230

About the Author

Prof. T. Veerarajan, Professor of Mathematics is currently heading the Department of Science and Humanities, Sree Sowdambika College of Engineering, Aruppukottai, Tamil Nadu. A Gold Medalist from Madras University, he has had a brilliant academic career all through. He has more than 42 years of teaching experience at undergraduate and postgraduate levels in various established engineering colleges in Tamil Nadu including Anna University, Chennai.

Other books by Prof. T Veerarajan in ASCENT Series

Engineering Mathematics, 2e (For First Year)
[ISBN: 0-07-053483-7]

Engineering Mathematics, 2e (For Third Semester)
[ISBN: 0-07-049501-7]

Trigonometry, Algebra and Calculus
[ISBN: 0-07-053507-8]

Analytical Geometry, Real and Complex Analysis
[ISBN: 0-07-053489-6]

Numerical Methods: 2E
[ISBN: 0-07-060161-5]

PROBABILITY, STATISTICS AND RANDOM PROCESSES

Second Edition

T Veerarajan

Professor of Mathematics and Head
Science and Humanities Department
Sree Sowdambika College of Engineering
Aruppukottai, Tamil Nadu

Tata McGraw-Hill Publishing Company Limited
NEW DELHI

McGraw-Hill Offices
New Delhi New York St Louis San Francisco Auckland Bogotá Caracas
Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal
San Juan Santiago Singapore Sydney Tokyo Toronto

7. $\text{Var}(X) = 68 - 64 = 4$

$p\{8 - c < X < 8 + c\} \geq 1 - \dfrac{4}{c^2}$; $c = 3$

Reqiured Probability $\geq \dfrac{5}{9}$

11. $\{E(XY)^2 \leq E(X^2)\,E(Y^2)$: Replace X by $X - \mu_x$ and Y by $Y - \mu_y$

$$\frac{\left[E(X-\mu_x)(Y-\mu_y)\right]^2}{E(X-\mu_x)^2\,E(Y-\mu_y)^2} \leq 1;\ p_{xy}^2 \leq 1;\ |p_{xy}| \leq 1$$

12. By Tchebycheff's inequality, $P\{|X - \mu| \leq \sqrt{3}\,\sigma\} \geq 1 - \dfrac{1}{3} = \left(\dfrac{2}{3}\right)$. Since $0.5 < \dfrac{2}{3}$, such a RV does not exist.
13. 21/25
14. 4/9; 0.134
15. 1/4; 1/4
16. 1/3
17. 19/24
20. 250
22. $c = 1$

Exercise 4(F)

8. If $X_1, X_2, \cdots, X_n$ be a sequence of independent and identically distributed RVs with $E(X_i) = \mu$ and $\text{Var}(X_i) = \sigma^2$ and if $\overline{X} = \dfrac{1}{n}(X_1 + X_2 + \cdots + X_n)$, then $\overline{X}$ follows $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$ as $n \to \infty$. This result is used in theory of sampling.
9. 0.9544
10. 41
11. 0.1112; 0.1915
12. 0.352
13. 0.1814
14. C.L.T does not hold good

Chapter **5**

Some Special Probability Distributions

Introduction

In Chapter 2, we have just stated the definitions of certain special probability distributions, both discrete and continuous. While constructing probabilistic models for observable phenomena, certain probability distributions arise more frequently than do others. We treat such distributions that play important roles in many engineering applications as special probability distributions. In this chapter we shall discuss a number of discrete as well as continuous (random variables) distributions in considerable detail.

Special Discrete Distributions

1. Binomial distribution

Definition: Let A be some event associated with a random experiment E, such that $P(A) = p$ and $P(\overline{A}) = 1 - p = q$. Assuming that p remains the same for all repetitions, if we consider n independent repetitions (or trials) of E and if the random variable (RV) X denotes the number of times the event A has occurred, then X is called a *binomial random variable* with parameters n and p or we say that X follows a *binomial distribution* with parameters n and p, or symbolically $B(n, p)$. Obviously the possible values that X can take, are $0, 1, 2, \ldots, n$.

By the theorem under Bernoulli's trials in chapter 1, the probability mass function of a binomial RV is given by

$$P(X = r) = nC_r\,p^r q^{n-r};\ r = 0, 1, 2, \ldots, n \text{ where } p + q = 1$$

Note

(i) *Binomial distribution is a legitimate probability distribution since*

$$\sum_{r=0}^{n} P(X=r) = \sum_{r=0}^{n} nC_r q^{n-r} p^r$$
$$= (q+p)^n = 1$$

(2) *The name 'binomial distribution' is given since the probabilities* $nC_r q^{n-r} p^r$ $(r = 0, 1, 2, \ldots, n)$ *are the successive terms in the expansion of the binomial expression* $(q+p)^n$.

(3) *If we assume that n trials constitute a set and if we consider N sets, the frequency function of the binomial distribution is given by* $f(r) = N\,p(r) = N \cdot nC_r q^{n-r} p^r$, $r = 0, 1, 2, \ldots, n$. *In other words, the number of sets in which we get exactly r successes (the occurrences of the event A)* $= N \cdot nC_r q^{n-r} p^r$; $r = 0, 1, 2, \ldots, n$.

Mean and Variance of the Binomial Distribution

We have already found out $E(X)$ and Var(X) for the binomial distribution $B(n, p)$ using the moment generating function in Example 3 in Worked Example 4 (b). Here we shall find them directly using the definitions of $E(X)$ and Var(X).

$$E(X) = \sum_r x_r p_r$$
$$= \sum_{r=0}^{n} r \cdot nC_r p^r q^{n-r}$$
$$= \sum_{r=0}^{n} r \cdot \frac{n!}{r!(n-r)!} p^r q^{n-r} \qquad (1)$$
$$= np \cdot \sum_{r=1}^{n} \frac{(n-1)!}{(r-1)!\{(n-1)-(r-1)\}!} p^{r-1} q^{(n-1)-(r-1)}$$
$$= np \sum_{r=1}^{n} (n-1)C_{r-1} \cdot p^{r-1} \cdot q^{(n-1)-(r-1)}$$
$$= np\,(q+p)^{n-1}$$
$$= np \qquad (2)$$

$$E(X^2) = \sum_r x_r^2 p_r = \sum_0^n r^2 p_r$$
$$= \sum_{r=0}^{n} \{r(r-1)+r\} \frac{n!}{r!(n-r)!} p^r q^{n-r}$$
$$= n(n-1)p^2 \sum_{r=2}^{n} (n-2)C_{r-2} p^{r-2} q^{n-r} + np, \qquad \text{[by (1) and (2)]}$$
$$= n(n-1)\,p^2 (q+p)^{n-2} + np$$

$$= n(n-1)\,p^2 + np$$
$$\text{Var}(X) = E(X^2) - \{E(X)\}^2$$
$$= n(n-1)\,p^2 + np - n^2p^2$$
$$= np\,(1-p)$$
$$= npq$$

Recurrence Formula for the Central Moments of the Binomial Distribution

By definition, the kth order central moment μ_k is given by $\mu_k = E\{X - E(X)\}^k$. For the binomial distribution $B(n, p)$,

$$\mu_k = \sum_{r=0}^{n} (r-np)^k\, nC_r p^r q^{n-r} \qquad (1)$$

Differentiating (1) with respect to p, we get

$$\frac{d\mu_k}{dp} = \sum_{r=0}^{n} nC_r \left[-nk(r-np)^{k-1} \cdot p^r q^{n-r} + (r-np)^k \{rp^{r-1} q^{n-r} + (n-r)\,p^r q^{n-r-1}(-1)\}\right]$$
$$= -nk\,\mu_{k-1} + \sum_{r=0}^{n} nC_r (r-np)^k p^{r-1} q^{n-r-1} \{rq - (n-r)p\}$$
$$= -nk\,\mu_{k-1} + \sum_{r=0}^{n} nC_r (r-np)^k p^{r-1} q^{n-r-1} (r-np) \quad (\because p+q=1)$$
$$= -nk\,\mu_{k-1} + \frac{1}{pq} \sum_{r=0}^{n} nC_r p^r q^{n-r} (r-np)^{k+1}$$
$$= -nk\,\mu_{k-1} + \frac{1}{pq}\,\mu_{k+1}$$

i.e.,
$$\mu_{k+1} = pq\left[\frac{d\mu_k}{dp} + nk\,\mu_{k-1}\right] \qquad (2)$$

Using recurrence relation (2), we may compute moments of higher order, provided we know moments of lower order.

Putting $k = 1$ in (2), we get

$$\mu_2 = pq\left[\frac{d\mu_1}{dp} + n\mu_0\right]$$
$$= npq \quad (\because \mu_0 = 1 \text{ and } \mu_1 = 0)$$

Putting $k = 2$ in (2), we get

$$\mu_3 = pq\left[\frac{d}{dp}\,\mu_2 + 2n\,\mu_1\right]$$

$$= pq \frac{d}{dp}[np(1-p)]$$
$$= npq[1-2p] = npq(q-p)$$

Putting $k = 3$ in (2), we get

$$\mu_4 = pq\left[\frac{d}{dp}\mu_3 + 3n\mu_2\right]$$
$$= npq\left[\frac{d}{dp}\{p(1-p)(1-2p)\} + 3npq\right]$$
$$= npq[1-6p+6p^2+3npq]$$
$$= npq[1-6pq+3npq]$$
$$= npq[1+3pq(n-2)]$$

Note *μ_2 is the variance, μ_3 is a measure of skewness and μ_4 is a measure of kurtosis. Sometimes the coefficients β_1 and β_2 are used to measure skewness and kurtosis respectively,*

where $\mu_1 = \dfrac{\mu_3^2}{\mu_2^3}$ *and* $\mu_2 = \dfrac{\mu_4}{\mu_2^2}$.

2. Poisson distribution

Definition: If X is a discrete RV that can assume the values 0, 1, 2, …, such that its probability mass function is given by

$$P(X=r) = \frac{e^{-\lambda}\lambda^r}{r!}; \quad r = 0, 1, 2, \ldots; \lambda > 0$$

then X is said to follow a *Poisson distribution* with parameter λ or symbolically X is said to follow $P(\lambda)$.

(**Note:** Poisson distribution is a legitimate probability distribution, since

$$\sum_{r=0}^{\infty} P(x=r) = \sum_{r=0}^{\infty}\frac{e^{-\lambda}\lambda^r}{r!}$$
$$= e^{-\lambda}e^{\lambda} = 1)$$

Poisson Distribution as Limiting Form of Binomial Distribution

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

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

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

(iii) $np (= \lambda)$ is finite or $p = \dfrac{\lambda}{n}$ and $q = 1 - \dfrac{\lambda}{n}$, where λ is a positive real number.

Proof

If X is a binomially distributed RV with parameters n and p, then

$$P(X=r) = nC_r p^r q^{n-r}; \; r = 0, 1, 2, \ldots, n.$$
$$= \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}p^r(1-p)^{n-r}$$
$$= \frac{n(n-1)\cdots(n-r+1)}{r!}\left(\frac{\lambda}{n}\right)^r\left(1-\frac{\lambda}{n}\right)^{n-r}$$

(on putting $p = \dfrac{\lambda}{n}$)

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

$$\therefore \lim_{\substack{n\to\infty \\ np=\lambda=\text{finite}}}[P(X=r)] = \frac{\lambda^r}{r!}\lim_{n\to\infty}\left[\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{r-1}{n}\right)\right]$$
$$\cdot \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-r}\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n$$
$$= \frac{\lambda^r}{r!}e^{-\lambda} \quad \left[\because \lim_{n\to\infty}\left(1-\frac{k}{n}\right) = 1, \text{ when } k \text{ is finite,}\right.$$
$$\left.\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-r} = 1 \text{ and } \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}\right]$$

But $P(X=r) = \dfrac{e^{-\lambda}\cdot\lambda^r}{r!}$, $\quad r = 0, 1, 2, \ldots$

is the probability mass function of the Poisson random variable.

Thus the limit of the binomial RV (distribution) is the Poisson RV (distribution).

Note *(1) The above result means that we may compute binomial probabilities approximately by using the corresponding Poisson probabilities, whenever n is large and p is small.*

(2) When an event occurs rarely, the number of occurrences of such an event may be assumed to follow a Poisson distribution. The following are some of the examples, which may be analysed using Poisson distribution:

(i) the number of alpha particles emitted by a radioactive source in a given time interval

(ii) the number of telephone calls received at a telephone exchange in a given time interval

(iii) the number of defective articles in a packet of 100

(iv) the number of printing errors at each page of a book

(v) the number of road accidents reported in a city per day.

Mean and Variance of Poisson Distribution

We have already found out $E(X)$ and $\mathrm{Var}(X)$ for the Poisson distribution $P(\lambda)$, using the characteristic function in Example 4 in Worked Example 4(b). Also, since the Poisson distribution is the limit of binomial distribution, the mean and variance of the Poisson distribution may be obtained as the limits of those of binomial distribution when $n \to \infty$.
i.e., if X is the Poisson RV

$$E(X) = \lim_{\substack{n\to\infty \\ np=\lambda}} (np) = \lambda$$

and $$\mathrm{Var}(X) = \lim_{\substack{p\to 0 \\ np=\lambda}} (npq) = \lim_{p\to 0} [\lambda(1-p)] = \lambda.$$

Now we shall find $E(X)$ and $\mathrm{Var}(X)$ for the Poisson distribution directly using the definitions.

$$E(X) = \sum_r x_r p_r$$

$$= \sum_{r=0}^{\infty} r \frac{e^{-\lambda}\cdot\lambda^r}{r!} \qquad (1)$$

$$= \lambda e^{-\lambda} \sum_{r=1}^{\infty} \frac{\lambda^{r-1}}{(r-1)!}$$

$$= \lambda e^{-\lambda} e^{\lambda} = \lambda \qquad (2)$$

$$E(X^2) = \sum_r x_r^2 p_r$$

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

$$= \lambda^2 e^{-\lambda} \sum_{r=2}^{\infty} \frac{\lambda^{r-2}}{(r-2)!} + \lambda \qquad \text{[by (1) and (2)]}$$

$$= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda$$

$$\mathrm{Var}(X) = E(X^2) - \{E(X)\}^2$$
$$= \lambda^2 + \lambda - \lambda^2 = \lambda$$

Recurrence Formula for the Central Moments of the Poisson Distribution

By definition, the kth order central moment μ_k is given by $\mu_k = E\{X - E(X)\}^k$.
For the Poisson distribution $P(\lambda)$,

$$\mu_k = \sum_{r=0}^{\infty} (r-\lambda)^k e^{-\lambda} \frac{\lambda^r}{r!} \qquad (1)$$

Differentiating (1) with respect to λ, we get

$$\frac{d\mu_k}{d\lambda} = \sum_{r=0}^{\infty} \frac{1}{r!} [-k(r-\lambda)^{k-1} e^{-\lambda}\lambda^r + (r e^{-\lambda}\lambda^{r-1} - e^{-\lambda}\lambda^r)(r-\lambda)^k]$$

$$= -k\,\mu_{k-1} + \sum_{r=0}^{\infty} \frac{1}{r!} e^{-\lambda}\lambda^{r-1}(r-\lambda)^{k+1}$$

$$= -k\mu_{k-1} + \frac{1}{\lambda}\mu_{k+1}$$

i.e., $$\mu_{k+1} = \lambda\left(\frac{d\mu_k}{d\lambda} + k\,\mu_{k-1}\right) \qquad (2)$$

Using recurrence relation (2), we may compute moments of higher order, provided we know moment of lower order.
Putting $k = 1$ in (2), we get

$$\mu_2 = \lambda\left(\frac{d\mu_1}{d\lambda} + \mu_0\right)$$
$$= \lambda \ (\because \mu_0 = 1 \text{ and } \mu_1 = 0)$$

Putting $k = 2$ in (2), we get

$$\mu_3 = \lambda\left(\frac{d\mu_2}{d\lambda} + 2\mu_1\right) = \lambda.$$

Putting $k = 3$ in (2), we get

$$\mu_4 = \lambda\left(\frac{d\mu_3}{d\lambda} + 3\mu_2\right) = \lambda(3\lambda+1)$$

Note *The interesting property of the Poisson distribution is the equality of its mean, variance and third-order central moment.*

3. Geometric distribution

Definition: Let the RV X denote the number of trials of a random experiment required to obtain the first success (occurrence of an event A). Obviously X can assume the values 1, 2, 3, ...

Now $X = r$, if and only if the first $(r-1)$ trials result in failure (occurrence of $\bar{A}$) and the rth trial results in success (occurrence of A). Hence

$$P(X=r) = q^{r-1}p; \quad r = 1, 2, 3, \ldots, \infty$$

where

$$P(A) = p \text{ and } P(\bar{A}) = q.$$

If X is a discrete RV that can assume the values 1, 2, 3, ..., ∞ such that its probability mass function is given by

$$P(X=r) = q^{r-1}p; \; r = 1, 2, \ldots, \infty \text{ where } p+q=1$$

then X is said to follow *a geometric distribution.*

Worked Example 5(A)

Example 1

Out of 800 families with 4 children each, how many families would be expected to have (i) 2 boys and 2 girls, (ii) at least 1 boy, (iii) at most 2 girls and (iv) children of both sexes. Assume equal probabilities for boys and girls.

Considering each child as a trial, $n = 4$. Assuming that birth of a boy is a success, $p = \frac{1}{2}$ and $q = \frac{1}{2}$. Let X denote the number of successes (boys).

(i) $P(\text{2 boys and 2 girls}) = P(X = 2)$

$$= 4C_2 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^{4-2}$$

$$= 6 \times \left(\frac{1}{2}\right)^4 = \frac{3}{8}$$

$\therefore$ No. of families having 2 boys and 2 girls

$= N \cdot (P(X = 2)$ (where N is the total no. of families considered)

$$= 800 \times \frac{3}{8}$$

$= 300.$

(ii) $P(\text{at least 1 boy}) = P(X \geq 1)$

$$= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

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

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

$$= 1 - \frac{1}{16} = \frac{15}{16}$$

$\therefore$ No. of families having at least 1 boy

$$= 800 \times \frac{15}{16} = 750.$$

(iii) $P(\text{at most 2 girls}) = P(\text{exactly 0 girl, 1 girl or 2 girls})$

$$= P(X = 4, X = 3 \text{ or } X = 2)$$

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

$$= 1 - \left\{ 4C_0 \cdot \left(\frac{1}{2}\right)^4 + 4C_1 \cdot \left(\frac{1}{2}\right)^4 \right\}$$

$$= \frac{11}{16}$$

$\therefore$ No. of families having at most 2 girls

$$= 800 \times \frac{11}{16} = 550.$$

(iv) $P(\text{children of both sexes})$

$$= 1 - P(\text{children of the same sex})$$

$$= 1 - \{P(\text{all are boys}) + P(\text{all are girls})\}$$

$$= 1 - \{P(X = 4) + P(X = 0)\}$$

$$= 1 - \left\{ 4C_4 \cdot \left(\frac{1}{2}\right)^4 + 4C_0 \cdot \left(\frac{1}{2}\right)^4 \right\}$$

$$= 1 - \frac{1}{8} = \frac{7}{8}$$

$\therefore$ No. of families having children of both sexes

$$= 800 \times \frac{7}{8} = 700.$$

Example 2

An irregular 6-faced die is such that the probability that it gives 3 even numbers in 5 throws is twice the probability that it gives 2 even numbers in 5 throws. How many sets of exactly 5 trials can be expected to give no even number out of 2500 sets?

Let the probability of getting an even number with the unfair die be p.

Let X denote the number of even numbers obtained in 5 trials (throws).

Given: $P(X = 3) = 2 \times P(X = 2)$

i.e., $5C_3\, p^3 q^2 = 2 \times 5C_2\, p^2 q^3$

i.e., $p = 2q = 2(1 - p)$

$\therefore$ $3p = 2$ or $p = \frac{2}{3}$ and $q = \frac{1}{3}$

Now $P(\text{getting no even number})$

$$= P(X = 0)$$

$$= 5C_0 \cdot p^0 \cdot q^5 = \left(\frac{1}{3}\right)^5 = \frac{1}{243}$$

$\therefore$ Number of sets having no success (even number) out of N sets $= N \times P(X = 0)$

$\therefore$ Required number of sets $= 2500 \times \frac{1}{243}$

$= 10$, nearly

Example 3

The probability of a successful rocket launching is p. If launching attempts are made until 3 successful launchings have occurred, what is the probability that exactly 5 attempts will be necessary? What is the probability that fewer than 5 attempts will be necessary?

If launching attempts are made until 3 consecutive successful launchings occur, what are the probabilities? [See example (1) in section 1 (c)]

(i) Exactly 5 attempts will be required to get 3 successes (successful launchings of rockets), if 2 successes occur in the first 4 attempts and third success occurs in the fifth attempt.

$\therefore P$(exactly 5 attempts are required)

$$= P(\text{2 successes in 4 attempts}) \times P(\text{success in the single 5th attempt})$$

$$= 4C_2\, p^2q^2 \times p$$

[$\because$ The no. of successes in the 4 independent attempts follow $B(4, p)$]

$$= 6\, p^3q^2$$

(ii) $P\{$fewer than 5 attemps are required$\}$

$$= P\{\text{exactly 3 or 4 attempts are required}\}$$

$$= [P\{\text{2 successes in the first 2 attempts}\} \times P(\text{success in the 3rd attempt})]$$

$$+ [P\{\text{2 successes in the first 3 attempts}\} \times P(\text{success in the 4th attempt})]$$

$$= 2C_2\, p^2q^0 \times p + 3C_2\, p^2q^1 \times p$$

$$= p^3 + 3p^3q = p^3(1 + 3q)$$

(iii) Five attempts will be required to get 3 consecutive successes, if the first 2 attempts result in failures and the last 3 attempts result in successes.

$\therefore$ Required probability $= q \cdot q \cdot p \cdot p \cdot p = p^3q^2$

(iv) Three attempts will be required to get 3 consecutive successes, if each attempt results in a success.

$\therefore$ Probability for this $= p^3$.

Four attempts will be required to get 3 consecutive successes, if the first attempt results in a failure and the remaining attempts result in a success each.

$\therefore$ Probability for this $= qp^3$

$\therefore$ P$\{$fewer than 5 attempts are required$\}$

$$= p^3 + qp^3 = p^3(1 + q).$$

Example 4

A communication system consists of n components, each of which will independently function with probability p. The total system will be able to operate effectively if at least one-half of its components function. For what values of p is a 5-component system more likely to operate effectively than a 3-component system?

Since the probability p of functioning of every component is a constant and the n components function independently, the no. of components X that function follow a binomial distribution with parameters n and p.

$$\therefore \quad P(X = r) = nC_r\, p^rq^{n-r};\ r = 0, 1, 2, \ldots, n.$$

P(5-component system functions effectively)

$$= P(X = 3 \text{ or } 4 \text{ or } 5)$$

$$= \sum_{r=3}^{5} 5C_r\, p^rq^{5-r} \quad (\because n = 5)$$

P(3-component system functions effectively)

$$= P(X = 2 \text{ or } 3)$$

$$= \sum_{r=2}^{3} 3C_r\, p^rq^{3-r} \quad (\because n = 3)$$

5-component system will function more effectively than the 3-component system,

if $\quad \sum_{r=3}^{5} 5C_r\, p^rq^{5-r} \geq \sum_{r=2}^{3} 3C_r\, p^rq^{3-r}$

i.e., $\quad 10\, p^3q^2 + 5p^4q + p^5 \geq 3p^2q + p^3$

i.e., $\quad 10p^3(1 - 2p + p^2) + 5p^4(1 - p) + p^5 \geq 3p^2(1 - p) + p^3$

i.e., $\quad 3p^2(2p^3 - 5p^2 + 4p - 1) \geq 0$

i.e., $\quad 3p^2(p - 1)^2(2p - 1) \geq 0$

i.e., $\quad (2p - 1) \geq 0, \quad [\text{since } 3p^2(p - 1)^2 \geq 0]$

i.e., $\quad p \geq \dfrac{1}{2}.$

Example 5

If the probability that a child is a boy is p, where $0 < p < 1$, find the expected number of boys in a family with n children, given that there is at least one boy.

Let X be the no. of boys (successes) out of n children (trials)

Then X follows a $B(n, p)$.

Required to find $E\{X/X \geq 1\}$.

$$E\{X/X \geq 1\} = \sum_r r \cdot P(X = r/X \geq 1)$$

$$= \sum_{r=1}^{n} r \cdot \frac{P(X = r)}{P(X \geq 1)}$$

$$= \sum_{r=1}^{n} \frac{r \cdot nC_r p^r q^{n-r}}{1 - P(X=0)}$$

$$= \frac{1}{1-q^n} \sum_{r=0}^{n} r \cdot nC_r p^r q^{n-r}$$

$$= \frac{np}{1-q^n}.$$

Example 6

Two dice are thrown 120 times. Find the average number of times in which the number on the first dice exceeds the number on the second dice.

The number on the first dice exceeds that on the second die, in the following combinations:

(2, 1); (3, 1), (3, 2); (4, 1), (4, 2), (4, 3); (5, 1), (5, 2), (5, 3); (5, 4); (6, 1), (6, 2), (6, 3), (6, 4), (6, 5),

where the numbers in the parentheses represent the numbers in the first and second dice respectively.

$\therefore$ P(success) = P(no. in the first die exceeds the no. in the second die)

$$= \frac{15}{36} = \frac{5}{12}$$

This probability remains the same in all the throws that are independent.

If X is the no. of successes, then X follows a binomial distribution with parameters $n\ (= 120)$ and $p\left(= \frac{5}{12}\right)$.

$\therefore$ $E(X) = np = 120 \times \frac{5}{12} = 50$

Example 7

Fit a binomial distribution for the following data:

x:	0	1	2	3	4	5	6	Total
f:	5	18	28	12	7	6	4	80

Fitting a binomial distribution means assuming that the given distribution is approximately binomial and hence finding the probability mass function and then finding the theoretical frequencies.

To find the binomial frequency distribution $N(q + p)^n$, which fits the given data, we require N, n and p. We assume N = total frequency = 80 and n = no. of trials = 6 from the given data.

To find p, we compute the mean of the given frequency distribution and equate it to np (mean of the binomial distribution).

x:	0	1	2	3	4	5	6	Total
f:	5	18	28	12	7	6	4	80
fx:	0	18	56	36	28	30	24	192

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{192}{80} = 2.4$$

i.e., $np = 2.4$ or $6p = 2.4$

$\therefore$ $p = 0.4$ and $q = 0.6$

If the given distribution is nearly binomial, the theoretical frequencies are given by the successive terms in the expansion of $80(0.6 + 0.4)^6$. Thus we get,

x:	0	1	2	3	4	5	6
Theoretical f:	3.73	14.93	24.88	22.12	11.06	2.95	0.33

Converting these values into whole numbers consistent with the condition that the total frequency is 80, the corresponding binomial frequency distribution is as follows:

x:	0	1	2	3	4	5	6	Total
f:	4	15	25	22	11	3	0	80

Example 8

The number of monthly breakdowns of a computer is a RV having a Poisson distribution with mean equal to 1.8. Find the probability that this computer will function for a month

(a) without a breakdown,
(b) with only one breakdown and
(c) with atleast one breakdown.

Let X denote the number of breakdowns of the computer in a month.

X follows a Poisson distribution with mean (parameter) $\lambda = 1.8$.

$$\therefore P\{X = r\} = \frac{e^{-\lambda} . \lambda^r}{r!} = \frac{e^{-1.8} . (1.8)^r}{r!}$$

(a) $P(X = 0) = e^{-1.8} = 0.1653$
(b) $P(X = 1) = e^{-1.8}(1.8) = 0.2975$
(c) $P(X \geq 1) = 1 - P(X = 0) = 0.8347$

Example 9

It is known that the probability of an item produced by a certain machine will be defective is 0.05. If the produced items are sent to the market in packets of 20, find the number of packets containing at least, exactly and at most 2 defective items in a consignment of 1000 packets using (i) binomial distribution and (ii) Poisson approximation to binomial distribution.

Use of binomial distribution

p = probability that an item is defective = 0.05, $q = 0.95$ and n = No. of independent items (trials) considered = 20.

Let X denote the number of defectives in the n items considered.

$$P(X = r) = nC_r\, p^r\, q^{n-r}$$

(i) $\therefore \quad P(X = 2) = 20\, C_2\, (.05)^2\, (0.95)^{18} = 0.1887$

If N is the number of sets (packets), each set (packet) containing 20 trials (items), then the number of sets containing exactly 2 successes (defectives) is given by

$$N(X = 2) = N \times P(X = 2) = 1000 \times 0.1887 = 189, \text{ nearly}$$

(ii) $P(\text{at least 2 defectives}) = P(X \geq 2)$

$$= 1 - \{P(X = 0) + P(X = 1)\}$$
$$= 1 - [20C_0\, (0.05)^0\, (0.95)^{20} + 20C_1\, (0.05)^1\, (0.95)^{19}]$$
$$= 1 - [0.3585 + 0.3774]$$
$$= 0.2641$$

$\therefore \quad N(X \geq 2) = N \times P(X \geq 2) = 1000 \times 0.2641 = 264$, nearly

(iii) $P(\text{at most 2 defectives}) = P(X \leq 2)$

$$= P(X = 0) + P(X = 1) + P(X = 2)$$
$$= \sum_{r=0}^{2} 20C_r\, (0.05)^r\, (0.95)^{20-r}$$
$$= 0.3585 + 0.3774 + 0.1887$$
$$= 0.9246$$

$\therefore \quad N(X \leq 2) = N \times P(X \leq 2) = 1000 \times 0.9246 = 925$, nearly

Use of Poisson distribution

Since $p = 0.05$ is very small and $n = 20$ is sufficiently large, binomial distribution may be approximated by Poisson distribution with parameter $\lambda = np = 1$.

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

(i) $P(X = 2) = \dfrac{e^{-1}}{2!} = 0.1839$

$\therefore N(X = 2) = 1000 \times 0.1839 = 184$, nearly

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

$$= 1 - \{e^{-1} + e^{-1}\} = 0.2642$$

$\therefore N(X \geq 2) = 1000 \times 0.2642 = 264$, nearly.

(iii) $P(X \leq 2) = \displaystyle\sum_{r=0}^{2} P(X = r) = \sum_{r=0}^{2} \frac{e^{-1}}{r!} = 0.9197$

$\therefore N(X \leq 2) = 920$, nearly.

Example 10

Prove the reproductive property of independent Poisson RVs. Hence find the probability of 5 or more telephone calls arriving in a 9 min period in a college switch-board, if the telephone calls that are received at the rate of 2 every 3 min follow a Poisson distribution.

Let X_1 and X_2 be independent RVs that follow Poisson distributions with parameters λ_1 and λ_2 respectively.

Let $\quad X = X_1 + X_2$

$$P\{X = n\} = P\{X_1 + X_2 = n\}$$
$$= \sum_{r=0}^{n} P\{X_1 = r\} \cdot P\{X_2 = n-r\},$$

(since X_1 and X_2 are independent)

$$= \sum_{r=0}^{n} \frac{e^{-\lambda_1} \cdot \lambda_1^r}{r!} \frac{e^{-\lambda_2} \cdot \lambda_2^{n-r}}{(n-r)!}$$
$$= \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} \lambda_1^r \cdot \lambda_2^{n-r}$$
$$= \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} \sum_{r=0}^{n} nC_r \cdot \lambda_1^r \cdot \lambda_2^{n-r}$$
$$= \frac{e^{-(\lambda_1 + \lambda_2)} \cdot (\lambda_1 + \lambda_2)^n}{n!}$$

Thus the sum of 2 independent Poisson variables with parameters λ_1 and λ_2 is also a Poisson variable with parameter $(\lambda_1 + \lambda_2)$.

This property, which can be extended to any finite number of independent Poisson variables is known as *the Reproductive Property* of Poisson RVs. [For an alternative proof, see example (9) in section 4(b)]

Let X_1, X_2, X_3 denote the number of telephone calls received in three consecutive 3-min periods.

Each of X_1, X_2, X_3 follows a Poisson distribution with parameter (mean) $\lambda = 2$.

$\therefore X = X_1 + X_2 + X_3$ follow a Poisson distrubution with parameter 6.

Clearly X represents the number of calls received in a 9-min period.

Now $P(X \geq 5) = 1 - P\{X \leq 4\}$

$$= 1 - \sum_{r=0}^{4} \frac{e^{-6} \cdot 6^r}{r!}$$

$$= 1 - (0.0025 + 0.0149 + 0.0446 + 0.0892 + 0.1339)$$

$$= 1 - 0.2851 = 0.7149$$

Example 11

If the number X of particles emitted during a 1-h period from a radioactive source has a Poisson distribution with parameter $\lambda = 4$ and that the probability that any emitted particle is recorded is $p = 0.9$, find the probability distribution of the number Y of the particles recorded in a 1-h period and hence the probability that no particle is recorded.

$$P\{Y = n\} = \sum_{r=0}^{\infty} P[X = n + r \text{ and } n \text{ of them are recorded}]$$

$$= \sum_{r=0}^{\infty} \frac{e^{-\lambda} \cdot \lambda^{n+r}}{(n+r)!} (n+r)C_n \cdot p^n q^r$$

$$= \sum_{r=0}^{\infty} \frac{e^{-\lambda} \cdot (\lambda p)^n}{(n+r)!} \frac{(n+r)!}{n!r!} (\lambda q)^r$$

$$= \frac{e^{-\lambda} \cdot (\lambda p)^n}{n!} \sum_{r=0}^{\infty} \frac{1}{r!} (\lambda q)^r$$

$$= \frac{e^{-\lambda} \cdot (\lambda p)^n}{n!} e^{\lambda q}$$

$$= \frac{e^{-\lambda(1-q)} \cdot (\lambda p)^n}{n!} = \frac{e^{-\lambda p} \cdot (\lambda p)^n}{n!}, \quad n = 0, 1, 2, \ldots, \infty.$$

Therefore Y, the number of recorded particles, follows a Poisson distribution with parameter λp.

Hence $P(Y = 0\} = e^{-\lambda p}$

$$= e^{-4 \times 0.9} = e^{-3.6}$$

$$= 0.0273.$$

Example 12

If X and Y are independent Poisson RVs, show that the conditional distribution of X, given the value of $X + Y$, is a binomial distribution.

Let X and Y follow Poisson distributions with parameters λ_1 and λ_2 respectively.

Now $P\{X = r/(X + Y) = n\}$

$$= \frac{P\{X = r \text{ and } (X+Y) = n\}}{P\{(X+Y) = n\}} = \frac{P[X = r; Y = n - r]}{P[X + Y = n]}$$

$$= \frac{P\{X = r\} \cdot P\{X = n - r\}}{P\{(X+Y) = n\}} \quad \text{(by independence of } X \text{ and } Y.)$$

$$= \frac{\{e^{-\lambda_1} \cdot \lambda_1^r / r!\}\{e^{-\lambda_2} \cdot \lambda_2^{n-r} / (n-r)!\}}{e^{-(\lambda_1 + \lambda_2)} \cdot (\lambda_1 + \lambda_2)^n / n!}$$

(by the reproductive property)

$$= \frac{n!}{r!(n-r)!} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^r \cdot \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-r}$$

$$= nC_r p^r q^{n-r}, \text{ where } p = \frac{\lambda_1}{\lambda_1 + \lambda_2} \text{ and } q = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

Hence the required result.

Example 13

Fit a Poisson distribution for the following distribution:

x:	0	1	2	3	4	5	Total
f:	142	156	69	27	5	1	400

Fitting a Poisson distribution for a given distribution means assuming that the given distribution is approximately Poisson and hence finding the probability mass function and then finding the theoretical frequencies.

To find the probability mass function

$$P\{X = r\} = \frac{e^{-\lambda} \cdot \lambda^r}{r!} \quad r = 0, 1, 2, \ldots, \infty$$

of the approximate Poisson distribution, we require λ, which is the mean of the Poisson distribution.

We find the mean of the given distribution and assume it as λ.

x :	0	1	2	3	4	5	Total
f :	142	156	69	27	5	1	400
fx:	0	156	138	81	20	5	400

$$\bar{x} = \frac{\sum f x}{\sum f} = \frac{400}{400} = 1 = \lambda$$

The theoretical frequencies are given by

$$\frac{N e^{-\lambda} \cdot \lambda^r}{r!} \quad \text{where } N = 400, \text{ obtained from the given distribution.}$$

$$= \frac{400\, e^{-1}}{r!}, \quad r = 0, 1, 2, \ldots, \infty$$

Thus, we get

	x:	0	1	2	3	4	5
Theoretical	f:	147.15	147.15	73.58	24.53	6.13	1.23

The theoretical frequencies for $x = 6, 7, 8, \ldots$ are very small and hence neglected.

Converting the theoretical frequencies into whole numbers consistent with the condition that the total frequency = 400, we get the following Poisson frequency distribution which fits the given distribution:

	x:	0	1	2	3	4	5
Theoretical	f:	147	147	74	25	6	1

Example 14

If the probability that an applicant for a driver's licence will pass the road test on any given trial is 0.8, what is the probability that he will finally pass the test (a) on the fourth trial and (b) in fewer than 4 trials?

Let X denote the number of trials required to achieve the first success. Then X follows a geometric distribution given by

$$P(X = r) = q^{r-1} p; \quad r = 1, 2, 3, \ldots, \infty$$

Here $p = 0.8$ and $q = 0.2$

(a) $P(X = 4) = 0.8 \times (0.2)^{4-1}$

$= 0.8 \times 0.008 = 0.0064.$

(b) $P(X < 4) = \sum_{r=1}^{3} 0.8 \times (0.2)^{r-1}$

$= 0.8\, [(0.2)^0 + (0.2)^1 + (0.2)^2 + (0.2)^3]$

$= 0.9984.$

Example 15

A and B shoot independently until each has hit his own target. The probabilities of their hitting the target at each shot are $\frac{3}{5}$ and $\frac{5}{7}$ respectively. Find the probability that B will require more shots than A.

Let X denote the number of trials required by A to get his first success. Then X follows a geometric distribution given by

$$P(X = r) = p_1 q_1^{\,r-1} = \frac{3}{5} \cdot \left(\frac{2}{5}\right)^{r-1}; \quad r = 1, 2, \ldots, \infty$$

Let Y denote the number of trials required by B to get his first success. Then Y follows a geometric distribution given by

$$P(Y = r) = p_2 \cdot q_2^{\,r-1} = \frac{5}{7} \cdot \left(\frac{2}{7}\right)^{r-1}; \quad r = 1, 2, \ldots, \infty$$

$P\{B$ requires more trials to get his first success than A requires to get his first success$\}$

$$= \sum_{r=1}^{\infty} P\{X = r \text{ and } Y = r+1 \text{ or } r+2, \ldots, \infty\}$$

$$= \sum_{r=1}^{\infty} [P\{X = r\} \cdot P\{Y = r+1 \text{ or } r+2, \ldots, \infty\}]$$

(by independence)

$$= \sum_{r=1}^{\infty} \frac{3}{5} \cdot \left(\frac{2}{5}\right)^{r-1} \cdot \sum_{k=1}^{\infty} \frac{5}{7} \cdot \left(\frac{2}{7}\right)^{r+k-1}$$

$$= \sum_{r=1}^{\infty} \frac{3}{7} \cdot \left(\frac{4}{35}\right)^{r-1} \sum_{k=1}^{\infty} \left(\frac{2}{7}\right)^{k}$$

$$= \sum_{r=1}^{\infty} \frac{3}{7} \cdot \left(\frac{4}{35}\right)^{r-1} \cdot \left(\frac{\frac{2}{7}}{1 - \frac{2}{7}}\right)$$

$$= \frac{6}{35} \sum_{r=1}^{\infty} \left(\frac{4}{35}\right)^{r-1} = \frac{6}{35} \cdot \frac{1}{1 - \frac{4}{35}} = \frac{6}{31}.$$

Example 16

A coin is tossed until the first head occurs. Assuming that the tosses are independent and the probability of a head occurring is p, find the value of p so that the probability that an odd number of tosses is required is equal to 0.6.

Can you find a value of p so that the probability is 0.5 that an odd number of tosses is required?

Let X denote the number of tosses required to get the first head (success). Then X follows a geometric distribution given by

$$P(X = r) = pq^{r-1}; \quad r = 1, 2, \ldots,$$

$\therefore P(X = \text{an odd number})\ P(X = 1 \text{ or } 3 \text{ or } 5, \ldots)$

$$= \sum_{r=1}^{\infty} P(X = 2r - 1) = \sum_{r=1}^{\infty} p\, q^{2r-2}$$

$$= \frac{p}{q^2} (q^2 + q^4 + q^6 + \ldots)$$

$\therefore$ Required probability

$$= 3C_1 p^2 q^2 + 4C_1 p^2 q^3 + 5C_1 p^2 q^4 + \ldots$$
$$= p^2 \{3q^2 + 4q^3 + 5q^4 + \ldots\}$$
$$= p^2 [(1-q)^{-2} - 1 - 2q]$$
$$= 1 - p^2 - 2p^2 q$$
$$= 1 - (0.02)^2 - 2 \times (0.02)^2 \times 0.98$$
$$= 0.9988$$

Example 23

Find the probability that a person tossing 3 fair coins get either all heads or all tails for the second time on the fifth trial.

$$p = P\text{ (3 heads or 3 tails in tossing 3 coins)}$$
$$= \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

and $\quad q = \frac{3}{4}$.

5th trial must result in the 2nd success.

$\therefore$ The first 4 trials must have resulted in 1 success and 3 failures.

$\therefore$ Required probability $= 4C_1\, p^1\, q^3 \times p$

$$= 4 \times \left(\frac{1}{4}\right)^2 \times \left(\frac{3}{4}\right)^3$$
$$= \frac{27}{256}$$

Exercise 5(A)

Part A (Short answer questions)

1. The mean and variance of a binomial distribution are 4 and $\frac{4}{3}$ respectively. Find $P(X \geq 1)$, if $n = 6$
2. If the recurrence relation for the central moments of the binomial distribution is $\mu_{r+1} = pq\left(nr\,\mu_{r-1} + \frac{d\mu_r}{dp}\right)$, find the value of β_1.
3. In 256 sets of 8 tosses of a coin, in how many sets one may expect heads and tails in equal numbers?
4. An experiment succeeds twice as often as it fails. Find the chance that in the next 4 trials, there shall be at least one success.
5. In a family of 4 children, what is the probability that there will be at least 1 boy and at least 1 girl, assuming equal probability for boys and girls.
6. If X has the distribution $B\left(25, \frac{1}{5}\right)$, find $P(X < \mu - 2\sigma)$ where μ and σ^2 are the mean and variance of the distribution.

7. Show that the largest value of the variance of a binomial distribution is $\frac{n}{4}$.
8. Find the mean and SD of the distribution whose moment generating functions is $(0.4\, e^t + 0.6)$.
9. When will the sum of 2 binomial variates having distributions $B(n_1, p_1)$ and $B(n_2, p_2)$ be also a binomial variate?
10. If X follows $B\left(3, \frac{1}{3}\right)$ and Y follows $B\left(5, \frac{1}{3}\right)$, find $P(X + Y \geq 1)$.
11. Write down the pmf of the Poisson distribution which is approximately equivalent to $B(100, 0.02)$.
12. If X is a Poisson variate such that $2P(X = 0) + P(X = 2) = 2P(X = 1)$, find $E(X)$.
13. If X is a poisson variate such that $E(X^2) = 6$, find $E(X)$
14. If X is a Poisson variate such that $P(X = 0) = 0.5$, find $\text{Var}(X)$.
15. If X is a Poisson variate with parameter $\lambda > 0$, prove that $E(X^2) = \lambda E(X + 1)$.
16. If X is a Poisson variate with parameter λ, prove that
$$E(X \text{ is even}) = \frac{1}{2}(1 + e^{-2\lambda}).$$
17. If the MGF of a discrete RV X is $e^{4(e^t - 1)}$, find $P(X = \mu + \sigma)$, where μ and σ are the mean and SD of X.
18. If X and Y are independent identical Poisson variates with mean 1, find $P(X + Y = 2)$.
19. Find the mean and variance of the discrete probability distribution given by
$$P(X = r) = e^{-t}(1 - e^{-t})^{r-1} \quad r = 1, 2, 3, \ldots, \infty.$$
20. If X is a geometric variate, taking values $1, 2, 3, \ldots, \infty$ find $P(X \text{ is odd})$.
21. Find the mean and variance of the distribution given by
$$P(X = r) = \frac{2}{3^r},\; r = 1, 2, \ldots, \infty.$$
22. For the geometric distribution of X, which represents the number of Bernoulli's trials required to get the first success $\text{Var}(X) = 2E(X)$. Find the pmf of the distribution.
23. Find the MGF of the geometric distribution, given by $P(X = r) = q^{r-1} p$, $r = 1, 2, \ldots, \infty$.
24. If the MGF of a discrete RV X, taking values $1, 2, \ldots, \infty$, is $e^t(5 - 4e^t)^{-1}$, find the mean and variance of X.
25. Define hypergeometric distribution and give an example for the situation where it arises.
26. Write down the mean and variance of the hypergeometric distribution given by
$$P(X = r) = kC_r(N - k)\, C_{n-r}/NC_n,\; r = 0, 1, 2, \ldots.$$

27. State the conditions under which the hypergeometric distribution tends to the binomial distribution. Hence deduce the mean and variance of the binomial distribution from those of the hypergeometric distribution.

Part B

28. It is known that diskettes produced by a certain company are defective with a probability 0.01 independently of each other. The company markets diskettes in packages of 10 and offers a money-back guarantee that atmost 1 of the 10 diskettes is defective. What proportion of diskettes are returned? If someone buys 3 diskettes, what is the probability that he will return exactly one of them?
29. Assuming that half the population is vegetarian and that 100 investigators each take 10 individuals to see whether they are vegetarians, how many would you expect to report that 3 people or less were vegetarians?
30. Show that, if 2 symmetric binomial distributions of degree n are so superposed that the r th term of the one coincides with the $(r + 1)$th term of the other, the distribution formed by adding superposed terms is a binomial distribution of degree $(n + 1)$.
[**Hint:** $nC_{r-1} + nC_r = (n + 1)C_r$]
31. A factory produces 10 articles daily. It may be assumed that there is a constant probability $p = 0.1$ of producing a defective article. Before these articles are stored, they are inspected and the defective ones are set aside. Suppose that there is a constant probability $r = 0.1$ that a defective article is misclassified. If X denotes the number of articles classified as defective at the end of a production day, find (a) $P(X = 3)$ and $P(X > 3)$.
[**Hint:** P(a defective article is classified as defective) = P(an article produced is defective) $\times$ P(it is classified as defective) $= 0.1 \times 0.9 = 0.09$]
32. A fair coin is tossed 4 times. If X denotes the number of heads obtained and Y denotes the excess of the number of heads over the number of tails, obtain the probability mass function of Y.
33. An irregular 6-faced die is thrown and the expectation that in 10 throws it will give 5 even numbers is twice the expectation that it will give 4 even numbers. How many times in 10,000 sets of 10 throws would you expect to give no even number?
34. If m things are distributed among a men and b women, show that the probability that the number of things received by men is odd is $\frac{1}{2}\left[\frac{(b+a)^m-(b-a)^m}{(b+a)^m}\right]$.
[**Hint:** P(a thing is received by men) $= p = \frac{a}{a+b}$ and $q = \frac{b}{a+b}$].
35. At least one half of an airplane's engines are required to function in order for it to operate. If each engine independently functions with probability p, for what values of p is a 4-engine plane to be preferred for operation to a 2-engine plane?

36. At least one half of an airplane's engines are required to function in order for it to operate. If each engine functions independently with probability of failure q, for what values of q is a 2-engine plane to be preferred for operation to a 4-engine plane?
37. If a fair coin is flipped an even number $2n$ times, show that the probability of getting more heads than tails is $\frac{1}{2}\left[1 - 2nC_n\left(\frac{1}{2}\right)^{2n}\right]$.
$\Big[$**Hint:** P(more heads than tails) $= P$(less heads than tails $= \frac{1}{2}$ [1 − P(equal number of heads and tails) $\Big]$.
38. If a fair coin is tossed at random 5 independent times, find the conditional probability of 5 heads relative to the hypothesis that there are at least 4 heads.
39. A factory has 10 machines which may need adjustment from time to time during the day. Three of these machines are old, each having a probability of $\frac{1}{11}$ of needing adjustment during the day and 7 are new, having the corresponding probability of $\frac{1}{21}$. Assuming that no machine needs adjustment twice on the same day, find the probabilities that on a particular day
(i) just 2 old and no new machine need adjustment and
(ii) just 2 machines that need adjustment are of the same type.
40. The probability of a man hitting a target is $\frac{1}{4}$. (i) If he fires 7 times, what is the probability of his hitting the target at least twice? and (ii) How many times must he fire so that the probability of his hitting the target at least once is greater than $\frac{2}{3}$?
41. A set of 6 similar coins are tossed 640 times with the following results:

Number of heads :	0	1	2	3	4	5	6
Frequency :	7	64	140	210	132	75	12

Calculate the binomial frequencies on the assumption that the coins are symmetrical.
42. Fit a binomial distribution for the following data and hence find the theoretical frequencies:

x:	0	1	2	3	4
f:	5	29	36	25	5

43. A car hire firm has 2 cars which it hires out day by day. The number of demands for a car on each day follows a Poisson distribution with mean 1.5. Calculate the proportion of days on which (i) neither car is used and (ii) some demand is not fulfilled.
44. The proofs of a 500-page book contains 500 misprints. Find the probability that there are at least 4 misprints in a randomly chosen page.
45. If the average number of claims handled daily by an insurance company is 5, what proportion of days will have less than 3 claims? What is the probability that there will be 4 claims in exactly 3 of the next 5 days. Assume that the number of claims on different days are independent.
46. In a certain factory producing razor blades, there is a small chance $\frac{1}{500}$ for any blade to be defective. The blades are supplied in packets of 10. Use Poisson distribution to calculate the approximate number of packets containing (i) no defective blade, (ii) at least 1 defective blade and (iii) at most 1 defective blade in a consignment of 10,000 packets.
47. An insurance company has discovered that only about 0.1% of the population is involved in a certain type of accident each year. If its 10,000 policy holders were randomly selected from the population, what is the probability that not more than 5 of its clients are involved in such an accident next year?
48. In a given city, 4% of all licenced drivers will be involved in at least 1 road accident in any given year. Determine the probability that among 150 licenced drivers randomly chosen in this city
    (i) only 5 will be involved in at least 1 accident in any given year and
    (ii) at most 3 will be involved in at least 1 accident in any given year.
49. A radioactive source emits on the average 2.5 particles per second. Find the probability that 3 or more particles will be emitted in an interval of 4s.
50. It has been established that the number of defective stereos produced daily at a certain plant is Poisson distributed with mean 4. Over a 2-day span, what is the probability that the number of defective stereos does not exceed 3?
51. In an industrial complex, the average number of fatal accidents per month is one-half. The number of accidents per month is adequately described by a Poisson distribution. What is the probability that 6 months will pass without a fatal accident?
52. If the numbers of telephone calls coming into a telephone exchange between 9 A.M. and 10 A.M. and between 10 A.M. and 11 A.M. are independent and follow Poisson distributions with parameters 2 and 6 respectively, what is the probability that more than 5 calls come between 9 A.M. and 11 A.M.?
53. Patients arrive randomly and independently at a doctor's consulting room from 5 P.M. at an average rate of one in 5 min. The waiting room can hold 12 persons. What is the probability that the room will be full, when the doctor arrives at 6 P.M.?

54. The number of blackflies on a broad bean leaf follows a Poisson distribution with mean 2. A plant inspector, however, records the number of flies on a leaf only if at least 1 fly is persent. What is the probability that he records 1 or 2 flies on a randomly chosen leaf? What is the expected number of flies recorded per leaf?
    [**Hint:** If X is the number of flies on a leaf, we have to find $P\{X = r/X \geq 1\}$ $r = 1, 2$, and add them.]
55. A radioactive source emits particles at a rate of 10 per minute in accordance with Poisson law. Each particle emitted has a probability of $\frac{2}{5}$ of being recorded. Find the probability that at least 4 particles are recorded in a 2-min period.
56. Fit a Poisson distribution for the following distribution and hence find the expected frequencies.

x:	0	1	2	3	4	5	6
f:	314	335	204	86	29	9	3

57. If the probability that a certain test yields a positive reaction equals 0.4 what is the probability that fewer than 5 negative reactions occur before the first positive one?
58. In a test a light switch is turned on and off until it fails. If the probability that the switch will fail any time it is turned on or off is 0.001, what is the probability that the switch will not fail during the first 800 times it is turned on or off?
59. An item is inspected at the end of each day to see whether it is still functioning properly. If it is found to fail at the 10th inspection and not earlier, what is the maximum value of the probability of its failure on any day?
60. If X and Y are 2 independent RVs, each representing the number of failures preceding the first success in a sequence of Bernoulli's trials with p as the probability of success in a single trial, show that $P(X = Y) = \frac{p}{1+q}$, where $p + q = 1$.
61. A throws 2 dice until he gets 6 and B throws independently 2 other dice until he gets 7. Find the probability that B will require more throws than A.
62. If 2 independent RVs X and Y have identical geometric distributions with parameter p, find the probability mass function of $(X + Y)$ and hence the expected value of $(X + Y)$.
63. As part of an air-pollution survey, an inspector decides to examine the exhaust of 6 of a company's 24 trucks. If four of the company's trucks emit excessive amounts of pollutants, what is the probability that none of them will be included in the inspector's sample?

Thus $\mu_{2n-1} = 0$ and $\mu_{2n} = \frac{1}{2n+1} \cdot \left(\frac{b-a}{2}\right)^{2n}$ for $n = 1, 2, 3, \ldots$ (3)

In particular, $\mu_2 =$ variance of $U(a, b) = \frac{1}{12}(b-a)^2$ (4)

$$\mu_3 = 0 \text{ and } \mu_4 = \frac{1}{80}(b-a)^4$$

The absolute central moments v_r of the uniform distribution $U(a, b)$ are given by

$$\begin{aligned}
v_r &= E\{|X - E(X)|^r\} \\
&= \int_a^b \frac{\left|x - \frac{1}{2}(b+a)\right|^r}{b-a} \, dx \\
&= \frac{1}{b-a} \int_{-c}^{c} |t|^r \, dt, \text{ on putting } t = x - \frac{1}{2}(b+a) \text{ and } c = \frac{1}{2}(b-a) \\
&= \frac{2}{b-a} \int_0^c t^r \, dt, \quad \text{(since the integrand is an even function of } t\text{)} \\
&= \frac{1}{r+1} \cdot \left(\frac{b-a}{2}\right)^r \qquad (5)
\end{aligned}$$

Definition: $E\{|X - E(X)|\}$ is called the *mean deviation* (MD) about the mean of the RV X or of the corresponding distribution.

Thus the MD about the mean of the distribution $U(a, b)$ is given by

$$v_1 = \frac{1}{4}(b-a)$$

2. Exponential distribution

Definitions: A continuous RV X is said to follow an *exponential distribution* or *negative exponential distribution* with parameter $\lambda > 0$, if its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

We note that $\int_0^\infty f(x)\, dx = \int_0^\infty \lambda e^{-\lambda x} \, dx = 1$ and hence $f(x)$ is a legitimate density function.

Mean and Variance of the Exponential Distribution

Raw moments μ'_r about the origin of the exponential distribution are given by

$$\begin{aligned}
\mu'_r &= E(X^r) \\
&= \int_0^\infty x^r \cdot \lambda e^{-\lambda x} \, dx \\
&= \frac{1}{\lambda^r} \int_0^\infty y^r e^{-y} \, dy, \text{ (on putting } y = \lambda x\text{)} \\
&= \frac{1}{\lambda^r} \overline{|(r+1)} \\
&= \frac{r!}{\lambda^r} \qquad (1)
\end{aligned}$$

$\therefore \quad E(X) =$ Mean of the exponential distribution

$$= \mu_1' = \frac{1}{\lambda}, \text{ [from (1)]}$$

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

$$\mu_2' = \frac{2}{\lambda^2}$$

$$\therefore \text{Var}(X) = E(X^2) - \{E(X)\}^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Memoryless Property of the Exponential Distribution

If X is exponentially distributed, then

$$P(X > s + t/X > s) = P(X > t), \text{ for any } s, t > 0$$

$$P(X > k) = \int_k^\infty \lambda e^{-\lambda x} \, dx = \left(-e^{-\lambda x}\right)_k^\infty = e^{-\lambda k} \qquad (1)$$

$$\begin{aligned}
\text{Now } P(X > s + t/X > s) &= \frac{P\{X > s+t \text{ and } X > s\}}{P\{X > s\}} \\
&= \frac{P\{X > s+t\}}{P\{X > s\}} \\
&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}, \text{ [by (1)]} \\
&= e^{-\lambda t} = P(X > t).
\end{aligned}$$

Note *The converse of this result is also true. That is, if $P(X > s + t/X > s) = P(X > t)$, then X follows an exponential distribution. See Example (8) in Worked Example 5(b).*

3. Erlang distribution or General Gamma distribution

Definition: A continuous RV X is said to follow an *Erlang distribution* or *General Gamma distribution* with parameters $\lambda > 0$ and $k > 0$, if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)}, & \text{for } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

We note that $\int_0^\infty f(x)\,dx = \dfrac{\lambda^k}{\Gamma(k)} \int_0^\infty x^{k-1} e^{-\lambda x}\,dx$

$$= \frac{1}{\Gamma(k)} \int_0^\infty t^{k-1} e^{-t}\,dt, \text{ [on putting } \lambda x = t]$$

$$= 1$$

Hence $f(x)$ is a legitimate density function.

Note 1. *When $\lambda = 1$, the Erlang distribution is called Gamma distribution or simple Gamma distribution with parameter k whose density function is $f(x) = \dfrac{1}{\Gamma(k)} x^{k-1} e^{-x}; x \geq 0; k > 0$.*

2. *When $k = 1$, the Erlang distribution reduces to the exponential distribution with parameter $\lambda > 0$.*
3. *Sometimes, the Erlang distribution itself is called Gamma distribution.*

Mean and Variance of Erlang Distribution

The raw moments μ_r' about the origin of the Erlang distribution are given by

$$\mu_r' = E(X^r)$$

$$= \int_0^\infty \frac{\lambda^k}{\Gamma(k)} x^{k+r-1} e^{-\lambda x}\,dx$$

$$= \frac{\lambda^k}{\Gamma(k)} \cdot \frac{1}{\lambda^{k+r}} \int_0^\infty t^{k+r-1} e^{-t}\,dt, \text{ (on putting } \lambda x = t)$$

$$= \frac{1}{\lambda^r} \frac{\Gamma(k+r)}{\Gamma(k)}$$

$$\therefore \quad \text{Mean} = E(X) = \frac{1}{\lambda} \cdot \frac{\Gamma(k+1)}{\Gamma(k)} = \frac{k}{\lambda}$$

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

$$= \frac{1}{\lambda^2} \cdot \frac{\Gamma(k+2)}{\Gamma(k)} - \left(\frac{k}{\lambda}\right)^2$$

$$= \frac{1}{\lambda^2} \{k(k+1) - k^2\} = \frac{k}{\lambda^2}$$

Reproductive Property of Gamma Distribution

The sum of a finite number of independent Erlang variables is also an Erlang variable. That is, if $X_1, X_2, \ldots, X_n$ are independent Erlang variables with parameters $(\lambda, k_1), (\lambda, k_2), \ldots, (\lambda, k_n)$, then $X_1 + X_2 + \ldots + X_n$ is also an Erlang variable with parameter $(\lambda, k_1 + k_2 + \ldots + k_n)$.

Let us first find the moment generating function of the Erlang variable X with parameters λ and k and use it to prove this property. MGF of X is given by

$$M_X(t) = E\{e^{tx}\}$$

$$= \int_0^\infty \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x} e^{tx}\,dx$$

$$= \frac{\lambda^k}{\Gamma(k)} \int_0^\infty x^{k-1} e^{-(\lambda - t)x}\,dx$$

$$= \frac{\lambda^k}{\Gamma(k)} \cdot \frac{1}{(\lambda - t)^k} \int_0^\infty y^{k-1} e^{-y}\,dy, \quad \text{(on putting } \lambda - t = y).$$

$$= \left(\frac{\lambda}{\lambda - t}\right)^k \quad [\because \text{ the integral} = \Gamma(k)]$$

$$= \left(1 - \frac{t}{\lambda}\right)^{-k}$$

Now $M_{(X_1 + X_2 + \ldots + X_n)}(t) = M_{X_1}(t)\, M_{X_2}(t) \ldots M_{X_n}(t)$ (since $X_1, X_2, \ldots, X_n$ are independent)

[Refer to property (4) of MGF given in section 4(b) of chapter (4)]

$$= \left(1 - \frac{t}{\lambda}\right)^{-k_1} \left(1 - \frac{t}{\lambda}\right)^{-k_2} \cdots \left(1 - \frac{t}{\lambda}\right)^{-k_n}$$

$$= \left(1 - \frac{t}{\lambda}\right)^{-(k_1 + k_2 + \ldots + k_n)}$$

which is the MGF of an Erlang variable with parameters $(\lambda, k_1 + k_2 + \ldots + k_n)$. Hence the reproductive property.

Relation Between the Distribution Functions (cdf) of the Erlang Distribution With $\lambda = 1$ (or Simple Gamma Distribution) and (Poisson Distribution)

If X is a Poisson random variable with mean λ,

$$\text{then } P(X \leq K) = \sum_{r=0}^{k} \frac{e^{-\lambda} \lambda^r}{r!} \tag{1}$$

Differentiating both sides with respect to λ, we get

$$\frac{d}{d\lambda} P(X \leq k) = \sum_{r=0}^{k} \frac{1}{r!} \{e^{-\lambda} \cdot r\lambda^{r-1} - e^{-\lambda} \cdot \lambda^r\}$$

$$= e^{-\lambda} \cdot \sum_{r=0}^{k} \left[\frac{\lambda^{r-1}}{(r-1)!} - \frac{\lambda^r}{r!}\right]$$

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

$$= -\frac{e^{-\lambda}\lambda^k}{k!} \qquad (2)$$

Integrating both sides of (2) with respect to λ from λ to ∞, we get

$$\left[\sum_{r=0}^{K}\frac{e^{-\lambda}\lambda^r}{r!}\right]_{\lambda}^{\infty} = -\int_{\lambda}^{\infty}\frac{1}{k!}e^{-\lambda}\lambda^k\,d\lambda$$

i.e., $$\sum_{r=0}^{K}\frac{e^{-\lambda}\lambda^r}{r!} = \int_{\lambda}^{\infty}\frac{1}{k!}e^{-y}y^k\,dy$$

i.e., $P(X \le k) = P(Y \ge \lambda)$,

[where Y is the Erlang variable with parameters 1 and $(k+1)$

or $P(X \le k) = 1 - P(Y \le \lambda)$

Note *The above relationship is valid only when the parameter k is a positive integer.*

4. Weibull Distribution

Definition: A continuous RV X is said to follow a *Weibull distribution* with parameters $\alpha, \beta > 0$, if the RV $Y = \alpha X^{\beta}$ follows the exponential distribution with density function $f_Y(y) = e^{-y}, y > 0$.

Density Function of the Weibull Distribution

Since $Y = \alpha \cdot X^{\beta}$, we have $y = \alpha \cdot x^{\beta}$.

By the transformation rule, derived in chapter 3, we have $f_X(x) = f_Y(y)\left|\frac{dy}{dx}\right|$,

where $f_X(x)$ and $f_Y(y)$ are the density functions of X and Y respectively.

$$\therefore \quad f_X(x) = e^{-y}\,\alpha\beta x^{\beta-1}$$
$$= \alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}};\ x > 0 \quad [\because y > 0]$$

Note *When $\beta = 1$, Weibull distribution reduces to the exponential distribution with parameter α.*

Mean and Variance of Weibull Distribution

The raw moments μ_r' about the origin of the Weibull distribution are given by

$$\mu_r' = E(X^r)$$
$$= \alpha\beta\int_0^{\infty} x^{r+\beta-1}e^{-\alpha x^{\beta}}\,dx$$
$$= \int_0^{\infty}\left(\frac{y}{\alpha}\right)^{\frac{r}{\beta}+1-\frac{1}{\beta}}e^{-y}\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}-1}dy,$$

on putting $y = \alpha x^{\beta}$ or $x = \left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}}$

$$= \alpha^{-r/\beta}\int_0^{\infty} y^{r/\beta}e^{-y}\,dy$$
$$= \alpha^{-r/\beta}\,\Gamma\!\left(\frac{r}{\beta}+1\right)$$

$$\therefore \quad \text{Mean} = E(X) = \mu_1' = \alpha^{-\frac{1}{\beta}}\,\Gamma\!\left(\frac{1}{\beta}+1\right)$$

$$\text{Var}(X) = E(X^2) - \{E(X)\}^2$$
$$= \alpha^{-2/\beta}\left[\Gamma\!\left(\frac{2}{\beta}+1\right) - \left\{\Gamma\!\left(\frac{1}{\beta}+1\right)\right\}^2\right]$$

Note *Weibull distribution finds frequent applications in Reliability Theory. It is assumed as the probability distribution of the time to failure (or length of life) of a component in a system. Other distributions used to describe the failure law are the exponential and normal distributions. See Example (19) in Worked Example 5(b).*

5. Normal (or Gaussian) distribution

Definition: A continuous RV X is said to follow a *normal distribution* or *Gaussian distribution* with parameters μ and σ, if its probability density function is given by

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

Symbolically 'X follows $N(\mu, \sigma)$'. Sometimes it is also given as $N(\mu, \sigma^2)$. We shall use only the notation $N(\mu, \sigma)$ as in the earlier chapters.

$f(x)$ is a legitimate density function, as

$$\int_{-\infty}^{\infty} f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\,dx$$
$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-t^2}\sigma\sqrt{2}\,dt, \quad \left(\text{on putting } t = \frac{x-\mu}{\sigma\sqrt{2}}\right)$$
$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^2}\,dt$$
$$= \frac{1}{\sqrt{\pi}}2\int_0^{\infty} e^{-t^2}\,dt = \frac{1}{\sqrt{\pi}}\Gamma\!\left(\frac{1}{2}\right) = \frac{1}{\sqrt{\pi}}\cdot\sqrt{\pi} = 1$$

Standard Normal Distribution

The normal distribution $N(0, 1)$ is called the standardised or simply the standard normal distribution, whose density function is given by

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

This is obtained by putting $\mu = 0$ and $\sigma = 1$ and by changing x and f respectively into z and ϕ. If X has distribution $N(\mu, \sigma)$ and if $Z = \frac{X-\mu}{\sigma}$, then we can prove that Z has distribution $N(0, 1)$.

[See the corollary under the property (6) of normal distribution]

The importance of $N(0, 1)$ is due to the fact that the values of $\phi(z)$ and $\int_0^z \phi(z)\,dz$ are tabulated.

Normal Probability Curve

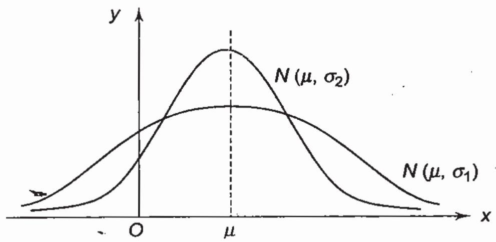

Fig. 5.1

The graph of $y = f(x)$, that is given above for $\sigma = \sigma_1$ and σ_2, is a well-known bell-shaped curve and is called the normal probability curve (Fig. 5.1).

The curve is symmetrical about the ordinate at $x = \mu$. The ordinate $f(x)$ decreases rapidly as x increases numerically, the maximum (occurring at $x = \mu$) given by $\frac{1}{\sigma\sqrt{2\pi}}$. The curve extends upto infinity on either side of $x = \mu$ and the x-axis is an asymptote to the curve.

The graph is concave downward at $x = \mu$ and it is concave upward for large numerical values of x. The points at which the concavity changes are called the points of inflexion of the curve. They are found by solving the equation $y'' = 0$ [i.e., $f''(x) = 0$]. We can prove that the points of inflexion of the normal probability curve occur at $x = \mu \pm \sigma$, that is, at points which are at a distance of σ on either side of $x = \mu$. Thus if σ is relatively large, the curve tends to be flat, while if σ is small, the curve tends to be peaked. Hence the steepness of the curve is determined by σ. The two curves given in the figure relate to $N(\mu, \sigma_1)$ and $N(\mu, \sigma_2)$, where $\sigma_1 > \sigma_2$.

Properties of the Normal Distribution $N(\mu, \sigma)$

1. If X follows $N(\mu, \sigma)$, then $E(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f(x)\,dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x\, e^{-(x-\mu)^2/2\sigma^2}\,dx \\
&= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\mu + \sqrt{2}\,\sigma t)\, e^{-t^2}\,dt, \quad \left(\text{on putting } t = \frac{x-\mu}{\sigma\sqrt{2}}\right) \\
&= \frac{\mu}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2}\,dt + \sqrt{\frac{2}{\pi}}\,\sigma \int_{-\infty}^{\infty} t\, e^{-t^2}\,dt \\
&= \frac{\mu}{\sqrt{\pi}} \cdot \sqrt{\pi} = \mu.
\end{aligned}$$

(since the integrand in the second integral is an odd function of t)

$$\begin{aligned}
E(X^2) &= \frac{1}{\sigma\sqrt{2}} \int_{-\infty}^{\infty} x^2 e^{-(x-\mu)^2/2\sigma^2}\,dx \\
&= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\mu + \sqrt{2}\,\sigma t)^2 e^{-t^2}\,dt, \quad \left(\text{on putting } t = \frac{x-\mu}{\sigma\sqrt{2}}\right) \\
&= \frac{1}{\sqrt{\pi}} \left[\int_{-\infty}^{\infty} \mu^2 e^{-t^2}\,dt + 2\sqrt{2}\,\mu\sigma \int_{-\infty}^{\infty} t\, e^{-t^2}\,dt + 2\sigma^2 \int_{-\infty}^{\infty} t^2 e^{-t^2}\,dt\right] \\
&= \mu^2 + 0 + \frac{2\sigma^2}{\sqrt{\pi}} \int_0^{\infty} t\, e^{-t^2}\, 2t\,dt, \quad (\because t^2 e^{-t^2} \text{ is even}) \\
&= \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \int_0^{\infty} u^{\frac{1}{2}} e^{-u}\,du, \quad (\text{on putting } u = t^2) \\
&= \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \cdot \Gamma\left(\frac{3}{2}\right) \\
&= \mu^2 + 2\,\frac{\sigma^2}{\sqrt{\pi}} \cdot \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) \\
&= \mu^2 + \sigma^2 \quad \left(\because \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}\right)
\end{aligned}$$

$$\begin{aligned}
\therefore \mathrm{Var}(X) &= E(X^2) - \{E(X)\}^2 \\
&= \sigma^2
\end{aligned}$$

2. *Median and mode of the normal distribution $N(\mu, \sigma)$*

Definition: If X is a continuous RV with density function $f(x)$, then M is called the median value of X, provided that

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

For the normal distribution $N(\mu, \sigma)$, the median M is given by

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

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

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

$$\left\{\text{since } \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}\,dx = 1 \text{ and the normal curve is symmetrical about } x = \mu\right\}$$

i.e., $$\int_{M}^{\mu} f(x)\,dx = 0$$

$\therefore$ $$M = \mu$$

Definition: Mode of a continuous RV X is defined as the value of x for which the density function $f(x)$ is maximum.

For the normal distribution $N(\mu, \sigma)$,

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

$\therefore$ $$\log f(x) = k - \frac{1}{2\sigma^2}(x-\mu)^2$$

Differentiating with respect to x,

$$\frac{f'(x)}{f(x)} = -\frac{1}{\sigma^2}(x-\mu)$$

i.e., $$f'(x) = -\frac{1}{\sigma^2}(x-\mu) f(x)$$

$$= 0, \text{ when } x = \mu$$

$$f''(x) = -\frac{1}{\sigma^2}\{(x-\mu) f'(x) + f(x)\}$$

$$\therefore [f''(x)]_{x=\mu} = -\frac{1}{\sigma^2} f(\mu) < 0$$

Therefore, $f(x)$ is maximum at $x = \mu$. That is, Mode of the distribution $N(\mu, \sigma) = \mu$.

Note *For the normal distribution, mean, median and mode are equal.*

3. Central moments of the normal distribution N(μ, σ)

Central moments μ_r of $N(\mu, \sigma)$ are given by $\mu_r = E(x-\mu)^r$

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

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\sqrt{2}\,\sigma t)^r e^{-t^2}\,dt$$

$$= \frac{2^{r/2}\sigma^r}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^r e^{-t^2}\,dt$$

Case (i): r is an odd integer, that is, $r = 2n+1$.

$$\therefore \quad \mu_{2n+1} = \frac{2^{(2n+1)/2}\sigma^{2n+1}}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^{2n+1} e^{-t^2}\,dt$$

$$= 0, \text{ (since the integrand is an odd function of } t)$$

Case (ii): r is an even integer, that is, $r = 2n$

$$\therefore \mu_{2n} = \frac{2^n\sigma^{2n}}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^{2n} e^{-t^2}\,dt$$

$$= \frac{2^n\sigma^{2n}}{\sqrt{\pi}} \cdot 2\int_{0}^{\infty} t^{2n} e^{-t^2}\,dt$$

{$\because$ the integrand is an even function of t)

$$= \frac{2^n\sigma^{2n}}{\sqrt{\pi}} \int_{0}^{\infty} u^{n-\frac{1}{2}} e^{-u}\,du, \qquad \text{[on putting } u = t^2)$$

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

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

$$= \frac{2^n\sigma^{2n}}{\sqrt{\pi}} \cdot \left(\frac{2n-1}{2}\right)\cdot\left(\frac{2n-3}{2}\right) \Gamma\left(\frac{2n-3}{2}\right)$$

$$= \frac{2^n\sigma^{2n}}{\sqrt{\pi}} \cdot \frac{2n-1}{2} \cdot \frac{2n-3}{2} \cdots \frac{1}{2}\Gamma\left(\frac{1}{2}\right)$$

$$= 1.3.5 \ldots (2n-1)\,\sigma^{2n}$$

From (1), we get,

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

From (1) and (2), we get

$$\frac{\mu_{2n}}{\mu_{2n-2}} = 2\sigma^2 (n - 1/2)$$

i.e., $$\mu_{2n} = (2n-1)\sigma^2 \mu_{2n-2} \qquad (3)$$

(3) gives a recurrence relation for the even order central moments of the normal distribution $N(\mu, \sigma)$.]

4. Mean Deviation about the mean of the normal distribution N(μ, σ)

Definition: The absolute (central) moment of the first order of a RV X is called the mean deviation about the mean of X.
i.e., MD about the mean = $E\{|x - E(X)|\}$
For the noraml distribution $N(\mu, \sigma)$,

$$\text{The MD about the mean} = \int_{-\infty}^{\infty} |x-\mu| \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx$$

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

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

$$= 2\sqrt{\frac{2}{\pi}}\,\sigma \int_0^{\infty} t e^{-t^2} dt, \text{ (since the integrand is an even function of } t)$$

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

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

5. Quartile deviation of the normal distribution N(μ, σ)

Definition: The first quartile Q_1 and the third quartile Q_3 of $N(\mu, \sigma)$ (or of any continuous random variable) are defined by the equations

$$\int_{-\infty}^{Q_1} f(x)\,dx = \frac{1}{4} \text{ and } \int_{-\infty}^{Q_3} f(x)\,dx = \frac{3}{4}$$

or equivalently

$$\int_{Q_1}^{\mu} f(x)\,dx = \frac{1}{4} \text{ and } \int_{\mu}^{Q_3} f(x)\,dx = \frac{1}{4},$$

[if the curve $y = f(x)$ is symmetrical about $x = \mu$]

Then the quartile deviation (QD) is defined as

$$QD = \frac{1}{2}(Q_3 - Q_1).$$

For the normal distribution $N(\mu, \sigma)$, Q_1 is given by

$$\int_{Q_1}^{\mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx = 0.25$$

i.e., $$\int_{(Q_1-\mu)/\sigma}^{0} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 0.25, \quad \left(\text{on putting } z = \frac{x-\mu}{\sigma}\right)$$

i.e., $$\int_0^{(\mu-Q_1)/\sigma} \phi(z)\,dz = 0.25,$$

$\left(\text{by symmetry of the normal curve and since } \frac{Q_1-\mu}{\sigma} < 0.\right)$

From the table of normal areas (areas under standard normal curve), we get

$$\int_0^{0.674} \phi(z)\,dz = 0.25$$

$$\therefore \quad \frac{\mu - Q_1}{\sigma} = 0.674$$

i.e., $Q_1 = \mu - 0.674\,\sigma$
By symmetry, $Q_3 = \mu + 0.674\,\sigma$

$$\therefore\ QD = \frac{1}{2}(Q_3 - Q_1) = 0.674\,\sigma = \frac{2}{3}\sigma \text{ (approximately)}$$

6. Moment generating function of N(0, 1) and N(μ, σ)

The moment generating function of $N(0, 1)$ is given by

$$M_Z(t) = E(e^{tZ})$$

$$= \int_{-\infty}^{\infty} e^{tz}\phi(z)\,dz,$$

[where $\phi(z)$ is the density function of $N(0, 1)$]

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} e^{tz} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z^2-2tz)/2} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\{(z-t)^2 - t^2\}/2} dz$$

$$= \frac{1}{\sqrt{2\pi}} e^{t^2/2} \sqrt{2} \int_{-\infty}^{\infty} e^{-u^2} du, \quad \left(\text{on putting } u = \frac{z-t}{\sqrt{2}} \right)$$

$$= \frac{1}{\sqrt{\pi}} e^{t^2/2} \Gamma\left(\frac{1}{2}\right) = e^{t^2/2} \left(\because \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \right)$$

The moment generating function of $N(\mu, \sigma)$ is given by

$$M_X(t) = M_{\sigma Z + \mu}(t), \quad \left(\text{since } Z = \frac{X-\mu}{\sigma} \right)$$

$$= e^{\mu t} \cdot \mu_Z(\sigma t), \quad \text{(by the property of MGF)}$$

$$= e^{\mu t} e^{\sigma^2 t^2/2}$$

$$= e^{t\left(\mu + \frac{\sigma^2 t}{2}\right)}$$

Now $\quad M_X(t) = 1 + \frac{t}{1!}\left(\mu + \frac{\sigma^2 t}{2}\right) + \frac{t^2}{2!}\left(\mu + \frac{\sigma^2 t}{2}\right)^2 + \frac{t^3}{3!}\left(\mu + \frac{\sigma^2 t}{2}\right)^3 + \frac{t^4}{4!}\left(\mu + \frac{\sigma^2 t}{2}\right)^4 + \ldots + \infty$

$\therefore \quad E(X) = \text{Coefficient of } \frac{t}{1!} = \mu$

$E(X^2) = \text{Coefficient of } \frac{t^2}{2!} = \sigma^2 + \mu^2$

$E(X^3) = \text{Coefficient of } \frac{t^3}{3!} = 3\mu\sigma^2 + \mu^3$ and

$E(X^4) = \text{Coefficient of } \frac{t^4}{4!} = 3\sigma^4 + 6\mu^2\sigma^2 + \mu^4$.

Using the relation $\mu_k = k$th order central moment $= E\{(X-\mu)^k\}$, we get

$$\mu_1 = 0, \mu_2 = \sigma^2, \mu_3 = 0 \ \mu_4 = 3\sigma^4$$

We could have got these values from the formulas $\mu_{2n+1} = 0$ and $\mu_{2n} = 1.3.5 \ldots (2n-1)\, \sigma^{2n}$, which we have derived already.

Corollary:

If X has the distribution $N(\mu, \sigma)$ then $Y = aX + b$ has the distribution $N(a\mu + b, a\sigma)$

$$M_X(t) = e^{\mu t + \sigma^2 t^2/2}$$

$$\therefore M_Y(t) = M_{aX+b}(t)$$

$$= e^{bt} M_X(at)$$

$$= e^{bt} e^{a\mu t + a^2\sigma^2 t^2/2}$$

$$= e^{(a\mu + b)t + (a^2\sigma^2)t^2/2}$$

which is the MGF of $N(a\mu + b, a\sigma)$.

In particular, if X has the distribution $N(\mu, \sigma)$, then $Z = \frac{X-\mu}{\sigma}$ has the distribution $N\left(\frac{1}{\sigma}\mu - \frac{\mu}{\sigma}, \frac{1}{\sigma}\cdot\sigma\right)$ that is, $N(0, 1)$.

7. Additive property of normal distribution

If X_i $(i = 1, 2, \ldots, n)$ be n independent normal RVs with mean μ_i and variance σ_i^2, then $\sum_{i=1}^{n} a_i X_i$ is also a normal RV with mean $\sum_{i=1}^{n} a_i \mu_i$ and variance $\sum_{i=1}^{n} a_i^2 \sigma_i^2$.

$$M_{\left(\sum_{i=1}^{n} a_i X_i\right)}(t) = M_{a_1 X_1}(t) \cdot M_{a_2 X_2}(t) \ldots M_{a_n X_n}(t), \quad \text{(by independence)}$$

$$= e^{a_1 \mu_1 t + a_1^2 \sigma_1^2 t^2/2} \times e^{a_2 \mu_2 t + a_2^2 \sigma_2^2 t^2/2} \ldots \times e^{a_n \mu_n t + a_n^2 \sigma_n^2 t^2/2}$$

$$= e^{(\Sigma a_i \mu_i)t + \Sigma a_i^2 \sigma_i^2 t^2/2}$$

which is the MGF of a normal RV with mean $\sum a_i \mu_i$ and variance $\sum a_i^2 \sigma_i^2$.

Hence the property.

Deductions:

1. Putting $a_1 = a_2 = 1$ and $a_3 = a_4 = \ldots = a_n = 0$, we get the following result, in particular:
   If X_1 is $N(\mu_1, \sigma_1)$ and X_2 is $N(\mu_2, \sigma_2)$, then $X_1 + X_2$ is $N(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$. Similarly, $X_1 - X_2$ is $N(\mu_1 - \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

2. Putting $a_1 = a_2 = \ldots = a_n = \frac{1}{n}$ and assuming that each X_i is $N(\mu, \sigma)$, then

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ has a normal distribution}$$

$$N\left\{\frac{1}{n}\sum_{i=1}^{n}\mu, \sqrt{\sum_{i=1}^{n}\frac{1}{n^2}\sigma^2}\right\}$$

i.e., $N\{\mu, \sigma/\sqrt{n}\}$

Thus if X_i $(i = 1, 2, \ldots, n)$ are independent and identically distributed normal variables with mean μ and standard deviation σ, then their mean

$\bar{X}$ is $N\{\mu, \sigma/\sqrt{n}\}$.

8. Normal distribution as limiting form of binomial distribution

When n is very large and neither p nor q is very small, the standard normal distribution can be regarded as the limiting form of the standardised binomial distribution.

When X follows the binomial distribution $B(n, p)$, the standardised binomial variable Z is given by $Z = \frac{X - np}{\sqrt{npq}}$. As X varies from 0 to n with step size 1, Z varies from $\frac{-np}{\sqrt{npq}}$ to $\frac{np}{\sqrt{npq}}$ with step size $\frac{1}{\sqrt{npq}}$. When neither p nor q is very small and n is very large, Z varies from $-\infty$ to ∞ with infinitesimally small step size. Hence, in the limit, the distribution of Z may be expected to be a continuous distribution extending from $-\infty$ to ∞ and having mean 0 and standard deviation 1. In fact the limiting form of the distribution of Z is standard normal distribution as seen below:

If X follows $B(n, p)$, then the MGF of X is given by $M_X(t) = (q + p\,e^t)^n$.

If $Z = \frac{X - np}{\sqrt{npq}}$, then

$$M_Z(t) = M_{\frac{1}{\sqrt{npq}}X - \frac{np}{\sqrt{npq}}}(t) = e^{\frac{-npt}{\sqrt{npq}}}\{q + p\,e^{t/\sqrt{npq}}\}^n$$

$$\therefore \quad \log M_Z(t) = -\frac{npt}{\sqrt{npq}} + n\log\{q + p\,e^{t/\sqrt{npq}}\}$$

$$= -\frac{npt}{\sqrt{npq}} + n\log\left[q + p\left\{1 + \frac{t}{\sqrt{npq}} + \frac{t^2}{2npq} + \frac{t^3}{6(npq)^{3/2}} + \cdots\right\}\right]$$

$$= -\frac{npt}{\sqrt{npq}} + n\log\left[1 + \left\{\frac{pt}{\sqrt{npq}} + \frac{pt^2}{2npq} + \frac{pt^3}{6(npq)^{3/2}} + \cdots\right\}\right]$$

$$= -\frac{npt}{\sqrt{npq}} + n\left[\frac{pt}{\sqrt{npq}}\left\{1 + \frac{t}{2\sqrt{npq}} + \frac{t^2}{6n^2p^2q^2} + \cdots\right\} - \frac{1}{2}\cdot\frac{p^2t^2}{npq}\left\{1 + \frac{t}{2\sqrt{npq}} + \frac{t^2}{6n^2p^2q^2} + \cdots\right\}^2 + \cdots\right]$$

$$= \frac{t^2}{2} + \text{terms containing } \frac{1}{\sqrt{n}} \text{ and lower powers of } n$$

$$\therefore \quad \lim_{n\to\infty} \log M_Z(t) = \frac{t^2}{2}$$

i.e., $$\log_e\left[\lim_{n\to\infty} M_Z(t)\right] = \frac{t^2}{2}$$

$$\therefore \quad \lim_{n\to\infty} M_Z(t) = e^{t^2/2},$$

which is the MGF of the standard normal distribution. Hence the limit of the standardised binomial distribution, as n tends to ∞, is the standard normal distribution.

Note *We recall De Moivre–Laplace approximation for the sum of a large number of terms of the form $nC_r p^r q^{n-r}$ in terms of the integral of standard normal density function, which was discussed in section 1(c). It was stated that*

$$\sum_{r=r_1}^{r_2} nC_r p^r q^{n-r} = \int_{z_1}^{z_2} \phi(z)\,dz$$

where $z_1 = \frac{r_1 - np - \frac{1}{2}}{\sqrt{npq}}$ *and* $z_2 = \frac{r_2 - np + \frac{1}{2}}{\sqrt{npq}}$ *and $\phi(z)$ is the density function of the standard normal distribution.*

Importance of Normal Distribution

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

(i) A large number of RVs, such as binomial and Poisson, occurring in many applications have a distribution closely resembling the normal distribution.

(ii) Many of the distributions of sample statistics, such as sample mean and sample variance, tend to normality for samples of large size. In particular, the sampling distributions like Student's t, Snedecor's F and Chi-square distributions tend to normality when the size of the sample is large.

(iii) Tests of significance for small samples are based on the assumption that samples have been drawn from normal populations.

(iv) Even if a variable is not normally distributed, it can sometimes be converted into a normal variable by simple transformation of the variable.

(v) Normal distribution is applied to a large extent in statistical Quality Control in industry.

Worked Example 5(B)

Example 1

If a string, 1 m long, is cut into 2 pieces at a random point along its length, what is the probability that the longer piece is at least twice the length of the shorter?

i.e., the required range of Y is

$$0.0025 \le Y \le 0.9025$$

Example 6

The mileage which car owners get with a certain kind of radial tire is a RV having an exponential distribution with mean 40,000 km. Find the probabilities that one of these tires will last (i) at least 20,000 km and (ii) at most 30,000 km.

Let X denote the mileage obtained with the tire

$$f(x) = \frac{1}{40{,}000} e^{-x/40{,}000} \quad x > 0$$

$$\text{(i) } P(X \ge 20{,}000) = \int_{20{,}000}^{\infty} \frac{1}{40{,}000} e^{-x/40{,}000}\, dx$$

$$= \left[-e^{-x/40{,}000}\right]_{20{,}000}^{\infty}$$

$$= e^{-0.5} = 0.6065$$

$$\text{(ii) } P(X \le 30{,}000) = \int_{0}^{30{,}000} \frac{1}{40{,}000} e^{-x/40{,}000}\, dx$$

$$= \left[-e^{-x/40{,}000}\right]_{0}^{30{,}000}$$

$$= 1 - e^{-0.75} = 0.5270$$

Example 7

If the time T to failure of a component is exponentially distributed with parameter λ and if n such components are installed, what is the probability that one-half or more of these components are still functioning at the end of t hours?

The density function of T is given by

$$f(t) = \lambda e^{-\lambda t}, \; t \ge 0$$

P (a component functions at the end of or after t hours)

$$= P(T \ge t) = \int_{t}^{\infty} \lambda e^{-\lambda t}\, dt = e^{-\lambda t}$$

If we consider a component functioning at the end or after t hours as a success in a single trial, we have $p = e^{-\lambda t}$ and $q = 1 - e^{-\lambda t}$.

Then the number X of successes in n independent trials follows a binomial distribution with parameters n and p.

$$\therefore \quad P(X = r) = nC_r p^r q^{n-r}; \; r = 0, 1, 2, \ldots, n$$

If n is even the required probability is given by

$$\sum_{r=\frac{n}{2}}^{n} P(X = r) = \sum_{r=\frac{n}{2}}^{n} nC_r e^{-\lambda rt} (1 - e^{-\lambda t})^{n-r}$$

If n is odd, the required probability is given by

$$\sum_{r=\frac{n+1}{2}}^{n} P(X = r) = \sum_{r=\frac{n+1}{2}}^{n} nC_r e^{-\lambda rt} (1 - e^{-\lambda t})^{n-r}$$

Example 8

If a continuous RV X (> 0) possesses memoryless property, that is $P(X > x + h) = P(X > x).\, P(X > h)$, then X follows an exponential distribution.

Let $G(x) = P(X > x)$

$\therefore$ The given condition means that

$$G(x + h) = G(x)G(h)$$

$$\therefore \frac{G(x+h) - G(x)}{h} = G(x)\left\{\frac{G(h) - 1}{h}\right\}$$

$$= G(x) \cdot \frac{\{G(h) - G(0)\}}{h} \quad [\because G(0) = P(X > 0) = 1, \text{ as } x > 0]$$

Taking limits on both sides as $h \to 0$, we have

$$G'(x) = G(x) \cdot G'(0)$$

$$= -\lambda \cdot G(X), \quad [\text{on putting } \lambda = -G'(0)] \qquad (1)$$

Solving the differential equation (i), we get

$$\log G(x) = -\lambda x + \log C$$

i.e., $\quad G(x) = C e^{-\lambda x}$ (2)

Using the fact that $G(0) = 1$ in (2), we get $C = 1$

Thus $\quad G(x) = P(X \ge x) = e^{-\lambda x}$

Now the distribution function $F(x)$ of X is given by $F(x) = P(X \le x)$

$$= 1 - P(X > x) \quad \{= 1 - G(x)\}$$

$$= 1 - e^{-\lambda x}$$

Therefore, the density function $f(x)$ of X is given by

$$f(x) = F'(x) = \lambda e^{-\lambda x}, \; x > 0$$

i.e., X follows an exponential distribution with parameter

$$\lambda = -G'(0) = F'(0) > 0$$

Example 9

The time (in hours) required to repair a machine is exponentially distributed with parameter $\lambda = 1/2$.

(a) What is the probability that the repair time exceeds 2 h?

(b) What is the conditional probability that a repair takes at least 10 h given that its duration exceeds 9 h?

If X represents the time to repair the machine, the density function of X is given by

$$f(x) = \lambda e^{-\lambda x} = \frac{1}{2} e^{-\frac{x}{2}}, x > 0$$

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

$$= \left(-e^{-\frac{x}{2}}\right)_2^\infty = e^{-1} = 0.3679$$

(b) $P\{X \geq 10/X > 9\} = P\{X > 1\}$, (by the memoryless property)

$$= \int_1^\infty \frac{1}{2} e^{-\frac{x}{2}} \, dx$$

$$= \left(-e^{-\frac{x}{2}}\right)_0^\infty = e^{-0.5} = 0.6065$$

Example 10

The life length X of an electronic component follows an exponential distribution. There are 2 processes by which the component may be manufactured. The expected life length of the component is 100 h. if process I is used to manufacture, while it is 150 h if process II is used. The cost of manufacturing a single component by process I is Rs. 10, while it is Rs. 20 for process II. Moreover if the component lasts less than the guaranteed life of 200 h, a loss of Rs. 50 is to be borne by the manufacturer. Which process is advantageous to the manufacturer?

If process I is used, the density function of X is given by

$$f(x) = \frac{1}{100} e^{-x/100}, x > 0.$$

$$\therefore P(X \geq 200) = \int_{200}^\infty \frac{1}{100} e^{-x/100} \, dx$$

$$= \left(-e^{-x/100}\right)_{200}^\infty = e^{-2}$$

$$\therefore P(X < 200) = 1 - e^{-2}$$

Similarly, if process II is used,

$$P(X \geq 200) = e^{-4/3} \text{ and } P(X < 200) = 1 - e^{-4/3}$$

Let C_1 and C_2 be the costs per component corresponding to the processes I and II respectively.

$$\text{Then} \quad C_1 = \begin{cases} 10, & X \geq 200 \\ 60, & X < 200 \end{cases}$$

$$\therefore E(C_1) = 10 \times P(X \geq 200) + 60 \cdot P(X < 200)$$
$$= 10 e^{-2} + 60 (1 - e^{-2})$$
$$= 60 - 50 e^{-2} = 53.235$$

$$\text{Now} \quad C_2 = \begin{cases} 20, & X \geq 200 \\ 70, & X < 200 \end{cases}$$

$$\therefore E(C_2) = 20 \times P(X \geq 200) + 70 \times P(X < 200)$$
$$= 20 e^{-4/3} + 70 (1 - e^{-4/3})$$
$$= 70 - 50 e^{-4/3} = 56.765$$

Since $E(C_1) < E(C_2)$, process I is advantageous to the manufacturer.

Example 11

If the density function of a continuous RV X is $f(x) = c\, e^{-b(x-a)}$, $a \leq x$, where a, b, c are constants. Show that $b = c = \frac{1}{\sigma}$ and $a = \mu - \sigma$, where $\mu = E(X)$ and $\sigma^2 = \text{Var}(X)$.

Since $f(x)$ is a density function, $\int_a^\infty f(x) \, dx = 1$.

i.e., $\int_a^\infty c\, e^{-b(x-a)} \, dx = 1$

i.e., $c \left\{ \frac{e^{-b(x-a)}}{-b} \right\}_a^\infty = 1$

i.e., $\frac{c}{b} = 1$ or $b = c$ (1)

Now $\mu = E(X) = \int_a^\infty bx\, e^{-b(x-a)} \, dx$

$$= b\, e^{ab} \left[x \cdot \left(\frac{e^{-bx}}{-b} \right) - \frac{e^{-bx}}{b^2} \right]_a^\infty$$

$$= b\, e^{ab}\left[\frac{a}{b}e^{-ab} + \frac{1}{b^2}e^{-ab}\right]$$

$$= a + \frac{1}{b} \qquad (2)$$

$$E(X^2) = \int_a^{\infty} bx^2 e^{-b(x-a)}\,dx$$

$$= b\, e^{ab}\left[x^2\left(\frac{e^{-bx}}{-b}\right) - 2x\left(\frac{e^{-bx}}{-b^2}\right) + 2\left(\frac{e^{-bx}}{-b^3}\right)\right]_a^{\infty}$$

$$= b\left[\frac{a^2}{b} + \frac{2a}{b^2} + \frac{2}{b^3}\right]$$

$$= \frac{1}{b^2}(a^2b^2 + 2ab + 2)$$

$$\text{Var}(X) = E(X^2) - \{E(X)\}^2$$

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

i.e., $\sigma^2 = \frac{1}{b^2}$ or $\sigma = \frac{1}{b}$ (3)

From (1) and (3), we get $b = c = \frac{1}{\sigma}$.

From (2) and (3), $\mu - \sigma = a$.

Example 12

In a certain city, the daily consumption of electric power in millions of kilowatt-hours can be treated as a RV having an Erlang distribution with parameters $\lambda = \frac{1}{2}$ and $k = 3$. If the power plant of this city has a daily capacity of 12 millions kilowatt-hours, what is the probability that this power supply will be inadequate on any given day.

Let X represent the daily consumption of electric power (in millions of kilo-watt-hours). Then the density function of X is given as

$$f(x) = \frac{\left(\frac{1}{2}\right)^3}{\Gamma(3)} x^2 e^{-x/2},\ x > 0$$

$P\{$the power supply is inadequate$\}$

$$= P(X > 12) = \int_{12}^{\infty} f(x)\,dx \qquad [\because \text{The daily capacity is only 12}]$$

$$= \int_{12}^{\infty} \frac{1}{\Gamma(3)}\cdot\frac{1}{2^3} x^2 e^{-x/2}\,dx$$

$$= \frac{1}{16}\left[x^2\left(\frac{e^{-x/2}}{-\frac{1}{2}}\right) - 2x\left(\frac{e^{-x/2}}{\frac{1}{4}}\right) + 2\left(\frac{e^{-x/2}}{-\frac{1}{8}}\right)\right]_{12}^{\infty}$$

$$= \frac{1}{16}e^{-6}(288 + 96 + 16)$$

$$= 25\, e^{-6} = 0.0625$$

Example 13

If a company employes n sales persons, its gross sales in thousands of rupees may be regarded as a RV having an Erlang distribution with $\lambda = \frac{1}{2}$ and $k = 80\sqrt{n}$. If the sales cost is Rs. 8000 per salesperson, how many salespersons should the company employ to maximise the expected profit?

Let X represent the gross sales (in Rupees) by n salespersons.

X follows the Erlang distribution with parameters $\lambda = \frac{1}{2}$ and $k = 80{,}000\sqrt{n}$.

$$\therefore E(X) = \frac{k}{\lambda} = 1{,}60{,}000\sqrt{n}$$

If y denotes the total expected profit of the company, then

$$y = \text{total expected sales} - \text{total sales cost}$$

$$= 1{,}60{,}000\sqrt{n} - 8000\,n$$

$$\frac{dy}{dn} = \frac{80{,}000}{\sqrt{n}} - 8000$$

$$= 0, \text{ when } \sqrt{n} = 10 \text{ or } n = 100$$

$$\frac{d^2y}{dn^2} = -\frac{40{,}000}{n^{3/2}} < 0, \text{ when } n = 100.$$

Therefore, y is maximum, when $n = 100$.

That is the company should employ 100 salespersons in order to maximise the total expected profit.

Example 14

Consumer demand for milk in a certain locality, per month, is known to be a general Gamma (Erlang) RV. If the average demand is a litres and the most likely demand is b litres ($b < a$), what is the variance of the demand?

Let X represent the monthly consumer demand of milk.

Average demand is the value of $E(X)$.

Most likely demand is the value of the mode of X or the value of X for which its density function is maximum.

If $f(x)$ is the density function of X, then

$$f(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x} \quad x > 0$$

$$f'(x) = \frac{\lambda^k}{\Gamma(k)} [(k-1) x^{k-2} e^{-\lambda x} - \lambda x^{k-1} e^{-\lambda x}]$$

$$= \frac{\lambda^k}{\Gamma(k)} x^{k-2} e^{-\lambda x} \{(k-1) - \lambda x\}$$

$$= 0, \text{ when } x = 0, x = \frac{k-1}{\lambda}$$

$$f''(x) = \frac{\lambda^k}{\Gamma(k)} [-\lambda x^{k-2} e^{-\lambda x} + \{(k-1) - \lambda x\} \frac{d}{dx} \{x^{k-2} e^{-\lambda x}\}]$$

$$< 0, \text{ when } x = \frac{k-1}{\lambda}$$

Therefore $f(x)$ is maximum, when $x = \frac{k-1}{\lambda}$.

i.e., Most likely demand $= \frac{k-1}{\lambda} = b$ (1)

and $E(X) = \frac{k}{\lambda} = a$ (2)

Now $\text{Var}(X) = \frac{k}{\lambda^2} = \frac{k}{\lambda} \cdot \frac{1}{\lambda}$

$= a(a-b),$ [from (1) and (2)]

Example 15

A random sample of size n is taken from a general Gamma (Erlang) distribution with parameters λ and k. Show that the mean $\overline{X}$ of the sample also follows a Gamma distribution with parameters $n\lambda$ and nk.

If X follows Erlang distribution with parameters λ and k, then the MGF of X is given by

$$M_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-k}$$

If $X_1, X_2, \ldots, X_n$ are the members of the sample drawn, then each X_i follows Erlang distribution with MGF equal to $\left(1 - \frac{t}{\lambda}\right)^{-k}$ and also they are independent.

Therefore, by the reproductive property,

$$M_{X_1 + X_2 + \ldots + X_n}(t) = \left(1 - \frac{t}{\lambda}\right)^{-nk}$$

$$\therefore \quad M_{\overline{X}}(t) = M_{\frac{1}{n}(X_1 + X_2 + \ldots + X_n)}(t)$$

$$= M_{X_1 + X_2 + \ldots + X_n}\left(\frac{t}{n}\right) \qquad [\because M_{aX}(t) = M_X(at)]$$

$$= \left(1 - \frac{t}{n\lambda}\right)^{-nk}$$

which is the MGF of an Erlang distribution with parameters $n\lambda$ and nk.

Therefore, $\overline{X}$ also follows an Erlang distribution with density function

$$\frac{(n\lambda)^{nk}}{\Gamma(nk)} \cdot x^{nk-1} \cdot e^{-n\lambda x}, x > 0.$$

Example 16

If the conditional distribution of Y, given $X = x$, is an exponential distribution with parameter x and if the unconditional distribution of X is an Erlang distribution with parameters $\lambda (> 0)$ and $k (> 2)$, prove that the conditional distribution of X, given $Y = y$, is an Erlang distribution with parameters $\lambda + y$ and $k + 1$.

Given: $F_{Y/X}(y) = x e^{-xy}, y > 0$ and $x > 0$

and $f_X(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x}, x > 0$

If $f(x, y)$ denotes the joint density function of (X, Y), then $f_{Y/X}(y) = \frac{f(x, y)}{f_X(x)}$

$$\therefore \quad f(x, y) = \frac{\lambda^k}{\Gamma(k)} x^k e^{-(\lambda + y)x}, x > 0, y > 0.$$

Now $f_Y(y)$ = the marginal density function of Y

$$= \int_0^\infty f(x, y)\, dx$$

$$= \frac{\lambda^k}{\Gamma(k)} \cdot \int_0^\infty x^k e^{-(\lambda + y)x}\, dx$$

$$= \frac{\lambda^k}{\Gamma(k)} \frac{1}{(\lambda + y)^{k+1}} \cdot \int_0^\infty t^k e^{-t}\, dt \text{ [on putting } (\lambda + y)x = t]$$

$$= \frac{\lambda^k}{\Gamma(k)} \frac{1}{(\lambda + y)^{k+1}} \cdot \Gamma(k+1) = \frac{k\lambda^k}{(\lambda + y)^{k+1}}, \quad y > 0$$

Now $$f_{X/Y}(x) = \frac{f(x, y)}{f_Y(y)}$$

$$= \frac{\frac{\lambda^k}{\Gamma(k)} x^k e^{-(\lambda+y)x}}{\frac{k\lambda^k}{(\lambda+y)^{k+1}}}, \quad x > 0 \text{ and } y > 0$$

$$= \frac{(\lambda + y)^{k+1}}{\Gamma(k+1)} \cdot x^k e^{-(\lambda+y)x}, \quad x > 0$$

This is the density function of an Erlang distribution with parameters $\lambda + y$ and $k + 1$.

Example 17

Each of the 6 tubes of a radio set has a life length (in years) which may be considered as a RV that follows a Weibull distribution with parameters $\alpha = 25$ and $\beta = 2$. If these tubes function independently of one another, what is the probability that no tube will have to be replaced during the first 2 months of service?

If X represents the life length of each tube, then its density function $f(x)$ is given by

$$f(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} \quad x > 0$$

i.e., $f(x) = 50x\, e^{-25x} \quad x > 0$

Now P(a tube is not to be replaced during the first 2 months)

$$= P\left(X > \frac{1}{6}\right)$$

$$= \int_{\frac{1}{6}}^\infty 50x\, e^{-25x^2}\, dx$$

$$= \left(-e^{-25x^2}\right)_{\frac{1}{6}}^\infty = e^{-25/36}$$

$\therefore$ P(all the 6 tubes are not to be replaced during the first 2 months)

$= (e^{-25/36})^6$ (by independence)

$= e^{-25/6}$

$= 0.0155$

Example 18

If the life X (in years) of a certain type of car has a Weibull distribution with the parameter $\beta = 2$, find the value of the parameter α, given that probability that the life of the car exceeds 5 years is $e^{-0.25}$. For these values of α and β, find the mean and variance of X.

The density function of X is given by

$$f(x) = 2\alpha x\, e^{-\alpha x^2}, x > 0 \quad [\because \beta = 2]$$

Now $$P(X > 5) = \int_5^\infty 2\alpha x\, e^{-\alpha x^2}\, dx$$

$$= \left(-e^{-\alpha x^2}\right)_5^\infty$$

$$= e^{-25\alpha}$$

Given that $P(X > 5) = e^{-0.25}$

$\therefore$ $e^{-25\alpha} = e^{-0.25}$

$\therefore$ $\alpha = \frac{1}{100}$

For the Weibull distribution with parameters α and β, $E(X) = \alpha^{-1/\beta}\, \Gamma\left(\frac{1}{\beta} + 1\right)$

$\therefore$ Required mean $= \left(\frac{1}{100}\right)^{-\frac{1}{2}} \cdot \Gamma\left(\frac{3}{2}\right)$

$$= 10 \times \frac{1}{2}\, \Gamma\left(\frac{1}{2}\right)$$

$$= 5\sqrt{\pi}.$$

$$\text{Var}(X) = \alpha^{-\frac{2}{\beta}} \left[\Gamma\left(\frac{2}{\beta} + 1\right) - \left\{\Gamma\left(\frac{1}{\beta} + 1\right)\right\}^2\right]$$

$$= \left(\frac{1}{100}\right)^{-1}\left[\Gamma(2) - \left\{\Gamma\left(\frac{3}{2}\right)\right\}^2\right]$$

$$= 100\left[1 - \left(\frac{1}{2}\sqrt{\pi}\right)^2\right]$$

$$= 100\left(1 - \frac{\pi}{4}\right)$$

Example 19

If the time T to failure of a component follows a Weibull distribution with parameters α and β, find the hazard rate or conditional failure rate at time t of the component.

Refer to Example (19) in Worked Example 2(A).

If $f(t)$ is the density function of T and $h(t)$ is the hazard rate at time t, then

$$h(t) = \frac{f(t)}{1 - F(t)}$$

where $F(t)$ is the distribution function of T.

Now $\quad f(t) = \alpha\beta \times t^{\beta-1} e^{-\alpha t^\beta} \quad t > 0$

$\therefore \quad F(t) = P(T \le t)$

$$= \int_0^t \alpha\beta t^{\beta-1}, e^{-\alpha t^\beta}\, dt$$

$$= \left[-e^{-\alpha t^\beta}\right]_0^t$$

$$= 1 - e^{-\alpha t^\beta}$$

$$\therefore \quad h(t) = \frac{\alpha\beta t^{\beta-1} e^{-\alpha t^\beta}}{e^{-\alpha t^\beta}}$$

$$= \alpha\beta t^{\beta-1}$$

Example 20

If Y is the smallest item of 3 independent observations X_1, X_2, X_3 from a Weibull distribution with parameters α and β, show that Y also has a Weibull distribution. What are its parameters?

Each of X_1, X_2, X_3 follows the Weibull distribution whose density function is given by

$$f(x) = \alpha\beta \times x^{\beta-1} e^{-\alpha x^\beta},\ x > 0$$

Now $\quad P(Y > y) = P[\min(X_1, X_2, X_3) > y]$

$= P(X_1 > y) \times P(X_2 > y) \times P(X_3 > y),$

(since X_1, X_2, X_3 are independent.)

$$= \{P(X_i > y)\}^3 \qquad (1)$$

Now $\quad P(X_i > y) = \int_y^\infty \alpha\beta x^{\beta-1} e^{-\alpha x^\beta}\, dx$

$$= \left(-e^{-\alpha x^\beta}\right)_y^\infty$$

$$= e^{-\alpha y^\beta} \qquad (2)$$

Using (2) in (1), we have

$$P(Y > y) = (e^{-\alpha y^\beta})^3 = e^{-3\alpha y^\beta}$$

Therefore, Y also has a Weibull distribution with parameters 3α and β.

Note *The result can be extended to n independent observations.*

Example 21

There are 400 students in the first year class of an engineering college. The probability that any student requires a copy of a particular Mathematics book from the college library on any day is 0.1. How many copies of the book should be kept in the library so that the probability may be greater than 0.95 that none of the students requiring a copy from the library has to come back disappointed? (Use normal approximation to the binomial distribution).

$p = P$(a student requires the book) $= 0.1$ and $q = 0.9$

$n =$ number of students $= 400$

If X represents the number of students requiring the book, then X follows a binomial distribution with mean $= np = 40$ and SD $= \sqrt{npq} = 6$.

As given in the problem, we may assume that X follows the distribution $N(40, 6)$.

Let m be the required number of books, satisfying the given condition.

i.e., $\quad P(X < m) > 0.95$

i.e., $\quad P\left(-\infty < \frac{X - 40}{6} < \frac{m - 40}{6}\right) > 0.95$

i.e., $\quad P\left(0 < Z < \frac{m - 40}{6}\right) > 0.45$

where Z is the standard normal variate.

From the table of areas under normal curve, we find that

$$P\{0 < Z < 1.65\} > 0.45$$

$$\therefore \quad \frac{m-40}{6} = 1.65$$

i.e., $m = 49.9$

Therefore, at least 50 copies of the book should be kept in the library.

Example 22

The marks obtained by a number of students in a certain subject are approximately normally distributed with mean 65 and standard deviation 5. If 3 students are selected at random from this group, what is the probability that at least 1 of them would have scored above 75?

If X represents the marks obtained by the students, X follows the distribution $N(65, 5)$.

P(a student scores above 75)

$$= P(X > 75) = P\left(\frac{75-65}{5} < \frac{X-65}{5} < \infty\right)$$

$= P(2 < Z < \infty)$, (where Z is the standard normal variate)

$= 0.5 - P(0 < Z < 2)$

$= 0.5 - 0.4772$, (from the table of areas)

$= 0.0228$

Let $p = P$(a student scores above 75) $= 0.0228$ then $q = 0.9772$ and $n = 3$.

Since p is the same for all the students, the number Y, of (successes) students scoring above 75, follows a binomial distribution.

P(at least 1 student scores above 75)

$= P$(at least 1 success)

$= P(Y \geq 1) = 1 - P(Y = 0)$

$= 1 - nC_0 \times P^0 q^n$

$= 1 - 3C_0\,(0.9772)^3$

$= 1 - 0.9333$

$= 0.0667$

Example 23

If the actual amount of instant coffee which a filling machine puts into '6-ounce' jars is a RV having a normal distribution with SD = 0.05 ounce and if only 3% of the jars are to contain less than 6 ounces of coffee, what must be the mean fill of these jars?

Let X be the actual amount of coffee put into the jars.

Then X follows $N(\mu, 0.05)$

Given: $P(X < 6) = 0.03$

$$\therefore P\left\{-\infty < \frac{X-\mu}{0.05} < \frac{6-\mu}{0.05}\right\} = 0.03$$

i.e., $$P\left\{-\infty < Z < \frac{6-\mu}{0.05}\right\} = 0.03$$

$\therefore$ $$P\left\{0 < Z < \frac{\mu-6}{0.05}\right\} = 0.47, \quad \text{(by symmetry)}$$

From the table of areas, we have

$P\{0 < Z < 1.808\} = 0.47$

$$\therefore \quad \frac{\mu-6}{0.05} = 1.808$$

$\therefore \quad \mu = 6.094$ ounces.

Example 24

In an engineering examination, a student is considered to have failed, secured second class, first class and distinction, according as he scores less than 45%, between 45% and 60%, between 60% and 75% and above 75% respectively. In a particular year 10% of the students failed in the examination and 5% of the students got distinction. Find the percentages of students who have got first class and second class. (Assume normal distribution of marks).

Let X represent the percentage of marks scored by the students in the examination.

Let X follow the distribution $N(\mu, \sigma)$.

Given: $P(X < 45) = 0.10$ and $P(X > 75) = 0.05$

i.e., $$P\left(-\infty < \frac{X-u}{\sigma} < \frac{45-\mu}{\sigma}\right) = 0.10 \text{ and}$$

$$P\left(\frac{75-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \infty\right) = 0.05$$

i.e., $$P\left(-\infty < Z < \frac{45-\mu}{\sigma}\right) = 0.10 \text{ and}$$

$$P\left(\frac{75-\mu}{\sigma} < Z < \infty\right) = 0.05$$

$\therefore$ $$P\left(0 < Z < \frac{\mu-45}{\sigma}\right) = 0.40 \text{ and}$$

$$P\left(0 < Z < \frac{75-\mu}{\sigma}\right) = 0.45$$

From the table of areas, we get

$$\frac{\mu-45}{\sigma} = 1.28 \text{ and } \frac{75-\mu}{\sigma} = 1.64$$

i.e., $\mu - 1.28\,\sigma = 45$ (1)

and $\mu + 1.64\,\sigma = 75$ (2)

Solving equations (1) and (2), we get

$\mu = 58.15$ and $\sigma = 10.28$

Now P (a student gets first class)

$$= P(60 < X < 75)$$
$$= P\left\{\frac{60 - 58.15}{10.28} < Z < \frac{75 - 58.15}{10.28}\right\}$$
$$= P\{0.18 < Z < 1.64\}$$
$$= P\{0 < Z < 1.64\} - P\{0 < Z < 0.18\}$$
$$= 0.4495 - 0.0714 = 0.3781$$

$\therefore$ Percentage of students getting first class = 38 (approximately)

Now percentage of students getting second class

= 100 – (sum of the percentages of students who have failed, got first class and got distinction)

= 100 – (10 + 38 + 5), approximately.

= 47 (approximately)

Example 25

The percentage X of a particular compound contained in a rocket fuel follows the distribution $N(33, 3)$, though the specification for X is that it should lie between 30 and 35. The manufacturer will get a net profit (per unit of the fuel) of Rs. 100, if $30 < X < 35$, Rs. 50, if $25 < X \le 30$ or $35 \le X < 40$ and incur a loss of Rs. 60 per unit of the fuel otherwise. Find the expected profit of the manufacturer. If he wants to increase his expected profit by 50% by increasing the net profit on that category of the fuel that meets the specification, what should be the new net profit per unit of the fuel of this category?

$$P(30 < X < 35) = P\left\{\frac{30 - 33}{3} < \frac{X - 33}{3} < \frac{35 - 33}{3}\right\}$$
$$= P\{-1 < Z < 0.67\}$$
$$= P\{0 < Z < 1\} + P\{0, < Z < 0.67\}$$
$$= 0.3413 + 0.2486, \quad \text{(using the table of areas)}$$
$$= 0.5899$$

$$P(25 < X \le 30) = P\left\{\frac{25 - 33}{3} < Z < \frac{30 - 33}{3}\right\}$$
$$= P(-2.67 < Z < -1)$$
$$= P(1 < Z < 2\ 67), \text{ (by symmetry)}$$
$$= P(0 < Z < 2.67) - P(0 < Z < 1)$$
$$= 0.4962 - 0.3413$$
$$= 0.1549$$

$$P(35 \le X < 40) = P\left\{\frac{35 - 33}{3} < Z < \frac{40 - 33}{3}\right\}$$
$$= P\{0.67 < Z < 2.33\}$$

$$= P\{0 < Z < 2.33\} - P\{0 < Z < 0.67\}$$
$$= 0.4901 - 0.2486$$
$$= 0.2415$$

$\therefore P\{(25 < X \le 30)) \text{ or } (35 \le X < 40)\}$

$$= P(25 < X \le 30) + P(35 \le X < 40),$$

(since the 2 ranges are mutually exclusive)

$$= 0.1549 + 0.2415 = 0.3964$$

$$P\{X < 25 \text{ or } X > 40\} = 1 - (0.5899 + 0.3964)$$
$$= 0.0137$$

Profit/unit	*Probability*
Rs. 100	0.5899
Rs. 50	0.3964
Rs. –60	0.0137

$\therefore E(\text{Profit per unit}) = \text{Rs } (100 \times 0.5899 + 50 \times 0.3964 - 60 \times 0.0137)$

= Rs 79, approximately

Let the revised net profit per unit of the first category fuel be k.

Then E(Revised profit per unit)

$$= \text{Rs. } (k \times 0.5899 + 50 \times 0.3964 - 60 \times 0.0137)$$
$$= \text{Rs. } (0.5899\,k + 18.998)$$

E(Revised profit per unit) = Rs. 79 + Rs. 39.5, as per the manufacturer's wish

$\therefore 0.5899\,k + 18.998 = 118.5$

$$\therefore \quad k = \frac{118.5 - 18.998}{0.5899}$$
$$= 168.68 \simeq \text{Rs. } 169 \text{ nearly.}$$

Example 26

The marks obtained by the students in Mathematics, Physics and Chemistry in an examination are normally distributed with the means 52, 50 and 48 and with standard deviations 10, 8 and 6 respectively. Find the probability that a student selected at random has secured a total of (i) 180 or above and (ii) 135 or less.

Let X, Y, Z denote the marks obtained by students in Mathematics, Physics and Chemistry respectively.

Given: X follows $N(52, 10)$, Y follows $N(50, 8)$ and Z follows $N(48, 6)$.

By the additive property of normal distribution, $T = X + Y + Z$ follows the distribution

$$N\{52 + 50 + 48, \sqrt{10^2 + 8^2 + 6^2}\}$$

i.e., $\quad N(150, 14.14)$

(i) $$P(T \ge 180) = P\left\{\frac{180 - 150}{14.14} < \frac{T - 150}{14.14} < \infty\right\}$$
$$= P\{2.12 < Z < \infty\}$$

$= 0.5 - P\{0 < Z < 2.12\}$
$= 0.5 - 0.4830,$ (from the table of areas)
$= 0.0170$

(ii) $P(T \le 135) = P\left\{\frac{T-150}{14.14} < \frac{135-150}{14.14}\right\}$
$= P(-\infty < Z < -1.06\}$
$= P\{1.06 < Z < \infty\},$ (by symmetry)
$= 0.5 - P\{0 < Z < 1.06\}$
$= 0.5 - 0.3554$
$= 0.1446$

Example 27

The independent RVs X and Y have distributions $N(45, 2)$ and $N(44, 1.5)$ respectively. What is the probability that randomly chosen values of X and Y differ by 1.5 or more?

X is $N(45, 2)$ and Y is $N(44, 1.5)$

$\therefore$ By the additive property,

$U = X - Y$ follows the distribution $N(1, \sqrt{4 + 2.25})$

i.e., $N(1, 2.5)$

Now $P\{X$ and Y differ by 1.5 or more$\}$

$= P\{|X - Y| \ge 1.5\}$
$= P\{|U| \ge 1.5\}$
$= 1 - P(|U| \le 1.5)$
$= 1 - P\{-1.5 \le U \le 1.5\}$
$= 1 - P\left\{\frac{-1.5-1}{2.5} \le \frac{U-1}{2.5} \le \frac{1.5-1}{2.5}\right\}$
$= 1 - P\{-1 \le Z \le 0.2\}$
$= 1 - \{P(0 \le z \le 1) + P(0 \le z \le 0.2)\}$
$= 1 - \{0.3413 + 0.0793\}$, (from the table of areas)
$= 0.5794.$

Example 28

If X and Y are independent RVs, each following $N(0, 3)$, what is the probability that the point (X, Y) lies between the lines $3X + 4Y = 5$ and $3X + 4Y = 10$?

X follows $N(0, 3)$ and Y follows $N(0, 3)$.

Therefore, by the additive property of normal distribution,

$U = 3X + 4Y$ follows $N[3 \times 0 + 4 \times 0, \sqrt{9 \times 9 + 16 \times 9}]$

i.e., $N(0, 15)$

Now $P\{$the point (X, Y) lies between the lines $3X + 4Y = 5$ and $3X + 4Y = 10\}$

$= P\{5 < 3X + 4Y < 10\}$
$= P\{5 < U < 10\}$
$= P\left\{\frac{5-0}{15} < \frac{U-0}{15} < \frac{10-0}{15}\right\}$
$= P\{0.33 < Z < 0.67\},$ where Z is the standard normal variable
$= P(0 < Z < 0.67) - P(0 < Z < 0.33)$
$= 0.2486 - 0.1293$, (from the table of areas)
$= 0.1193.$

Example 29

If X and Y are independent RVs following $N(8, 2)$ and $N(12, 4\sqrt{3})$ respectively, find the value of λ such that

$P(2X - Y \le 2\lambda) = P(X + 2Y \ge \lambda)$

By the additive property of normal distribution

$U = 2X - Y$ follows $N\{2 \times 8 - 12, \sqrt{4 \times 4 + 1 \times 48}\}$

i.e., $N(4, 8)$

and $V = X + 2Y$ follows $N\{8 + 2 \times 12, \sqrt{4 + 4 \times 48}\}$

i.e., $N(32, 14)$

Now $P(2X - Y \le 2\lambda) = P(X + 2Y \ge \lambda)$

i.e., $P(U \le 2\lambda) = P(V \ge \lambda)$

i.e., $P\left(\frac{U-4}{8} \le \frac{2\lambda-4}{8}\right)$

$= P\left(\frac{V-32}{14} \ge \frac{\lambda-32}{14}\right)$

i.e., $P\left(Z \le \frac{2\lambda-4}{8}\right) = P\left(Z \ge \frac{\lambda-32}{14}\right)$, where Z is the standard normal variable.

$\therefore$ $\frac{2\lambda-4}{8} = -\left(\frac{\lambda-32}{14}\right)$

i.e., $28\lambda - 56 = 256 - 8\lambda$

$\therefore$ $\lambda = \frac{26}{3}$

Example 30

Fit a normal distribution to the following frequency distribution and hence find the theoretical frequencies:

x:	125,	135,	145,	155,	165,	175,	185,	195,	205	Total
f:	1,	1,	14,	22,	25,	19,	13,	3,	2,	100

To fit a normal distribution for the given data, we require the density function of the normal distribution which involves the mean and SD. Let us now compute the mean $\bar{x}$ and SD s of the given distribution and assume them as μ and σ of the approximate normal distribution.

x	f	$d = \frac{x-165}{10}$	fd	fd^2
125	1	−4	−4	16
135	1	−3	−3	9
145	14	−2	−28	56
155	22	−1	−22	22
165	25	0	0	0
175	19	1	19	19
185	13	2	26	52
195	3	3	9	27
205	2	4	8	32
Total:	100	–	5	233

$$\bar{x} = A + \frac{c}{N}\sum fd = 165 + \frac{10}{100} \times 5 = 165.5$$

$$s^2 = c^2\left\{\frac{1}{N}\sum fd^2 - \left(\frac{1}{N}\sum fd\right)^2\right\}$$

$$= 10^2\,(2.33 - 0.0025)$$

$$= 232.75$$

$\therefore$ $$s = 15.26$$

Therefore, the density function of the approximate normal distribution that fits the given distribution is

$$f(x) = \frac{1}{15.26\sqrt{2\pi}}\, e^{-(x-165.5)^2/465.5} \qquad -\infty < x < \infty$$

To find the theoretical frequency of the class $120 \le X \le 130$, whose mid-value is 125, we first get $P(120 \le X \le 130) = P\left\{\frac{120-165.5}{15.26} \le Z \le \frac{130-165.5}{15.26}\right\}$ and multiply this probability by the total frequency. Proceeding likewise, we get all the theoretical frequencies. The computations are shown in the table given in the next page.

Class mid-value (x)	Class $(X_L \le X \le X_R)$	$Z_L \le Z \le Z_R$	$P(Z_L \le Z \le Z_R)$	Theoretical frequency	Corrected frequency
125	$120 \le X \le 130$	$-2.98 \le Z \le -2.33$	0.0085	0.85	1
135	$130 \le X \le 140$	$-2.33 \le Z \le -1.67$	0.0376	3.76	4
145	$140 \le X \le 150$	$-1.67 \le Z \le -1.02$	0.1064	10.64	11
155	$150 \le X \le 160$	$-1.02 \le Z \le -0.36$	0.2055	20.55	21
165	$160 \le X \le 170$	$-0.36 \le Z \le 0.29$	0.2547	25.47	25
175	$170 \le X \le 180$	$0.29 \le Z \le 0.95$	0.2148	21.48	21
185	$180 \le X \le 190$	$0.95 \le Z \le 1.61$	0.1174	11.74	12
195	$190 \le X \le 200$	$1.61 \le Z \le 2.26$	0.0418	4.18	4
205	$200 \le X \le 210$	$2.26 \le Z \le 2.92$	0.0102	1.02	1
					Total: 100

Exercise 5(B)

Part A (Short answer questions)

1. If X has uniform distribution in $(-3, 3)$, find $P(|X-2|<2)$.
2. If X has uniform distribution in $(-a, a)$, $a>0$, find 'a' such that $P(|X|<1) = P(|X|>1)$.
3. If the MGF of a continuous RV X is $\frac{1}{t}(e^{5t}-e^{4t})$, $t \neq 0$, what is the distribution of X? What are its mean and variance?
4. A continuous RV X has the density function $c\, e^{-x/5}$; $x>0$. Find c, $E(X)$ and $\text{Var}(X)$.
5. What do you mean by memoryless property of the exponential distribution?
6. If X and Y are independent identically distributed RVs, each with density function e^{-x}, $x>0$, find the density function of $(X+Y)$.
7. Define Erlang distribution and also give its mean and variance.
8. Write down the MGF of simple Gamma distribution and hence find its mean and variance.
9. Give the values of β_1 and β_2 coefficients of the Erlang distribution with parameters $(1, k)$.
10. Find where the maximum occurs for the Erlang density function.
11. If X has uniform distribution in $(0, 2)$ and Y has exponential distribution with parameter λ, find λ such that $P(X<1) = P(Y<1)$.
12. If X has uniform distribution in $(-1, 3)$ and Y has exponential distribution with parameter λ, find λ such that $\text{Var}(X) = \text{Var}(Y)$.
13. Define Weibull distribution and also give its mean and variance.
14. Find the value of k, mean and variance of the normal distribution whose density function is $k \cdot 2^{-x^2}$ $-\infty < x < \infty$.
15. If X follows $N(30, 5)$ and Y follows $N(15, 10)$ show that $P(26 \le X \le 40) = P(7 \le Y \le 35)$.
16. If X follows $N(3, 2)$, find the value of k such that $P(|X-3|>k) = 0.05$.
17. If $\log_{10} X$ follows $N(4, 2)$, find $P(1.202 < X < 83180000)$, given that $\log_{10}(1202) = 3.08$ and $\log_{10}(8318) = 3.92$.
18. For a certain normal distribution, the first moment about 10 is 40 and the fourth moment about 50 is 48. What are its mean and SD?
19. Show that, for a normal distribution, the quartile deviation, the mean deviation and the standard deviation are in the ratio 10 : 12 : 15.
20. If 2 normal universes A and B have the same total frequency, but the SD of A is k times the SD of B, prove that the maximum frequency of A is $\frac{1}{k}$ times that of B.
21. State the reproductive property of normal distribution.
22. If X and Y are independent RVs having $N(1, 2)$ and $N(2, 2)$ respectively find the density function of $(X+2Y)$.

23. Why is normal distribution considered an important distribution?

Part B

24. X is uniformly distributed with mean 1 and variance $\frac{4}{3}$. If 3 independent observations of X are made, what is the probability that all of them are negative?
25. A point D is chosen on the line AB whose length is 1 and whose mid-point is C. If the distance X from D to A is a RV having a uniform distribution in $(0, a)$, what is the probability that AD, BD and AC will form a triangle?
26. A passenger arrives at a bus stop at 10 A.M., knowing that the bus will arrive at some time uniformly distributed between 10 A.M. and 10.30 A.M. What is the probability that he will have to wait longer than 10 min? If at 10.15 A.M. the bus has not yet arrived, what is the probability that he will have to wait at least 10 additional minutes?
27. A man and a woman agree to meet at a certain place between 10 A.M. and 11 A.M. They agree that the one arriving first will wait 15 min for the other to arrive. Assuming that the arrival times are independent and uniformly distributed, find the probability that they meet.
28. The RVs a and b are independently and uniformly distributed in the intervals $(0, 6)$ and $(0, 9)$ respectively. Find the probability that the roots of the equation $x^2 - ax + b = 0$ are real.
29. If a, b, c are randomly chosen between 0 and 1, find the probability that the quadratic equation $ax^2 + bx + c = 0$ has real roots.
30. X, Y, Z are independent RVs, each following a uniform distribution in $(0, 1)$. If U = maximum of X, Y, Z and V = minimum of X, Y, Z, find
    (i) $P\left(U \le \frac{1}{2}\right)$, (ii) $P\left(V \ge \frac{1}{3}\right)$ and (iii) $P\left\{U \le \frac{1}{2} \text{ and } V \ge \frac{1}{3}\right\}$.
31. If the number of kilometres that a car can run before its battery wears out is exponentially distributed with an average value of 10,000 km and if the owner desires to take a 5000 km trip, what is the probability that he will be able to complete his trip without having to replace the car battery. Assume that the car has been used for some time. What is the probability, when the distribution is not exponential?
32. If X is exponentially distributed, prove that the probability that X exceeds its expected value is less than 0.5.
33. The amount of time that a watch will run without having to be reset is a RV having an exponential distribution with mean 120 days. Find the probability that such a watch will
    (i) have to be set in less than 24 days and
    (ii) not have to be reset in at least 180 days.
34. The daily consumption of milk in excess of 20,000 litres in a town is approximately exponentially distributed with parameter 1/3000. The town has a daily stock of 35,000L. What is the probability that of 2 days selected at random, the stock is insufficient for both days?

[**Hint:** If Y denotes the daily consumption of milk, then $X = Y - 20{,}000$ follows the exponential distribution.]

35. The length of the shower on a tropical island during rainy season has an exponential distribution with parameter 2, time being measured in minutes. What is the probability that a shower will last more than 3 min? If a shower has already lasted for 2 min, what is the probability that it will last for at least one more minute?
36. If X is exponentially distributed with parameter λ, find the value of k such that

$$P(X > k)/P(X \le K) = a.$$

37. If X is exponentially distributed with parameter λ, prove that the RV $Y = e^{-\lambda X}$ is uniformly distributed in (0, 1).
38. If X_1, X_2, X_3 are independent RVs having exponential distributions with parameters $\lambda_1, \lambda_2, \lambda_3$ respectively, prove that $Y = \text{minimum}\ (X_1, X_2, X_3)$ follows an exponential distribution with parameter $(\lambda_1 + \lambda_2 + \lambda_3)$.
[**Hint:** Find the distribution function of $Y = F(y) = 1 - P\{\min(X_1, X_2, X_3) > y\}$.]
39. The daily consumption of milk in a town in excess of 20,000L is approximately distributed as an Erlang variate with parameters

$\lambda = \dfrac{1}{10{,}000}$ and $k = 2$. The town has a daily stock of 30,000L. What is the probability that the stock is insufficient on a particular day?

40. Find the probabilities that the value of a RV will exceed 4, if it has an Erlang distribution with

(i) $\lambda = \dfrac{1}{3}$ and $k = 2$ and (ii) $\lambda = \dfrac{1}{4}$ and $k = 3$.

41. Show that for the Erlang distribution with parameters λ, k, (mean–mode)/

SD $= \dfrac{1}{\sqrt{k}}$.

[**Hint:** If $f(x)$ is the Erlang density function, the mode is the value of x for which $f(x)$ is maximum.]

42. If X follows the Erlang distribution with parameters λ and k, prove that the

expected value of the positive square root of X is $\Gamma\left(k+\frac{1}{2}\right)\Big/\sqrt{\lambda}\,\Gamma k$.

43. If $X_1, X_2, \ldots, X_n$ are independent RVs, each following the same exponential distribution with parameter λ, prove that $X_1 + X_2 + \ldots + X_n$ follows an Erlang distribution with parameters λ and n.
[**Hint:** Use moment generating function. Also see example (7) in section 4(b).]
44. A random sample of size n is taken from a population which is exponentially distributed with parameter λ. If $\overline{X}$ is the sample mean, show that $n\lambda\overline{X}$ follows a simple Gamma distribution with parameter n.
[**Hint:** Use moment generating function.]

45. If the service life, in hours, of a semiconductor is a RV having a Weibull distribution with the parameters $\alpha = 0.025$ and $\beta = 0.5$,
    (i) how long can such a semiconductor be expected to last? and
    (ii) what is the probability that such a semiconductor will still be in operating condition after 4000 h?
46. Find the mode of the Weibull distribution with parameters α and β, when $\alpha > 1$.
47. If the hazard rate at time t of a system is given by $h(t) = \alpha\beta t^{\beta-1}$, prove that the time to failure of the system follows a Weibull distribution with parameters α and β.
48. If the RV X follows an exponential distribution with parameter 2, prove that $Y = X^3$ follows a Weibull distribution with parameters 2 and $\dfrac{1}{3}$.
49. Find the probability of failure-free performance of roller-bearings over a period of 10^4 h if the life expectancy of the bearings is defined by Weibull distribution with parameters $\alpha = 10^{-7}$ and $\beta = 1.5$.
[**Hint:** P(failure-free performance over a period t) $= P$(the component does not fail in $(0, t)$] $= P(T \ge t)$, where T is the life expectancy or time to failure of the component.]
50. The time when a country bus passes a certain point is distributed normally with a mean 9.25 A.M. and a SD of 3 min. What is the least time one could arrive at this point and still have a probability of 0.99 of catching the bus?
[**Hint:** If T is the time in minutes past 9 A.M., then T follows $N(25, 3)$.]
51. The marks obtained by a number of students in a certain subject are assumed to be approximately normally distributed with mean 55 and a SD of 5. If 5 students are taken at random from this set, what is the probability that 3 of them would have scored marks above 60?
52. The life lengths in hours of 2 electronic devices A and B have distributions $N(40, 6)$ and $N(45, 3)$ respectively. If the electronic device is to be used for a 45-h period, which device is to be preferred? If it is to be used for a 48-h period, which device is to be preferred?
53. The mean and SD of a certain group of 1000 high schoold grades, that are normally distributed are 78% and 11% respectively.
    (i) Find how many grades were above 90%?
    (ii) What was the highest grade of the lowest 10?
    (iii) What was the semi-interquartile range (Quartile deviation)?
    (iv) Within what limits did the middle 900 lie?
54. The local authorities in a certain city instal 10,000 electric lamps in the streets of the city. If these lamps have an average life of 1,000 burning hours with a standard deviation of 200 h, how many lamps might be expected to fail (i) in the first 800 burning hours? (ii) between 800 and 1200 burning hours? After how many burning hours would you expect (iii) 10% of the lamps to fail? (iv) 10% of the lamps to be still burning? Assume that the life of lamps is normally distributed.

55. In a normal population with mean 15 and SD 3.5, it is found that 647 observations exceed 16.25. What is the total number of observations in the population?
56. A RV has a normal distribution with SD 10. If the probability that the RV will take on a value less than 82.5 is 0.8212, what is the probability that it will take on a value greater than 58.3?
57. In a normal distribution, 7% of the items are under 35 and 89% are under 63. What are the mean and standard deviation of the distribution? What percentage of items are under 49?
58. A normal population has coefficient of variation equal to 2% and 8% of the population lies above 120 cm. What percentage of the population lies below 115 cm?
59. The breaking strength X of a certain kind of rope (in kg) has distribution $N(45, 1.8)$. Each 50 metre coil of rope brings a profit of Rs 1000, provided $X > 43$. If $X \le 43$, the rope may be used for a different purpose and a profit of Rs. 400 per coil is realised. Find the expected profit per coil.
60. The mean and standard deviation of marks in Mathematics are 45 and 10 respectively. The corresponding values for computer science are 50 and 15 respectively. Assuming that the marks in the two subjects are independent normal variates, find the probability that a student scores a total of marks lying between 100 and 120 in the 2 subjects.
61. If $\log_{10} X$ has the distribution $N(7, \sqrt{3})$ and $\log_{10} Y$ has the distribution $N(3, 1)$, find $P\{1.202 < \frac{X}{Y} < 8318 \times 10^4\}$, given that X and Y are independent.
[**Hint:** Find $P\{\log(1.202) < (\log X - \log Y) < \log(10^4 \times 8318\}$.]
62. If X and Y are independent RVs having normal distributions with a common mean μ, but with variances 4 and 48 respectively, such that $P(X + 2Y \le 3) = P(2X - Y \ge 4)$, determine μ.
63. Fit a normal distribution to the following distribution and hence find the theoretical frequencies:

Class :	60–65	65–70	70–75	75–80	80–85	85–90	90–95	95–100	Total
Frequency :	3	21	150	335	326	135	26	4	1000

ANSWERS

Exercise 5(A)

1. $p = \frac{2}{3}, q = \frac{1}{3}; n = 6; P(X \ge 1) = 1 - nC_0\, p^0 q^n = \frac{728}{729}$

2. $\mu_3 = npq\,(q-p);\ \beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(q-p)^2}{npq}$
3. Reqd. no. $= 256 \times 8C_4 \left(\frac{1}{2}\right)^8 = 70$
4. $p = \frac{2}{3}, q = \frac{1}{3}; P(X \ge 1) = 1 - 4C_0\, p^0 q^4 = \frac{80}{81}$
5. Reqd. probability $= \sum_{r=1}^{3} 4C_r \left(\frac{1}{2}\right)^4 = \frac{7}{8}$
6. $\mu = 5, \sigma^2 = 4; P(X < \mu - 2\sigma) = P(X < 1) = \left(\frac{4}{5}\right)^{25}$
7. $V = np(1-p); \frac{dV}{dp} = n(1-2p) = 0$, when $p = \frac{1}{2}$ and $\frac{d^2V}{dp^2} < 0$
8. The MGF of $B(n, p)$ is $(q + pe^t)^n$. The given MGF is that of $B(6, 0.4)$. Hence mean = 2.4 and SD = 1.2
9. When $p_1 = p_2$, the sum is also a binomial variate
10. $(X + Y)$ follows $B\left(8, \frac{1}{3}\right)$; $P(X + Y \ge 1) = 1 - \left(\frac{2}{3}\right)^8 = \frac{6305}{6561}$
11. $P(X = r) = e^{-2}\, 2^r/r!$
12. $2e^{-\lambda} + e^{-\lambda}\frac{\lambda^2}{2} = 2e^{-\lambda}\lambda$. $\therefore (\lambda - 2)^2 = 0$ or $\lambda = 2$
13. If λ is the parameter of the Poisson distribution, $\text{Var}(X) = E(X^2) - E^2(X)$.
i.e., $\lambda = 6 - \lambda^2$
$\therefore \lambda^2 + \lambda - 6 = 0 \therefore \lambda = E(X) = 2$, since $\lambda > 0$
14. If λ is the parameter, $e^{-\lambda} = 0.5 \therefore \text{Var}(X) = \lambda = \log 2$
15. $E(X^2) = \lambda^2 + \lambda = \lambda(\lambda + 1) = \lambda E(X + 1)$
16. $P(X \text{ is even}) = e^{-\lambda}\left\{1 + \frac{\lambda^2}{2!} + \frac{\lambda^4}{4!} + \cdots\right\} = e^{-\lambda}\cosh\lambda = \frac{1}{2}e^{-\lambda}(e^{\lambda} + e^{-\lambda})$
$= \frac{1}{2}(1 + e^{-2\lambda})$
17. $e^{4(e^t - 1)}$ is the MGF of a Poisson distribution with parameter 4. $\therefore \mu = 4$ and $\sigma = 2$.
$\therefore P(X = 6) = e^{-4} \cdot 4^6/6!$
18. $(X + Y)$ is a Poisson variate with parameter 2.
$\therefore P(X + Y = 2) = e^{-2}\, 2^2/2! = \frac{2}{e}$