CHAPTER – 14

THEORETICAL DISTRIBUTIONS

LEARNING OBJECTIVES

The Students will be introduced in this chapter to the techniques of developing discrete and continuous probability distributions and its applications.

14.1 INTRODUCTION

In chapter ten, it may be recalled, we discussed frequency distribution. In a similar manner, we may think of a probability distribution where just like distributing the total frequency to different class intervals, the total probability (i.e. one) is distributed to different mass points in case of a discrete random variable or to different class intervals in case of a continuous random variable. Such a probability distribution is known as Theoretical Probability Distribution, since such a distribution exists only in theory. We need study theoretical probability distribution for the following important factors:

(a) An observed frequency distribution, in many a case, may be regarded as a sample i.e. a representative part of a large, unknown, boundless universe or population and we may be interested to know the form of such a distribution. By fitting a theoretical probability distribution to an observed frequency distribution of, say, the lamps produced by a manufacturer, it may be possible for the manufacturer to specify the length of life of the lamps produced by him up to a reasonable degree of accuracy. By studying the effect of a particular type of missiles, it may be possible for our scientist to suggest the number of such missiles necessary to destroy an army position. By knowing the distribution of smokers, a social activist may warn the people of a locality about the nuisance of active and passive smoking and so on.

(b) Theoretical probability distribution may be profitably employed to make short term projections for the future.

(c) Statistical analysis is possible only on the basis of theoretical probability distribution. Setting confidence limits or testing statistical hypothesis about population parameter(s) is based on the probability distribution of the population under consideration.

A probability distribution also possesses all the characteristics of an observed distribution. We define population mean (μ), population median (μ_0), population mode (μ_0), population standard deviation (σ) etc. exactly same way we have done earlier. These characteristics are known as population parameters. Again a probability distribution may be either a discrete probability distribution or a Continuous probability distribution depending on the random variable under study. **Two important discrete probability distribution are (a) Binomial Distribution and (b) Poisson distribution.**

Some important continuous probability distributions are

(a) Normal Distribution

(b) Chi-square Distribution

(c) Students-Distribution

(d) F-Distribution

14.2 BINOMIAL DISTRIBUTION

One of the most important and frequently used discrete probability distribution is Binomial Distribution. It is derived from a particular type of random experiment known as Bernoulli process after the famous mathematician Bernoulli. Noting that a 'trial' is an attempt to produce a particular outcome which is neither certain nor impossible, the characteristics of Bernoulli trials are stated below:

(i) Each trial is associated with two mutually exclusive and exhaustive outcomes, the occurrence of one of which is known as a 'success' and as such its non occurrence as a 'failure'. As an example, when a coin is tossed, usually occurrence of a head is known as a success and its non–occurrence i.e. occurrence of a tail is known as a failure.

(ii) The trials are independent.

(iii) The probability of a success, usually denoted by p, and hence that of a failure, usually denoted by q = 1–p, remain unchanged throughout the process.

(iv) The number of trials is a finite, positive integer.

A discrete random variable x is defined to follow binomial distribution with parameters n and p, to be denoted by x ~ B (n, p), if the probability mass function of x is given by

$$f(x) = p(X = x) = {}^{n}c_{x}\, p^{x} q^{n-x} \quad \text{for } x = 0, 1, 2, \ldots, n$$

$$= 0, \text{ otherwise} \quad \ldots\ldots\ldots (14.1)$$

We may note the following important points in connection with binomial distribution:

(a) As n >0, p, q ≥ 0, it follows that f(x) ≥ 0 for every x

Also $\sum_x f(x) = f(0) + f(1) + f(2) + \ldots + f(n) = 1$ **………(14.2)**

(b) Binomial distribution is known as biparametric distribution as it is characterised by two parameters n and p. This means that if the values of n and p are known, then the distribution is known completely.

(c) The mean of the binomial distribution is given by $\mu = np$ **…. (14.3)**

(d) Depending on the values of the two parameters, binomial distribution may be unimodal or bi- modal. μ_0**, the mode of binomial distribution, is given by** μ_0 **= the largest integer contained in (n+1)p if (n+1)p is a non-integer = (n+1)p and (n+1)p - 1 if (n+1)p is an integer ….(14.4)**

(e) The variance of the binomial distribution is given by

$$\sigma^2 = npq \quad \ldots\ldots\ldots\ldots (14.5)$$

Since p and q are numerically less than or equal to 1, npq < np

$\Rightarrow$ **variance of a binomial variable is always less than its mean.**

Also variance of X attains its maximum value at p = q = 0.5 and this maximum value is n/4.

(f) Additive property of binomial distribution.

If X and y are two independent variables such that

$X \sim \beta\ (n_1, P)$

and $y \sim \beta\ (n_2, P)$

Then $(X+y) \sim \beta\ (n_1 + n_2 +, P)$ .... (14.6)

Applications of Binomial Distribution

Binomial distribution is applicable when the trials are independent and each trial has just two outcomes success and failure. It is applied in coin tossing experiments, sampling inspection plan, genetic experiments and so on.

Example 14.1: A coin is tossed 8 times. Assuming the coin to be unbiased, what is the probability of getting?

(i) 4 heads

(ii) at least 4 heads

(iii) at most 3 heads

Solution: We apply binomial distribution as the tossing are independent of each other. With every tossing, there are just two outcomes either a head, which we call a success or a tail, which we call a failure and the probability of a success (or failure) remains constant throughout.

Let X denotes the no. of heads. Then X follows binomial distribution with parameter n = 8 and p = 1/2 (since the coin is unbiased). Hence q = 1 – p = 1/2

The probability mass function of X is given by

$f(x) = {}^{n}c_{x}\ p^{x}\ q^{n-x}$

$= {}^{10}c_{x} \cdot (1/2)^{x} \cdot (1/2)^{10-x}$

$$= \frac{{}^{10}c_{x}}{2^{10}}$$

$= {}^{10}c_{x} / 1024$ for x = 0, 1, 2, ...10

(i) probability of getting 4 heads

$= f(4)$

$= {}^{10}c_{4} / 1024$

$= 210 / 1024$

$= 105 / 512$

(ii) probability of getting at least 4 heads

$= P (X \geq 4)$

$= P (X = 4) + P (X = 5) + P (X = 6) + P(X = 7) + P (X = 8)$

$= {}^{10}c_4 / 1024 + {}^{10}c_5 / 1024 + {}^{10}c_6 / 1024 + {}^{10}c_7 / 1024 + {}^{10}c_8 /1024$

$$= \frac{210 + 252 + 210 + 120 + 45}{1024}$$

$= 837 / 1024$

(iii) probability of getting at most 3 heads

$= P (X \leq 3)$

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

$= f (0) + f (1) + f (2) + f (3)$

$= {}^{10}c_0 / 1024 + {}^{10}c_1 / 1024 + {}^{10}c_2 / 1024 + {}^{10}c_3 / 1024$

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

$= 176 / 1024$

$= 11/64$

Example 14.2 : If 15 dates are selected at random, what is the probability of getting two Sundays?

Solution: If X denotes the number at Sundays, then it is obvious that X follows binomial distribution with parameter n = 15 and p = probability of a Sunday in a week = 1/7 and q = 1 – p = 6 / 7.

Then $f(x) = {}^{15}c_x (1/7)^x . (6/7)^{15-x}$.

for x = 0, 1, 2,……….. 15.

Hence the probability of getting two Sundays

$= f(2)$

$= {}^{15}c_2 (1/7)^2 . (6/7)^{15-2}$

$$= \frac{105 \times 6^{13}}{7^{15}}$$

$\cong 0.29$

Example 14.3 : The incidence of occupational disease in an industry is such that the workmen have a 10% chance of suffering from it. What is the probability that out of 5 workmen, 3 or more will contract the disease?

Solution: Let X denote the number of workmen in the sample. X follows binomial with parameters n = 5 and p = probability that a workman suffers from the occupational disease = 0.1

Hence q = 1 – 0.1 = 0.9.

Thus $f(x) = {}^5c_x \, (0.1)^x . (0.9)^{5-x}$

For x = 0, 1, 2,…….,5.

The probability that 3 or more workmen will contract the disease

$= P(x \geq 3)$

$= f(3) + f(4) + f(5)$

$= {}^5c_3 (0.1)^3 (0.9)^{5-3} + {}^5c_4 (0.1)^4 . (0.9)^{5-4} + {}^5c_5 (0.1)^5$

$= 10 \times 0.001 \times 0.81 + 5 \times 0.0001 \times 0.9 + 1 \times 0.00001$

$= 0.0081 + 0.00045 + 0.00001$

$\cong 0.0086.$

Example 14.4 : Find the probability of a success for the binomial distribution satisfying the following relation 4 P (x = 4) = P (x = 2) and having the other parameter as six.

Solution : We are given that n = 6. The probability mass function of x is given by

$f(x) = {}^nc_x \, p^x \, q^{n-x}$

$= {}^6c_x \, p^x \, q^{n-x}$

for x = 0, 1, …… ,6.

Thus P (x = 4) = f (4):

$= {}^6c_4 \, p^4 \, q^{6-4}$

$= 15 \, p^4 \, q^2$

and P (x = 2) = f (2)

$= {}^6c_2 \, p^2 \, q^{6-2}$

$= 15p^2 \, q^4$

Hence 4 P (x = 4) = P (x = 2)

$\Rightarrow \quad 60 \, p^4 q^2 = 15 \, p^2 q^4$

$\Rightarrow \quad 15 \, p^2 q^2 (4p^2 - q^2) = 0$

$\Rightarrow \quad 4p^2 - q^2 = 0$ (as $p \neq 0, q \neq 0$)

$\Rightarrow \quad 4p^2 - (1 - p)^2 = 0$ (as $q = 1 - p$)

$\Rightarrow \quad (2p + 1 - p) = 0$ or $(2p - 1 + p) = 0$

$\Rightarrow \quad p = -1$ or $p = 1/3$

Thus p = 1/3 (as $p \neq -1$)

Example 14.5 : Find the binomial distribution for which mean and standard deviation are 6 and 2 respectively.

Solution : Let x ~ B (n, p)

Given that mean of x = np = 6 … (1)

and SD of x = 2

$\Rightarrow$ variance of x = npq = 4 ….. (2)

Dividing (2) by (1), we get $q = \frac{2}{3}$

Hence $p = 1 - q = \frac{1}{3}$

Replacing p by $\frac{1}{3}$ in equation (1), we get $n \times \frac{1}{3} = 6$

$\Rightarrow n = 18$

Thus the probability mass function of x is given by

$f(x) = {}^{n}c_{x}\, p^{x}\, q^{\,n-x}$

$= {}^{18}c_{x}\, (1/3)^{x} \cdot (2/3)^{18-x}$

for x = 0, 1, 2,…… ,18

Example 14.6 : Fit a binomial distribution to the following data:

x:	0	1	2	3	4	5
f:	3	6	10	8	3	2

Solution: In order to fit a theoretical probability distribution to an observed frequency distribution it is necessary to estimate the parameters of the probability distribution. There are several methods of estimating population parameters. One rather, convenient method is 'Method of Moments'. This comprises equating p moments of a probability distribution to p moments of the observed frequency distribution, where p is the number of parameters to be estimated. Since n = 5 is given, we need estimate only one parameter p. We equate the first moment about origin i.e. AM of the probability distribution to the AM of the given distribution and estimate p.

i.e. $n\hat{p} = \bar{x}$

$\Rightarrow \hat{p} = \frac{\bar{x}}{n}$ ($\hat{p}$ is read as p hat)

The fitted binomial distribution is then given by

$f(x) = {}^{n}c_{x}\, \hat{p}^{\,x} (1 - \hat{p})^{n-x}$

For x = 0, 1, 2, …… n

On the basis of the given data, we have

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

$$= \frac{3\times0+6\times1+10\times2+8\times3+3\times4+2\times5}{3+6+10+8+3+2} = 2.25$$

Thus $\hat{p} = \bar{x}/n = \frac{2.25}{n} = 0.45$

and $\hat{q} = 1 - \hat{p} = 0.55$

The fitted binomial distribution is

$f(x) = {}^5c_x\ (0.45)^x\ (0.55)^{5-x}$

For x = 0, 1, 2, 3, 4, 5.

Table 14.1

Fitting Binomial Distribution to an Observed Distribution

X	f (x)	Expected frequency	Observed frequency
	$= {}^5c_x\ (0.4)^x\ (0.6)^{5-x}$	Nf (x) = 32 f (x)	
0	0.07776	2.49 ≅ 3	3
1	0.25920	8.29 ≅ 8	6
2	0.34560	11.06 ≅ 11	10
3	0.23040	7.37 ≅ 7	8
4	0.07680	2.46 ≅ 3	3
5	0.01024	0.33 ≅ 0	2
Total	1.000 00	32	32

A look at table 14.1 suggests that the fitting of binomial distribution to the given frequency distribution is satisfactory.

Example 14.7 : 6 coin are tossed 512 times. Find the expected frequencies of heads. Also, compute the mean and SD of the number of heads.

Solution : If x denotes the number of heads, then x follows binomial distribution with parameters

n = 6 and p = prob. of a head = ½, assuming the coins to be unbiased. The probability mass function of x is given by

$f(x) = {}^6c_x\ (1/2)^x.\ (1/2)^{6-x}$

$= {}^6c_x/2^6$

for x = 0, 1, …..6.

The expected frequencies are given by Nf (x).

TABLE 14.2

Finding Expected Frequencies when 6 coins are tossed 512 times

x	f (x)	Nf (x) Expected frequency	x f (x)	x^2f (x)
0	1/64	8	0	0
1	6/64	48	6/64	6/64
2	15/64	120	30/64	60/64
3	20/64	160	60/64	180/64
4	15/64	120	60/64	240/64
5	6/64	48	30/64	150/64
6	1/64	8	6/64	36/64
Total	1	512	3	10.50

Thus mean = $\mu = \sum_x xf(x) = 3$

$$E(x^2) = \sum_x x^2 f(x) = 10.50$$

$$\text{Thus } \sigma^2 = \sum_x x^2 f(x) - \mu^2$$

$$= 10.50 - 3^2 = 1.50$$

$$\therefore SD = \sigma = \sqrt{1.50} \cong 1.22$$

Applying formula for mean and SD, we get

$\mu = np = 6 \times 1/2 = 3$

and $\sigma = \sqrt{npq} = \sqrt{6 \times 1/2 \times 1/2} = \sqrt{1.50} \cong 1.22$

Example 14.8 : An experiment succeeds thrice as after it fails. If the experiment is repeated 5 times, what is the probability of having no success at all ?

Solution: Denoting the probability of a success and failure by p and q respectively, we have,

$p = 3q$

$\Rightarrow p = 3(1 - p)$

$\Rightarrow p = 3/4$

$\therefore q = 1 - p = 1/4$

when n = 5 and p = 3/4**,** we have

$f(x) = {}^5c_x\ (3/4)^x\ (1/4)^{5-x}$

for n = 0, 1, ..., 5.

So probability of having no success

$= f(0)$

$= {}^5c_0\ (3/4)^0\ (1/4)^{5-0}$

$= 1/1024$

Example 14.9 : What is the mode of the distribution for which mean and SD are 10 and $\sqrt{5}$ respectively.

Solution: As given np = 10 .... (1)

and $\sqrt{npq} = \sqrt{5}$

$\Rightarrow$ npq = 5 .... (2)

on solving (1) and (2), we get n = 20 and p = 1/2

Hence mode = Largest integer contained in (n+1)p

= Largest integer contained in $(20+1) \times 1/2$

= Largest integer contained in 10.50

= 10.

Example 14.10 : If x and y are 2 independent binomial variables with parameters 6 and 1/2 and 4 and 1/2 respectively, what is $P(x + y \geq 1)$?

Solution: Let z = x + y.

It follows that z also follows binomial distribution with parameters

(6 + 4) and 1/2

i.e. 10 and 1/2

Hence $P(z \geq 1)$

$= 1 - P(z < 1)$

$= 1 - P(z = 0)$

$= 1 - {}^{10}c_0\ (1/2)^0 . (1/2)^{10-0}$

$= 1 - 1/2^{10}$

$= 1023 / 1024$

14.3 POISSON DISTRIBUTION

Poisson distribution is a theoretical discrete probability distribution which can describe many processes. Simon Denis Poisson of France introduced this distribution way back in the year 1837.

Poisson Model

Let us think of a random experiment under the following conditions:

I. The probability of finding success in a very small time interval (t, t + dt) is kt, where k (>0) is a constant.

II. The probability of having more than one success in this time interval is very low.

III. The probability of having success in this time interval is independent of t as well as earlier successes.

The above model is known as Poisson Model. The probability of getting x successes in a relatively long time interval T containing m small time intervals t i.e. T = mt. is given by

$$\frac{e^{-kt}.(kt)^x}{x!}$$

for x = 0, 1, 2, ...∞ ...... (14.7)

Taking kT = m, the above form is reduced to

$$\frac{e^{-m}.m^x}{x!}$$

for x = 0, 1, 2, ... ∞ .... (14.8)

Definition of Poisson Distribution

A random variable X is defined to follow Poisson distribution with parameter λ, to be denoted by X ~ P (λ) if the probability mass function of x is given by

$$f(x) = P(X = x) = \frac{e^{-m}.m^x}{x!} \text{ for } x = 0, 1, 2, \ldots \infty$$

$$= 0 \quad \text{otherwise} \qquad \ldots\ldots (14.9)$$

Here e is a transcendental quantity with an approximate value as 2.71828.

It is wiser to remember the following important points in connection with Poisson distribution:

(i) Since $e^{-m} = 1/e^m > 0$, whatever may be the value of m, m > 0, it follows that f (x) ≥ 0 for every x.

Also it can be established that $\sum_x f(x) = 1$ i.e. f(0) + f(1) + f(2) +....... = 1.... (14.10)

(ii) Poisson distribution is known as a uniparametric distribution as it is characterised by only one parameter m.

(iii) The mean of Poisson distribution is given by m i,e μ = m. (14.11)

(iv) The variance of Poisson distribution is given by σ^2 = m (14.12)

(v) Like binomial distribution, Poisson distribution could be also unimodal or bimodal depending upon the value of the parameter m.

We have μ_0 = The largest integer contained in m if m is a non-integer

= m and m–1 if m is an integer .... (14.13)

(vi) Poisson approximation to Binomial distribution

If n, the number of independent trials of a binomial distribution, tends to infinity and p, the probability of a success, tends to zero, so that m = np remains finite, then a binomial distribution with parameters n and p can be approximated by Poisson distribution with parameter m (= np).

In other words when n is rather large and p is rather small so that m = np is moderate then

β (n, p) $\cong$ P (m). .... (14.14)

(vii) Additive property of Poisson distribution

If X and y are two independent variables following Poisson distribution with parameters m_1 and m_2 respectively, then z = X + y also follows Poisson distribution with parameter $(m_1 + m_2)$.

i.e. if x ~ p (m_1)

and y ~ p (m_2)

and X and y are independent, then

z = X + y ~ p $(m_1 + m_2)$ .... (14.15)

Application of Poisson distribution

Poisson distribution is applied when the total number of events is pretty large but the probability of occurrence is very small. Thus we can apply Poisson distribution, rather profitably, for the following cases:

a) The distribution of the no. of printing mistakes per page of a large book.
b) The distribution of the no. of road accidents on a busy road per minute.
c) The distribution of the no. of radio-active elements per minute in a fusion process.
d) The distribution of the no. of demands per minute for health centre and so on.

Example 14.11 : Find the mean and standard deviation of x where x is a Poisson variate satisfying the condition P (x = 2) = P (x = 3).

Solution: Let x be a Poisson variate with parameter m. The probability max function of x is then given by

$$f(x) = \frac{e^{-m} . m^x}{x!} \qquad \text{for } x = 0, 1, 2, \ldots\ldots \infty$$

now, P (x = 2) = P (x = 3)

$\Rightarrow$ f(2) = f(3)

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

$$\Rightarrow \frac{e^{-m}.m^2}{2}(1 - m/3) = 0$$

$\Rightarrow$ 1 – m / 3 = 0 (as $e^{-m} > 0$, m > 0)

$\Rightarrow$ m = 3

Thus the mean of this distribution is m = 3 and standard deviation = $\sqrt{3} \cong 1.73$.

Example 14.12 : The probability that a random variable x following Poisson distribution would assume a positive value is $(1 - e^{-2.7})$. What is the mode of the distribution?

Solution : If x ~ P (m), then its probability mass function is given by

$$f(x) = \frac{e^{-m}.m^2}{x!} \quad \text{for } x = 0, 1, 2, ... \infty$$

The probability that x assumes a positive value

= P (x > 0)

= 1– P (x ≤ 0)

= 1 – P (x = 0)

= 1 – f(0)

= $1 - e^{-m}$

As given,

$1 - e^{-m} = 1 - e^{-2.7}$

$\Rightarrow e^{-m} = e^{-2.7}$

$\Rightarrow$ m = 2.7

Thus μ_0 = largest integer contained in 2.7

= 2

Example 14.13 : The standard deviation of a Poisson variate is 1.732. What is the probability that the variate lies between –2.3 to 3.68?

Solution: Let x be a Poisson variate with parameter m.

Then SD of x is $\sqrt{m}$.

As given $\sqrt{m} = 1.732$

$\Rightarrow m = (1.732)^2 \cong 3$.

The probability that x lies between –2.3 and 3.68

$= P(-2.3 < x < 3.68)$

$= f(0) + f(1) + f(2) + f(3)$ (As x can assume 0, 1, 2, 3, 4 .....)

$$= \frac{e^{-3}.3^0}{0!} + \frac{e^{-3}.3^1}{1!} + \frac{e^{-3}.3^2}{2!} + \frac{e^{-3}.3^3}{3!}$$

$= e^{-3} (1 + 3 + 9/2 + 27/6)$

$= 13e^{-3}$

$$= \frac{13}{e^3}$$

$$= \frac{13}{(2.71828)^3} \text{ (as e = 2.71828)}$$

$\cong 0.65$

Example 14.14 : X is a Poisson variate satisfying the following relation:

$P(X = 2) = 9P(X = 4) + 90P(X = 6).$

What is the standard deviation of X?

Solution: Let X be a Poisson variate with parameter m. Then the probability mass function of X is

$$P(X = x) = f(x) = \frac{e^{-m}.m^x}{x!} \text{ for } x = 0, 1, 2, \ldots\ldots \infty$$

Thus $P(X = 2) = 9P(X = 4) + 90P(X = 6)$

$\Rightarrow f(2) = 9 f(4) + 90 f(6)$

$$\Rightarrow \frac{e^{-m} m^2}{2!} = \frac{9e^{-m}.m^4}{4!} + \frac{90.e^{-m} m^6}{6!}$$

$$\Rightarrow \frac{e^{-m} m^2}{2}\left(\frac{90m^4}{360} + \frac{9m^2}{12} - 1\right) = 0$$

$$\Rightarrow \frac{e^{-m} m^2}{8}(m^4 + 3m^2 - 4) = 0$$

$\Rightarrow e^{-m}.m^2 (m^2 + 4)(m^2 - 1) = 0$

$\Rightarrow m^2 - 1 = 0$ (as $e^{-m} > 0$ $m > 0$ and $m^2 + 4 \neq 0$)

$\Rightarrow m = 1$ (as $m > 0$, $m \neq -1$)

Thus the standard deviation of X is $\sqrt{1} = 1$

Example 14.15 : Between 9 and 10 AM, the average number of phone calls per minute coming into the switchboard of a company is 4. Find the probability that during one particular minute, there will be,

1. no phone calls
2. at most 3 phone calls (given $e^{-4} = 0.018316$)

Solution: Let X be the number of phone calls per minute coming into the switchboard of the company. We assume that X follows Poisson distribution with parameters m = average number of phone calls per minute = 4.

1. The probability that there will be no phone call during a particular minute

$= P(X = 0)$

$= \frac{e^{-4}.4^0}{0!}$

$= e^{-4}$

$= 0.018316$

2. The probability that there will be at most 3 phone calls

$= P(X \leq 3)$

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

$= \frac{e^{-4}.4^0}{0!} + \frac{e^{-4}.4^1}{1!} + \frac{e^{-4}.4^2}{2!} + \frac{e^{-4}.4^3}{3!}$

$= e^{-4}(1 + 4 + 16/2 + 64/6)$

$= e^{-4} \times 71/3$

$= 0.018316 \times 71/3$

$\cong 0.43$

Example 14.16 : If 2 per cent of electric bulbs manufactured by a company are known to be defectives, what is the probability that a sample of 150 electric bulbs taken from the production process of that company would contain

1. exactly one defective bulb?
2. more than 2 defective bulbs?

Solution: Let x be the number of bulbs produced by the company. Since the bulbs could be either defective or non-defective and the probability of bulb being defective remains the same, it follows that x is a binomial variate with parameters n = 150 and p = probability of a bulb being defective = 0.02. However since n is large and p is very small, we can approximate this binomial distribution with Poisson distribution with parameter m = np = 150 x 0.02 = 3.

1. The probability that exactly one bulb would be defective

$= P (X = 1)$

$= \frac{e^{-3}.3^1}{1!}$

$= e^{-3} \times 3$

$= \frac{3}{e^3}$

$= 3/(2.71828)^3$

$\cong 0.15$

(ii) The probability that there would be more than 2 defective bulbs

$= P (X > 2)$

$= 1 - P (X \leq 2)$

$= 1 - [f (0) + f (1) + f (2)]$

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

$= 1 - 8.5 \times e^{-3}$

$= 1 - 0.4232$

$= 0.5768 \cong 0.58$

Example 14.17 : The manufacturer of a certain electronic component is certain that two per cent of his product is defective. He sells the components in boxes of 120 and guarantees that not more than two per cent in any box will be defective. Find the probability that a box, selected at random, would fail to meet the guarantee? Given that $e^{-2.40} = 0.0907$.

Solution: Let x denote the number of electric components. Then x follows binomial distribution with n = 120 and p = probability of a component being defective = 0.02. As before since n is quite large and p is rather small, we approximate the binomial distribution with parameters n and p by a Poisson distribution with parameter m = n.p = 120 × 0.02 = 2.40. Probability that a box, selected at random, would fail to meet the specification = probability that a sample of 120 items would contain more than 2.40 defective items.

$= P (X > 2.40)$

$= 1 - P (X \leq 2.40)$

$= 1 - [P (X = 0) + P (X = 1) + P (X = 2)]$

$= 1 - [e^{-2.40} + e^{-2.40} \times 2.4 + e^{-2.40} \times \left(\frac{2.40}{2}\right)^2]$

$$= 1 - e^{-2.40}\left(1 + 2.40 + \frac{(2.40)^2}{2}\right)$$

$$= 1 - 0.0907 \times 6.28$$

$$\cong 0.43$$

Example 14.18 : A discrete random variable x follows Poisson distribution. Find the values of

(i) P (X = at least 1)

(ii) $P (X \leq 2/ X \geq 1)$

You are given E (x) = 2.20 and $e^{-2.20} = 0.1108$.

Solution: Since X follows Poisson distribution, its probability mass function is given by

$$f(x) = \frac{e^{-m}.m^x}{x!} \text{ for } x = 0, 1, 2, \ldots\ldots \infty$$

(i) P (X = at least 1)

$= P (X \geq 1)$

$= 1 - P (X < 1)$

$= 1 - P (X = 0)$

$= 1 - e^{-m}$

$= 1 - e^{-2.20}$ (as E (x) = m = 2.20, given)

$= 1 - 0.1108$ (as $e^{-2.20} = 0.1108$ as given)

$\cong 0.89$.

(ii) $P (x \leq 2 / x \geq 1)$

$$= P\frac{[(X \leq 2) \cap (X \geq 1)]}{P(X \geq 1)} \qquad \left(\text{as } P(A/B) = P\frac{(A \cap B)}{P(B)}\right)$$

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

$$= \frac{f(1) + f(2)}{1 - f(0)}$$

$$= \frac{e^{-m}.m + e^{-m}.m^2/2}{1 - e^{-m}}$$

$$= \frac{e^{-2.20} \times 2.2 + e^{-2.20} \times (2.20)^2/2}{1 - e^{-2.20}} \qquad (\because m = 2.2)$$

$$= \frac{0.5119}{0.8892}$$

$$\cong 0.58$$

Fitting a Poisson distribution

As explained earlier, we can apply the method of moments to fit a Poisson distribution to an observed frequency distribution. Since Poisson distribution is uniparametric, we equate m, the parameter of Poisson distribution, to the arithmetic mean of the observed distribution and get the estimate of m.

i.e. $\hat{m} = \bar{x}$

The fitted Poisson distribution is then given by

$$\hat{f}(x) = \frac{e^{-\hat{m}} . (\hat{m})^x}{x!} \qquad \text{for } x = 0, 1, 2 \ldots\ldots\ldots\ldots\ldots\infty$$

Example 14.19: Fit a Poisson distribution to the following data :

Number of death:	0	1	2	3	4
Frequency:	122	46	23	8	1

(Given that $e^{-0.6} = 0.5488$)

Solution: The mean of the observed frequency distribution is

$$\bar{x} = \frac{\Sigma f_i x_i}{N}$$

$$= -\frac{122 \times 0 + 46 \times 1 + 23 \times 2 + 8 \times 3 + 1 \times 4}{122 + 46 + 23 + 8 + 1}$$

$$= \frac{120}{200}$$

$$= 0.6$$

Thus $\hat{m} = 0.6$

Hence $\hat{f}(0) = e^{-\hat{m}} = e^{-0.6} = 0.5488$

$$\hat{f}(1) = \frac{e^{-\hat{m}} \times m}{1!} = 0.6 \times e^{-0.6} = 0.3293$$

$$\frac{(0.6)^2}{2!} \times 0.5488 = 0.0988$$

$$\frac{(0.6)^3}{3!} \times 0.5488 = 0.0198$$

Lastly $P(X \geq 4) = 1 - P(X < 4)$.

Table 14.3

Fitting Poisson Distribution to an Observed Frequency Distribution of Deaths

X	f (x)	Expected frequency $N \times f(x)$	Observed frequency
0	0.5488	109.76 = 110	122
1	0.6 x 0.5488 = 0.3293	65.86 = 65	46
2	$(0.6)^2/2$ x 0.5488 = 0.0.0988	19.76 = 20	23
3	$(0.6)^3/3$ x 0.5488 = 0.0.0198	3.96 = 4	8
4 or more	0.0033 (By subtraction)	0.66 = 1	1
Total	1	200	200

14.4 NORMAL OR GAUSSIAN DISTRIBUTION

The two distributions discussed so far, namely binomial and Poisson, are applicable when the random variable is discrete. In case of a continuous random variable like height or weight, it is impossible to distribute the total probability among different mass points because between any two unequal values, there remains an infinite number of values. Thus a continuous random variable is defined in term of its probability density function f (x), provided, of course, such a function really exists f (x) satisfies the following condition:

$f(x) \geq 0$ for $x \in (\alpha, \beta)$

and $\int_{\alpha}^{\beta} f(x) = 1$ (α, β), $\beta > \alpha$, being the domain of the continuous variable x.

The most important and universally accepted continuous probability distribution is known as normal distribution. Though many mathematicians like De-Moivre, Laplace etc. contributed towards the development of normal distribution, Karl Gauss was instrumental for deriving normal distribution and as such normal distribution is also referred to as Gaussian Distribution.

A continuous random variable x is defined to follow normal distribution with parameters μ and σ^2, to be denoted by

$X \sim N(\mu, \sigma^2)$ ………………. (14.16)

If the probability density function of the random variable x is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-(\bar{x}-u)^2/2\sigma^2}$$

for $-\infty < x < \infty$ ……… (14.17)

Some important points relating to normal distribution are listed below:

(a) The name Normal Distribution has its origin some two hundred years back as the then mathematician were in search for a normal model that can describe the probability distribution of most of the continuous random variables.

(b) If we plot the probability function y = f (x), then the curve, known as probability curve, takes the following shape:

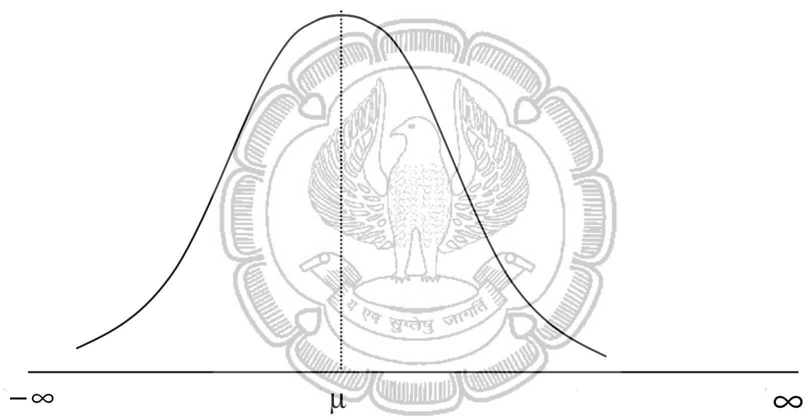

Figure 14.1

Showing Normal Probability Curve

A quick look at figure 14.1 reveals that the normal curve is bell shaped and has one peak, which implies that the normal distribution has one unique mode. The line drawn through x = μ has divided the normal curve into two parts which are equal in all respect. Such a curve is known as symmetrical curve and the corresponding distribution is known as Symmetrical distribution. Thus, we find that the normal distribution is symmetrical about x = μ. It may also be noted that the binomial distribution is also symmetrical about p = 0.5. We next note that the two tails of the normal curve extend indefinitely on both sides of the curve and both the left and right tails never touch the horizontal axis. The total area of the normal curve or for that any probability curve is taken to be unity i.e. one. Since the vertical line drawn through x = μ divides the curve into two equal halves, it automatically follows that, The area between $-\infty$ to μ = the area between μ to ∞ = 0.5

When the mean is zero, we have

The area between $-\infty$ to 0 = the area between 0 to ∞ = 0.5

(c) If we take μ = 0 and σ = 1 in (14.17), we have

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad \text{for } -\infty < x < \infty \quad \text{........ (14.18)}$$

The random variable x is known as standard normal variate (or variable) or standard normal deviate. The probability that a standard normal variate X would take a value less than or equal to a particular value say X = x is given by

$\phi(x) = p(X \le x)$ .... (14.19)

$\phi(x)$ is known as the cumulative distribution function.

We also have $\phi(0) = P(X \le 0)$ = Area of the standard normal curve between $-\infty$ and 0 = 0.5 .... (14.20)

(d) **The normal distribution is known as biparametric distribution as it is characterised by two parameters μ and σ^2. Once the two parameters are known, the normal distribution is completely specified.**

Properties of Normal Distribution

1. Since π = 22/7 , $e^{-\theta} = 1 / e^{\theta} > 0$, whatever θ may be,

   it follows that f (x) $\ge$ 0 for every x.

   It can be shown that

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

2. The mean of the normal distribution is given by μ. Further, since the distribution is symmetrical about $x = \mu$, it follows that the mean, median and mode of a normal distribution coincide, all being equal to μ.

3. The standard deviation of the normal distribution is given by σ.

   Mean deviation of normal distribution is

   $\sigma\sqrt{2\pi} \cong 0.8\sigma$ .... (14.21)

   The first and third quartiles are given by

   $q_1 = \mu - 0.675\,\sigma$ .... (14.22)

   and $q_3 = \mu + 0.675\,\sigma$ .... (14.23)

   so that, quartile deviation = 0.675 σ .... (14.24)

4. The normal distribution is symmetrical about $x = \mu$. As such, its skewness is zero i.e. the normal curve is neither inclined move towards the right (negatively skewed) nor towards the left (positively skewed).

5. The normal curve y = f (x) has two points of inflexion to be given by $x = \mu - \sigma$ and $x = \mu + \sigma$ i.e. at these two points, the normal curve changes its curvature from concave to convex and from convex to concave.

6. If $x \sim N(\mu, \sigma^2)$ then $z = x - \mu/\sigma \sim N(0, 1)$, z is known as standardised normal variate or normal deviate.

   We also have $P(z \le k) = \phi(k)$ ……………. (14.25)

   The values of $\phi(k)$ for different k are given in a table known as "Biometrika."

   Because of symmetry, we have

   $\phi(-k) = 1 - \phi(k)$ …………………. (14.26)

   We can evaluate the different probabilities in the following manner:

   $P(x < a) = P(x - \mu/\sigma < a - \mu/\sigma)$

   $= P(z < k), (k = a - \mu/\sigma)$

   $= \phi(k)$ ………………….. (14.27)

   Also $P(x \le a) = P(x < a)$ as x is continuous.

   $P(x > b) = 1 - P(x \le b)$

   $= 1 - \phi(b - \mu/\sigma)$ …………. (14.28)

   and $P(a < x < b) = \phi(b - \mu/\sigma) - \phi(a - \mu/\sigma)$ …. (14.29)

   ordinate at x = a is given by

   $(1/\sigma)\,\phi(a - \mu/\sigma)$ …………….. (14.30)

   Also, $\phi(-k) = \phi(k)$ ……….. (14.31)

   The values of $\phi(k)$ for different k are also provided in the Biometrika Table.

7. Area under the normal curve is shown in the following figure :

$\mu - 3\sigma$	$\mu - 2\sigma$	$\mu - \sigma$	$x = \mu$	$\mu + \sigma$	$\mu + 2\sigma$	$\mu + 3\sigma$
$(z = -3)$	$(z = -2)$	$(z = -1)$	$(z = 0)$	$(z = 1)$	$(z = 2)$	$(z = 3)$

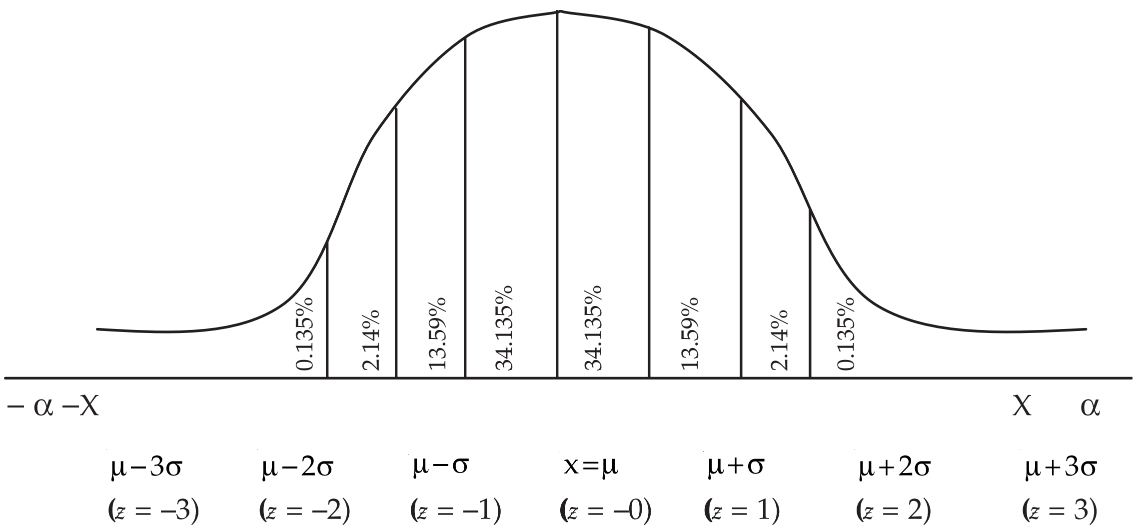

Figure 14.2

Area Under Normal Curve

From this figure, we find that

$P(\mu - \sigma < x < \mu) = P(\mu < x < \mu + \sigma) = 0.34135$

or alternatively, $P(-1 < z < 0) = P(0 < z < 1) = 0.34135$

$P(\mu - 2\sigma < x < \mu) = P(\mu < x < \mu + 2\sigma) = 0.47725$

i.e. $P(-2 < z < 1) = P(1 < z < 2) = 0.47725$

$P(\mu - 3\sigma < x < \mu) = P(\mu < x < \mu + 3\sigma) = 0.49865$

i.e. $P(-3 < z < 0) = P(0 < z < 3) = 0.49865$

……. (14.32)

combining these results, we have

$P(\mu - \sigma < x < \mu + \sigma) = 0.6828$

$\Rightarrow P(-1 < z < 1) = 0.6828$

$P(\mu - 2\sigma < x < \mu + 2\sigma) = 0.9546$

$\Rightarrow P(-2 < z < 2) = 0.9546$

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

$\Rightarrow P(-3 < z < 3) = 0.9973$.

………… (14.33)

We note that 99.73 per cent of the values of a normal variable lies between $(\mu - 3\sigma)$ and $(\mu + 3\sigma)$. Thus the probability that a value of x lies outside that limit is as low as 0.0027.

8. If x and y are independent normal variables with means and standard deviations as μ_1 and μ_2 and σ_1, and σ_2 respectively, then z = x + y also follows normal distribution with mean $(\mu_1 + \mu_2)$ and SD = $\sqrt{\sigma_1^2+\sigma_2^2}$ respectively.

   i.e. If $x \sim N(\mu_1, \sigma_1^2)$

   and $y \sim N(\mu_2, \sigma_2^2)$ and x and y are independent,

   then $z = x + y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

   …… (14.34)

Applications of Normal Distribution

The applications of normal distributions is not restricted to statistics only. Many science subjects, social science subjects, management, commerce etc. find many applications of normal distributions. Most of the continuous variables like height, weight, wage, profit etc. follow normal distribution. If the variable under study does not follow normal distribution, a simple transformation of the variable, in many a case, would lead to the normal distribution of the changed variable. When n, the number of trials of a binomial distribution, is large and p, the probability of a success, is moderate i.e. neither too large nor too small then the binomial distribution, also, tends to normal distribution. Poisson distribution, also for large value of m approaches normal distribution. Such transformations become necessary as it is easier to compute probabilities under the assumption of a normal distribution. Not only the distribution of discrete random variable, the probability distributions of t, chi-square and F also tend to normal distribution under certain specific conditions. In order to infer about the unknown universe, we take recourse to sampling and inferences regarding the universe is made possible only on the basis of normality assumption. Also the distributions of many a sample statistic approach normal distribution for large sample size.

Example 14.20: For a random variable x, the probability density function is given by

$$f(x) = \frac{e^{-(x-4)^2}}{\sqrt{\pi}}$$

for $-\infty < x < \infty$.

Identify the distribution and find its mean and variance.

Solution: The given probability density function may be written as

$$f(x) = \frac{1}{1/\sqrt{2} \times \sqrt{2}\,\pi} e^{-(x-4)^2/2\times 1/2} \qquad \text{for } -\infty < x < \infty$$

$$= \frac{1}{\sigma \times \sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad \text{for } -\infty < x < \infty$$

with $\mu = 4$ and $\sigma^2 = ½$

Thus the given probability density function is that of a normal distribution with $\mu = 4$ and variance = ½.

Example 14.21: If the two quartiles of a normal distribution are 47.30 and 52.70 respectively, what is the mode of the distribution? Also find the mean deviation about median of this distribution.

Solution: The 1st and 3rd quartiles of $N(\mu, \sigma^2)$ are given by $(\mu - 0.675\sigma)$ and $(\mu + 0.675\sigma)$ respectively. As given,

$\mu - 0.675\sigma = 47.30$ …. (1)

$\mu + 0.675\sigma = 52.70$ …. (2)

Adding these two equations, we get

$2\mu = 100$ or $\mu = 50$

Thus Mode = Median = Mean = 50. Also $\sigma = 4$.

Also Mean deviation about median

= mean deviation about mode

= mean deviation about mean

$\cong 0.80\ \sigma$

= 3.20

Example 14.22: Find the points of inflexion of the normal curve

$$f(x) = \frac{1}{4\sqrt{2\pi}} . e^{-(x-10)^2/32}$$

for $-\infty < x < \infty$

Solution: Comparing f (x) to the probability densities function of a normal variable x , we find that $\mu = 10$ and $\sigma = 4$.

The points of inflexion are given by

$\mu - \sigma$ and $\mu + \sigma$

i.e. 10 – 4 and 10 + 4

i.e. 6 and 14.

Example 14.23 : If x is a standard normal variable such that

$P(0 \le x \le b) = a$, what is the value of $P(|x| \ge b)$?

Solution : $P((x) \ge b)$

$= 1 - P(|x| \le b)$

$= 1 - P(-b \le x \le b)$

$= 1 - [P(0 \le x \le b) - P(-b \le x \le 0)]$

$= 1 - [P (0 \le x \le b) + P (0 \le x \le b)]$

$= 1 - 2a$

Example 14.24: X follows normal distribution with mean as 50 and variance as 100. What is P $(x \ge 60)$? Given $\phi (1) = 0.8413$

Solution: We are given that $x \sim N (\mu, \sigma^2)$ where

$\mu = 50$ and $\sigma^2 = 100 => \sigma = 10$

Thus $P (x \ge 60)$

$= 1 - P (x \le 60)$

$$= 1 - P\left(\frac{x-50}{10} \le \frac{60-50}{10}\right) = 1 - P (z \le 1)$$

$= 1 - \phi (1)$ (From 14.27)

$= 1 - 0.8413$

$\cong 0.16$

Example 14.25: If a random variable x follows normal distribution with mean as 120 and standard deviation as 40, what is the probability that $P (x \le 150 / x > 120)$?

Given that the area of the normal curve between $z = 0$ to $z = 0.75$ is 0.3734.

Solution: $P (x \le 150 / x > 120)$

$$= \frac{P(120 < x \le 150)}{P(x > 120)}$$

$$= \frac{P(120 < x \le 150)}{1 - P(x \le 120)}$$

$$= \frac{P\left(\frac{120-120}{40} \le \frac{x-120}{40} \le \frac{150-120}{40}\right)}{1 - P\left(\frac{x-120}{40} \le \frac{120-120}{40}\right)}$$

$$= \frac{P(0 < z \le 0.75)}{1 - P(z \le 0)}$$

$$= \frac{\phi(0.75) - \phi(0)}{1 - \phi(0)}$$ (From 14.29)

$$= \frac{0.8734 - 0.50}{1 - 0.50}$$

$\cong 0.75$ (ϕ (0.75) = Area of the normal curve between $z = -\infty$ to $z = 0.75$ = area between $-\infty$ to 0 + Area between 0 to 0.75 = 0.50 + 0.3734 = 0.8734)

Example 14.26: X is a normal variable with mean = 5 and SD 10. Find the value of b such that the probability of the interval [2 5, b] is 0.4772 given ϕ (2) = 0.9772.

Solution: We are given that x ~ N (μ, σ^2) where μ = 25 and σ = 10

and P [25 < x < b] = 0.4772

$$\Rightarrow \left[\frac{25-25}{10} < \frac{x-25}{10} < \frac{b-25}{10}\right] = 0.4772$$

$$\Rightarrow P\left[0 < z < \frac{b-25}{10}\right] = 0.4772$$

$$\Rightarrow \phi\left(\frac{b-25}{10}\right) - \phi(0) = 0.4772$$

$$\Rightarrow \phi\left(\frac{b-25}{10}\right) - 0.50 = 0.4772$$

$$\Rightarrow \phi\left(\frac{b-25}{10}\right) = 0.9772$$

$$\Rightarrow \phi \frac{b-25}{10} = \phi(2) \qquad \text{(as given)}$$

$$\Rightarrow \frac{b-25}{10} = 2$$

$\Rightarrow$ b = 25 + 2 × 10 = 45.

Example 14.27: In a sample of 500 workers of a factory, the mean wage and SD of wages are found to be Rs. 500 and Rs. 48 respectively. Find the number of workers having wages:

(i) more than Rs. 600

(ii) less than Rs. 450

(iii) between Rs. 548 and Rs. 600.

Solution: Let X denote the wage of the workers in the factory. We assume that X is normally distributed with mean wage as Rs. 500 and standard deviation of wages as Rs. 48 respectively.

(i) Probability that a worker selected at random would have wage more than Rs. 600

$= P(X > 600)$

$= 1 - P(X \le 600)$

$$= 1 - P\left(\frac{X - 500}{48} \le \frac{600 - 500}{48}\right)$$

$= 1 - P(z \le 2.08)$

$= 1 - \phi(2.08)$

$= 1 - 0.9812$ (From Biometrika Table)

$= 0.0188$

Thus the number of workers having wages less than Rs. 600

$= 500 \times 0.0188$

$= 9.4$

$\cong 9$

(ii) Probability of a worker having wage less than Rs. 450

$= P(X < 450)$

$$= P\left(\frac{X - 500}{48} < \frac{450 - 500}{48}\right)$$

$= P(z < -1.04)$

$= \phi(-1.04)$

$= 1 - \phi(1.04)$ (from 14.26)

$= 1 - 0.8508$ (from Biometrika Table)

$= 0.1492$

Hence the number of workers having wages less than Rs. 450

$= 500 \times 0.1492$

$\cong 75$

(iii) Probability of a worker having wage between Rs. 548 and Rs. 600.

$= P(548 < x < 600)$

$$= P\left(\frac{548 - 500}{48} < \frac{x - 500}{48} < \frac{600 - 500}{48}\right)$$

$= P (1 < z < 2.08)$

$= \phi (2.08) - \phi (1)$

$= 0.9812 - 0.8413$ (consulting Biometrika)

$= 0.1399$

So the number of workers with wages between Rs. 548 and Rs. 600

$= 500 \times 0.1399$

$\cong 70.$

Example 14.28: The distribution of wages of a group of workers is known to be normal with mean Rs. 500 and SD Rs. 100. If the wages of 100 workers in the group are less than Rs. 430, what is the total number of workers in the group?

Solution : Let X denote the wage. It is given that X is normally distributed with mean as Rs. 500 and SD as Rs. 100 and P (X < 430) = 100/N, N being the total no. of workers in the group

$$\Rightarrow P\left(\frac{X-500}{100} < \frac{430-500}{100}\right) = \frac{100}{N}$$

$$\Rightarrow P(z < -0.70) = \frac{100}{N}$$

$$\Rightarrow \phi(-0.70) = \frac{100}{N}$$

$$\Rightarrow 1-\phi(0.70) = \frac{100}{N}$$

$$\Rightarrow 1-0.758 = \frac{100}{N}$$

$$\Rightarrow 0.242 = \frac{100}{N}$$

$\Rightarrow N \cong 413.$

Example 14.29: The mean height of 2000 students at a certain college is 165 cms and SD 9 cms. What is the probability that in a group of 5 students of that college, 3 or more students would have height more than 174 cm?

Solution: Let X denote the height of the students of the college. We assume that X is normally distributed with mean (μ) 165 cms and SD (σ) as 9 cms. If p denotes the probability that a student selected at random would have height more than 174 cms., then p = P (X > 174)

= 1 – P (X ≤ 174)

$$=1-P\left(\frac{X-165}{9}\le\frac{174-165}{9}\right)$$

= 1 – P (z ≤ 1)

= 1 – ϕ (1)

= 1 – 0.8413

= 0.1587

If y denotes the number of students having height more than 174 cm. in a group of 5 students then y ~ β (n, p) where n = 5 and p = 0.1587. Thus the probability that 3 or more students would be more than 174 cm.

= p (y ≥ 3)

= p (y = 3) + p (y = 4) + p (y = 5)

$= 5_{C_3}(0.1587)^3 . (0.8413)^2 + 5_{C_4}(0.1587)^4 \times (0.8413) + 5_{C_5}(0.1587)^5$

= 0.02829 + 0.002668 + 0.000100

= 0.03106.

Example 14.30: The mean of a normal distribution is 500 and 16 per cent of the values are greater than 600. What is the standard deviation of the distribution?

(Given that the area between z = 0 to z = 1 is 0.34)

Solution : Let σ denote the standard deviation of the distribution.

We are given that

P (X > 600) = 0.16

⇒ 1 – P (X ≤ 600) = 0.16

⇒ P (X ≤ 600) = 0.84

$$\Rightarrow P\left(\frac{X-500}{\sigma}\le\frac{600-500}{\sigma}\right)=0.84$$

$$\Rightarrow P\left(z\le\frac{100}{\sigma}\right)=0.84$$

$$\Rightarrow \phi\left(\frac{100}{\sigma}\right)=\phi(1)$$

$$\Rightarrow \frac{(100)}{\sigma} = 1$$

$$\Rightarrow \sigma = 100.$$

Example 14.31: In a business, it is assumed that the average daily sales expressed in rupees follows normal distribution.

Find the coefficient of variation of sales given that the probability that the average daily sales is less than Rs. 124 is 0.0287 and the probability that the average daily sales is more than Rs. 270 is 0.4599.

Solution: Let us denote the average daily sales by x and the mean and SD of x by μ and σ respectively. As given,

$P (x < 124) = 0.0287$ ……..(1)

$P (x > 270) = 0.4599$ ……..(2)

From (1), we have

$$P\left(\frac{X-\mu}{\sigma} < \frac{124-\mu}{\sigma}\right) = 0.0287$$

$$\Rightarrow P(z < \frac{124-\mu}{\sigma}) = 0.0287$$

$$\Rightarrow \phi\left(\frac{124-\mu}{\sigma}\right) = 0.0287$$

$$\Rightarrow 1-\phi\left(\frac{\mu-124}{\sigma}\right) = 0.0287$$

$$\Rightarrow \phi\left(\frac{\mu-124}{\sigma}\right) = 0.9713$$

$$\Rightarrow \phi\left(\frac{\mu-124}{\sigma}\right) = \phi\,(2.085) \text{ (From Biometrika)}$$

$$\Rightarrow \left(\frac{\mu-124}{\sigma}\right) = 2.085 \text{ …….(3)}$$

From (2) we have,

$$1 - P (x \le 270) = 0.4599$$

$$\Rightarrow P\left(\frac{X-\mu}{\sigma} \le \frac{270-\mu}{\sigma}\right) = 0.5401$$

$$\Rightarrow \phi\left(\frac{270-\mu}{\sigma}\right) = 0.5401$$

$$\Rightarrow \phi\left(\frac{270-\mu}{\sigma}\right) = \phi\ (0.1)$$

$$\Rightarrow \left(\frac{270-\mu}{\sigma}\right) = 0.1 \quad .....(4)$$

Dividing (3) by (4), we get

$$\frac{\mu-124}{270-\mu} = 20.85$$

$\Rightarrow \mu - 124 = 5629.50 - 20.85\ \mu$

$\Rightarrow \mu = 5753.50/21.85$

$= 263.32$

Substituting this value of μ in (3), we get

$$\frac{263.32 - 124}{\sigma} = 2.085$$

$\Rightarrow \sigma = 66.82$

Thus the coefficient of variation of sales

$= \sigma/\mu \times 100$

$$= \frac{66.82}{263.32} \times 100$$

$= 25.38$

Example 14.32: x and y are independent normal variables with mean 100 and 80 respectively and standard deviation as 4 and 3 respectively. What is the distribution of (x + y)?

Solution: We know that if $x \sim N(\mu_1, \sigma_1^2)$ and $y \sim N(\mu_2, \sigma_2^2)$ and they are independent, then $z = x + y$ follows normal with mean $(\mu_1 + \mu_2)$ and

SD = $\sqrt{\sigma_1^2+\sigma_1^2}$ respectively.

Thus the distribution of (x + y) is normal with mean (100 + 80) or 180

and SD $\sqrt{4^2+3^2} = 5$

14.5 CHI-SQUARE DISTRIBUTION, T-DISTRIBUTION AND F – DISTRIBUTION

We are going to study statistical inference in the concluding chapter. For statistical inference, we need some basic ideas about three more continuous theoretical probability distributions, namely, chi-square distribution, t – distribution and F – distribution. Before discussing this distribution, let us review standard normal distribution.

Standard Normal Distribution

If a continuous random variable z follows standard normal distribution, to be denoted by $z \sim N(0, 1)$, then the probability density function of z is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad \text{for } -\infty < z < \infty \quad \ldots\ldots (14.35)$$

Some important properties of z are listed below :

(i) z has mean, median and mode all equal to zero.

(ii) The standard deviation of z is 1. Also the approximate values of mean deviation and quartile deviation are 0.8 and 0.675 respectively.

(iii) The standard normal distribution is symmetrical about $z = 0$.

(iv) The two points of inflexion of the probability curve of the standard normal distribution are –1 and 1.

(v) The two tails of the standard normal curve never touch the horizontal axis.

(vi) The upper and lower p per cent points of the standard normal variable z are given by

$P(Z > z_p) = p$ …….. (14.36)

And $P(Z < z_{1-p}) = p$

i.e. $P(Z < -z_p) = p$ respectively … (14.37)

(since for a standard normal distribution $z_{1-p} = -z_p$)

Selecting P = 0.005, 0.025, 0.01 and 0.05 respectively,

We have $\mathbf{z_{0.005} = 2.58}$

$\mathbf{z_{0.025} = 1.96}$

$\mathbf{z_{0.01} = 2.33}$

$\mathbf{z_{0.05} = 1.645}$ ………… (14.38)

These are shown in fig 14.3.

(vii) If $\bar{x}$ denotes the arithmetic mean of a random sample of size n drawn from a normal population then,

$$Z = \frac{\sqrt{n}\,(\bar{x} - \mu)}{\sigma} \sim N\,(0, 1) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \; 14.39$$

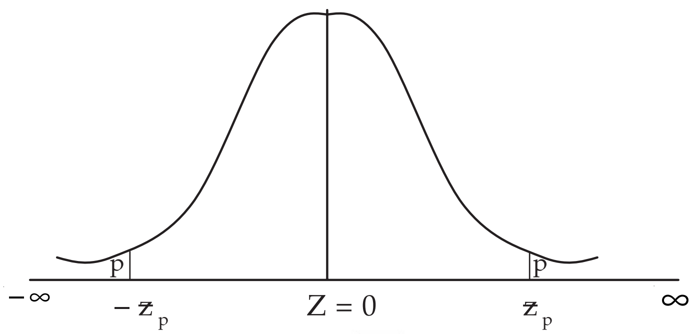

Fig . 14.3

Showing upper and lower p % points of the standard normal variable.

Chi–square distribution: (χ^2 – distribution)

If a continuous random variable x follows Chi–square distribution with n degrees of freedom (df) i.e. n independent condition without any restriction or constraints, to be denoted by $x \sim X_n^2$ then the probability density function of x is given by

$f(x) = k \,.\, e^{-x/2}\, x^{n/2 - 1}$

(Where k is a constant) for $0 < x < \infty$ ……….. (14.40)

The important properties of χ^2 (chi-square) distribution are mentioned below:

(i) Mean of the chi-square distribution = n

(ii) Standard deviation of chi–square distribution = $\sqrt{2n}$

(iii) Additive property of chi–square distribution.

If x and y are two independent chi-square distribution with m and n degrees of freedom, then (x + y) also follows chi-square distribution with (m + n) df.

i.e., if $x \sim \chi_m^2$

and $y \sim \chi_m^2$

and x and y are independent,

then $\mu = x + y \sim \chi_{m+n}^2$ ……….. (14.41)

(iv) For large n, $\sqrt{2x^2} - \sqrt{2n-1}$ follows as approximate standard normal distribution.

(v) The upper and lower p per cent points of chi-square distribution with n df are given by

$P(\chi^2 > \chi^2_{p,n}) = p$

and $P(\chi^2 < \chi^2_{1-p}, n) = p$ ……….. (14.42)

(vi) If $z_1, z_2, z_3 \ldots\ldots z_n$ are n independent standard normal variables, then

$\mu = \sum_1^n zi^2 \sim \chi^2_n$ Similarly, if $x_1, x_2, x_3 \ldots\ldots x_n$ are n independent normal variables, with a common mean μ and common variables σ^2, then $\mu = \sum (xi - \mu/\sigma) 2 \sim \chi^2_n$ …. (14.43)

Lastly if a random sample of size n is taken from a normal population with mean μ and variance σ^2, then

$$\mu = \frac{\sum (x_i - \bar{x})^2}{\sigma^2} \sim \chi^2_{n-1} \quad \ldots\ldots\ldots\ldots (14.44)$$

(vii) Chi-square distribution is positively skewed i.e. the probability curve of the chi–square distribution is inclined move on the right.

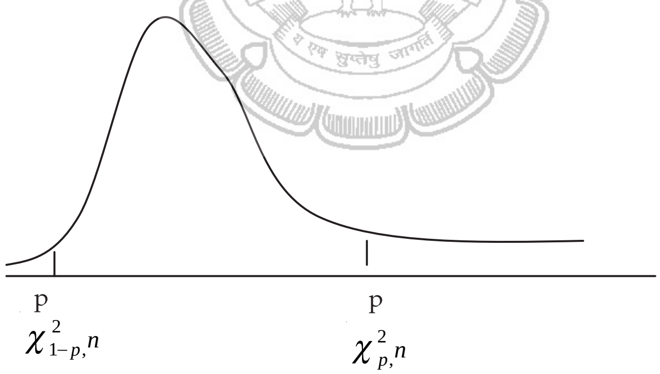

Figure 14.4

Showing the upper and lower p per cent point of chi-square distribution with n df.

t – distribution: If a continuous random variable t follows t – distribution with n df, then its probability density function is given by

$$f(t) = k\left[1 + t^2/n\right]^{-(n+1)/2}$$

(where k is a constant) for $-\infty < t < \infty$ ……… 14.45

This is denoted by $t \sim t_n$.

The important properties of t-distribution are mentioned below:

(i) Mean of t-distribution is zero.

(ii) Standard deviation of t-distribution $\sqrt{n/(n-2)}$, n > 2

(iii) t-distribution is symmetrical about t = 0.

(iv) For large n (> 30), t-distribution tends to the standard normal distribution.

(v) The upper and lower p per cent points of t-distribution are given by

$$P (t > t_{p}, n) = p$$

And $P (t < t_{p}, n) = p$ ………… (14.46)

(vi) If y and z are two independent random variables such that $y \sim \chi_n^2$ and $Z \sim N (0 , 1)$, then

$$t = \frac{\sqrt{n}\, z}{\sqrt{y}} \sim t_n \quad \text{…………(14.47)}$$

Similarly, if a random sample of size n is taken from a normal distribution with mean m and SD σ, then

$$t = \frac{\sqrt{n-1}(\bar{x} - \mu)}{S} : t_{n-1} \quad \text{…………(14.48)}$$

Here $\bar{x}$ and S denote the sample mean and sample SD respectively.

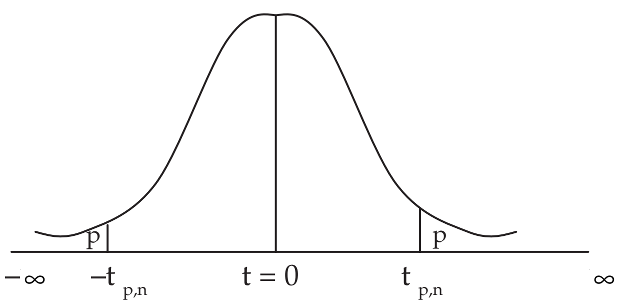

Figure 14.5

Showing the upper and lower p per cent point pf t – distribution with n df.

F – Distribution

If a continuous random variable F follows F – distribution with n_1 and n_2 degrees of freedom, to be denoted by $F \sim F_{n_1, n_2}$, then its probability density function is given by

$f(F) = k \cdot F^{n_1/2-1} \cdot (1 + n_1 F / n)^{-(n_1+n_2)/2}$

(where k is a constant) for $0 < F < \infty$ …………….(14.49)

Important properties of F – distribution

1. Mean of the F – distribution = $\frac{n_2}{n_2-2}$, $n_2 > 2$
2. Standard deviation of the F – distribution

$$= \frac{n_2}{n_2-2}\sqrt{\frac{2(n_1+n_2-2)}{n_1(n_2-4)}},\ n_2 > 4$$

and for large n_1 and n_2, SD = $\sqrt{\frac{2(n_1+n_2)}{n_1 n_2}}$

3. F – distribution has a positive skewness.
4. The upper and lower p per cent points of F – distribution are given by

$P = (F > F_p, (n_1, n_2)) = p$

and $P\left(F < \frac{1}{F_p(n_2, n_1)}\right) = p$ ………. (14.50)

5. If U and V are two independent random variables such that $U \sim \chi^2_{n_1}$

and $V \sim \chi^2_{n_2}$ then

$$F = \frac{U/n_1}{V/n_2} \sim F_{n_1, n_2} \text{ ....} \quad (14.51)$$

6. For large values of n_1 and n_2, F – distribution tends to normal distribution with mean, and

SD = $\sqrt{\frac{2(n_1+n_2)}{n_1 n_2}}$

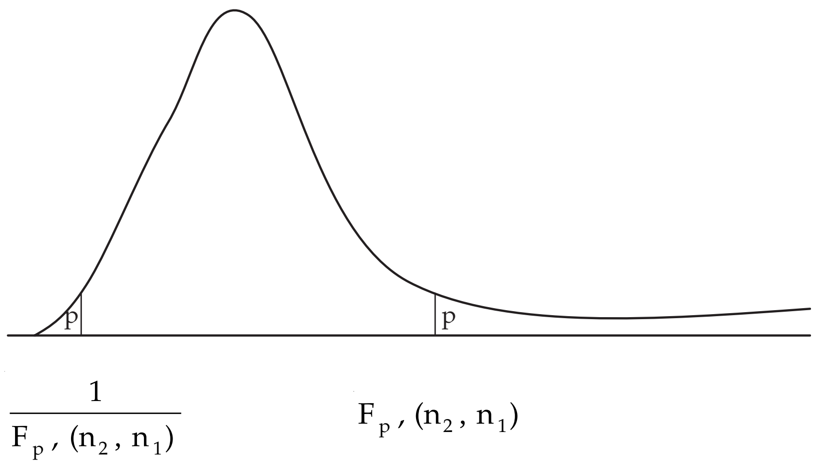

Figure 14.6

Showing the upper and lower p per cent points of F–distribution with n_1 and n_2 degree of freedom.

EXERCISE

Set : A

Write down the correct answers. Each question carries 1 mark.

1. A theoretical probability distribution.
   (a) does not exist. (b) exists only in theory.
   (c) exists in real life. (d) both (b) and (c).
2. Probability distribution may be
   (a) discrete. (b) continuous. (c) infinite. (d) both (a) and (b).
3. An important discrete probability distribution is
   (a) Poisson distribution. (b) Normal distribution.
   (c) Cauchy distribution. (d) Log normal distribution.
4. An important continuous probability distribution
   (a) Binomial distribution. (b) Poisson distribution.
   (c) Geometric distribution. (d) Chi-square distribution.
5. Parameter is a characteristic of
   (a) population. (b) sample. (c) probability distribution. (d) both (a) and (b).
6. An example of a parameter is
   (a) sample mean. (b) population mean.
   (c) binomial distribution. (d) sample size.

7. A trial is an attempt to
   (a) make something possible.
   (b) make something impossible.
   (c) prosecute an offender in a court of law.
   (d) produce an outcome which is neither certain nor impossible.
8. The important characteristic(s) of Bernoulli trials
   (a) each trial is associated with just two possible outcomes.
   (b) trials are independent.
   (c) trials are infinite.
   (d) both (a) and (b).
9. The probability mass function of binomial distribution is given by
   (a) $f(x) = p^x q^{n-x}$.
   (b) $f(x) = {}^nc_x p^x q^{n-x}$.
   (c) $f(x) = {}^nc_x q^x p^{n-x}$.
   (d) $f(x) = {}^nc_x p^{n-x} q^x$.
10. If x is a binomial variable with parameters n and p, then x can assume
   (a) any value between 0 and n.
   (b) any value between 0 and n, both inclusive.
   (c) any whole number between 0 and n, both inclusive.
   (d) any number between 0 and infinity.
11. A binomial distribution is
   (a) never symmetrical.
   (b) never positively skewed.
   (c) never negatively skewed.
   (d) symmetrical when p = 0.5.
12. The mean of a binomial distribution with parameter n and p is
   (a) n (1– p).
   (b) np (1 – p).
   (c) np.
   (d) $\sqrt{np(1-p)}$.
13. The variance of a binomial distribution with parameters n and p is
   (a) $np^2 (1 - p)$.
   (b) $\sqrt{np(1-p)}$.
   (c) nq (1 – q).
   (d) $n^2p^2 (1- p)^2$.
14. An example of a bi-parametric discrete probability distribution is
   (a) binomial distribution.
   (b) poisson distribution.
   (c) normal distribution.
   (d) both (a) and (b).
15. For a binomial distribution, mean and mode
   (a) are never equal.
   (b) are always equal.
   (c) are equal when q = 0.50.
   (d) do not always exist.

16. The mean of binomial distribution is

    (a) always more than its variance.
    (b) always equal to its variance.
    (c) always less than its variance.
    (d) always equal to its standard deviation.

17. For a binomial distribution, there may be

    (a) one mode.
    (b) two mode.
    (c) (a).
    (d) (a) or (b).

18. The maximum value of the variance of a binomial distribution with parameters n and p is

    (a) n/2.
    (b) n/4.
    (c) np (1 – p).
    (d) 2n.

19. The method usually applied for fitting a binomial distribution is known as

    (a) method of least square.
    (b) method of moments.
    (c) method of probability distribution.
    (d) method of deviations.

20. Which one is not a condition of Poisson model?

    (a) the probability of having success in a small time interval is constant.
    (b) the probability of having success more than one in a small time interval is very small.
    (c) the probability of having success in a small interval is independent of time and also of earlier success.
    (d) the probability of having success in a small time interval (t, t + dt) is kt for a positive constant k.

21. Which one is uniparametric distribution?

    (a) Binomial.
    (b) Poisson.
    (c) Normal.
    (d) Hyper geometric.

22. For a Poisson distribution,

    (a) mean and standard deviation are equal.
    (b) mean and variance are equal.
    (c) standard deviation and variance are equal.
    (d) both (a) and (b).

23. Poisson distribution may be

    (a) unimodal.
    (b) bimodal.
    (c) Multi-modal.
    (d) (a) or (b).

24. Poisson distribution is

    (a) always symmetric.
    (b) always positively skewed.
    (c) always negatively skewed.
    (d) symmetric only when m = 2.

25. A binomial distribution with parameters m and p can be approximated by a Poisson distribution with parameter m = np is

    (a) $m \to \propto$.
    (b) $p \to 0$.
    (c) $m \to \propto$ and $p \to 0$.
    (d) $m \to \propto$ and $p \to 0$ so that mp remains finite..

26. For Poisson fitting to an observed frequency distribution,

(a) we equate the Poisson parameter to the mean of the frequency distribution.

(b) we equate the Poisson parameter to the median of the distribution.

(c) we equate the Poisson parameter to the mode of the distribution.

(d) none of these.

27. The most important continuous probability distribution is known as

(a) Binomial distribution. (b) Normal distribution.

(c) Chi-square distribution. (d) sampling distribution.

28. The probability density function of a normal variable x is given by

(a) $f(x) = \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$ for $-\propto < x < \propto$.

(b) $f(x) = f(x) = \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{-(x-\mu)^2}{2\sigma^2}}$ for $0 < x < \propto$.

(c) $f(x) = \frac{1}{\sqrt{2\pi\sigma}}.e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ for $-\propto < x < \propto$.

(d) none of these.

29. The total area of the normal curve is

(a) one. (b) 50 per cent.

(c) 0.50. (d) any value between 0 and 1.

30. The normal curve is

(a) Bell-shaped. (b) U- shaped.

(c) J- shaped. (d) Inverted J – shaped.

31. The normal curve is

(a) positively skewed. (b) negatively skewed.

(c) Symmetrical. (d) all these.

32. Area of the normal curve is

(a) between $-\propto$ to μ is 0.50. (b) between μ to $\propto$ is 0.50.

(c) between $-\propto$ to $\propto$ is 0.50. (d) both (a) and (b).

33. The cumulative distribution function of a random variable X is given by

(a) $F(x) = P(X \le x)$.
(b) $F(X) = P(X \le x)$.
(c) $F(x) = P(X \ge x)$.
(d) $F(x) = P(X = x)$.

34. The mean and mode of a normal distribution

(a) may be equal.
(b) may be different.
(c) are always equal.
(d) (a) or (b).

35. The mean deviation about median of a standard normal variate is

(a) $0.675\ \sigma$.
(b) 0.675 .
(c) $0.80\ \sigma$.
(d) 0.80.

36. The quartile deviation of a normal distribution with mean 10 and SD 4 is

(a) 0.675.
(b) 67.50.
(c) 2.70.
(d) 3.20.

37. For a standard normal distribution, the points of inflexion are given by

(a) $\mu - \sigma$ and $\mu + \sigma$.
(b) $-\sigma$ and σ.
(c) –1 and 1.
(d) 0 and 1.

38. The symbol ϕ (a) indicates the area of the standard normal curve between

(a) 0 to a.
(b) a to ∞.
(c) $-\infty$ to a.
(d) $-\infty$ to ∞.

39. The interval $(\mu - 3\sigma, \mu + 3\sigma)$ covers

(a) 95% area of a normal distribution.
(b) 96% area of a normal distribution.
(c) 99% area of a normal distribution.
(d) all but 0.27% area of a normal distribution.

40. Number of misprints per page of a thick book follows

(a) Normal distribution.
(b) Poisson distribution.
(c) Binomial distribution.
(d) Standard normal distribution.

41. The result of ODI matches between India and Pakistan follows

(a) Binomial distribution.
(b) Poisson distribution.
(c) Normal distribution.
(d) (b) or (c).

42. The wage of workers of a factory follow

(a) Binomial distribution.
(b) Poisson distribution.
(c) Normal distribution .
(d) Chi-square distribution.

43. If X and Y are two independent random variables such that $X \sim \chi^2 m$ and $Y \sim \chi^2 n$, then the distribution of (X +Y) is

(a) normal.
(b) standard normal.
(c) T.
(d) chi-square.

Set B :

Write down the correct answers. Each question carries 2 marks.

1. What is the standard deviation of the number of recoveries among 48 patients when the probability of recovering is 0.75?

   (a) 36. (b) 81. (c) 9. (d) 3.

2. X is a binomial variable with n = 20. What is the mean of X if it is known that x is symmetric?

   (a) 5. (b) 10. (c) 2. (d) 8.

3. If X ~ B (n, p), what would be the least value of the variance of x when n = 16?

   (a) 2. (b) 4. (c) 8. (d) $\sqrt{5}$.

4. If x is a binomial variate with parameter 15 and 1/3, what is the value of mode of the distribution

   (a) 5 and 6. (b) 5. (c) 5.50. (d) 6.

5. What is the no. of trials of a binomial distribution having mean and SD as 3 and 1.5 respectively?

   (a) 2. (b) 4. (c) 8. (d) 12.

6. What is the probability of getting 3 heads if 6 unbiased coins are tossed simultaneously?

   (a) 0.50. (b) 0.25. (c) 0.3125. (d) 0.6875.

7. If the overall percentage of success in an exam is 60, what is the probability that out of a group of 4 students, at least one has passed?

   (a) 0.6525. (b) 0.9744. (c) 0.8704. (d) 0.0256.

8. What is the probability of making 3 correct guesses in 5 True – False answer type questions?

   (a) 0.3125. (b) 0.5676. (c) 0.6875. (d) 0.4325

9. If the standard deviation of a Poisson variate X is 2, what is P (1.5 < X < 2.9)?

   (a) 0.231. (b) 0.158. (c) 0.15. (d) 0.144.

10. If the mean of a Poisson variable X is 1, what is P (X = at least one)?

    (a) 0.456. (b) 0.821. (c) 0.632. (d) 0.254.

11. If X ~ P (m) and its coefficient of variation is 50, what is the probability that X would assume only non-zero values?

    (a) 0.018. (b) 0.982. (c) 0.989. (d) 0.976.

12. If 1.5 per cent of items produced by a manufacturing units are known to be defective, what is the probability that a sample of 200 items would contain no defective item?

    (a) 0.05. (b) 0.15. (c) 0.20. (d) 0.22.

13. For a Poisson variate X, P (X = 1) = P (X = 2). What is the mean of X?

(a) 1.00. (b) 1.50. (c) 2.00. (d) 2.50.

14. If 1 per cent of an airline's flights suffer a minor equipment failure in an aircraft, what is the probability that there will be exactly two such failures in the next 100 such flights?

(a) 0.50. (b) 0.184. (c) 0.265. (d) 0.256.

15. If for a Poisson variable X, f(2) = 3 f(4), what is the variance of X?

(a) 2. (b) 4. (c) $\sqrt{2}$. (d) 3.

16. What is the coefficient of variation of x, characterised by the following probability density function: $f(x) = \frac{1}{\sqrt[4]{2\pi}} e - \frac{(x-10)^2}{3^2}$ for $-\infty < x < \infty$

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

17. What is the first quartile of X having the following probability density function?

$f(x) = \frac{1}{\sqrt{72\pi}} e - \frac{-(x-10)^2}{72}$ for $-\infty < x < \infty$

(a) 4. (b) 5. (c) 5.95. (d) 6.75.

18. If the two quartiles of N (μ , σ^2) are 14.6 and 25.4 respectively, what is the standard deviation of the distribution?

(a) 9. (b) 6. (c) 10. (d) 8.

19. If the mean deviation of a normal variable is 16, what is its quartile deviation?

(a) 10.00. (b) 13.50. (c) 15.00. (d) 12.05.

20. If the points of inflexion of a normal curve are 40 and 60 respectively, then its mean deviation is

(a) 40. (b) 45. (c) 50. (d) 60.

21. If the quartile deviation of a normal curve is 4.05, then its mean deviation is

(a) 5.26. (b) 6.24. (c) 4.24. (d) 4.80.

22. If the Ist quartile and mean deviation about median of a normal distribution are 13.25 and 8 respectively, then the mode of the distribution is

(a) 20. (b) 10. (c) 15. (d) 12.

23. If the area of standard normal curve between z = 0 to z = 1 is 0.3413, then the value of ϕ (1) is

(a) 0.5000. (b) 0.8413. (c) –0.5000. (d) 1.

24. If X and Y are 2 independent normal variables with mean as 10 and 12 and SD as 3 and 4, then (X+Y) is normally distributed with

(a) mean = 22 and SD = 7. (b) mean = 22 and SD = 25.

(c) mean = 22 and SD = 5. (d) mean = 22 and SD = 49.

Set : C

Answer the following questions. Each question carries 5 marks.

1. If it is known that the probability of a missile hitting a target is 1/8, what is the probability that out of 10 missiles fired, at least 2 will hit the target?

(a) 0.4258. (b) 0.3968. (c) 0.5238. (d) 0.3611.

2. X is a binomial variable such that 2 P(X = 2) = P(X = 3) and mean of X is known to be 10/3. What would be the probability that X assumes at most the value 2?

(a) 16/81. (b) 17/81. (c) 47/243. (d) 46/243.

3. Assuming that one-third of the population are tea drinkers and each of 1000 enumerators takes a sample of 8 individuals to find out whether they are tea drinkers or not, how many enumerators are expected to report that five or more people are tea drinkers?

(a) 100. (b) 95. (c) 88. (d) 90.

4. If a random variable X follows binomial distribution with mean as 5 and satisfying the condition 10 P (X = 0) = P (X = 1), what is the value of P (X ≥ / x > 0)?

(a) 0.67. (b) 0.56. (c) 0.99. (d) 0.82.

5. Out of 128 families with 4 children each, how many are expected to have at least one boy and one girl?

(a) 100. (b) 105. (c) 108. (d) 112.

6. In 10 independent rollings of a biased die, the probability that an even number will appear 5 times is twice the probability that an even number will appear 4 times. What is the probability that an even number will appear twice when the die is rolled 8 times?

(a) 0.0304. (b) 0.1243. (c) 0.2315. (d) 0.1926.

7. If a binomial distribution is fitted to the following data:

x:	0	1	2	3	4
f:	16	25	32	17	10

then the sum of the expected frequencies for x = 2, 3 and 4 would be

(a) 58. (b) 59. (c) 60. (d) 61.

8. If X follows normal distribution with $\mu = 50$ and $\sigma = 10$, what is the value of P $(x \leq 60 / x > 50)$?

   (a) 0.8413. (b) 0.6828. (c) 0.1587. (d) 0.7256.

9. X is a Poisson variate satisfying the following condition 9 P (X = 4) + 90 P (X = 6) = P (X = 2). What is the value of P (X £ 1)?

   (a) 0.5655 (b) 0.6559 (c) 0.7358 (d) 0.8201

10. A random variable x follows Poisson distribution and its coefficient of variation is 50. What is the value of P (x > 1 / x > 0)?

    (a) 0.1876 (b) 0.2341 (c) 0.9254 (d) 0.8756

11. A renowned hospital usually admits 200 patients every day. One per cent patients, on an average, require special room facilities. On one particular morning, it was found that only one special room is available. What is the probability that more than 3 patients would require special room facilities?

    (a) 0.1428 (b) 0.1732 (c) 0.2235 (d) 0.3450

12. A car hire firm has 2 cars which is hired out everyday. The number of demands per day for a car follows Poisson distribution with mean 1.20. What is the proportion of days on which some demand is refused? (Given $e^{1.20} = 3.32$).

    (a) 0.25 (b) 0.3012 (c) 0.12 (d) 0.03

13. If a Poisson distribution is fitted to the following data:

Mistake per page	0	1	2	3	4	5
No. of pages	76	74	29	17	3	1

    Then the sum of the expected frequencies for x = 0, 1 and 2 is

    (a) 150. (b) 184. (c) 165. (d) 148.

14. The number of accidents in a year attributed to taxi drivers in a locality follows Poisson distribution with an average 2. Out of 500 taxi drivers of that area, what is the number of drivers with at least 3 accidents in a year?

    (a) 162 (b) 180 (c) 201 (d) 190

15. In a sample of 800 students, the mean weight and standard deviation of weight are found to be 50 Kg and 20 Kg respectively. On the assumption of normality, what is the number of students weighing between 46 Kg and 62 Kg? Given area of the standard normal curve between $z = 0$ to $z = 0.20 = 0.0793$ and area between $z = 0$ to $z = 0.60 = 0.2257$.

    (a) 250 (b) 244 (c) 240 (d) 260

16. The salary of workers of a factory is known to follow normal distribution with an average salary of Rs. 10,000 and standard deviation of salary as Rs. 2,000. If 50 workers receive salary more than Rs. 14,000, then the total no. of workers in the factory is

    (a) 2,193 (b) 2,000 (c) 2,200 (d) 2,500

17. For a normal distribution with mean as 500 and SD as 120, what is the value of k so that the interval [500, k] covers 40.32 per cent area of the normal curve? Given ϕ (1.30) = 0.9032.

(a) 740 (b) 750 (c) 760 (d) 800

18. The average weekly food expenditure of a group of families has a normal distribution with mean Rs. 1,800 and standard deviation Rs. 300. What is the probability that out of 5 families belonging to this group, at least one family has weekly food expenditure in excess of Rs. 1,800? Given ϕ (1) = 0.84.

(a) 0.418 (b) 0.582 (c) 0.386 (d) 0.614

19. If the weekly wages of 5000 workers in a factory follows normal distribution with mean and SD as Rs. 700 and Rs. 50 respectively, what is the expected number of workers with wages between Rs. 660 and Rs. 720?

(a) 2,050 (b) 2,200 (c) 2,218 (d) 2,300

20. 50 per cent of a certain product have weight 60 Kg or more whereas 10 per cent have weight 55 Kg or less. On the assumption of normality, what is the variance of weight?

Given ϕ (1.28) = 0.90.

(a) 15.21 (b) 9.00 (c) 16.00 (d) 22.68

ANSWERS

Set : A

1.	(a)	2.	(d)	3.	(a)	4.	(d)	5.	(a)	6.	(b)	7.	(d)	8.	(d)
9.	(a)	10.	(c)	11.	(d)	12.	(c)	13.	(c)	14.	(a)	15.	(c)	16.	(a)
17.	(c)	18.	(b)	19.	(b)	20.	(a)	21.	(b)	22.	(b)	23.	(d)	24.	(b)
25.	(d)	26.	(a)	27.	(b)	28.	(a)	29.	(a)	30.	(a)	31.	(c)	32	(d)
33.	(a)	34.	(c)	35.	(d)	36.	(c)	37.	(c)	38.	(c)	39.	(d)	40.	(b)
41.	(a)	42.	(c)	43.	(d)										

Set : B

1.	(d)	2.	(b)	3.	(a)	4.	(b)	5.	(d)	6.	(c)	7.	(b)	8.	(a)
9.	(d)	10.	(c)	11.	(b)	12.	(a)	13.	(c)	14.	(b)	15.	(a)	16.	(c)
17.	(c)	18.	(d)	19.	(b)	20.	(a)	21.	(d)	22.	(a)	23.	(b)	24.	(c)

Set : C

1.	(d)	2.	(b)	3.	(c)	4.	(c)	5.	(d)	6.	(a)	7.	(d)	8.	(b)
9.	(c)	10.	(c)	11.	(a)	12.	(d)	13.	(b)	14.	(a)	15.	(b)	16.	(a)
17.	(c)	18.	(b)	19.	(c)	20.	(a)								

ADDITIONAL QUESTION BANK

1. When a coin is tossed 10 times then

   (a) Normal Distribution　　(b) Poisson Distribution
   (c) Binomial Distribution　　(d) None is used

2. In Binomial Distribution 'n' means

   (a) No. of trials of the experiment　　(b) the probability of getting success
   (c) no. of success　　(d) none

3. Binomial Distribution is a

   (a) Continuous　　(b) discrete
   (c) both　　(d) none probability distribution.

4. When there are a fixed number of repeated trial of any experiments under identical conditions for which only one of two mutually exclusive outcomes, success or failure can result in each trial then

   (a) Normal Distribution　　(b) Binomial Distribution
   (c) Poisson Distribution　　(d) None is used

5. In Binomial Distribution 'p' denotes Probability of

   (a) Success　　(b) Failure　　(c) Both　　(d) None

6. When 'p' = 0.　　5, the binomial distribution is

   (a) asymmetrical　　(b) symmetrical　　(c) Both　　(d) None

7. When 'p' is larger than 0. 5, the binomial distribution is

   (a) asymmetrical　　(b) symmetrical　　(c) Both　　(d) None

8. Mean of Binomial distribution is

   (a) npq　　(b) np　　(c) both　　(d) none

9. Variance of Binomial distribution is

   (a) npq　　(b) np　　(c) both　　(d) none

10. When p = 0.1 the binomial distribution is skewed to the

    (a) left　　(b) right　　(c) both　　(d) none

11. If in Binomial distribution np = 9 and npq = 2. 25 then q is equal to

    (a) 0.25　　(b) 0.75　　(c) 1　　(d) none

12. In Binomial Distribution

    (a) mean is greater than variance　　(b) mean is less than variance
    (c) mean is equal to variance　　(d) none

13. Standard deviation of binomial distribution is

(a) square of npq (b) square root of npq
(c) square of np (d) square root of np

14. _________ distribution is a limiting case of Binomial distribution

(a) Normal (b) Poisson (c) Both (d) none

15. When the no. of trials is large then

(a) Normal (b) Poisson
(c) Binomial (d) none distribution is used

16. In Poisson Distribution, probability of success is very close to

(a) 1 (b) – 1 (c) 0 (d) none

17. In Poisson Distribution np is

(a) finite (b) infinite (c) 0 (d) none

18. In ________________ distribution, mean = variance

(a) Normal (b) Binomial (c) Poisson (d) none

19. In Poisson distribution mean is equal to

(a) npq (b) np (c) square root mp (d) square root mpq

20. In Poisson distribution standard deviation is equal to

(a) square root of np (b) square of np (c) square root of npq (d) square mpq

21. For continuous events _________________ distribution is used.

(a) Normal (b) Poisson (c) Binomial (d) none

22. Probability density function is associated with

(a) discrete cases (b) continuous cases (c) both (d) none

23. Probability density function is always

(a) greater than 0 (b) greater than equal to 0

(c) less than 0 (d) less than equal to 0

24. In continuous cases probability of the entire space is

(a) 0 (b) –1 (c) 1 (d) none

25. In discrete case the probability of the entire space is

(a) 0 (b) 1 (c) –1 (d) none

26. Binomial distribution is symmetrical if

(a) $p > q$ (b) $p < q$ (c) $p = q$ (d) none

27. The Poisson distribution tends to be symmetrical if the mean value is

(a) high (b) low (c) zero (d) none

28. The curve of ____________ distribution has single peak

(a) Poisson (b) Binomial (c) Normal (d) none

29. The curve of _________ distribution is unimodal and bell shaped with the highest point over the mean

(a) Poisson (b) Normal (c) Binomial (d) none

30. Because of the symmetry of Normal distribution the median and the mode have the ______ value as that of the mean

(a) greater (b) smaller (c) same (d) none

31. For a Normal distribution, the total area under the normal curve is

(a) 0 (b) 1 (c) 2 (d) –1

32. In Normal distribution the probability has the maximum value at the

(a) mode (b) mean (c) median (d) none

33. In Normal distribution the probability decreases gradually on either side of the mean but never touches the axis.

(a) True (b) false (c) both (d) none

34. Whatever may be the parameter of __________ distribution, it has same shape.

(a) Normal (b) Binomial (c) Poisson (d) none

35. In Standard Normal distribution

(a) mean=1, S.D=0 (b) mean=1, S.D=1
(c) mean = 0, S.D = 1 (d) mean=0, S. D=0

36. The no. of methods for fitting the normal curve is

(a) 1 (b) 2 (c) 3 (d) 4

37. ____________ distribution is symmetrical around t = 0

(a) Normal (b) Poisson (c) Binomial (d) t

38. As the degree of freedom increases, the ________ distribution approaches the Standard Normal distribution

(a) T (b) Binomial (c) Poisson (d) Normal

39. _________ distribution is asymptotic to the horizontal axis.

(a) Binomial (b) Normal (c) Poisson (d) t

40. ________ distribution has a greater spread than Normal distribution curve

(a) T (b) Binomial (c) Poisson (d) none

41. In Binomial Distribution if n is infinitely large, the probability p of occurrence of event' is close to _______ and q is close to _________

(a) 0 , 1 (b) 1 , 0 (c) 1 , 1 (d) none

42. Poisson distribution approaches a Normal distribution as n

(a) increase infinitely (b) decrease (c) increases moderately (d) none

43. If neither p nor q is very small but n sufficiently large, the Binomial distribution is very closely approximated by _________ distribution

(a) Poisson (b) Normal (c) t (d) none

44. For discrete random variable x, Expected value of x (i.e E(x)) is defined as the sum of products of the different values and the corresponding probabilities.

(a) True (b) false (c) both (d) none

45. For a probability distribution, ____________ is the expected value of x.

(a) median (b) mode (c) mean (d) none

46. _________ is the expected value of $(x - m)^2$, where m is the mean.

(a) median (b) variance (c) standard deviation (d) mode

47. The probability distribution of x is given below :

value of x :	1	0	Total
probability :	p	1–p	1

Mean is equal to

(a) p (b) 1–p (c) 0 (d) 1

48. For n independent trials in Binomial distribution the sum of the powers of p and q is always n , whatever be the no. of success.

(a) True (b) false (c) both (d) none

49. In Binomial distribution parameters are

(a) n and q (b) n and p (c) p and q (d) none

50. In Binomial distribution if n = 4 and p = 1/3 then the value of variance is

(a) 8/3 (b) 8/9 (c) 4/3 (d) none

51. In Binomial distribution if mean = 20, S.D.= 4 then q is equal to

(a) 2/5 (b) 3/8 (c) 1/5 (d) 4/5

52. If in a Binomial distribution mean = 20 , S.D.= 4 then p is equal to

(a) 2/5 (b) 3/5 (c) 1/5 (d) 4/5

53. If is a Binomial distribution mean = 20 , S.D.= 4 then n is equal to

(a) 80 (b) 100 (c) 90 (d) none

54. Poisson distribution is a ___________ probability distribution .

(a) discrete (b) continuous (c) both (d) none

55. No. of radio- active atoms decaying in a given interval of time is an example of

(a) Binomial distribution (b) Normal distribution
(c) Poisson distribution (d) None

56. __________ distribution is sometimes known as the "distribution of rare events".

(a) Poisson (b) Normal (c) Binomial (d) none

57. The probability that x assumes a specified value in continuous probability distribution is

(a) 1 (b) 0 (c) –1 (d) none

58. In Normal distribution mean, median and mode are

(a) equal (b) not equal (c) zero (d) none

59. In Normal distribution the quartiles are equidistant from

(a) median (b) mode (c) mean (d) none

60. In Normal distribution as the distance from the ___________ increases, the curve comes closer and closer to the horizontal axis.

(a) median (b) mean (c) mode (d) none

61. A discrete random variable x follows uniform distribution and takes only the values 6, 8, 11, 12, 17

The probability of P(x = 8) is

(a) 1/5 (b) 3/5 (c) 2/8 (d) 3/8

62. A discrete random variable x follows uniform distribution and takes the values 6, 9, 10, 11, 13

The probability of P(x = 12) is

(a) 1/5 (b) 3/5 (c) 4/5 (d) 0

63. A discrete random variable x follows uniform distribution and takes the values 6, 8, 11, 12, 17

The probability of P(x $\leq$ 12) is

(a) 3/5 (b) 4/5 (c) 1/5 (d) none

64. A discrete random variable x follows uniform distribution and takes the values 6, 8, 10, 12, 18

The probability of P(x < 12) is

(a) 1/5 (b) 4/5 (c) 3/5 (d) none

65. A discrete random variable x follows uniform distribution and takes the values 5, 7, 12, 15, 18

The probability of P(x > 10) is

(a) 3/5 (b) 2/5 (c) 4/5 (d) none

66. The probability density function of a continuous random variable is defined as follows :

f(x) = c when $-1 \leq x \leq 1$ = 0 , otherwise The value of c is

(a) 1 (b) –1 (c) 1/2 (d) 0

67. A continuous random variable x has the probability density fn.f(x) = ½ –ax , $0 \leq x \leq 4$ When 'a' is a constant. The value of ' a' is

(a) 7/8 (b) 1/8 (c) 3/16 (d) none

68. A continuous random variable x follows uniform distribution with probability density function
f(x) = ½, $(4 \leq x \leq 6)$. Then $P(4 \leq x \leq 5)$

(a) 0.1 (b) 0.5 (c) 0 (d) none

69. An unbiased die is tossed 500 times.The mean of the no. of 'Sixes' in these 500 tosses is

(a) 50/6 (b) 500/6 (c) 5/6 (d) none

70. An unbiased die is tossed 500 times. The Standard deviation of the no. of 'sixes' in these 500 tossed is

(a) 50/6 (b) 500/6 (c) 5/6 (d) none

71. A random variable x follows Binomial distribution with mean 2 and variance 1.2.Then the value of n is

(a) 8 (b) 2 (c) 5 (d) none

72. A random variable x follows Binomial distribution with mean 2 and variance 1.6 then the value of p is

(a) 1/5 (b) 4/5 (c) 3/5 (d) none

73. "The mean of a Binomial distribution is 5 and standard deviation is 3"

(a) True (b) false (c) both (d) none

74. The expected value of a constant k is the constant

(a) k (b) k–1 (c) k+1 (d) none

75. The probability distribution whose frequency function f(x)= 1/n($x = x_1, x_2, \ldots, x_n$) is known as

(a) Binomial distribution (b) Poisson distribution
(c) Uniform distribution (d) Normal distribution

76. Theoretical distribution is a

(a) Random distribution (b) Standard distribution
(c) Probability distribution (d) None

77. Probability function is known as

(a) frequency function (b) continuous function
(c) discrete function (d) none

78. The no. of points obtained in a single throw of an unbiased die follow :

(a) Binomial distribution (b) Poisson distribution
(c) Uniform distribution (d) None

79. The no of points in a single throw of an unbiased die has frequency function

(a) f(x)=1/4 (b) f(x)= 1/5 (c) f(x) = 1/6 (d) none

80. In uniform distribution random variable x assumes n values with

(a) equal probability (b) unequal probability (c) zero (d) none

81. In a discrete random variable x follows uniform distribution and assumes only the values 8 , 9, 11, 15, 18, 20. Then P(x = 9) is

(a) 2/6 (b) 1/7 (c) 1/5 (d) 1/6

82. In a discrete random variable x follows uniform distribution and assumes only the values 8 , 9, 11, 15, 18, 20. Then P(x = 12) is

(a) 1/6 (b) 0 (c) 1/7 (d) none

83. In a discrete random variable x follows uniform distribution and assumes only the values 8, 9, 11, 15, 18, 20. Then $P(x < 15)$ is

(a) 1/2 (b) 2/3 (c) 1 (d) none

84. In a discrete random variable x follows uniform distribution and assumes only the values 8 , 9, 11, 15, 18, 20. Then $P(x \leq 15)$ is

(a) 2/3 (b) 1/3 (c) 1 (d) none

85. In a discrete random variable x follows uniform distribution and assumes only the values 8, 9, 11, 15, 18, 20. Then $P(x > 15)$ is

(a) 2/3 (b) 1/3 (c) 1 (d) none

86. In a discrete random variable x follows uniform distribution and assumes only the values 8, 9, 11, 15, 18, 20. Then $P(|x - 14| < 5)$ is

(a) 1/3 (b) 2/3 (c) 1/2 (d) 1

87. When f(x)= 1/n then mean is

(a) (n–1)/2 (b) (n+1)/2 (c) n/2 (d) none

88. In continuous probability distribution $P(x \leq t)$ means

(a) Area under the probability curve to the left of the vertical line at t.

(b) Area under the probability curve to the right of the vertical line at t.

(c) both (d) none

89. In continuous probability distribution F(x) is called.

(a) frequency distribution function (b) cumulative distribution function
(c) probability density function (d) none

90. The probability density function of a continuous random variable is
$y = k(x-1)$, ($1 \le x \le 2$) then the value of the constant k is

(a) –1 (b) 1 (c) 2 (d) 0

ANSWERS

1	(c)	2	(a)	3	(b)	4	(b)	5	(a)
6	(b)	7	(a)	8	(b)	9	(a)	10	(b)
11	(b)	12	(a)	13	(b)	14	(b)	15	(b)
16	(c)	17	(a)	18	(c)	19	(b)	20	(a)
21	(a)	22	(b)	23	(b)	24	(c)	25	(b)
26	(c)	27	(a)	28	(c)	29	(b)	30	(c)
31	(b)	32	(b)	33	(a)	34	(a)	35	(c)
36	(b)	37	(d)	38	(a)	39	(d)	40	(a)
41	(a)	42	(a)	43	(b)	44	(a)	45	(c)
46	(b)	47	(a)	48	(a)	49	(b)	50	(b)
51	(d)	52	(c)	53	(b)	54	(a)	55	(c)
56	(a)	57	(b)	58	(a)	59	(c)	60	(b)
61	(a)	62	(d)	63	(b)	64	(c)	65	(a)
66	(c)	67	(b)	68	(b)	69	(b)	70	(a)
71	(c)	72	(a)	73	(b)	74	(a)	75	(c)
76	(c)	77	(a)	78	(c)	79	(c)	80	(a)
81	(d)	82	(b)	83	(a)	84	(a)	85	(b)
86	(c)	87	(b)	88	(a)	89	(b)	90	(c)