = \frac{(3-0.5)-5}{1.58} = \frac{-2.5}{1.58} = -1.58$$

$$z_2 = \frac{\left(x_2 + \frac{1}{2}\right) - \mu}{\sigma} = \frac{(5+0.5)-5}{1.58} = \frac{0.5}{1.58} = 0.32$$

$$\text{Hence the required probability} = \int_{z_1}^{z_2} \phi(z)\,dz = \int_{-1.58}^{0.32} \phi(z)\,dz$$

$$= P(-1.58 \le z \le 0.32)$$

$$= A(0.32) + A(1.58)$$

$$= 0.1256 + 0.4429$$

$$= 0.5685$$

5.10 FITTING OF NORMAL DISTRIBUTION

Consider the following frequency distribution

Class-interval (x)	Frequency (f)
$l_1 - u_1$	f_1
$l_2 - u_2$	f_2
$l_3 - u_3$	f_3
...	...
...	...
$l_i - u_i$	f_i
...	...
$l_n - u_n$	f_n

Then the normal curve (i.e., Normal Distribution) can be fitted to it using

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

Working Procedure :

1. Calculate the mean μ and standard deviation σ from the given frequency distribution.

2. Compute the standard normal variate corresponding to the lower limits of each of the class interval. That is, find $z_i = \frac{x_i' - \mu}{\sigma}$, where x_i' is the lower limit of the ith class interval.
3. Compute the areas under the normal curve from 0 to z_i using normal tables.
4. Areas for the successive class intervals are obtained by taking the difference between the successive areas calculated in Step 3 (when z_i's are of opposite sign, add the successive areas).
5. Expected or Theoretical normal frequencies are obtained by multiplying the areas in Step 4 by $N(= \Sigma f_i)$.

SOLVED EXAMPLES

Example 1 : Fit a normal distribution to the following data :

Class-interval	60-62	63-65	66-68	69-71	72-74
Frequency (f)	5	18	42	27	8

Solution : Here $N = \Sigma f_i = 100$, mean $= \mu = 67.45$ and standard deviation $= \sigma = 2.92$.

The calculations are arranged as follows :

S. No.	Class interval	Observed frequency (f)	True lower class limit x_i	Standard variate $z_i = (x_i - 67.45)/2.92$	Area from 0 to z_i	Area for each class (A)	Expected frequency = NA = 100A
1	60-62	5	59.5	−2.72	0.4967	0.0413	4.13
2.	63-65	18	62.5	−1.70	0.4554		
						0.2068	20.68
			65.5	−0.67	0.2486		
						0.3892	38.92
3.	66-68	42	68.5	0.36	0.1406		
						0.2771	27.71
4.	69-71	27	71.5	1.39	0.4177		
						0.0743	7.43
5.	72-74	8	74.5	2.41	0.4920		

Hence the expected frequencies corrected to nearest integers are 4, 21, 39, 28, 7 which agree with the observed frequencies.

REVIEW QUESTIONS

1. Define Normal Distribution and derive Normal Distribution as a limiting form of Binomial distribution.
2. Write the uses of Normal Distribution
3. List the properties of Normal Distribution.
4. Define a Normal distribution and derive the Mean, Median and Mode of a Normal distribution. **[JNTU (H) Dec. 2014]**

ANSWERS

1. (i) 0.465 (ii) 0.7734 (iii) 0.6514 (iv) 0.0401	3. (i) 0.044 (ii) 0.1598 (iii) 0.5328
2. 0.044	5. 0.3174
4. 0.0919	7. 294
6. $\mu = 50, \sigma = 10$	9. (i) 67 (ii) 184 (iii) 1909
8. (i) 0.6826 (ii) 0.0228 (iii) 0	11. 3.85
10. 84	13. 0.1841
12. 28	15. 0.7530
14. 0.1887	17. 0.02872
16. (i) 0.0648 (ii) 0.8964 (iii) 0.0575	22. 1, 4, 6, 4, 1
18. 155, 235	24. 1, 4, 6, 4, 1
23. 2, 1, 3, 36, 16, 3, 0	

5.11 EXPONENTIAL DISTRIBUTION

Exponential distribution is also a continuous distribution useful for describing business data and its probability density function is given by $f(x) = \lambda e^{-\lambda x}$ for $x > 0$, where λ is a parameter.

In other words, a random variable X is said to have an exponential distribution with parameter $\lambda > 0$, if its probability density function is given by

$$f(x,\lambda) = \begin{cases} \lambda e^{-\lambda x}, 0 < x < \infty, \lambda > 0 \\ 0, x \le 0 \end{cases}$$

Constants of Exponential Distribution :

(i) Mean of the Exponential distribution is λ and

(ii) Standard Deviation is also $\frac{1}{\lambda}$.

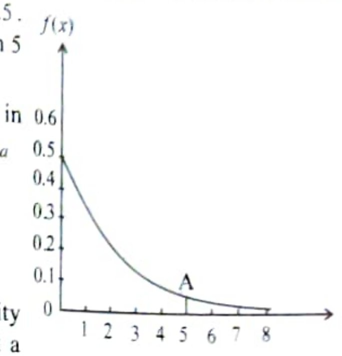
(iii) Its Median is $\frac{1}{\lambda}\log_e 2$ and mode is at $x = 0$.

Examples :

1. The time between two successive breakdowns of a machine will be exponentially distributed.
2. The time gap between two successive arrivals to a waiting line, known as inter arrival time, will be exponentially distributed.

SOLVED EXAMPLES

Example 1 : The length of time (in days) between sales for an automobile sales man is modeled as an exponential random variable with $\lambda = 0.5$. What is the probability that the salesman goes more than 5 days without a sale.



Solution : The area A to the right of a number 'a' in exponential distribution is given by $A = P(x \ge a) = e^{-\lambda a}$

Here $\lambda = 0.5$ and $a = 5$

$\therefore A = P(x \ge 5) = e^{-(0.5)(5)} = e^{-0.25} = 0.0821$

Hence the automobile salesman has a probability of about 0.0821 of going more than 5 days without a sale.

Example 2 : The average number of accidents in an industry during a year is estimated to be 5. If the distribution of time between two consecutive accidents is known to be exponential, find the probability that there will be no accident during the next two months.

Solution : The number of accidents in an industry during a month $= \frac{5}{12}$.

So $\lambda = \frac{5}{12}$.

$\therefore P(t > 2) = e^{-\lambda t} = e^{-\frac{5}{12}\times 2} = e^{-0.833} = 0.4347$ (From $e^{-\lambda}$ tables)

Example 3 : For an exponential distribution $\lambda = 1.2$, find

(i) $P(x \ge 0.5)$ (ii) $P(1 \le x \le 2)$. Also find its mean and variance.

Solution : Here $\lambda = 1.2$. $\therefore f(x) = \lambda e^{-\lambda x} = (1.2). e^{-1.2x}. x \ge 0$

(i) $P(x \ge 0.5) = e^{-\lambda x} = e^{-1.2\times 5} = e^{-0.6} = 0.5488$

(ii) $P(1 \le x \le 2) = \int_1^2 f(x)\,dx = \int_1^2 1.2e^{-1.2x}\,dx$

$$= (1.2)\left[\frac{e^{-1.2x}}{(-1.2)}\right]_1^2 = -(e^{-1.2x})_1^2$$

$$= -(e^{-2.4} - e^{-1.2}) = e^{-1.2} - e^{-2.4}$$

$$= 0.2105$$

Now Mean $= \frac{1}{\lambda} = \frac{1}{1.2} = 0.8333$ and Variance $= \frac{1}{\lambda^2} = \left(\frac{1}{1.2}\right)^2 = \frac{1}{1.44} = 0.6944$

REVIEW QUESTIONS

1. Define Exponential distribution.
2. Derive mean of the Exponential distribution.
3. Write the mean, median and variance of the Exponential distribution.

EXERCISE 5(B)

1. The distribution of life, in hours, of a bulb is known to be exponential with mean life of 600 hours. What is the probability that
   (*i*) it will not last for more than 500 hours
   (*ii*) it will last for 700 hours.
2. A telephone operator attends on an average 150 telephone calls per minute. Assuming that the distribution of consecutive calls follows an exponential distribution, find the probability that
   (*i*) the time between two consecutive calls is less than 2 minutes.
   (*ii*) the next call will be received only after 3 minutes.

ANSWERS

1. (*i*) 0.5653 (*ii*) 0.3114 **2.** (*i*) 0.9933 (*ii*) 0.006

5.12 GAMMA DISTRIBUTION [JNTU (H) Dec. 2019 (R18)]

Definition. A random variable X which is distributed with probability density function

$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & \text{for } x > 0, \alpha > 0, \beta > 0 \\ 0, & \text{otherwise} \end{cases}$$

is called a Gamma variate with parameters α and β and its distribution is called the Gamma distribution. Here $\Gamma(\alpha)$ denote the Gamma variate with one parameter α and $\Gamma(\alpha)$ is a value of the Gamma function, defined by

$$\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1} dx$$

Constants of the Gamma distribution :

(*i*) **Mean :** We know that the mean denoted by μ is the expected value of a random variable having the probability density $f(x)$.

The mean of the Gamma distribution is given by

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x) dx = \int_0^\infty x\, f(x) dx$$

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty x \cdot e^{-x/\beta} x^{\alpha-1} dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty e^{-x/\beta} x^\alpha dx$$

Putting $y = \frac{x}{\beta}$, we get

$$\mu = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty e^{-y} \cdot (\beta y)^\alpha \beta\, dy = \frac{\beta}{\Gamma(\alpha)} \int_0^\infty e^{-y} y^\alpha dy$$

$$= \frac{\beta}{\Gamma(\alpha)} \int_0^\infty e^{-y} y^{(\alpha+1)-1} dy = \frac{\beta \Gamma(\alpha+1)}{\Gamma(\alpha)}, \text{ by definition of } \Gamma \text{ function}$$

$$= \frac{\beta \cdot \alpha \Gamma(\alpha)}{\Gamma(\alpha)} \quad [\because \Gamma(n+1) = n \cdot \Gamma(n)]$$

$$= \alpha\beta$$

(*ii*) **Variance :** Similarly, the variance of the Gamma distribution is $\alpha\beta^2$

i.e. $\sigma^2 = \alpha\beta^2$

(*iii*) **Mode :** The mode of the probability curve for the Gamma variate is at $x = \alpha - 1$ if $\alpha > 1$.

Note. When $\alpha = 1$, the Gamma distribution reduces to

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & \text{for } x > 0, \beta > 0 \\ 0, & \text{otherwise} \end{cases}$$

This is the Exponential distribution and whose mean and variance are β and β^2 respectively.

So the Gamma distribution with parameter unity is an exponential distribution.

REVIEW QUESTIONS

1. Define Gamma distribution. [JNTU (H) Dec. 2019 (R18)]
2. Write the mean and variance of the Gamma distribution.