Dr.P.K. SIVAKUMARAN
ASSISTANT PROFESSOR OF STATISTICS
PROBABILITY AND STATISTICS(20MCA25C)

DISCRETE DISTRIBUTIONS

BINOMIAL DISTRIBUTION

Introduction:

In this chapter we will discuss the theoretical discrete distributions in which variables are distributed according to some definite probability law, which can be expressed mathematically. The Binomial distribution is a discrete distribution expressing the probability of a set of dichotomous alternative i.e., success or failure. This distribution has been used to describe a wide variety of process in business and social sciences as well as other areas.

Bernoulli Distribution:

A random variable X which takes two values 0 and 1 with probabilities q and p i.e., P (x = 1) = p and P (x = 0) = q, q = 1 – p, is called a Bernoulli variate and is said to be a Bernoulli Distribution, where p and q takes the probabilities for success and failure respectively. It is discovered by Swiss Mathematician James Bernoulli (1654-1705).

Examples of Bernoulli' s Trails are:

1) Toss of a coin (head or tail)

2) Throw of a die (even or odd number)

3) Performance of a student in an examination (pass or fail)

Binomial Distribution:

A random variable X is said to follow binomial distribution, if its probability mass function is given by

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

Here, the two independent constants n and p are known as the ‘ parameters’ of the distribution. The distribution is completely determined if n and p are known. x refers the number of successes. If we consider N sets of n independent trials, then the number of times we get x success is $N(nC_x\ p^x\ q^{n-x})$. It follows that the terms in the expansion of $N\ (q + p)^n$ gives the frequencies of the occurrences of 0, 1, 2,..., x, ..., n success in the N sets of independent trials.

1.1.1 Condition for Binomial Distribution:

We get the Binomial distribution under the following experimental conditions.

1) The number of trials ‘ n’ is finite.

2) The trials are independent of each other.

3) The probability of success ‘ p’ is constant for each trial.

4) Each trial must result in a success or a failure.

The problems relating to tossing of coins or throwing of dice or drawing cards from a pack of cards with replacement lead to binomial probability distribution.

1.1.2 Characteristics of Binomial Distribution:

1. Binomial distribution is a discrete distribution in which the random variable X (the number of success) assumes the values 0,1, 2, ….n, where n is finite.

2. Mean = np, variance = npq and

   Standard deviation $\sigma = \sqrt{npq}$

   Coefficient of skewness $= \dfrac{q - p}{\sqrt{npq}}$,

   Coefficient of kurtosis $= \dfrac{1 - 6pq}{npq}$, clearly each of the probabilities is non-negative and sum of all probabilities is 1 ($p < 1$, $q < 1$ and $p + q = 1$, $q = 1 - p$).

3. The mode of the binomial distribution is that value of the variable which occurs with the largest probability. It may have either one or two modes.

4. If two independent random variables X and Y follow binomial distribution with parameter (n_1, p) and (n_2, p) respectively, then their sum (X + Y) also follows Binomial distribution with parameter $(n_1 + n_2, p)$.

5. If n independent trials are repeated N times, N sets of n trials are obtained and the expected frequency of x success is $N(nC_x\ p^x\ q^{n-x})$. The expected frequencies of 0, 1, 2… n success are the successive terms of the binomial distribution of $N(p + q)^n$.

Example 1:

Comment on the following: “ The mean of a binomial distribution is 5 and its variance is 9”

Solution:

The parameters of the binomial distribution are n and p

We have mean $\Rightarrow np = 5$

Variance $\Rightarrow npq = 9$

$$\therefore\ q = \frac{npq}{np} = \frac{9}{5}$$

$$q = \frac{9}{5} > 1$$

Which is not admissible since q cannot exceed unity. Hence the given statement is wrong.

Example :

Eight coins are tossed simultaneously. Find the probability of getting atleast six heads.

Solution:

Here number of trials, n = 8, p denotes the probability of getting a head.

$$\therefore\ p = \frac{1}{2} \text{ and } q = \frac{1}{2}$$

If the random variable X denotes the number of heads, then the probability of a success in n trials is given by

$$P(X = x) = nC_x p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

$$= 8C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{8-x} = 8C_x \left(\frac{1}{2}\right)^8$$

$$= \frac{1}{2^8}\ 8C_x$$

Probability of getting atleast six heads is given by

$$P(x \geq 6) = P(x = 6) + p(x = 7) + P(x = 8)$$

$$= \frac{1}{2^8} 8C_6 + \frac{1}{2^8} 8C_7 + \frac{1}{2^8} 8C_8$$

$$= \frac{1}{2^8}[8C_6 + 8C_7 + 8C_8]$$

$$= \frac{1}{2^8}[28 + 8 + 1] = \frac{37}{256}$$

$x = 0, 1, 2, ..., n$

Example 3:

Ten coins are tossed simultaneously. Find the probability of getting (i) atleast sevenheads (ii) exactly seven heads (iii) atmost seven heads

Solution:

p = Probability of getting a head $= \frac{1}{2}$

q = Probability of not getting a head $= \frac{1}{2}$

The probability of getting x heads throwing 10 coins simultaneously is given by

$$P(X = x) = nC_x p^x q^{n-x},$$

$$= 10C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x} = \frac{1}{2^{10}} 10C_x$$

i) Probability of getting atleast seven heads

$$P(x \geq 7) = P(x = 7) + P(x = 8) + P(x = 9) + P(x = 10)$$

$$= \frac{1}{2^{10}} [10C_7 + 10C_8 + 10C_9 + 10C_{10}]$$

$$= \frac{1}{1024}[120 + 45 + 10 + 1] = \frac{176}{1024}$$

ii) Probability of getting exactly 7 heads

$$P(x = 7) = \frac{1}{2^{10}} 10C_7 = \frac{1}{2^{10}} (120)$$

$$= \frac{120}{}$$

$1024 \qquad \overline{2^{10}}$

ii) Probability of getting atmost 7 heads

$$P(x \le 7) = 1 - P(x > 7)$$
$$= 1 - \{P(x=8) + P(x=9) + P(x=10)\}$$
$$= 1 - \frac{1}{2^{10}} \{10C_8 + 10C_9 + 10C_{10}\}$$
$$= 1 - \frac{1}{2^{10}}[45 + 10 + 1]$$
$$= 1 - \frac{56}{1024}$$
$$= \frac{968}{1024}$$

Example :

20 wrist watches in a box of 100 are defective. If 10 watches are selected at random, find the probability that (i) 10 are defective (ii) 10 are good (iii) at least one watch is defective (iv) at most 3 are defective.

Solution:

20 out of 100 wrist watches are defective

Probability of defective wrist watch , $p = \frac{20}{100} = \frac{1}{5}$

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

Since 10 watches are selected at random, n =10

$$P(X = x) = nC_x p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

$$= 10C_x \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{10-x}$$

i) Probability of selecting 10 defective watches

$$P(x = 10) = 10C_{10} \left(\frac{1}{5}\right)^{10} \left(\frac{4}{5}\right)^0$$

$$= 1.\frac{1}{5^{10}}.1 = \frac{1}{5^{10}}$$

ii) Probability of selecting 10 good watches (i.e. no defective)

$$P(x = 0) = 10C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^{10}$$

$$= 1.1 \left(\frac{4}{5}\right)^{10} = \left(\frac{4}{5}\right)^{10}$$

iii) Probability of selecting at least one defective watch

$$P(x \geq 1) = 1 - P(x < 1)$$

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

$$= 1 - 10C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^{10}$$

$$= 1 - \left(\frac{4}{5}\right)^{10}$$

iv) Probability of selecting at most 3 defective watches

$$P(x \le 3) = P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)$$

$$= 10C_0\left(\frac{1}{5}\right)^0\left(\frac{4}{5}\right)^{10} + 10C_1\left(\frac{1}{5}\right)^1\left(\frac{4}{5}\right)^9 + 10C_2\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^8 + 10C_3\left(\frac{1}{5}\right)^3\left(\frac{4}{5}\right)^7$$

$$= 1.1.\left(\frac{4}{5}\right)^{10} + 10\left(\frac{1}{5}\right)^1\left(\frac{4}{5}\right)^9 + \frac{10.9}{1.2}\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^8 + \frac{10.9.8}{1.2.3}\left(\frac{1}{5}\right)^3\left(\frac{4}{5}\right)^7$$

$$= 1.(0.107) + 10(0.026) + 45(0.0062) + 120(0.0016)$$

$$= 0.859 \text{(approx)}$$

Example :

With the usual notation find p for binomial random variable X if n = 6 and 9P(X = 4) = P(X = 2)

Solution:

The probability mass function of binomial random variable X is given by

$P(X = x) = nC_x\, p^x\, q^{n-x}$, x = 0 , 1, 2, ...,n

Here n = 6 $\quad \therefore\ P(X = x) = 6C_x\, p^x\, q^{6-x}$

$$P(x = 4) = 6C_4\, p^4\, q^2$$

$$P(x = 2) = 6C_2\, p^2\, q^4$$

Given that,

$$9.\ P(x = 4) = P(x = 2)$$

$$9.\ 6C_4\, p^4\, q^2 = 6C_2\, p^2\, q^4$$

$$\Rightarrow 9 \times 15p^2 = 15q^2$$

$$9p^2 = q^2$$

Taking positive square root on both sides we get,

$$3p = q$$

$$= 1 - p$$

$$4p = 1$$

$$\therefore p = \frac{1}{4} = 0.25$$

Fitting of Binomial Distribution:

When a binomial distribution is to be fitted to an observed data, the following procedure is adopted.

1. Find Mean = $\bar{x} = \frac{\sum fx}{} = np$

$\sum f$

$\Rightarrow p = \dfrac{\bar{x}}{n}$ where n is number of trials.

2. Determine the value, $q = 1 - p$.
3. The probability function is $P(x) = nC_x\, p^x\, q^{n-x}$ put $x = 0$, we set $P(0) = q^n$ and $f(0) = N \times P(0)$
4. The other expected frequencies are obtained by using the recurrence formula is given by $f(x+1) = \dfrac{n-x}{x+1}\,\dfrac{p}{q}\, f(x)$

Example :

A set of three similar coins are tossed 100 times with the following results

Number of heads :	0	1	2	3
Frequency :	36	40	22	2

Fit a Binomial distribution.

Solution :

X	f	fx
0	36	0
1	40	40
2	22	44
3	2	6
	$\sum f = 100$	$\sum fx = 90$

$$\text{Mean} = \bar{x} = \frac{\sum fx}{\sum f} = \frac{90}{100} = 0.9$$

$$p = \frac{\bar{x}}{n}$$

$$= \frac{0.9}{3} = 0.3$$

$$q = 1 - 0.3$$

$$= 0.7$$

The probability function is $P(x) = nCx\ p^x\ q^{n-x}$

Here $n = 3$, $p = 0.3$ $q = 0.7$

$$\therefore\ P(x) = 3C_x\ (0.3)^x\ (0.7)^{3-x}$$

$$P(0) = 3C_0\ (0.3)^0\ (0.7)^3$$

$$= (0.7)^3 = 0.343$$

$\therefore\ f(0) = N \times P(0) = 0.343 \times 100 = 34.3$

The other frequencies are obtained by using the recurrence formula $f(x+1) = \frac{n-x}{x+1}\left(\frac{p}{q}\right) f(x)$.

By putting $x = 0, 1, 2$ the expected frequencies are calculated as follows.

$$f(1) = \frac{3-0}{0+1}\left(\frac{p}{q}\right) \times 34.3$$

$$= 3 \times (0.43) \times 34.3 = 44.247$$

$$f(2) = \frac{3-1}{1+1}\left(\frac{p}{q}\right) f(1)$$

$$= \frac{2}{2}(0.43) \times 44.247$$

$$= 19.03$$

$$f(3) = \frac{3-2}{2+1}\left(\frac{p}{q}\right) f(2)$$

$$= \frac{1}{3}(0.43) \times 19.03$$

$$= 2.727$$

The observed and theoretical (expected) frequencies are tabulated below:

					Total
Observed frquencies	36	40	22	2	100
Expected frequencies	34	44	19	3	100

Example 7:

4 coins are tossed and number of heads noted. The experiment is repeated 200 times and the following distribution is obtained .

x: Number of heads	0	1	2	3	4
f: frequencies	62	85	40	11	2

Solution :

X	0	1	2	3	4	Total
f	62	85	40	11	2	200
fx	0	85	80	33	8	206

$$\text{Mean} = \bar{x} = \frac{\sum fx}{\sum f} = \frac{206}{200} = 1.03$$

$$p = \frac{\bar{x}}{n} = \frac{1.03}{4} = 0.2575$$

$$\therefore\ q = 1 - 0.2575 = 0.7425$$

Here $n = 4$, $p = 0.2575$; $q = 0.7425$

The probability function of binomial distribution is $P(x) = nC_x p^x q^{n-x}$

The binomial probability function is

$$P(x) = 4C_x\,(0.2575)^x\,(0.7425)^{4-x}$$

$$P(0) = (0.7425)^4$$

$$= 0.3039$$

$$\therefore\ f(0) = NP(0)$$

$$= 200 \times 0.3039$$

$$= 60.78$$

The other frequencies are calculated using the recurrence formula $f(x+1) = \left(\frac{n-x}{x+1}\right)\left(\frac{p}{q}\right) f(x)$.

By putting x = 0,1, 2, 3 then the expected frequencies are calculated as follows:

Put x = 0, we get

$$f(1) = \frac{4-0}{0+1}(0.3468)(60.78)$$

$$= 84.3140$$

$$f(2) = \frac{4-1}{1+1}(0.3468)(84.3140)$$

$$= 43.8601$$

$$f(3) = \frac{4-2}{2+1}(0.3468)(43.8601)$$

$$= 10.1394$$

$$f(4) = \frac{4-3}{3+1}(0.3468)(10.1394)$$

$$= 0.8791$$

The theoretical and expected frequencies are tabulated below:

						Total
Observed frquencies	62	85	40	11	2	200
Expected frequencies	61	84	44	10	1	200

POISSON DISTRIBUTION:

Introduction:

Poisson distribution was discovered by a French Mathematician-cum-Physicist Simeon Denis Poisson in 1837. Poisson distribution is also a discrete distribution. He derived it as a limiting case of Binomial distribution. For n-trials the binomial distribution is $(q + p)^n$; the probability of x successes is given by $P(X = x) = nC_x\, p^x\, q^{n-x}$. If the number of trials n is very large and the probability of success ‘ p’ is very small so that the product np = m is non – negative and finite.

The probability of x success is given by

$$P(X = x) = \begin{cases} \dfrac{e^{-m} m^{x}}{x!} & \text{for } x = 0,1,2.... \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Here m is known as parameter of the distribution so that m >0

Since number of trials is very large and the probability of success p is very small, it is clear that the event is a rare event. Therefore Poisson distribution relates to rare events.

Note:

1) e is given by $e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + ... = 2.71828$

2) $P(X = 0) = \dfrac{e^{-m} m^{0}}{0!}$, $0! = 1$ and $1! = 1$

3) $P(X = 1) = \dfrac{e^{-m} m^1}{1!}$

Some examples of Poisson variates are :

1. The number of blinds born in a town in a particular year.
2. Number of mistakes committed in a typed page.
3. The number of students scoring very high marks in all subjects
4. The number of plane accidents in a particular week.
5. The number of defective screws in a box of 100, manufactured by a reputed company.
6. Number of suicides reported in a particular day.

Conditions:

Poisson distribution is the limiting case of binomial distribution under the following conditions:

1. The number of trials n is indefinitely large i.e., $n \to \infty$
2. The probability of success ‘ p’ for each trial is very small; i.e., $p \to 0$
3. $np = m$ (say) is finite , $m > 0$

Characteristics of Poisson Distribution:

The following are the characteristics of Poisson distribution

1. Discrete distribution: Poisson distribution is a discrete distribution like Binomial distribution, where the random variable assume as a countably infinite number of values 0,1,2 ….
2. The values of p and q: It is applied in situation where the probability of success p of an event is very small and that of failure q is very high almost equal to 1 and n is very large.
3. The parameter: The parameter of the Poisson distribution is m. If the value of m is known, all the probabilities of the Poisson distribution can be ascertained.
4. Values of Constant: Mean = m = variance; so that standard deviation = $\sqrt{m}$

   Poisson distribution may have either one or two modes.
5. Additive Property: If X and Y are two independent Poisson distribution with parameter m_1 and m_2 respectively. Then (X+Y) also follows the Poisson distribution with parameter $(m_1 + m_2)$.
6. As an approximation to binomial distribution: Poisson distribution can be taken as a limiting form of Binomial distribution when n is large and p is very small in such a

way that product np = m remains constant.

7. Assumptions: The Poisson distribution is based on the following assumptions.

   i) The occurrence or non- occurrence of an event does not influence the occurrence or non-occurrence of any other event.

   ii) The probability of success for a short time interval or a small region of space is proportional to the length of the time interval or space as the case may be.

   iii) The probability of the happening of more than one event is a very small interval is negligible.

Example :

Suppose on an average 1 house in 1000 in a certain district has a fire during a year. If there are 2000 houses in that district, what is the probability that exactly 5 houses will have a fire during the year? [given that $e^{-2} = 0.13534$]

$$\text{Mean, } \bar{x} = np, \; n = 2000 \;\text{ and } p = \frac{1}{1000}$$

$$= 2000 \times \frac{1}{1000}$$

$$m = 2$$

The Poisson distribution is $\frac{e^{-2}\,2^5}{5!}$

$$P(X = x) = \frac{e^{-m} m^x}{x!}$$

$$\therefore P(X = 5) =$$

$$= \frac{(0.13534)\times 32}{120}$$

$$= 0.036$$

(Note: The values of e^{-m} are given in Appendix)

Example :

In a Poisson distribution 3P(X = 2) = P(X = 4) Find the parameter 'm' .

Solution:

Poisson distribution is given by $P(X = x) = \frac{e^{-m} m^x}{x!}$

Given that $3P(x = 2) = P(x = 4)$

$$3.\frac{e^{-m} m^2}{2!} = \frac{e^{-m} m^4}{4!}$$

$$m^2 = \frac{3\times 4!}{2!}$$

$$\therefore m = \pm 6$$

Since mean is always positive $\therefore$ m = 6

Example :

If 2% of electric bulbs manufactured by a certain company are defective. Find the probability that in a sample of 200 bulbs i) less than 2 bulbs ii) more than 3 bulbs are defective.[$e^{-4} = 0.0183$]

Solution:

The probability of a defective bulb = $p = \frac{2}{100} = 0.02$

Given that n = 200 since p is small and n is large

We use the Poisson distribution

mean, m = np = 200 X 0.02 = 4

Now, Poisson Probability function, $P(X = x) = \dfrac{e^{-m}\, m^{x}}{x!}$

i) Probability of less than 2 bulbs are defective

$$= P(X < 2)$$
$$= P(x = 0) + P(x = 1)$$
$$= \frac{e^{-4}4^0}{0!} + \frac{e^{-4}4^1}{1!}$$
$$= e^{-4} + e^{-4}(4)$$
$$= e^{-4}(1+4) = 0.0183 \times 5$$
$$= 0.0915$$

ii) Probability of getting more than 3 defective bulbs

$$P(x > 3) = 1 - P(x \le 3)$$
$$= 1 - \{P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)\}$$
$$= 1 - e^{-4}\{1 + 4 + \frac{4^2}{2!} + \frac{4^3}{3!}\}$$
$$= 1 - \{0.0183 \times (1 + 4 + 8 + 10.67)\}$$
$$= 0.567$$

Fitting of Poisson Distribution:

The process of fitting of Poisson distribution for the probabilities of x = 0, 1,2,... success are given below :

i) First we have to calculate the mean = $\bar{x} = \frac{\sum fx}{\sum f} = m$

ii) The value of e^{-m} is obtained from the table (see Appendix)

iii) By using the formula $P(X = x) = \frac{e^{-m}.m^x}{x!}$

Substituting x = 0, $P(0) = e^{-m}$

Then $f(0) = N \times P(0)$

The other expected frequencies will be obtained by using the recurrence formula

$$f(x+1) = \frac{m}{x+1} f(x);\ x = 0,1,2,.......$$

Example :

The following mistakes per page were observed in a book.

Number of mistakes (per page)	0	1	2	3	4
Number of pages	211	90	19	5	0

Fit a Poisson distribution to the above data.

Solution:

x_i	f_i	$f_i x_i$
0	211	0
1	90	90
2	19	38
3	5	15
4	0	0
	N = 325	$\sum fx = 143$

$$\text{Mean} = \bar{x} = \frac{\sum fx}{N}$$

$$= \frac{143}{325} = 0.44 = m$$

Then $e^{-m} \Rightarrow e^{-0.44} = 0.6440$

Probability mass function of Poisson distribution is

$$P(x) = e^{-m} \frac{m^x}{x!}$$

Put $x = 0$, $P(0) = e^{-0.44} \frac{44^0}{0!}$

$$= e^{-0.44}$$

$$= 0.6440$$

$$\therefore f(0) = N\ P(0)$$

$$= 325 \times 0.6440$$

$$= 209.43$$

The other expected frequencies will be obtained by using the recurrence formula

$f(x+1) = \frac{m}{x+1} f(x)$. By putting x = 0,1,2,3 we get the expected frequencies and are calculated as follows.

$$f(1) = 0.44 \times 209.43 = 92.15$$

$$f(2) = \frac{0.44}{2} \times 92.15 = 20.27$$

$$f(3) = \frac{0.44}{3} \times 20.27 = 2.97$$

$$f(4) = \frac{0.44}{4} \times 2.97 = 0.33$$

						Total
Observed frquencies	211	90	19	5	0	325
Expected frequencies	210	92	20	3	0	325

Example :

Find mean and variance to the following data which gives the frequency of the number of deaths due to horse kick in 10 corps per army per annum over twenty years.

X	0	1	2	3	4	Total
F	109	65	22	3	1	200

Solution :

Let us calculate the mean and variance of the given data

x_i	f_i	$f_i x_i$	$f_i x_i^2$
0	109	0	0
1	65	65	65
2	22	44	88
3	3	9	27
4	1	4	16
Total	N = 200	$\sum fx = 122$	$\sum fx^2 = 196$

$$\text{Mean} = \bar{x} = \frac{\sum f_i x_i}{N}$$

$$= \frac{122}{200}$$

$$= 0.61$$

$$\text{Variance} = \sigma^2 = \frac{\sum f_i x_i^2}{N} - (\bar{x}^{\ 2}$$

$$= \frac{196}{200} - (0.61)^2$$

$$= 0.61$$

Hence, mean = variance = 0.61

Example :

100 car radios are inspected as they come off the production line and number of defects per set is recorded below

No. of defects	0	1	2	3	4
No. of sets	79	18	2	1	0

Fit a Poisson distribution and find expected frequencies

Solution :

x	f	fx
0	79	0
1	18	18
2	2	4
3	1	3
4	0	0
	N = 100	$\sum fx = 25$

$$\text{Mean} = \bar{x} = \frac{\sum fx}{N}$$

$$= \frac{25}{100}$$

$$\therefore m = 0.25$$

Then $e^{-m} = e^{-0.25} = 0.7788 = 0.779$

Poisson probability function is given by

$$P(x) = \frac{e^{-m} m^x}{x!}$$

$$P(0) = \frac{e^{-0.25}(0.25)^0}{0!} = (0.779)$$

$\therefore$ f(0) = N.P(0) = 100 × (0.779) = 77.9

Other frequencies are calculated using the recurrence formula

$$f(x+1) = \frac{m}{x+1} f(x).$$

By putting x = 0,1,2,3, we get the expected frequencies and are calculated as follows.

$$f(1) = f(0+1) = \frac{m}{0+1} f(0)$$

$$f(1) = \frac{0.25}{1}(77.9)$$

$$= 19.46$$

$$f(2) = \frac{0.25}{2}(19.46)$$
$$= 2.43$$

$$f(3) = \frac{0.25}{3}(2.43)$$
$$= 0.203$$

$$f(4) = \frac{0.25}{4}(0.203)$$
$$= 0.013$$

Observed frequencies	79	18	2	1	0	100
Expected frequencies	78	20	2	0	0	100

Example:

Assuming that one in 80 births in a case of twins, calculate the probability of 2 or more sets of twins on a day when 30 births occurs. Compare the results obtained by using (i) the binomial and (ii) Poisson distribution.

Solution:

(i) Using Binomial distribution

$$\text{Probability of twins birth} = p = \frac{1}{80} = 0.0125$$

$$\therefore q = 1 - p = 1 - 0.0125$$
$$= 0.9875$$
$$n = 30$$

Binomial distribution is given by

$$P(x) = nC_x\, p^x\, q^{n-x}$$

$$P(x \geq 2) = 1 - P(x < 2)$$

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

$$= 1 - \{30C_0(0.0125)^0\,(0.9875)^{30} + 30C_1(0.0125)^1(0.9875)^{29}\}$$

$$= 1 - \{1.1(0.9875)^{30} + 3\,(0.125)\,(0.9875)^{29}\}$$

$$= 1 - \{\,0.6839 + 0.2597\}$$

$$= 1 - 0.9436$$

$$P(x \geq 2) = 0.0564$$

(ii) By using Poisson distribution:

The probability mass function of Poisson distribution is given by

$$P(x) = \frac{e^{-m} m^x}{x!}$$

$$\text{Mean} = m = np$$
$$= 30(0.0125) = 0.375$$
$$P(x \geq 2) = 1 - P(x < 2)$$
$$= 1 - \{P(x = 0) + P(x = 1)\}$$
$$= 1 - \left\{ \frac{e^{-0.375}(0.375)^0}{0!} + \frac{e^{-0.375}(0.375)^1}{1!} \right\}$$
$$= 1 - e^{-0.375}(1 + 0.375)$$
$$= 1 - (0.6873)(1.375) = 1 - 0.945 = 0.055$$

CONTINUOUS DISTRIBUTIONS:

NORMAL DISTRIBUTION:

Introduction:

In the preceding sections we have discussed the discrete distributions, the Binomial and Poisson distribution.

In this section we deal with the most important continuous distribution, known as normal probability distribution or simply normal distribution. It is important for the reason that it plays a vital role in the theoretical and applied statistics.

The normal distribution was first discovered by DeMoivre (English Mathematician) in 1733 as limiting case of binomial distribution. Later it was applied in natural and social science by Laplace (French Mathematician) in 1777. The normal distribution is also known as Gaussian distribution in honour of Karl Friedrich Gauss(1809).

Definition:

A continuous random variable X is said to follow normal distribution with mean μ and standard deviation σ, if its probability density function

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

Note:

The mean μ and standard deviation σ are called the parameters of Normal distribution. The normal distribution is expressed by $X \sim N(\mu, \sigma^2)$

Condition of Normal Distribution:

i) Normal distribution is a limiting form of the binomial distribution under the following conditions.

a) n, the number of trials is indefinitely large ie., $n \rightarrow \infty$ and

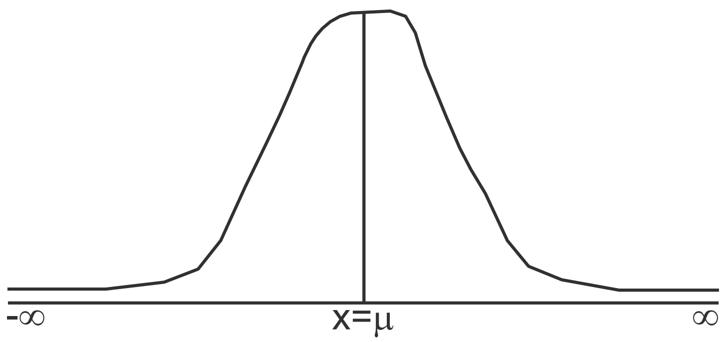
b) Neither p nor q is very small.

ii) Normal distribution can also be obtained as a limiting form of Poisson distribution with parameter $m \rightarrow \infty$

iii) Constants of normal distribution are mean = μ, variation = σ^2, Standard deviation = σ.

Normal probability curve:

The curve representing the normal distribution is called the normal probability curve. The curve is symmetrical about the mean (μ), bell-shaped and the two tails on the right and left sides of the mean extends to the infinity. The shape of the curve is shown in the following figure.



Properties of normal distribution:

1. The normal curve is bell shaped and is symmetric at $x = \mu$.
2. Mean, median, and mode of the distribution are coincide

   i.e., Mean = Median = Mode = μ
3. It has only one mode at $x = \mu$ (i.e., unimodal)
4. Since the curve is symmetrical, Skewness = $\beta_1 = 0$ and Kurtosis = $\beta_2 = 3$.
5. The points of inflection are at $x = \mu \pm \sigma$
6. The maximum ordinate occurs at $x = \mu$ and its value is = $\dfrac{1}{\sigma\sqrt{2\pi}}$
7. The x axis is an asymptote to the curve (i.e. the curve continues to approach but never touches the x axis)
8. The first and third quartiles are equidistant from median.

9. The mean deviation about mean is 0.8 σ

10. Quartile deviation = 0.6745 σ

11. If X and Y are independent normal variates with mean μ_1 and μ_2, and variance σ_1^2 and σ_2^2 respectively then their sum (X + Y) is also a normal variate with mean $(\mu_1 + \mu_2)$ and variance $(\sigma_1^2 + \sigma_2^2)$

12. Area Property $P(\mu - \sigma < x < \mu + \sigma) = 0.6826$

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

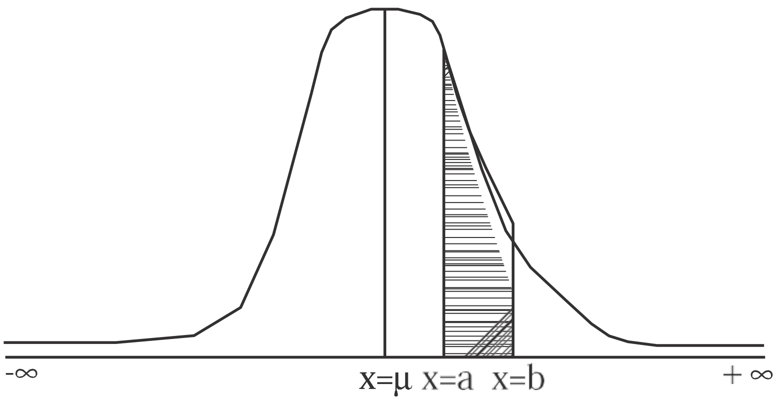
$P(\mu - 3\sigma < x < \mu + 3\sigma) = 0.9973$

Standard Normal distribution:

Let X be random variable which follows normal distribution with mean μ and variance σ^2 .The standard normal variate is defined as $Z = \dfrac{X-\mu}{\sigma}$ which follows standard normal distribution with mean 0 and standard deviation 1 i.e., $Z \sim N(0,1)$. The standard normal distribution is given by $\phi(z) = \dfrac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}z^2}$; $-\infty < z < \infty$ The advantage of the above function is that it doesn' t contain any parameter. This enable us to compute the area under the normal probability curve.

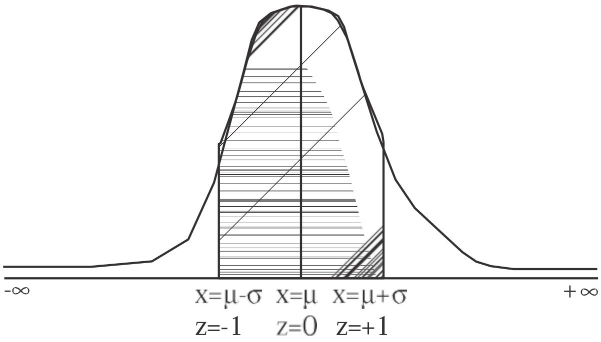
Area properties of Normal curve:

The total area under the normal probability curve is 1. The curve is also called standard probability curve. The area under the curve between the ordinates at x = a and x = b where a < b, represents the probabilities that x lies between x = a and x = b i.e., $P(a \le x \le b)$



To find any probability value of x, we first standardize it by using $Z = \dfrac{X-\mu}{\sigma}$, and use the area probability normal table. (given in the Appendix).

For Example: The probability that the normal random variable x to lie in the interval $(\mu - \sigma , \mu + \sigma)$ is given by



$P(\mu - \sigma < x < \mu + \sigma) \quad = P(-1 \leq z \leq 1)$

$= 2P(0 < z < 1)$

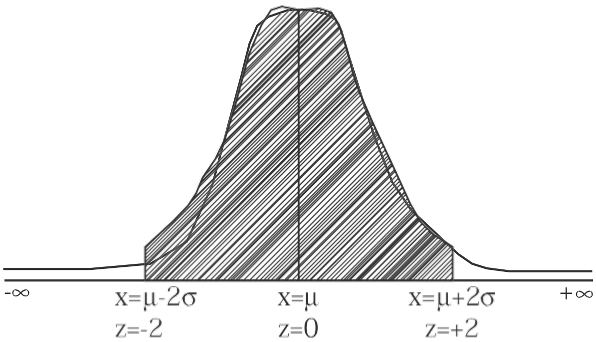
$= 2\,(0.3413)$ (from the area table)

$= 0.6826$

$P(\mu - 2\sigma < x < \mu + 2\sigma) \quad = P(-2 < z < 2)$

$= 2P(0 < z < 2)$

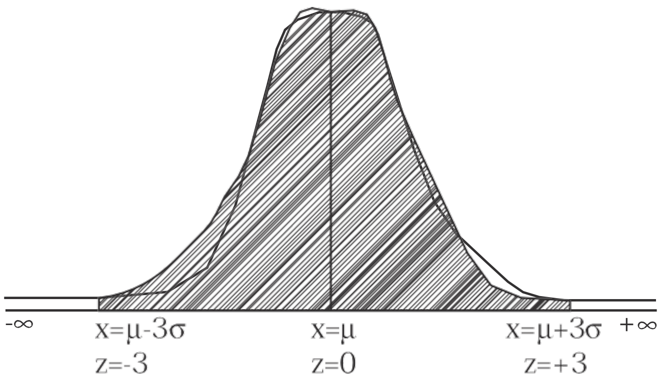
$= 2\,(0.4772) = 0.9544$



$P(\mu - 3\sigma < x < \mu + 3\sigma) = P(-3 < z < 3)$

$= 2P(0 < z < 3)$

$= 2\,(0.49865) = 0.9973$



The probability that a normal variate x lies outside the range $\mu \pm 3\sigma$ is given by

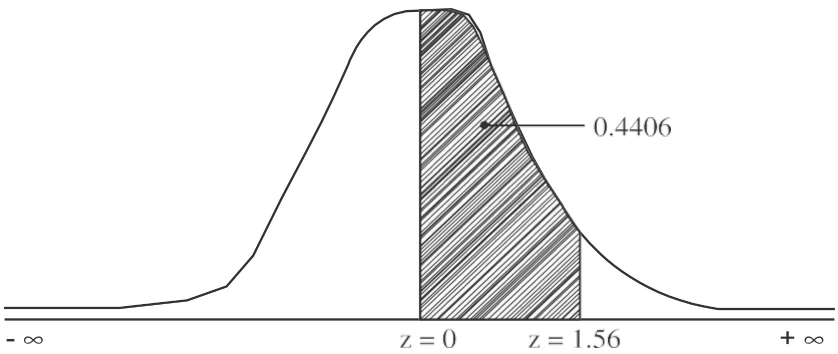
$$P(|x - \mu| > 3\sigma) = P(|z| > 3)$$
$$= 1 - P(-3 \leq z \leq 3)$$
$$= 1 - 0.9773 = 0.0027$$

Thus we expect that the values in a normal probability curve will lie between the range $\mu \pm 3\sigma$, though theoretically it range from $-\infty$ to ∞.

Example:

Find the probability that the standard normal variate lies between 0 and 1.56

Solution:



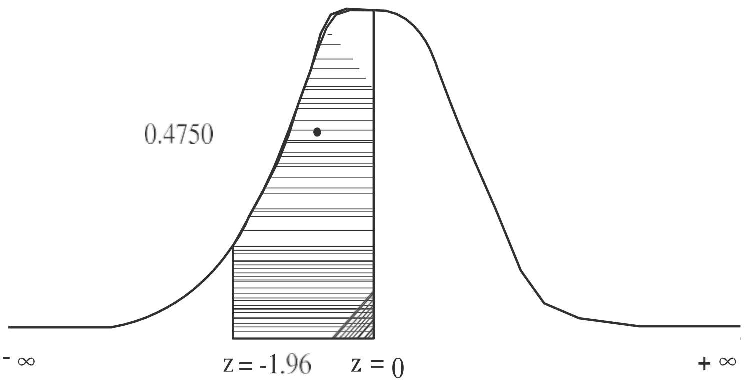
$P(0 < z < 1.56)$ = Area between $z = 0$ and $z = 1.56$

$= 0.4406$ (from table)

Example:

Find the area of the standard normal variate from -1.96 to 0.

Solution:



Area between $z = 0$ & $z = 1.96$ is same as the area $z = -1.96$ to $z = 0$

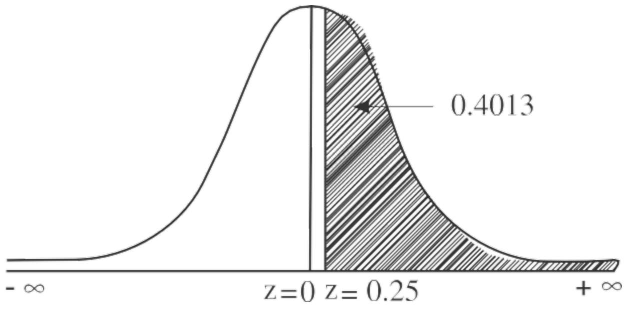
$P(-1.96 < z < 0) = P(0 < z < 1.96)$ (by symmetry)

$= 0.4750$ (from the table)

Example:

Find the area to the right of $z = 0.25$

Solution:



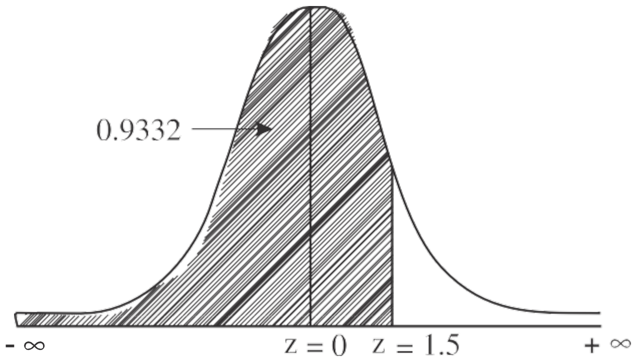
$$P(z > 0.25) = P(0 < z < \infty) - P(0 < z < 0.25)$$

$$= 0.5000 - 0.0987 \text{ (from the table)} = 0.4013$$

Example:

Find the area to the left of z = 1.5

Solution:



$$P(z < 1.5) = P(-\infty < z < 0) + P(0 < z < 1.5)$$

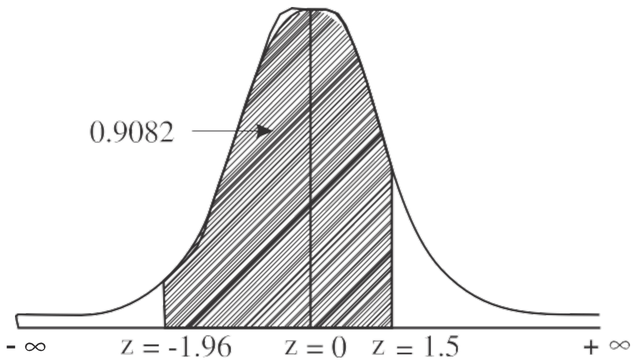
$$= 0.5 + 0.4332 \text{ (from the table)}$$

$$= 0.9332$$

Example:

Find the area of the standard normal variate between –1.96 and 1.5

Solution:



$$P(-1.96 < z < 1.5) = P(-1.96 < z < 0) + P(0 < z < 1.5)$$

$$= P(0 < z < 1.96) + P(0 < z < 1.5)$$

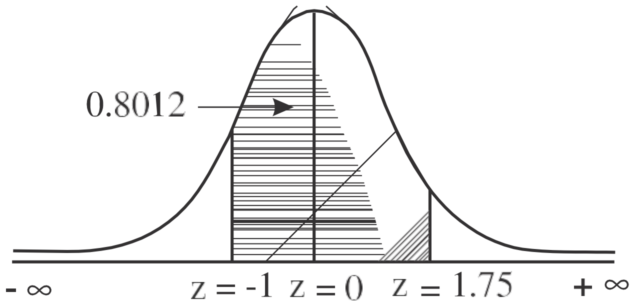
$$= 0.4750 + 0.4332 \text{ (from the table)}$$

$$= 0.9082$$

Example :

Given a normal distribution with $\mu = 50$ and $\sigma = 8$, find the probability that x assumes a value between 42 and 64.

Solution:



Given that $\mu = 50$ and $\sigma = 8$

The standard normal variate $z = \frac{X-\mu}{\sigma}$

If $X = 42$, $Z_1 = \frac{42-50}{8} = \frac{-8}{8} = -1$

If $X = 64$, $Z_2 = \frac{64-50}{8} = \frac{14}{8} = 1.75$

$\therefore$ $P(42 < x < 64) = P(-1 < z < 1.75)$

$= P(-1 < z < 0) + P(0 < z < 1.95)$

$= P(0 < z < 1) + P(0 < z < 1.75)$ (by symmetry)

$= 0.3413 + 0.4599$ (from the table)

$= 0.8012$

Example:

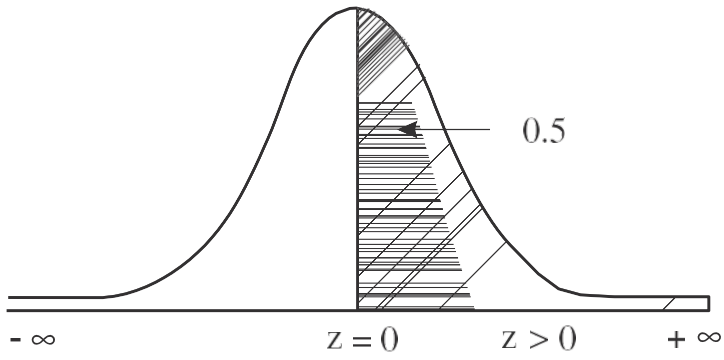
Students of a class were given an aptitude test. Their marks were found to be normally distributed with mean 60 and standard deviation 5. What percentage of students scored.

i) More than 60 marks (ii) Less than 56 marks (iii) Between 45 and 65 marks

Solution:

Given that mean = $\mu = 60$ and standard deviation = $\sigma = 5$

i) The standard normal varaiate $Z = \frac{X-\mu}{\sigma}$



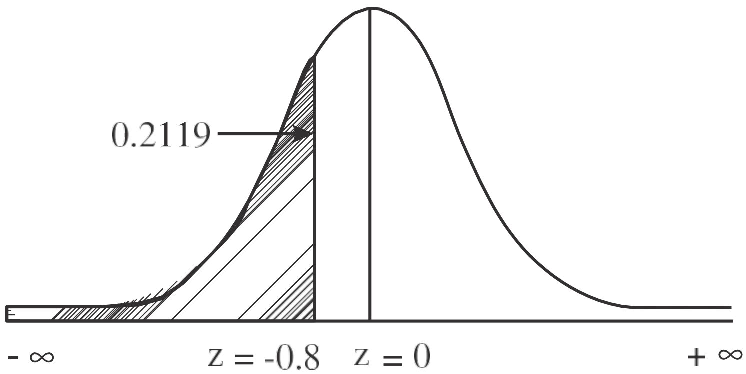
If $X = 60$, $Z = \dfrac{x-\mu}{\sigma} = \dfrac{60-60}{5} = 0$

$\therefore$ $P(x > 60) = P(z > 0)$

$= P(0 < z < \infty) = 0.5000$

Hence the percentage of students scored more than 60 marks is 0.5000(100) = 50 %

ii) If $X = 56$, $Z = \dfrac{56-60}{5} = \dfrac{-4}{5} = -0.8$



$P(x < 56) = P(z < -0.8)$

$= P(-\infty < z < 0) - P(-0.8 < z < 0)$ (by symmetry)

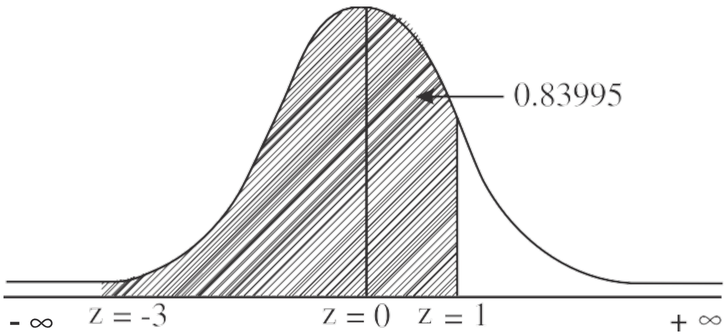
$= P(0 < z < \infty) - P(0 < z < 0.8)$

$= 0.5 - 0.2881$ (from the table)

$= 0.2119$

Hence the percentage of students score less than 56 marks is 0.2119(100) = 21.19 %

iii) If $X = 45$, $Z = \dfrac{45-60}{5} = \dfrac{-15}{5} = -3$



$X = 65$ then $z = \dfrac{65-60}{5} = \dfrac{5}{5} = 1$

$$
\begin{aligned}
P(45 < x < 65) &= P(-3 < z < 1) \\
&= P(-3 < z < 0) + P(0 < z < 1) \\
&= P(0 < z < 3) + P(0 < z < 1) \qquad \text{(by symmetry)} \\
&= 0.4986 + 0.3413 \qquad \text{(from the table)} \\
&= 0.8399
\end{aligned}
$$

Hence the percentage of students scored between 45 and 65 marks is 0.8399 (100) = 83.99 %

Example:

X is normal distribution with mean 2 and standard deviation 3. Find the value of the variable x such that the probability of the interval from mean to that value is 0.4115

Solution:

Given $\mu = 2$, $\sigma = 3$

Suppose z_1 is required standard value,

Thus $P(0 < z < z_1) = 0.4115$

From the table the value corresponding to the area 0.4115 is 1.35 that is $z_1 = 1.35$

$$
\begin{aligned}
\text{Here } z_1 &= \frac{x-\mu}{\sigma} \\
1.35 &= \frac{x-2}{3} \\
x &= 3(1.35) + 2 \\
&= 4.05 + 2 = 6.05
\end{aligned}
$$

Example:

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

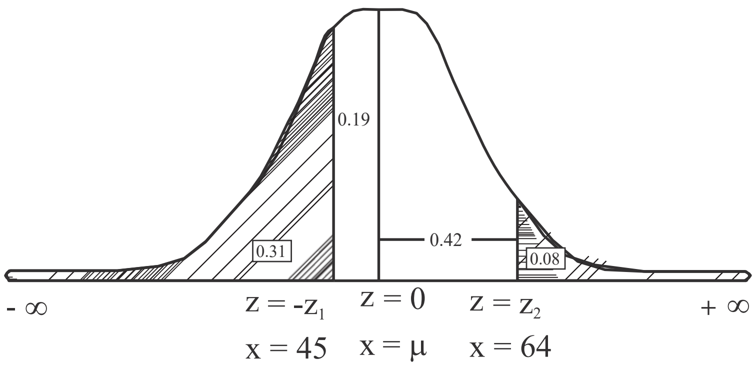
Solution:

Let x denotes the items are given and it follows the normal distribution with mean μ and standard deviation σ

The points x = 45 and x = 64 are located as shown in the figure.

i) Since 31 % of items are under x = 45, position of x into the left of the ordinate $x = \mu$

ii) Since 8 % of items are above x = 64 , position of this x is to the right of ordinate $x = \mu$



When $x = 45,\ z = \dfrac{x-\mu}{\sigma} = \dfrac{45-\mu}{\sigma} = -z_1$ (say)

Since x is left of $x = \mu$, z_1 is taken as negative

When $x = 64,\ z = \dfrac{64-\mu}{\sigma} = z_2$ (say)

From the diagram $P(x < 45) = 0.31$

$$P(z < -z_1) = 0.31$$

$$P(-z_1 < z < 0) = P(-\infty < z < 0) - p(-\infty < z < z_1)$$

$$= 0.5 - 0.31 = 0.19$$

$$P(0 < z < z_1) = 0.19 \text{ (by symmetry)}$$

$$z_1 = 0.50 \text{ (from the table)}$$

Also from the diagram $p(x > 64) = 0.08$

$$P(0 < z < z_2) = P(0 < z < \infty) - P(z_2 < z < \infty)$$

$$= 0.5 - 0.08 = 0.42$$

$$z_2 = 1.40 \text{ (from the table)}$$

Substituting the values of z_1 and z_2 we get

$$\frac{45-\mu}{\sigma} = -0.50 \text{ and } \frac{64-\mu}{\sigma} = 1.40$$

Solving $\mu - 0.50\,\sigma = 45$ ------(1)

$\mu + 1.40\,\sigma = 64$ ----- (2)

$(2) - (1) \Rightarrow 1.90\,\sigma = 19 \Rightarrow \sigma = 10$

Substituting $\sigma = 10$ in (1) $\quad \mu = 45 + 0.50\,(10)$

$= 45 + 5.0 = 50.0$

Hence mean = 50 and variance = $\sigma^2 = 100$

Exercise

I. Choose the best answer:

1. Binomial distribution applies to

   (a) rare events (b) repeated alternatives (c) three events (d) impossible events

2. For Bernoulli distribution with probability p of a success and q of a failure, the relation between mean and variance that hold is

   (a) mean < variance (b) mean > variance

   (c) mean = variance (d) mean ≤ variance

3. The variance of a binomial distribution is

   (a) npq (b) np (c) $\sqrt{npq}$ (d) 0

4. The mean of the binomial distribution $15C_x\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{15-x}$ in which $p = \frac{2}{3}$ is

   (a) 5 (b) 10 (c) 15 (d) 3

5. The mean and variance of a binomial distribution are 8 and 4 respectively. Then P(x = 1) is equal to

   (a) $\frac{1}{2^{12}}$ (b) $\frac{1}{2^4}$ (c) $\frac{1}{2^6}$ (d) $\frac{1}{2^8}$

6. The mean of a binomial distribution is 10 and the number of trials is 30 then probability of failure of an event is

(a) 0.25 (b) 0.333 (c) 0.666 (d) 0.9

7. The variance of a binomial distribution is 2. Its standard deviation is

(a) 2 (b) 4 (c) 1/2 (d) $\sqrt{2}$

8. In a binomial distribution if the numbers of independent trials is n, then the probability of n success is

(a) $nC_x p^x q^{n-x}$ (b) 1 (c) p^n (d) q^n

9. The binomial distribution is completely determined if it is known

(a) p only (b) q only (c) p and q (d) p and n

10. The trials in a binomial distribution are

(a) mutually exclusive (b) non-mutually exclusive

(c) independent (d) non-independent

11. If two independent variables x and y follow binomial distribution with parameters, (n_1, p) and (n_2, p) respectively, their sum $(x + y)$ follows binomial distribution with parameters

(a) $(n_1 + n_2, 2p)$ (b) (n, p) (c) $(n_1 + n_2, p)$ (d) $(n_1 + n_2, p + q)$

12. For a Poisson distribution

(a) mean > variance (b) mean = variance

(c) mean < variance (d) mean $\leq$ variance

13. Poisson distribution correspondents to

(a) rare events (b) certain event

(c) impossible event (d) almost sure event

14. If the Poisson variables X and Y have parameters m_1 and m_2 then X + Y is a Poisson variable with parameter.

(a) $m_1 m_2$ (b) $m_1 + m_2$ (c) $m_1 - m_2$ (d) m_1/m_2

15. Poisson distribution is a

(a) Continuous distribution (b) discrete distribution

(c) either continuous or discrete (d) neither continue nor discrete

16. Poisson distribution is a limiting case of Binomial distribution when

(a) $n \to \infty$; $p \to 0$ and $np = \sqrt{m}$ (b) $n \to 0$; $p \to \infty$ and $p=1/m$

(c) $n \to \infty$; $p \to \infty$ and $np = m$ (d) $n \to \infty$; $p \to 0$,$np = m$

17. If the expectation of a Poisson variable (mean) is 1 then $P(x < 1)$ is

(a) e^{-1} (b) $1-2e^{-1}$ (c) $1- 5/2e^{-1}$ (d) none of these

18. The normal distribution is a limiting form of Binomial distribution if

(a) $n \to \infty$ $p \to 0$ (b) $n \to 0$, $p \to q$

(c) $n \to \infty$, $p \to n$ (d) $n \to \infty$ and neither p nor q is small.

19. In normal distribution, skewness is

(a) one (b) zero (c) greater than one (d) less than one

20. Mode of the normal distribution is

(a) σ (b) $\frac{1}{\sqrt{2\pi}}$ (c) μ (d) 0

21. The standard normal distribution is represented by

(a) N (0,0) (b) N(1,1) (c) N(1,0) (d) N(0,1)

22. Total area under the normal probability curve is

(a) less than one (b) unity (c) greater than one (d) zero

23. The probability that a random variable x lies in the interval $(\mu - 2\sigma , \mu + 2\sigma)$ is

(a) 0.9544 (b) 0.6826 (c) 0.9973 (d) 0.0027

24. The area $P(- \infty < z < 0)$ is equal to

(a) 1 (b) 0.1 (c) 0.5 (d) 0

25. The standard normal distribution has

(a) $\mu =1, \sigma = 0$ (b) $\mu = 0, \sigma = 1$ (c) $\mu = 0, \sigma = 0$ (d) $\mu =1, \sigma = 1$

26. The random variable x follows the normal distribution $f(x) = C.e^{-\frac{1}{2}\frac{(x-100)^2}{25}}$ then the value of C is

(a) $5\sqrt{2\pi}$ (b) $\frac{1}{5\sqrt{2\pi}}$ (c) 1 (d) 5

$\sqrt{2\pi}$

27. Normal distribution has

(a) no mode (b) only one mode (c) two modes (d) many mode

28. For the normal distribution

(a) mean = median = mode (b) mean < median < mode

(c) mean > median > mode (d) mean > median < mode

29. Probability density function of normal variable $P(X = x) = \frac{1}{5\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-30)^2}{25}}$; $-\alpha < x < \alpha$ then mean and variance are

(a) mean = 30 variance = 5 (b) mean = 0, variance = 25

(c) mean = 30 variance = 25 (d) mean = 30, variance = 10

30. The mean of a Normal distribution is 60, its mode will be

(a) 60 (b) 40 (c) 50 (d) 30

31. If x is a normal variable with $\mu = 100$ and $\sigma^2 = 25$ then $P(90 < x < 120)$ is same as

(a) $P(-1 < z < 1)$ (b) $P(-2 < z < 4)$

(c) $P(4 < z < 4.1)$ (d) $P(-2 < z < 3)$

33. If x is N(6, 1.2) and $P(0 \le z \le 1) = 0.3413$ then $P(4.8 \le x \le 7.2)$ is

(a) 0.3413 (b) 0.6587 (c) 0.6826 (d) 0.3174

II. fill in the blanks:

34. The probability of getting a head in successive throws of a coin is _________.

35. If the mean of a binomial distribution is 4 and the variance is 2 then the parameter is __________.

36. $\left(\frac{2}{3}+\frac{1}{3}\right)^9$ refers the binomial distribution and its standard deviation is__________.

37. In a binomial distribution if number of trials to be large and probability of success bezero, then the distribution becomes_______.

38. The mean and variance are________in Poisson distribution.

39. The mean of Poisson distribution is 0.49 and its standard deviation is_________.

40. In Poisson distribution, the recurrence formula to calculate expected frequencies is _______.

41. The formula $\frac{\Sigma fx^2}{N} - (\bar{x})^2$ is used to find_________.

42. In a normal distribution, mean takes the values from___________to_________.

43. When $\mu = 0$ and $\sigma = 1$ the normal distribution is called_________.

44. $P(-\infty < z < 0)$ covers the area_______.

45. If $\mu = 1200$ and $\sigma = 400$ then the standard normal variate z for $x = 800$ is__________.

46. At $x = \mu \pm \sigma$ are called as___________in a normal distribution.

47. $P(-3 < z < 3)$ takes the value ___________

48. X axis be the_________to the normal curve.

III. Answer the following

49. Comment the following :

“ For a binomial distribution mean = 7 and variance = 16

50. Find the binomial distribution whose mean is 3 and variance 2.

51. In a binomial distribution the mean and standard deviation are 12 and 2 respectively. Findn and p.

52. A pair of dice is thrown 4 times. If getting a doublet is considered a success, find the probability of 2 success.

53. Explain a binomial distribution.

54. State the characteristics of a binomial distribution.

55. State the conditions for a binomial variate.

56. Explain the fitting of a binomial distribution.

57. For the binomial distribution $(0.68 + 0.32)^{10}$ find the probability of 2 success.

58. Find the mean of binomial distribution of the probability of occurrence of an event is

1/5and the total number of trials is 100.

59. If on an average 8 ships out of 10 arrive safely at a port, find the mean and standard deviation of the number of ships arriving safely out of total of 1600 ships.

60. The probability of the evening college student will be a graduate is 0.4. Determine theprobability that out of 5 students (i) none (ii) one (iii) atleast one will be a graduate.

61. Four coins are tossed simultaneously. What is the probability of getting i) 2 heads and 2tails ii) atleast 2 heads iii) atleast one head.

62. 10% of the screws manufactured by an automatic machine are found to be defective.20 screws are selected at random. Find the probability that i) exactly 2 are defective
ii) atmost 3 are defective iii) atleast 2 are defective.

63. 5 dice are thrown together 96 times. The numbers of getting 4, 5 or 6 in the experiment is given below. Calculate the expected frequencies and compare the standard deviation of the expected frequencies and observed frequencies.

Getting 4 ,5 or 6	:	0	1	2	3	4	5
Frequency	:	1	10	24	35	18	8

64. Fit a binomial distribution for the following data and find the expected frequencies.

X :	0	1	2	3	4
f	18	35	30	13	4

65. Eight coins are tossed together 256 times. Number of heads observed at each toss is recorded and the results are given below. Find the expected frequencies. What are the theoretical value of mean and standard deviation? Calculate also mean and standard deviation of the observed frequencies.

Number of heads:		0	1	2	3	4	5	6	7	8
Frequencies	:	2	6	39	52	67	56	32	10	1

66. Explain Poisson distribution.

67. Give any two examples of Poisson distribution.

68. State the characteristics of Poisson distribution.

69. Explain the fitting of a Poisson distribution

70. A variable x follows a Poisson distribution with mean 6 calculate i) $P(x = 0)$ ii) $P(x = 2)$

71. The variance of a Poisson Distribution is 0.5. Find $P(x = 3)$. [$e^{-0.5} = 0.6065$]

72. If a random variable X follows Poisson distribution such that P(x =1) = P(x = 2) find (a) the mean of the distribution and P(x = 0). [e^{-2} = 0.1353]

73. If 3% of bulbs manufactured by a company are defective then find the probability in a sample of 100 bulbs exactly five bulbs are defective.

74. It is known from the past experience that in a certain plant there are on the average 4 industrial accidents per month. Find the probability that in a given year there will be less than 3 accidents. Assume Poisson distribution.[e^{-4} = 0.0183]

75. A manufacturer of television sets known that of an average 5% of this product is defective. He sells television sets in consignment of 100 and guarantees that not more than 4 sets will be defective. What is the probability that a television set will fail to meet the guaranteed quality? [e^{-5} = 0.0067]

76. One fifth percent of the blades produced by a blade manufacturing factory turns out to be a defective. The blades are supplied in pockets of 10. Use Poisson distribution to calculate the approximate number of pockets containing i) no defective (ii) all defective (iii) two defective blades respectively in a consignment of 1,00,000 pockets.

77. A factory employing a huge number of workers find that over a period of time, average absentee rate is three workers per shift. Calculate the probability that in a given shift i) exactly 2 workers (ii) more than 4 workers will be absent.

78. A manufacturer who produces medicine bottles finds that 0.1 % of the bottles are defective. They are packed in boxes containing 500 bottles. A drag manufactures buy 100 boxes from the producer of bottles. Using Poisson distribution find how many boxes will contain (i) no defective ii) exactly 2 (iii) atleast 2 defective.

79. The distribution of typing mistakes committed by a typist is given below:

Mistakes per page :	0	1	2 3 4 5			
No. of pages :	142	156	69	57	5	1

Fit a Poisson distribution.

80. Fit a Poisson distribution to the following data:

x :	0	1	2	3	4	5	6	7	8	Total
f :	229	325	257	119	50	17	2	1	0	1000

81. The following tables given that number of days in a 50, days period during which automatically accidents occurred in city. Fit a Poisson distribution to the data

No of accidents :	0	1	2	3	4

No of days	:	21	18	7	3	1

82. Find the probability that standard normal variate lies between 0.78 and 2.75.

83. Find the area under the normal curve between $z = 0$ and $z = 1.75$.

84. Find the area under the normal curve between $z = -1.5$ and $z = 2.6$.

85. Find the area to the left side of $z = 1.96$

86. Find the area under the normal curve which lies to the right of $z = 2.70$.

87. A normal distribution has mean = 50 and standard deviation is 8. Find the probability that x assumes a value between 34 and 62.

88. A normal distribution has mean = 20 and S.D = 10. Find area between $x = 15$ and $x = 40$.

89. Given a normal curve with mean 30 and standard deviation 5. Find the area under the curve between 26 and 40

90. The customer accounts of a certain departmental store have an average balance of Rs.1200 and a standard deviation of Rs.400. Assuming that the account balances are normally distributed. (i) what percentage of the accounts is over Rs.1500? (ii) What percentage of the accounts is between Rs.1000 and Rs.1500? iii) What percentage of the accounts is below Rs.1500?

91. The weekly remuneration paid to 100 lecturers coaching for professional entrance examinations are normally distributed with mean Rs.700 and standard deviation Rs.50.

    Estimate the number of lecturers whose remuneration will be i) between Rs.700 andRs.720 ii) more than Rs.750 iii) less than Rs.630.

92. x is normally distributed with mean 12 and standard deviation 4. Find the probability of the following i) $x \geq 20$ ii) $x \leq 20$ iii) $0 < x < 12$

93. A sample of 100 dry cells tested to find the length of life produced the following results $\mu = 12$ hrs, $\sigma = 3$ hrs. Assuming the data, to be normally distributed. What percentage of battery cells are expressed to have a life. i) more than 15 hrs ii) between 10 and 14 hrs as iii) less than 6 hrs?.

94. Find the mean and standard deviation of marks in an examination where 44 % of thecandidates obtained marks below 55 and 6 % got above 80 marks.

95. In a normal distribution 7 % of the items are under 35 and 89 % of the items are under 63. Find its mean as standard deviation.

Note: For fitting a binomial distribution in the problem itself, if it is given that the coin is unbiased, male and female births are equally probable, then we consider $p = q = ½$. All other cases we have to find the value of p from the mean value of the given data.

Variance

9·3. RECTANGULAR (OR UNIFORM) DISTRIBUTION

Definition. *A random variable X is said to have a continuous rectangular (uniform) distribution over an interval (a, b), i.e., ($-\infty < a < b < \infty$), if its p.d.f. is given by :*

$$f(x\,;a,b) = \begin{cases} \dfrac{1}{b-a}, & \text{if } a < x < b \\ 0, & \text{otherwise} \end{cases} \qquad \ldots (9\cdot19)$$

Remarks 1. a and b, $(a < b)$ are the two parameters of the distribution. The distributin is called uniform distribution on (a, b) since it assumes a constant (uniform) value for all x in (a, b).

2. The distribution is also known as rectangular distribution, since the curve $y = f(x)$ describes a rectangle over the x-axis and between the ordinates at $x = a$ and $x = b$.

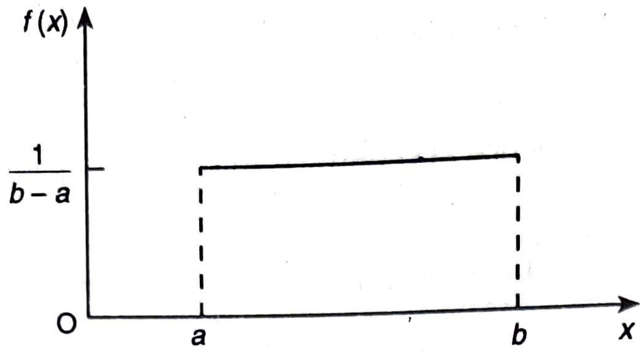
3. A uniform or rectangular variate X on the interval (a, b) is written as : $X \sim U[a, b]$ or $X \sim R[a, b]$.

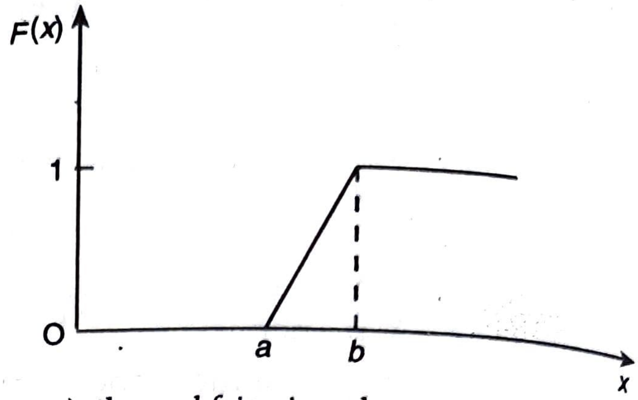
4. The cumulative distribution function $F(x)$ is given by :

$$F(x) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a < x < b \\ 1, & x \ge b \end{cases} \qquad \ldots (9\cdot19a)$$

Since $F(x)$ is not continuous at $x = a$ and $x = b$, it is not differentiable at these points. Thus $\dfrac{d}{dx}F(x) = f(x) = \dfrac{1}{b-a} \neq 0$, exists everywhere except at the points $x = a$ and $x = b$ and consequently $p.d.f.$ $f(x)$ is given by (9·19).

5. The graphs of uniform *p.d.f.* $f(x)$ and the corresponding distribution function $F(x)$ are given below.





6. For a rectangular or uniform variate X in $(-a, a)$, the *p.d.f.* is given by :

$$f(x) = \begin{cases} \dfrac{1}{2a}, & -a < x < a \\ 0, & \text{otherwise} \end{cases} \qquad \ldots (9{\cdot}19b)$$

9·3·1. Moments of Rectangular Distribution. Let $X \sim U[a, b]$.

$$\mu'_r = \int_a^b x^r f(x)\, dx = \frac{1}{b-a} \int_a^b x^r\, dx = \frac{1}{b-a}\left(\frac{b^{r+1} - a^{r+1}}{r+1}\right) \qquad \ldots (9{\cdot}20)$$

In particular

$$\text{Mean} = \mu_1' = \frac{1}{b-a}\left(\frac{b^2 - a^2}{2}\right) = \frac{b+a}{2} \qquad \ldots (9{\cdot}20a)$$

and

$$\mu_2' = \frac{1}{b-a}\left(\frac{b^3 - a^3}{3}\right) = \frac{1}{3}(b^2 + ab + a^2)$$

$$\therefore \quad \text{Variance} = \mu_2' - \mu_1'^2 = \frac{1}{3}(b^2 + ab + a^2) - \left\{\frac{1}{2}(b+a)\right\}^2 = \frac{1}{12}(b-a)^2 \qquad \ldots (9{\cdot}20b)$$

9·3·2. M.G.F. of Rectangular Distribution is given by :

$$M_X(t) = \int_a^b e^{tx} f(x)\, dx = \int_a^b \frac{e^{tx}}{b-a}\, dx = \frac{e^{bt} - e^{at}}{t(b-a)}, \; t \neq 0 \qquad \ldots (9{\cdot}20c)$$

9·3·3. Characteristic Function of Rectangular Distribution is given by :

$$\phi_X(t) = \int_a^b e^{itx}\, dx = \frac{e^{ibt} - e^{iat}}{it(b-a)}, \; t \neq 0. \qquad \ldots (9{\cdot}20d)$$

9·3·4. Mean Deviation about Mean, η of Rectangular Distribution is given by :

$$\eta = E\left|X - \text{Mean}\right| = \int_a^b \left|x - \text{Mean}\right| f(x)\, dx$$

$$= \frac{1}{b-a}\int_a^b \left|x - \frac{a+b}{2}\right| dx = \frac{1}{b-a}\int_{-(b-a)/2}^{(b-a)/2} |t|\, dt, \text{ where } t = x - \frac{a+b}{2}$$

$$= \frac{1}{b-a} \cdot 2\int_0^{(b-a)/2} t\, dt = \frac{b-a}{4} \qquad \ldots (9{\cdot}20e)$$

Example 9·21. *If X is uniformly distributed with mean 1 and variance $\frac{4}{3}$, find $P(X < 0)$.*

Solution. Let $X \sim U[a, b]$, so that $p(x) = \frac{1}{b-a}, a < x < b$. We are given :

$$\text{Mean} = \frac{1}{2}(b+a) = 1 \Rightarrow b + a = 2 \text{ and } \text{Var}(X) = \frac{1}{12}(b-a)^2 = \frac{4}{3} \Rightarrow b - a = \pm 4.$$

Solving, we get $a = -1$ and $b = 3$; $(a < b)$. $\quad \therefore \; p(x) = \frac{1}{4}; -1 < x < 3$

$$P(X < 0) = \int_{-1}^{0} p(x)\, dx = \frac{1}{4} \left| x \right|_{-1}^{0} = \frac{1}{4}.$$

Example 9·22. *Subway trains on a certain line run every half hour between mid-night and six in the morning. What is the probability that a man entering the station at a random time during this period will have to wait at least twenty minutes ?*

Solution. Let the *r.v.* X denote the waiting time (in minutes) for the next train. Under the assumption that a man arrives at the station at random, X is distributed uniformly on (0, 30), with *p.d.f.*,

$$f(x) = \begin{cases} \frac{1}{30}, & 0 < x < 30 \\ 0, & \text{otherwise} \end{cases}$$

The probability that he has to wait at least 20 minutes is given by :

$$P(X \geq 20) = \int_{20}^{30} f(x)\, dx = \frac{1}{30}\int_{20}^{30} 1.\, dx = \frac{1}{30}(30 - 20) = \frac{1}{3}.$$

Example 9·23. *If X has a uniform distribution in [0, 1], find the distribution (p.d.f.) of $-2 \log X$. Identify the distribution also.*

Solution. Let $Y = -2 \log X$. Then the distribution function G of Y is given by :

$$G_Y(y) = P(Y \leq y) = P(-2 \log X \leq y) = P(\log X \geq -y/2) = P(X \geq e^{-y/2})$$

$$= 1 - P(X \leq e^{-y/2}) = 1 - \int_0^{e^{-y/2}} f(x)\, dx = 1 - \int_0^{e^{-y/2}} 1.\, dx = 1 - e^{-y/2} \quad \ldots (*)$$

$$\therefore \quad g_Y(y) = \frac{d}{dy} G(y) = \frac{1}{2} e^{-y/2}, 0 < y < \infty$$

[$\because$ as X ranges in (0, 1), $Y = -2 \log X$ ranges from 0 to ∞]

Remark. This example illustrates that if $X \sim U[0, 1]$, then $Y = -2 \log X$, has an exponential distribution with parameter $\theta = \frac{1}{2}$. [*c.f.* § 9·8] or $Y = -2 \log X$, has chi-square distribution with $n = 2$ degrees of freedom [*c.f.* Chapter 15].

Example 9·24. *Show that for rectangular distribution : $f(x) = \frac{1}{2a}, -a < x < a$, m.g.f. about origin is $\frac{1}{at}$ (sinh at). Also show that moments of even order are given by :*

$$\mu_{2n} = \frac{a^{2n}}{(2n+1)}.$$