Contents

20%

14 Marks

14 Marks

December

GTU. Winter 2019

Chap = 1, chap. 2 → 14 Marks
Chap 3, chap 4 → 14 Marks
Chap = 5 → 18 Marks
Chap = 6 → 17 Marks
Chap = 7 → 7 Marks

70 Marks.

from :- D.G. BORAD

-: Shreenathji Engineering Zone :

DzPatel

CHAPTER 5

Some Special Probability Distributions

Chapter Outline

5.1 INTRODUCTION

There are some specific distributions that are used in practice. There is a random experiment behind each of these distributions. Since these random experiments model a lot of real life phenomenon, these special distributions are used frequently in different applications. Often a random experiment that we encounter in practice is such that we are interested in the associated random variable X with such a standard distribution. This chapter discusses special random variables and their distributions. These include binomial distribution, Poisson distribution, normal distribution, exponential distribution and gamma distribution.

Scanned with CamScanner

5.2 BINOMIAL DISTRIBUTION

Consider n independent trials of a random experiments which results in either success or failure. Let p be the probability of success remaining constant every time and $q = 1 - p$ be the probability of failure. The probability of x successes and $n - x$ failures is given by $p^x q^{n-x}$ (multiplication theorem of probability). But these x successes and $n - x$ failures can occur in any of the ${}^nC_x$ ways in each of which the probability is same. Hence, the probability of x successes is ${}^nC_x p^x q^{n-x}$.

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

A random variable X is said to follow the binomial distribution if the probability of x is given by

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

The two constants n and p are called the parameters of the distribution.

5.2.1 Examples of Binomial Distribution

(i) Number of defective bolts in a box containing n bolts.
(ii) Number of post-graduates in a group of n people.
(iii) Number of oil wells yielding natural gas in a group of n wells test drilled.
(iv) Number of machines lying idle in a factory having n machines.

5.2.2 Conditions for Binomial Distribution

The binomial distribution holds under the following conditions:
(i) The number of trials n is finite.
(ii) There are only two possible outcomes, success or failure.
(iii) The trials are independent of each other.
(iv) The probability of success p is constant for each trial.

5.2.3 Constants of the Binomial Distribution

1. Mean of the Binomial Distribution

$$E(X) = \sum x\, p(x)$$
$$= \sum_{x=0}^{n} x\, {}^nC_x\, p^x\, q^{n-x}$$
$$= 0 \cdot {}^nC_0\, p^0\, q^n + 1 \cdot {}^nC_1\, p q^{n-1} + 2 \cdot {}^nC_2\, p^2\, q^{n-2} + \cdots + n\, p^n$$
$$= np\,[q^{n-1} + {}^{(n-1)}C_1\, q^{n-2}\, p + {}^{(n-1)}C_2\, q^{n-3}\, p^2 + \cdots + p^{n-1}]$$
$$= np\,(q + p)^{n-1}$$
$$= np \qquad [\because\ p + q = 1]$$

2. Variance of the Binomial Distribution

$$\text{Var}(X) = E(X^2) - \mu^2$$
$$= \sum_{x=0}^{n} x^2\, p(x) - \mu^2$$
$$= \sum_{x=0}^{n} x^2\, {}^nC_x\, p^x\, q^{n-x} - \mu^2$$
$$= \sum_{x=0}^{n} [x + x(x-1)]\, {}^nC_x\, p^x\, q^{n-x} - \mu^2$$
$$= \sum_{x=0}^{n} x\, {}^nC_x\, p^x\, q^{n-x} + \sum_{x=0}^{n} x(x-1)\, {}^nC_x\, p^x\, q^{n-x} - \mu^2$$
$$= np + \sum_{x=0}^{n} x(x-1) \frac{n(n-1)}{x(x-1)} \cdot {}^{(n-2)}C_{x-2}\, p^x\, q^{n-x} - \mu^2$$
$$= np + \sum_{x=0}^{n} n(n-1) \cdot {}^{(n-2)}C_{x-2}\, p^2 p^{x-2}\, q^{n-x} - \mu^2$$
$$= np + n(n-1)\, p^2 \sum_{x=0}^{n} {}^{(n-2)}C_{x-2}\, p^{x-2}\, q^{n-x} - \mu^2$$
$$= np + n(n-1)p^2 \cdot (q+p)^{n-2} - \mu^2$$
$$= np + n(n-1)p^2 - \mu^2 \qquad [\because\ p + q = 1]$$
$$= np\,[1 + (n-1)p] - \mu^2$$
$$= np\,[1 - p + np] - \mu^2$$
$$= np\,(q + np) - \mu^2 \qquad [\because\ 1 - p = q]$$
$$= np\,(q + np) - (np)^2$$
$$= npq$$

3. Standard Deviation of the Binomial Distribution

$$\text{SD} = \sqrt{\text{Variance}} = \sqrt{npq}$$

4. Mode of the Binomial Distribution

Mode of the binomial distribution is the value of x at which $p(x)$ has maximum value.

Mode = integral part of $(n + 1)p$, if $(n + 1)p$ is not an integer
$= (n + 1)\, p$ and $(n + 1)\, p - 1$, if $(n + 1)\, p$ is an integer.

Scanned with CamScanner

5.2.4 Recurrence Relation for the Binomial Distribution

For the binomial distribution,

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

$$P(X=x+1) = {}^nC_{x+1}\, p^{x+1}\, q^{n-x-1}$$

$$\frac{P(X=x+1)}{P(X=x)} = \frac{{}^nC_{x+1}\, p^{x+1}\, q^{n-x-1}}{{}^nC_x\, p^x\, q^{n-x}}$$

$$= \frac{n!}{(x+1)!\,(n-x-1)!} \times \frac{x!\,(n-x)!}{n!} \cdot \frac{p}{q}$$

$$= \frac{(n-x)\,(n-x-1)!\,x!}{(x+1)\,x!\,(n-x-1)!} \cdot \frac{p}{q}$$

$$= \frac{n-x}{x+1} \cdot \frac{p}{q}$$

$$P(X=x+1) = \frac{n-x}{x+1} \cdot \frac{p}{q} \cdot P(X=x)$$

5.2.5 Binomial Frequency Distribution

If n independent trials constitute one experiment and this experiment is repeated N times, the frequency of x successes is $N\,P(X=x)$, i.e., $N\,{}^nC_x\, p^x\, q^{n-x}$. This is called expected or theoretical frequency $f(x)$ of a success.

$$\sum_{x=0}^{n} f(x) = N \sum_{x=0}^{n} P(X=x) = N \qquad \left[\because \sum_{x=0}^{n} P(X=x) = 1\right]$$

The expected or theoretical frequencies $f(0), f(1), f(2), \ldots, f(n)$ of $0, 1, 2, \ldots, n$ successes are respectively the first, second, third, ..., $(n+1)^{\text{th}}$ term in the expansion of $N(q+p)^n$. The possible number of successes and their frequencies is called a binomial frequency distribution. In practice, the expected frequencies differ from observed fequencies due to chance factor.

Example 1

The mean and standard deviation of a binomial distribution are 5 and 2. Determine the distribution.

Solution

$$\mu = np = 5$$

$$\text{SD} = \sqrt{npq} = 2$$

$$npq = 4$$

$$\frac{npq}{np} = \frac{4}{5}$$

$$\therefore\ q = \frac{4}{5}$$

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

$$np = 5$$

$$n\left(\frac{1}{5}\right) = 5$$

$$\therefore\ n = 25$$

Hence, the binomial distribution is

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

$$= {}^{25}C_x \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{25-x}, \quad x = 0, 1, 2, \ldots, 25$$

Example 2

The mean and variance of a binomial variate are 8 and 6. Find $P(X \geq 2)$.

Solution

$$\mu = np = 8$$

$$\sigma^2 = npq = 6$$

$$\frac{npq}{np} = \frac{6}{8} = \frac{3}{4}$$

$$\therefore\ q = \frac{3}{4}$$

$$p = 1 - q = 1 - \frac{3}{4} = \frac{1}{4}$$

$$np = 8$$

$$n\left(\frac{1}{4}\right) = 8$$

$$\therefore\ n = 32$$

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

$$= {}^{32}C_x \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{32-x}, \quad x = 0, 1, 2, \ldots, 32$$

Scanned with CamScanner

$$P(X \ge 2) = 1 - P(X < 2)$$
$$= 1 - [P(X=0) + P(X=1)]$$
$$= 1 - \sum_{x=0}^{1} P(X = x)$$
$$= 1 - \sum_{x=0}^{1} {}^{32}C_x \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{32-x}$$
$$= 0.9988$$

Example 3

Suppose $P(X = 0) = 1 - P(X = 1)$. If $E(X) = 3\ Var\ (X)$, find $P(X = 0)$.

Solution

$$E(X) = 3\,\text{Var}\,(X)$$
$$np = 3npq$$
$$1 = 3q$$
$$\therefore\ q = \frac{1}{3}$$
$$p = 1 - q = 1 - \frac{1}{3} = \frac{2}{3}$$

Let $P(X = 1) = p$

$$P(X = 0) = 1 - P(X = 1)$$
$$= 1 - p$$
$$= 1 - \frac{2}{3}$$
$$= \frac{1}{3}$$

Example 4

The mean and variance of a binomial distribution are 4 and $\frac{4}{3}$ respectively. Find $P(X \ge 1)$.

Solution

$$\mu = np = 4$$
$$\sigma^2 = npq = \frac{4}{3}$$

$$\frac{npq}{np} = \frac{\frac{4}{3}}{4} = \frac{1}{3}$$
$$\therefore\ q = \frac{1}{3}$$
$$p = 1 - q = 1 - \frac{1}{3} = \frac{2}{3}$$
$$np = 4$$
$$n\left(\frac{2}{3}\right) = 4$$
$$\therefore\ n = 6$$
$$P(X = x) = {}^nC_x\, p^x q^{n-x}$$
$$= {}^6C_x \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{6-x}, \quad x = 0, 1, 2, \ldots, 6$$
$$P(X \ge 1) = 1 - P(X < 1)$$
$$= 1 - P(X = 0)$$
$$= 1 - {}^6C_0 \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^6$$
$$= 0.9986$$

Example 5

A discrete random variable X has mean 6 and variance 2. If it is assumed that the distribution is binomial, find the probability that $5 \le X \le 7$.

Solution

$$\mu = np = 6$$
$$\sigma^2 = npq = 2$$
$$\frac{npq}{np} = \frac{2}{6} = \frac{1}{3}$$
$$\therefore\ q = \frac{1}{3}$$
$$p = 1 - q = 1 - \frac{1}{3} = \frac{2}{3}$$
$$np = 6$$
$$n\left(\frac{2}{3}\right) = 6$$
$$\therefore\ n = 9$$

Scanned with CamScanner

$$P(X=x)={}^nC_x\, p^x q^{n-x}$$
$$={}^9C_x\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{9-x},\quad x=0,1,2,\ldots,9$$
$$P(5\le X\le 7)=P(X=5)+P(X=6)+P(X=7)$$
$$=\sum_{x=5}^{7}P(X=x)$$
$$=\sum_{x=5}^{7}{}^9C_x\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{9-x}$$
$$=\frac{4672}{6561}$$
$$=0.7121$$

Example 6

With the usual notation, find p for a binomial distribution if n = 6 and $9P(X=4)=P(X=2)$.

Solution

For the binomial distribution,

$$P(X=x)={}^nC_x\, p^x q^{n-x},\quad x=0,1,2,\ldots,n$$
$$n=6$$
$$9P(X=4)=P(X=2)$$
$$9\,{}^6C_4\, p^4 q^2={}^6C_2\, p^2q^4$$
$$9p^2=q^2=(1-p)^2$$
$$9p^2=1-2p+p^2$$
$$8p^2+2p-1=0$$
$$p=\frac{-2\pm\sqrt{4+32}}{2\times 8}=\frac{-2\pm 6}{16}=-\frac{1}{2},\frac{1}{4}$$

Since probability cannot be negative, $p=\frac{1}{4}$.

Example 7

In a binomial distribution consisting of 5 independent trials, the probability of 1 and 2 successes are 0.4096 and 0.2048 respectively. Find the parameter p of the distribution.

Solution

$$n=5,\quad P(X=1)=0.4096,\quad P(X=2)=0.2048$$

Probability of getting x successes out of 5 trials

$$P(X=x)={}^nC_x\, p^x q^{n-x}={}^5C_x\, p^x q^{5-x},\; x=0,1,2,\ldots,5$$
$$P(X=1)={}^5C_1\, pq^4=0.4096 \qquad \ldots(1)$$
$$P(X=2)={}^5C_2\, p^2q^3=0.2048 \qquad \ldots(2)$$

Dividing Eq. (2) by Eq. (1),

$$\frac{{}^5C_2\, p^2q^3}{{}^5C_1\, pq^4}=\frac{0.2048}{0.4096}$$
$$\frac{10p}{5q}=\frac{1}{2}$$
$$\frac{p}{q}=\frac{1}{4}$$
$$4p=q=1-p$$
$$5p=1$$
$$p=\frac{1}{5}$$

Example 8

In a binomial distribution, the sum and product of the mean and variance are $\frac{25}{3}$ and $\frac{50}{3}$ respectively. Determine the distribution.

Solution

For the binomial distribution,

$$np+npq=\frac{25}{3}$$
$$np(1+q)=\frac{25}{3} \qquad \ldots(1)$$

and

$$np(npq)=\frac{50}{3}$$
$$n^2p^2q=\frac{50}{3} \qquad \ldots(2)$$

Scanned with CamScanner

Squaring Eq. (1) and then dividing by Eq. (2),

$$\frac{n^2 p^2 (1+q)^2}{n^2 p^2 q} = \frac{\frac{625}{9}}{\frac{50}{3}}$$

$$\frac{1+2q+q^2}{q} = \frac{25}{6}$$

$$6(q^2+2q+1) = 25q$$

$$6q^2 - 13q + 6 = 0$$

$$(2q-3)(3q-2) = 0$$

$$q = \frac{3}{2} \text{ or } q = \frac{2}{3}$$

Since q can not be greater than 1,

$$q = \frac{2}{3}$$

$$p = 1 - q = 1 - \frac{2}{3} = \frac{1}{3}$$

From Eq. (1),

$$n\left(\frac{1}{3}\right)\left(1+\frac{2}{3}\right) = \frac{25}{3}$$

$$\therefore \quad n = 15$$

Hence, the binomial distribution is

$$P(X=x) = {}^{15}C_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{15-x}, \quad x = 0, 1, 2, \ldots 15$$

Example 9

If the probability of a defective bolt is $\frac{1}{8}$, find the (i) mean, and (ii) variance for the distribution of 640 defective bolts.

Solution

$$p = \frac{1}{8}, \quad n = 640$$

$$\mu = np = \frac{640}{8} = 80$$

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

Variance of the distribution $= npq = 640\left(\frac{1}{8}\right)\left(\frac{7}{8}\right) = 70$

Example 10

In eight throws of a die, 5 or 6 is considered as a success. Find the mean number of success and the standard deviation.

Solution

Let p be the probability of success.

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

$$q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$$

$$n = 8$$

$$\mu = np = 8\left(\frac{1}{3}\right) = \frac{8}{3}$$

$$\text{SD} = \sqrt{npq} = \sqrt{8\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)} = \frac{4}{3}$$

Example 11

4 coins are tossed simultaneously. What is the probability of getting (i) 2 heads? (ii) at least 2 heads? (iii) at most 2 heads?

Solution

Let p be the probability of getting a head in the toss of a coin.

$$p = \frac{1}{2}, \quad q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}, \quad n = 4$$

The probability of getting x heads when 4 coins are tossed

$$P(X=x) = {}^nC_x\, p^x q^{n-x} = {}^4C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{4-x}, \quad x = 0, 1, 2, 3, 4$$

(i) Probability of getting 2 heads when 4 coins are tossed

$$P(X=2) = {}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = \frac{3}{8}$$

Scanned with CamScanner

(ii) Probability of getting at least two heads when 4 coins are tossed

$$P(X \geq 2) = P(X=2)+P(X=3)+P(X=4)$$
$$= \sum_{x=2}^{4} P(X=x)$$
$$= \sum_{x=2}^{4} {}^4C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{4-x}$$
$$= \frac{11}{16}$$

(iii) Probability getting at most 2 heads when 4 coins are tossed

$$P(X \leq 2) = P(X=0)+P(X=1)+P(X=2)$$
$$= \sum_{x=0}^{2} P(X=x)$$
$$= \sum_{x=0}^{2} {}^4C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{4-x}$$
$$= \frac{11}{16}$$

Example 12

Two dice are thrown five times. Find the probability of getting the sum as 7 (i) at least once, (ii) two times, and (iii) $P(1 < X < 15)$.

Solution

In a single throw of two dice, a sum of 7 can occur in 6 ways out of $6 \times 6 = 36$ ways.

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

Let p be the probability of getting the sum as 7 in a single throw of a pair of dice.

$$p = \frac{6}{36} = \frac{1}{6}, \quad q = 1-p = 1-\frac{1}{6} = \frac{5}{6}, \quad n = 5$$

Probability of getting the sum x times in 5 throws of a pair of dice

$$P(X=x) = {}^nC_x\, p^x q^{n-x} = {}^5C_x \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{5-x}, \quad x = 0, 1, 2, \ldots, 5$$

(i) Probability of getting the sum as 7 at least once in 5 throws of two dice

$$P(X \geq 1) = 1 - P(X=0)$$
$$= 1 - {}^5C_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^5$$

$$= 1 - \frac{3125}{7776}$$
$$= \frac{4651}{7776}$$

(ii) Probability of getting the sum as 7 two times in 5 throws of two dice

$$P(X=2) = {}^5C_2 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3 = \frac{625}{3888}$$

(iii) Probability of getting the sum as 7 for $P(1 < X < 5)$ in 5 throws of two dice

$$P(1 < X < 5) = P(X=2)+P(X=3)+P(X=4)$$
$$= \sum_{x=2}^{4} P(X=x)$$
$$= \sum_{x=2}^{4} {}^5C_x \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{5-x}$$
$$= \frac{1525}{7776}$$

Example 13

If 10% of the screws produced by a machine are defective, find the probability that out of 5 screws chosen at random, (i) none is defective, (ii) one is defective, and (iii) at most two are defective.

Solution

Let p be the probability of defective screws.

$$p = 0.1, \quad q = 1-p = 1-0.1 = 0.9, \quad n = 5$$

Probability that x screws out of 5 screws are defective

$$P(X=x) = {}^nC_x\, p^x q^{n-x} = {}^5C_x (0.1)^x (0.9)^{5-x}, \quad x = 0, 1, 2, \ldots, 5$$

(i) Probability that none of the screws out of 5 screws is defective

$$P(X=0) = {}^5C_0 (0.1)^0 (0.9)^5 = 0.5905$$

(ii) Probability that one screw out of 5 screws is defective

$$P(X=1) = {}^5C_1 (0.1)^1 (0.9)^4 = 0.3281$$

(iii) Probability that at most 2 screws out of 5 screws are defective

$$P(X \leq 2) = P(X=0)+P(X=1)+P(X=2)$$
$$= \sum_{x=0}^{2} P(X=x)$$
$$= \sum_{x=0}^{2} {}^5C_x (0.1)^x (0.9)^{5-x}$$
$$= 0.9914$$

Scanned with CamScanner

Example 14

A multiple-choice test consists of 8 questions with 3 answers to each question (of which only one is correct). A student answers each question by rolling a balanced die and checking the first answer if he gets 1 or 2, the second answer if he gets 3 or 4, and the third answer if he gets 5 or 6. To get a distinction, the student must secure at least 75% correct answers. If there is no negative making, what is the probability that the student secures a distinction? **[Summer 2015]**

Solution

Let p be the probability of getting an answer to a question correctly. There are three answers to each question, out of which only one is correct.

$$p=\frac{1}{3}, \quad q=1-p=1-\frac{1}{3}=\frac{2}{3}, \quad n=8$$

Probability of getting x correct answers in an 8 questions test

$$P(X=x)={}^nC_x\, p^x\, q^{n-x}={}^8C_x\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{8-x}, \quad x=0,1,2,\ldots,8$$

Probability of securing a distinction, i.e., getting at least 6 correct answers out of the 8 questions

$$P(X\le 6)=P(X=6)+P(X=7)+P(X=8)$$
$$=\sum_{x=6}^{8}P(X=x)$$
$$=\sum_{x=6}^{8}{}^8C_x\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{8-x}$$
$$=\frac{43}{2187}$$
$$=0.0197$$

Example 15

A and B play a game in which their chances of winning are in the ratio 3:2. Find A's chance of winning at least three games out of the five games played.

Solution

Let p be the probability that A wins the game.

$$p=\frac{3}{3+2}=\frac{3}{5}, \quad q=1-p=1-\frac{3}{5}=\frac{2}{5}, \quad n=5$$

Probability that A wins x games out of 5 games

$$P(X=x)={}^nC_x\, p^x\, q^{n-x}={}^5C_x\left(\frac{3}{5}\right)^x\left(\frac{2}{5}\right)^{5-x}, \quad x=0,1,2,\ldots,5$$

Probability that A wins at least 3 games

$$P(X\ge 3)=P(X=3)+P(X=4)+P(X=5)$$
$$=\sum_{x=3}^{5}P(X=x)$$
$$=\sum_{x=3}^{5}{}^5C_x\left(\frac{3}{5}\right)^x\left(\frac{2}{5}\right)^{5-x}$$
$$=\frac{2133}{3125}$$
$$=0.6826$$

Example 16

It has been claimed that in 60% of all solar heat installations the utility bill is reduced by at least one-third. Accordingly, what are the probabilities that the utility bill will be reduced by at least one third in (i) four of five installations? (ii) at least four of five installations?

Solution

Let p be the probability that the utility bill is reduced by one-third in the solar heat installations.

$$p=60\%=0.6, \quad q=1-p=1-0.6=0.4, \quad n=5$$

Probability that the utility bill is reduced by one-third in x installations out of 5 installations

$$P(X=x)={}^nC_x\, p^x\, q^{n-x}={}^5C_x(0.6)^x(0.4)^{5-x}, \quad x=0,1,2,\ldots,5$$

Probability that the utility bill is reduced by one-third in 4 of 5 installations

$$P(X=5)={}^5C_4\,(0.6)^4\,(0.4)^1=\frac{162}{625}$$

Probability that the utility bill is reduced by one-third in at least 4 of 5 installations

$$P(X\ge 4)=P(X=4)+P(X=5)$$
$$=\sum_{x=4}^{5}P(X=x)$$
$$=\sum_{x=4}^{5}{}^5C_x\,(0.6)^x(0.4)^{5-x}$$
$$=\frac{1053}{3125}$$
$$=0.337$$

Scanned with CamScanner

Example 17

The incidence of an occupational disease in an industry is such that the workers have a 20% chance of suffering from it. What is the probability that out of 6 workers chosen at random, four or more will suffer from the disease?

Solution

Let p be the probability of a worker suffering from the disease.

$$p = 0.2, \quad q = 1 - p = 1 - 0.2 = 0.8, \quad n = 6$$

Probability that x workers will suffer from the disease

$$P(X = x) = {}^nC_x\, p^x\, q^{n-x} = {}^6C_x\,(0.2)^x\,(0.8)^{6-x}, \quad x = 0, 1, 2, \ldots, 6$$

Probability that 4 or more workers will suffer from the disease

$$P(X \ge 4) = P(X = 4) + P(X = 5) + P(X = 6)$$
$$= \sum_{x=4}^{6} P(X = x)$$
$$= \sum_{x=4}^{6} {}^6C_x\,(0.2)^x(0.8)^{6-x}$$
$$= \frac{53}{3125}$$
$$= 0.017$$

Example 18

The probability that a man aged 60 will live up to 70 is 0.65. What is the probability that out of 10 such men now at 60 at least 7 will live up to 70?

Solution

Let p be the probability that a man will live up to 70.

$$p = 0.65, \quad q = 1 - p = 1 - 0.65 = 0.35, \quad n = 10$$

Probability that x men out of 10 will live up to 70

$$P(X = x) = {}^nC_x\, p^x\, q^{n-x} = {}^{10}C_x(0.65)^x\,(0.35)^{10-x}, \; x = 0, 1, 2, \ldots, 10$$

Probability that at least 7 men out of 10 will live up to 70

$$P(X \ge 7) = P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10)$$
$$= \sum_{x=7}^{10} P(X = x)$$

$$= \sum_{x=7}^{10} {}^{10}C_x\,(0.65)^x\,(0.35)^{10-x}$$
$$= 0.5138$$

Example 19

In a multiple-choice examination, there are 20 questions. Each question has 4 alternative answers following it and the student must select one correct answer. 4 marks are given for a correct answer and 1 mark is deducted for a wrong answer. A student must secure at least 50% of the maximum possible marks to pass the examination. Suppose a student has not studied at all, so that he answers the questions by guessing only. What is the probability that he will pass the examination?

Solution

Since there are 20 questions and each carries with 4 marks, the maximum marks are 80. If the student solves 12 questions correctly and 8 questions wrongly, he gets 48 − 8 = 40 marks required for passing. If he gets more than 12 correct answers, he gets more than 40 marks. Let p be the probability of getting a correct answer.

$$\therefore \; p = \frac{1}{4}, \quad q = 1 - p = 1 - \frac{1}{4} = \frac{3}{4}, \quad n = 20$$

Probability of getting x correct answers out of 20 answers

$$P(X = x) = {}^nC_x\, p^x\, q^{n-x} = {}^{20}C_x\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{20-x}, \quad x = 0, 1, 2, \ldots, 20$$

Probability of passing the examination. i.e., probability of getting at least 12 correct answers out of 20 answers

$$P(X \ge 12) = \sum_{x=12}^{20} P(X = x)$$
$$= \sum_{x=12}^{20} {}^{20}C_x\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{20-x}$$
$$= 9.3539 \times 10^{-4}$$

Example 20

The probability of a man hitting a target is $\frac{1}{3}$. (i) If he fires 5 times, what is the probability of his hitting the target at least twice? (ii) How many times must he fire so that the probability of his hitting the target at least once is more than 90%?

Scanned with CamScanner

Solution

Let p be probability of hitting a target.

$$p = \frac{1}{3}, \quad q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}, \quad n = 5$$

Probability of hitting the target x times out of 5 times

$$P(X = x) = {}^nC_x\, p^x\, q^{n-x} = {}^5C_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{5-x}, \quad x = 0, 1, 2, \ldots, 5$$

(i) Probability of hitting the target at least twice out of 5 times

$$P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$$
$$= \sum_{x=2}^{5} P(X = x)$$
$$= \sum_{x=2}^{5} {}^5C_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{5-x}$$
$$= \frac{131}{243}$$
$$= 0.5391$$

(ii) Probability of hitting the target at least once out of 5 times

$$P(X \geq 1) > 0.9$$
$$1 - P(X = 0) > 0.9$$
$$1 - {}^nC_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^n > 0.9$$
$$1 - \left(\frac{2}{3}\right)^n > 0.9$$

For $n = 6$, $1 - \left(\frac{2}{3}\right)^6 = 0.9122$

Hence, the man must fire 6 times so that the probability of hitting the target at lest once is more than 90%.

Example 21

In sampling a large number of parts manufactured by a machine, the mean number of defectives in a sample of 20 is 2. Out of 1000 such samples, how many would be expected to contain exactly two defective parts? [Summer 2015]

Solution

Let p be the probability of parts being defective.

$$\mu = np = 2, \quad n = 20, \quad N = 1000$$
$$np = 2$$
$$20(p) = 2$$
$$\therefore \quad p = 0.1$$
$$q = 1 - p = 1 - 0.1 = 0.9$$

Probability that the samples contain x defective parts out of 20 parts

$$P(X = x) = {}^nC_x p^x q^{n-x} = {}^{20}C_x (0.1)^x (0.9)^{20-x}, \quad x = 0, 1, 2, \ldots, 20$$

Probability that the samples contain exactly 2 defective parts

$$P(X = 2) = {}^{20}C_2 (0.1)^2 (0.9)^{18} = 0.2852$$

Expected number of samples to contain exactly 2 defective parts $= N\,P(X = 2)$

$$= 1000\,(0.2852)$$
$$= 285.2$$
$$\approx 285$$

Example 22

An irregular 6-faced die is thrown 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?

Solution

Let p be the probability of getting an even number in a throw of a die.

$$n = 5, \quad N = 2500$$

Probability of getting x even numbers in 5 throws of a die

$$P(X = x) = {}^nC_x\, p^x\, q^{n-x} = {}^5C_x\, p^x q^{5-x}, \quad x = 0, 1, 2, \ldots, 5$$

$$P(X = 3) = 2\,P(X = 2)$$
$${}^5C_3\, p^3\, q^2 = 2\left({}^5C_2\, p^2\, q^3\right)$$
$$10\, p^3\, q^2 = 20\, p^2\, q^3$$
$$p = 2q$$
$$p = 2(1 - p) = 2 - 2p$$
$$\therefore \quad p = \frac{2}{3}$$
$$q = 1 - p = 1 - \frac{2}{3} = \frac{1}{3}$$

Scanned with CamScanner

Probability of getting no even number in 5 throws of a die

$$P(X=0)={}^5C_0\left(\frac{2}{3}\right)^0\left(\frac{1}{3}\right)^5=\frac{1}{243}$$

Expected number of sets $= NP(X=0)$

$$=\frac{2500}{243}$$

Example 23

Out of 800 families with 5 children each, how many would you expect to have (i) 3 boys? (ii) 5 girls? (iii) either 2 or 3 boys? (iv) at least one boy? Assume equal probabilities for boys and girls.

Solution

Let p be the probability of having a boy in each family.

$$p=\frac{1}{2},\quad q=1-\frac{1}{2}=1-\frac{1}{2}=\frac{1}{2},\quad n=5,\quad N=800$$

Probability of having x boys out of 5 children in each family

$$P(X=x)={}^nC_x\, p^x q^{n-x}={}^5C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{5-x},\ x=0,1,2,\ldots,5$$

(i) Probability of having 3 boys out of 5 children in each family

$$P(X=3)={}^5C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2=\frac{5}{16}$$

Expected number of families having 3 boys out of 5 children $= N\,P(X=3)$

$$=800\left(\frac{5}{16}\right)$$

$$=250$$

(ii) Probability of having 5 girls, i.e., no boys out of 5 children in each family

$$P(X=0)={}^5C_0\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^5=\frac{1}{32}$$

Expected number of families 5 girls out of 5 children $= NP(X=0)$

$$=800\left(\frac{1}{32}\right)$$

$$=25$$

(iii) Probability of having either 2 or 3 boys out of 5 children in each family

$$P(X=2)+P(X=3)=\sum_{x=2}^{3}P(X=x)$$

$$=\sum_{x=2}^{3}{}^5C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{5-x}$$

$$=\frac{5}{8}$$

Expected number of families having either 2 of 3 boys out of 5 children

$$=N[P(X=2)+P(X=3)]$$

$$=800\left(\frac{5}{8}\right)$$

$$=500$$

(iv) Probability of having at least one boy out of 5 children in each family

$$P(X\ge 1)=P(X=1)+P(X=2)+P(X=3)+P(X=4)+P(X=5)$$

$$=\sum_{x=1}^{5}P(X=x)$$

$$=\sum_{x=1}^{5}{}^5C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{5-x}$$

$$=\frac{31}{32}$$

Expected number of families having at least-one boy out of 5 children

$$=NP(X\ge 1)$$

$$=800\left(\frac{31}{32}\right)$$

$$=775$$

Example 24

If hens of a certain breed lay eggs on 5 days a week on an average, find how many days during a season of 100 days a will poultry keeper with 5 hens of this breed expect to receive at least 4 eggs.

Solution

Let p be the probability of hen laying an egg on any day of a week.

$$p=\frac{5}{7},\quad q=1-p=1-\frac{5}{7}=\frac{2}{7},\quad n=5,\quad N=100$$

Probability of x hens laying eggs on any day of a week

$$P(X=x)={}^nC_x\, p^x q^{n-x}={}^5C_x\left(\frac{5}{7}\right)^x\left(\frac{2}{7}\right)^{5-x},\ x=0,1,2,\ldots,5$$

Scanned with CamScanner

Probability of receiving at least 4 eggs on any day of a week

$$P(X \ge 4) = P(X=4) + P(X=5)$$
$$= \sum_{x=4}^{5} P(X=x)$$
$$= \sum_{x=4}^{5} {}^5C_x \left(\frac{5}{7}\right)^x \left(\frac{2}{7}\right)^{5-x}$$
$$= 0.5578$$

Expected number of days during a season of 100 days, a poultry keeper with 5 hens of this breed will receive at least 4 eggs $= N\,P(X \ge 4)$

$$= 100\,(0.5578)$$
$$= 55.78$$
$$\approx 56$$

Example 25

Seven unbiased coins are tossed 128 times and the number of heads obtained is noted as given below:

No. of heads	0	1	2	3	4	5	6	7
Frequency	7	6	19	35	30	23	7	1

Fit a binomial distribution to the data.

Solution

Since the coin is unbiased,

$$p = \frac{1}{2}, \quad q = \frac{1}{2}, n = 7, N = 128$$

For binomial distribution,

$$P(X=x) = {}^nC_x\, p^x\, q^{n-x} = {}^7C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{7-x}, x = 0, 1, 2, \ldots, 7$$

Theoretical or expected frequency $f(x) = N\,P(X=x)$

$$f(x) = 128\ {}^7C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{7-x} = 128\ {}^7C_x \left(\frac{1}{2}\right)^7$$
$$f(0) = 128\ {}^7C_0 \left(\frac{1}{2}\right)^7 = 1$$
$$f(1) = 128\ {}^7C_1 \left(\frac{1}{2}\right)^7 = 7$$

$$f(2) = 128\ {}^7C_2 \left(\frac{1}{2}\right)^7 = 21$$
$$f(3) = 128\ {}^7C_3 \left(\frac{1}{2}\right)^7 = 35$$
$$f(4) = 128\ {}^7C_4 \left(\frac{1}{2}\right)^7 = 35$$
$$f(5) = 128\ {}^7C_5 \left(\frac{1}{2}\right)^7 = 21$$
$$f(6) = 128\ {}^7C_6 \left(\frac{1}{2}\right)^7 = 7$$
$$f(7) = 128\ {}^7C_7 \left(\frac{1}{2}\right)^7 = 1$$

Binomial distribution

No. of heads x	0	1	2	3	4	5	6	7
Expected binomial frequency $f(x)$	1	7	21	35	35	21	7	1

Example 26

Fit a binomial distribution to the following data:

x	0	1	2	3	4	5
f	2	14	20	34	22	8

Solution

$$\text{Mean} = \frac{\sum fx}{\sum f}$$
$$= \frac{2(0)+14(1)+20(2)+34(3)+22(4)+8(5)}{2+14+20+34+22+8}$$
$$= \frac{284}{100}$$
$$= 2.84$$

For binomial distribution,

$$n = 5$$

Scanned with CamScanner

$$\mu = np = 2.84$$
$$5p = 2.84$$
$$\therefore\ p = 0.568$$
$$q = 1 - p = 1 - 0.568 = 0.432$$
$$P(X = x) = {}^nC_x\, p^x\, q^{n-x} = {}^5C_x (0.568)^x (0.432)^{5-x},\ x = 0, 1, 2, \ldots, 5$$
$$N = \sum f = 100$$

Theoretical or expected frequency $f(x) = N\,P(X = x)$

$$f(x) = 100\ {}^5C_x\ (0.568)^x (0.432)^{5-x}$$
$$f(0) = 100\ {}^5C_0\ (0.568)^0\ (0.432)^5 = 1.505 \approx 1.5$$
$$f(1) = 100\ {}^5C_1\ (0.568)^1\ (0.432)^4 = 9.89 \approx 10$$
$$f(2) = 100\ {}^5C_2\ (0.568)^2\ (0.432)^3 = 26.01 \approx 26$$
$$f(3) = 100\ {}^5C_2\ (0.568)^3\ (0.432)^2 = 34.2 \approx 34$$
$$f(4) = 100\ {}^5C_2\ (0.568)^4\ (0.432)^1 = 22.48 \approx 22$$
$$f(5) = 100\ {}^5C_2\ (0.568)^5\ (0.432)^0 = 5.91 \approx 6$$

Binomial Distribution

x	0	1	2	3	4	5
Expected binomial frequency	1.5	10	26	34	22	6

EXERCISE 5.1

1. Find the fallacy if any in the following statements:
   (a) The mean of a binomial distribution is 6 and SD is 4.
   (b) The mean of a binomial distribution is 9 and its SD is 4.

   $\left[\text{Ans.: (a) False, } q = \dfrac{8}{3} \text{ is impossible} \quad \text{(b) False, } q = \dfrac{19}{9} \text{ is impossible}\right]$

2. The mean and variance of a binomial distribution are 3 and 1.2 respectively. Find n, p, and $P(X < 4)$.

   $\left[\text{Ans.: } 5, 0.6, \dfrac{2068}{3125}\right]$

3. Find the binomial distribution if the mean is 5 and the variance is $\dfrac{10}{3}$. Find $P(X = 2)$.

   $\left[\text{Ans.: } P(X = x) = {}^{25}C_x \left(\dfrac{1}{3}\right)^x \left(\dfrac{2}{3}\right)^{25-x}, 0.003\right]$

4. In a binomial distribution, the mean and variance are 4 and 3 respectively. Find $P(X \geq 1)$.

   [Ans.: 0.9899]

5. The odds in favour of X winning a game against Y are 4:3. Find the probability of Y winning 3 games out of 7 played.

   [Ans.: 0.0929]

6. On an average, 3 out of 10 students fail in an examination. What is the probability that out of 10 students that appear for the examination none will fail?

   [Ans.: 0.0282]

7. If on the average rain falls on 10 days in every thirty, find the probability (i) that the first three days of a week will be fine and remaining wet, and (ii) that rain will fall on just three days of a week.

   $\left[\text{Ans.: (i) } \dfrac{8}{2187} \text{ (ii) } \dfrac{280}{2187}\right]$

8. Two unbiased dice are thrown three times. Find the probability that the sum nine would be obtained (i) once, and (ii) twice.

   [Ans.: (i) 0.26 (ii) 0.03]

9. For special security in a certain protected area, it was decided to put three lightbulbs on each pole. If each bulb has probability p of burning out in the first 100 hours of service, calculate the probability that at least one of them is still good after 100 hours. If $p = 0.3$, how many bulbs would be needed on each pole to ensure with 99% safety that at least one is good after 100 hours?

   [Ans.: (i) $1 - p^3$ (ii) 4]

10. It is known from past records that 80% of the students in a school do their homework. Find the probability that during a random check of 10 students, (i) all have done their homework, (ii) at the most two have not done their homework, and (iii) at least one has not done the homework.

   [Ans.: (i) 0.1074 (ii) 0.6778 (iii) 0.8926]

Scanned with CamScanner

11. An insurance salesman sells policies to 5 men, all of identical age and good health. According to the actuarial tables, the probability that a man of this particular age will be alive 30 years hence is $\frac{2}{3}$. Find the probability that 30 years hence (i) at least 1 man will be alive, (ii) at least 3 men will be alive, and (iii) all 5 men will be alive.

$\left[\text{Ans.: (i) } \frac{242}{243} \text{ (ii) } \frac{64}{81} \text{ (iii) } \frac{32}{243}\right]$

12. A company has appointed 10 new secretaries out of which 7 are trained. If a particular executive is to get three secretaries selected at random, what is the chance that at least one of them will be untrained?

[Ans.: 0.7083]

13. The overall pass rate in a university examination is 70%. Four candidates take up such an examination. What is the probability that (i) at least one of them will pass? (ii) all of them will pass the examination?

[Ans.: (i) 0.9919 (ii) 0.7599]

14. The normal rate of infection of a certain disease in animals is known to be 25%. In an experiment with a new vaccine, it was observed that none of the animals caught the infection. Calculate the probability of the observed result.

$\left[\text{Ans.: } \frac{729}{4096}\right]$

15. Suppose that weather records show that on the average, 5 out of 31 days in October are rainy days. Assuming a binomial distribution with each day of October as an independent trial, find the probability that the next October will have at most three rainy days.

[Ans.: 0.2403]

16. Assuming that half the population of a village is female and assuming that 100 samples each of 10 individuals are taken, how many samples would you expect to have 3 or less females?

[Ans.: 17]

17. Assuming that half the population of a town is vegetarian so that the chance of an individual being vegetarian is $\frac{1}{2}$, and assuming that 100 investigators can take a sample of 10 individuals to see whether they are vegetarians, how many investigators would you expect to report that three people or less in the sample were vegetarians?

[Ans.: 17]

18. The probability of failure in a physics practical examination is 20%. If 25 batches of 6 students each take the examination, in how many batches of 4 or more students would pass?

[Ans.: 23]

19. A lot contains 1% defective items. What should be the number of items in a lot so that the probability of finding at least one defective item in it is at least 0.95?

[Ans.: 299]

20. The probability that a bomb will hit the target is 0.2. Two bombs are required to destroy the target. If six bombs are used, find the probability that the target will be destroyed.

[Ans.: 0.3447]

21. Out of 1000 families with 4 children each, how many would you expect to have (i) 2 boys and 2 girls? (ii) at least one boy? (iii) no girl? (iv) at most 2 girls?

[Ans.: (i) 375 (ii) 938 (iii) 63 (iv) 69]

22. In a sampling of a large number of parts produced by a machine, the mean number of defectives in a sample of 20 is 2. Out of 1000 such samples, how many samples would you expect to contain at least 3 defectives?

[Ans.: 323]

23. Five pair coins are tossed 3200 times, find the frequency distribution of the number of heads obtained. Also, find the mean and SD.

[Ans.: (i) 100, 500, 1000, 1000, 500, 100 (ii) 1600 (iii) 28.28]

24. Fit a binomial distribution to the following data:

x	0	1	2	3	4
f	12	66	109	59	10

[Ans.: 17, 67, 96, 61, 15]

5.3 POISSON DISTRIBUTION

A random variable X is said to follow poisson distribution if the probability of x is given by

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

where λ is called the *parameter of the distribution*.

Scanned with CamScanner

5.3.1 Poisson Approximation to the Binomial Distribution

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

(i) The number of trials should be infinitely large, i.e., $n \to \infty$.
(ii) The probability of successes p for each trial should be very small, i.e., $p \to 0$.
(iii) $np = \lambda$ should be finite where λ is a constant.

The binomial distribution is

$$P(X = x) = {}^nC_x\, p^x q^{n-x}$$
$$= {}^nC_x \left(\frac{p}{q}\right)^x q^n$$
$$= {}^nC_x \left(\frac{p}{1-p}\right)^x (1-p)^n$$

Putting $p = \frac{\lambda}{n}$,

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

Since $\lim\limits_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n-x} = e^{-\lambda}$

and $\lim\limits_{n\to\infty}\left(1-\frac{1}{n}\right) = \lim\limits_{n\to\infty}\left(1-\frac{2}{n}\right) = 1$

Taking the limits of both the sides as $n \to \infty$,

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

5.3.2 Examples of Poisson Distribution

(i) Number of defective bulbs produced by a reputed company
(ii) Number of telephone calls per minute at a switchboard
(iii) Number of cars passing a certain point in one minute
(iv) Number of printing mistakes per page in a large text
(v) Number of persons born blind per year in a large city

5.3.3 Conditions of Poisson Distribution

The Poisson distribution holds under the following conditions:

(i) The random variable X should be discrete.
(ii) The numbers of trials n is very large.
(iii) The probability of success p is very small (very close to zero).
(iv) $\lambda = np$ is finite.
(v) The occurrences are rare.

5.3.4 Constants of the Poisson Distribution

1. Mean of the Poisson Distribution

$$E(X) = \sum_{x=0}^{\infty} x\, p(x)$$
$$= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$$
$$= \sum_{x=0}^{\infty} \frac{x e^{-\lambda}\,\lambda\,\lambda^{x-1}}{x!}$$
$$= e^{-\lambda}\cdot\lambda \sum_{x=1}^{\infty}\frac{x\,\lambda^{x-1}}{x!}$$
$$= \lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} \qquad \left[\because \frac{x}{x!} = \frac{1}{(x-1)!}\right]$$
$$= \lambda e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2!}+\cdots\right)$$
$$= \lambda e^{-\lambda} e^{\lambda}$$
$$= \lambda$$

2. Variance of the Poisson Distribution

$$\begin{aligned}
\text{Var}(X) &= E(X^2)-\mu^2\\
&= \sum_{x=0}^{\infty} x^2\, p(x)-\mu^2\\
&= \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!}-\lambda^2\\
&= \sum_{x=0}^{\infty} x[(x-1)+x]\frac{e^{-\lambda}\lambda^x}{x!}-\lambda^2\\
&= \sum_{x=0}^{\infty} \frac{x(x-1)e^{-\lambda}\lambda^x}{x!}+\sum_{x=0}^{\infty}\frac{xe^{-\lambda}\lambda^x}{x!}-\lambda^2\\
&= \sum_{x=0}^{\infty} \frac{x(x-1)e^{-\lambda}\lambda^{x-2}\lambda^2}{x(x-1)(x-2)\cdots 1}+\lambda-\lambda^2\\
&= e^{-\lambda}\lambda^2\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!}+\lambda-\lambda^2\\
&= e^{-\lambda}\lambda^2\left(1+\lambda+\frac{\lambda^2}{2!}+\cdots\right)+\lambda-\lambda^2\\
&= -e^{\lambda}e^{-\lambda}\lambda^2+\lambda-\lambda^2\\
&= \lambda^2+\lambda-\lambda^2\\
&= \lambda
\end{aligned}$$

3. Standard Deviation of the Poisson Distribution

$$\text{SD}=\sqrt{\text{Variance}}=\sqrt{\lambda}$$

4. Mode of the Poisson Distribution

Mode is the value of x for which the probability $p(x)$ is maximum.

$$p(x)\ge p(x+1) \text{ and } p(x)\ge p(x-1)$$

When $p(x)\ge p(x+1)$,

$$\frac{e^{-\lambda}\lambda^x}{x!}\ge\frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!}$$

$$1\ge\frac{\lambda}{x+1}$$

$$(x+1)\ge\lambda$$

$$x\ge\lambda-1 \qquad \text{...(5.1)}$$

Similarly, for $p(x)\ge p(x-1)$,

$$x\le\lambda \qquad \text{...(5.2)}$$

Combining Eqs (5.1) and (5.2),

$$\lambda-1\le x\le\lambda$$

Hence, the mode of the Poisson distribution lies between $\lambda-1$ and λ.

Case I If λ is an integer then $\lambda-1$ is also an integer. The distribution is bimodal and the two modes are $\lambda-1$ and λ.

Case II If λ is not an integer, the distribution is unimodal and the mode of the Poisson distribution is an integral part of λ. The mode is the integer between $\lambda-1$ and λ.

5.3.5 Recurrence Relation for the Poisson Distribution

For the Poisson distribution,

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

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

$$\frac{p(x+1)}{p(x)}=\frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!}\cdot\frac{x!}{e^{-\lambda}\lambda^x}=\frac{\lambda}{x+1}$$

$$p(x+1)=\frac{\lambda}{x+1}p(x)$$

Example 1

Find out the fallacy if any in the statement. "The mean of a Poisson distribution is 2 and the variance is 3."

Solution

In a Poisson distribution, the mean and variance are same. Hence, the above statement is false.

Example 2

If the mean of the Poisson distribution is 4, find

$$P(\lambda-2\sigma<X<\lambda+2\sigma).$$

Solution

For a Poisson distribution,

Variance $=\lambda$

Mean $= \lambda = 4, \quad \sigma = 2$

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-4}4^x}{x!}, \quad x = 0, 1, 2, \ldots$$

$$\begin{aligned} P(\lambda - 2\sigma < X < \lambda + 2\sigma) &= P(0 < X < 8) \\ &= \sum_{x=1}^{7} P(X = x) \\ &= \sum_{x=1}^{7} \frac{e^{-4}4^x}{x!} \\ &= 0.9306 \end{aligned}$$

Example 3

If the mean of a Poisson variable is 1.8, find (i) $P(X > 1)$, (ii) $P(X = 5)$, and (iii) $P(0 < X < 5)$.

Solution

For a Poisson distribution,

$$\lambda = 1.8$$

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-1.8}1.8^x}{x!}, \quad x = 0, 1, 2, \ldots$$

(i) $$\begin{aligned} P(X > 1) &= 1 - P(X \le 1) \\ &= 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - \sum_{x=0}^{1} P(X = x) \\ &= 1 - \sum_{x=0}^{1} \frac{e^{-1.8}1.8^x}{x!} \\ &= 0.5372 \end{aligned}$$

(ii) $$P(X = 5) = \frac{e^{-1.8}1.8^5}{5!} = 0.026$$

(iii) $$\begin{aligned} P(0 < X < 5) &= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) \\ &= \sum_{x=1}^{4} P(X = x) \\ &= \sum_{x=1}^{4} \frac{e^{-1.8}1.8^x}{x!} \\ &= 0.7983 \end{aligned}$$

Example 4

If a random variable has a Poisson distribution such that $P(X = 1) = P(X = 2)$, find (i) the mean of the distribution, (ii) $P(X = 4)$, (iii) $P(X \ge 1)$, and (iv) $P(1 < X < 4)$.

Solution

For a Poisson distribution,

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

(i) $$\begin{aligned} P(X = 1) &= P(X = 2) \\ \frac{e^{-\lambda}\lambda^1}{1!} &= \frac{e^{-\lambda}\lambda^2}{2!} \\ \lambda^2 &= 2\lambda \\ \lambda^2 - 2\lambda &= 0 \\ \lambda(\lambda - 2) &= 0 \\ \lambda = 0 &\text{ or } \lambda = 2 \end{aligned}$$

Since $\lambda \ne 0, \quad \lambda = 2$

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

(ii) $$P(X = 4) = \frac{e^{-2}2^4}{4!} = 0.9022$$

(iii) $$\begin{aligned} P(X \ge 1) &= 1 - P(X < 1) \\ &= 1 - P(X = 0) \\ &= 1 - \frac{e^{-2}2^0}{0!} \\ &= 0.8647 \end{aligned}$$

(iv) $$\begin{aligned} P(1 < X < 4) &= P(X = 2) + P(X = 3) \\ &= \sum_{x=2}^{3} P(X = x) \\ &= \sum_{x=2}^{3} \frac{e^{-2}2^x}{x!} \\ &= 0.4511 \end{aligned}$$

Example 5

If X is a Poisson variate such that $P(X = 0) = P(X = 1)$, find $P(X = 0)$ and using recurrence relation formula, find the probabilities at $x = 1, 2, 3, 4,$ and 5.

Solution

For a Poisson distribution,

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

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

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

$$\lambda = 1$$

Hence, $P(X = x) = \dfrac{e^{-1}1^x}{x!}, \quad x = 0, 1, 2, \ldots$

(i) $P(X = 0) = \dfrac{e^{-\lambda}\lambda^0}{0!} = 0.3678$

(ii) By recurrence relation,

$$p(x+1) = \frac{\lambda}{x+1}p(x)$$

$$p(x+1) = \frac{1}{x+1}p(x) \qquad [\because \lambda = 1]$$

$$p(1) = p(0) = 0.3678$$

$$p(2) = \frac{1}{2}p(1) = \frac{1}{2}(0.3678) = 0.1839$$

$$p(3) = \frac{1}{3}p(2) = \frac{1}{3}(0.1839) = 0.0613$$

$$p(4) = \frac{1}{4}p(3) = \frac{1}{4}(0.0613) = 0.015325$$

$$p(5) = \frac{1}{5}p(4) = \frac{1}{5}(0.015325) = 0.003065$$

Example 6

If the variance of a Poisson variate is 3, find the probability that (i) $X = 0$, (ii) $0 < X \le 3$, and (iii) $1 \le X < 4$.

Solution

For a Poisson distribution,

Variance = Mean = $\lambda = 3$

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

(i) $P(X = 0) = \dfrac{e^{-3}3^0}{0!} = 0.0498$

(ii) $P(0 < X \le 3) = P(X = 1) + P(X = 2) + P(X = 3)$

$$= \sum_{x=1}^{3} P(X = x)$$

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

$$= 0.5974$$

(iii) $P(1 \le X < 4) = P(X = 1) + P(X = 2) + P(X = 3)$

$$= \sum_{x=1}^{3} P(X = x)$$

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

$$= 0.5974$$

Example 7

If a Poisson distribution is such that $\frac{3}{2}P(X = 1) = P(X = 3)$, find (i) $P(X \ge 1)$, (ii) $P(X \le 3)$, and (iii) $P(2 \le X \le 5)$.

Solution

For a Poisson distribution,

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

$$\frac{3}{2}P(X = 1) = P(X = 3)$$

$$\frac{3}{2}\frac{e^{-\lambda}\lambda^1}{1!} = \frac{e^{-\lambda}\lambda^3}{3!}$$

$$\frac{3}{2}\lambda = \frac{\lambda^3}{6}$$

$$\lambda^3 - 9\lambda = 0$$

$$\lambda(\lambda^2-9)=0$$
$$\lambda=0,3,-3$$

Since $\lambda>0$, $\lambda=3$

Hence, $P(x=x)=\dfrac{e^{-3}3^x}{x!}$, $x=0,1,2,\ldots$

(i) $P(X\geq 1)=1-P(X<1)$
$=1-P(X=0)$
$=1-\dfrac{e^{-3}3^0}{0!}$
$=0.9502$

(ii) $P(X\leq 3)=P(X=0)+P(X=1)+P(X=2)+P(X=3)$
$=\sum_{x=0}^{3}P(X=x)$
$=\sum_{x=0}^{3}\dfrac{e^{-3}3^x}{x!}$
$=0.6472$

(iii) $P(2\leq X\leq 5)=P(X=2)+P(X=3)+P(X=4)+P(X=5)$
$=\sum_{x=2}^{5}P(X=x)$
$=\sum_{x=2}^{5}\dfrac{e^{-3}3^x}{x!}$
$=0.7169$

Example 8

If X is a Poisson variate such that
$$P(X=2)=9\,P(X=4)+90\,P(X=6)$$
Find (i) the mean of X, (ii) the variance of X, (iii) $P(X<2)$, (iv) $P(X>4)$, and (v) $P(X\geq 1)$.

Solution

For a Poisson distribution,
$$P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!},\qquad x=0,1,2,\ldots$$
$$P(X=2)=9P(X=4)+90P(X=6)$$

$$\frac{e^{-\lambda}\lambda^2}{2!}=9\frac{e^{-\lambda}\lambda^4}{4!}+90\frac{e^{-\lambda}\lambda^6}{6!}$$
$$=e^{-\lambda}\lambda^2\left(\frac{9\lambda^2}{4!}+\frac{90\lambda^4}{6!}\right)$$
$$\frac{1}{2}=\frac{9\lambda^2}{4!}+\frac{90\lambda^4}{6!}$$
$$\frac{1}{2}=\frac{3\lambda^2}{8}+\frac{\lambda^4}{8}$$
$$\lambda^4+3\lambda^2-4=0$$
$$\lambda^2=\frac{-3\pm\sqrt{9+16}}{2}=\frac{-3\pm 5}{2}=1,-4$$

Since $\lambda>0$, $\lambda^2=1$

(i) Mean $=\lambda=1$

(ii) Variance $=\lambda=1$

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

(iii) $P(X<2)=P(X=0)+P(X=1)$
$=\sum_{x=0}^{1}\dfrac{e^{-1}1^x}{x!}$
$=0.7358$

(iv) $P(X>4)=1-P(X\leq 4)$
$=1-[P(X=0)+(X=1)+P(X=2)+P(X=3)+P(X=4)]$
$=1-\sum_{x=0}^{4}\dfrac{e^{-1}1^x}{x!}$
$=0.00366$

(v) $P(X\geq 1)=1-P(X=0)$
$=1-\dfrac{e^{-1}1^0}{1!}$
$=0.6321$

Example 9

If a Poisson distribution is such that $\frac{3}{2}P(X=1)=P(X=3)$, find (i) $P(X\geq 1)$, (ii) $P(X\leq 3)$, and (iii) $P(2\leq X\leq 5)$.

Solution

$$\frac{3}{2}P(X=1)=P(X=3)$$

$$\frac{3}{2}\frac{e^{-\lambda}\lambda^1}{1!}=\frac{e^{-\lambda}\lambda^3}{3!}$$

$$\frac{3}{2}=\frac{\lambda^2}{6}$$

$$\lambda^2=9$$

$$\lambda=\pm 3$$

Since $\lambda>0$, $\lambda=3$

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

(i) $P(X\geq 1)=1-P(X<1)$

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

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

$$=0.9502$$

(ii) $P(X\leq 3)=P(X=0)+P(X=1)+P(X=2)+P(X=3)$

$$=\sum_{x=0}^{3}P(X=x)$$

$$=\sum_{x=0}^{3}\frac{e^{-3}3^x}{x!}$$

$$=0.6472$$

(iii) $P(2\leq X\leq 5)=P(X=2)+P(X=3)+P(X=4)+P(X=5)$

$$=\sum_{x=2}^{5}P(X=x)$$

$$=\sum_{x=2}^{5}\frac{e^{-3}3^x}{x!}$$

$$=0.7169$$

Example 10

If X is a Poisson variate such that

$$3P(X=4)=\frac{1}{2}P(X=2)+P(X=0)$$

Find (i) the mean of X, and (ii) $P(X\leq 2)$.

Solution

(i) For a Poisson distribution,

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

$$3P(X=4)=\frac{1}{2}P(X=2)+P(X=0)$$

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

$$\lambda^4-2\lambda^2-8=0$$

$$(\lambda^2-4)(\lambda^2+2)=0$$

$$\lambda=\pm 2 \qquad (\because \lambda \text{ is real})$$

$$\lambda=2 \qquad (\because \lambda>0)$$

Mean $=\lambda=2$

Hence, $P(X=x)=\dfrac{e^{-2}2^x}{x!}, \qquad x=0,1,2,\ldots$

(ii) $P(X\leq 2)=P(X=0)+P(X=1)+P(X=2)$

$$=\sum_{x=0}^{2}P(X=x)$$

$$=\sum_{x=0}^{2}\frac{e^{-2}2^x}{x!}$$

$$=0.6766$$

Example 11

A manufacturer of cotterpins knows that 5% of his products are defective. If he sells cotterpins in boxes of 100 and guarantees that not more than 10 pins will be defective, what is the approximate probability that a box will fail to meet the guaranteed quality?

Solution

Let p be the probability of a pin being defective.

$$p=5\%=0.05, \qquad n=100$$

Since p is very small and n is large, Poisson distribution is used.

$$\lambda=np=100\times 0.05=5$$

Let X be the random variable which denotes the number of defective pins in a box of 100.

Probability of x defective pins in a box of 100

$$P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}=\frac{e^{-5}5^x}{x!}, \quad x=0,1,2,...$$

Probability that a box will fail to meet the guaranteed quality

$$\begin{aligned}P(X>10)&=1-P(X\le 10)\\&=1-\sum_{x=0}^{10}P(X=x)\\&=1-\sum_{x=0}^{10}\frac{e^{-5}5^x}{x!}\\&=0.0137\end{aligned}$$

Example 12

A car-hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is distributed as a Poisson distribution with a mean of 1.5. Calculate the proportion of days on which (i) neither car is used, and (ii) the proportion of days on which some demand is refused.

Solution

$$\lambda=1.5$$

Let X be the random variable which denotes the number of demands for a car on each day.

Probability of days on which there are x demands for a car

$$P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}=\frac{e^{-1.5}1.5^x}{x!}, \quad x=0,1,2,...$$

(i) Proportion or probability of days on which neither car is used

$$P(X=0)=\frac{e^{-1.5}1.5^0}{0!}=0.2231$$

(ii) Proportion or probability of days on which some demand is refused

$$\begin{aligned}P(X>2)&=1-P(X\le 2)\\&=1-\sum_{x=0}^{2}P(X=x)\\&=1-\sum_{x=0}^{2}\frac{e^{-1.5}1.5^x}{x!}\\&=0.1912\end{aligned}$$

Example 13

Six coins are tossed 6400 times. Using the Poisson distribution, what is the approximate probability of getting six heads 10 times?

Solution

Let p be the probability of getting one head with one coin.

$$p=\frac{1}{2}$$

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

$$n=6400$$

$$\lambda=np=6400\left(\frac{1}{64}\right)=100$$

Probability of getting x heads

$$P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}=\frac{e^{-100}100^x}{x!}, \quad x=0,1,2,...$$

Probability of getting 6 heads 10 times

$$P(X=10)=\frac{e^{-100}100^{10}}{10!}=1.025\times10^{-30}$$

Example 14

If 2% of lightbulbs are defective, find the probability that (i) at least one is defective, and (ii) exactly 7 are defective. Also, find $P(1<X<8)$ in a sample of 100.

Solution

Let p be the probability of defective bulb.

$$p=2\%=0.02$$

$$n=100$$

Since p is very small and n is large, Poisson distribution is used.

$$\lambda=np=100(0.02)=2$$

Let X be the random variable which denotes the number of defective bulbs in a sample of 100.

Probability of x defective bulb in a sample of 100

$$P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}=\frac{e^{-2}2^x}{x!}, \quad x=0,1,2,...$$

(i) Probability that at least one bulb is defective

$$P(X \geq 1) = 1 - P(X = 0)$$
$$= 1 - \frac{e^{-2} 2^0}{0!}$$
$$= 0.8647$$

(ii) Probability that exactly 7 bulbs are defective

$$P(X = 7) = \frac{e^{-2} 2^7}{7!} = 0.0034$$

(iii) $P(1 < X < 8) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7)$

$$= \sum_{x=2}^{7} P(X = x)$$
$$= \sum_{x=2}^{7} \frac{e^{-2} 2^x}{x!}$$
$$= 0.5929$$

Example 15

An insurance company insured 4000 people against loss of both eyes in a car accident. Based on previous data, the rates were computed on the assumption that on the average, 10 persons in 100000 will have car accidents each year that result in this type of injury. What is the probability that more than 3 of the insured will collect on their policy in a given year?

Solution

Let p be the probability of loss of both eyes in a car accident.

$$p = \frac{10}{100000} = 0.0001$$
$$n = 4000$$

Since p is very small and n is large, Poisson distribution is used.

$$\lambda = np = 4000(0.0001) = 0.4$$

Let X be the random variable which denotes the number of car accidents in a group of 4000 people.

Probability of x car accidents in a group of 4000 people

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} = \frac{e^{-0.4} 0.4^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Probability that more than 3 of the insured will collect on their policy, i.e., probability of more than 3 car accidents in a group of 4000 people

$$P(X > 3) = 1 - P(X \leq 3)$$
$$= 1 - [P(X = 0) + (X = 1) + P(X = 2) + P(X = 3)]$$
$$= 1 - \sum_{x=0}^{3} P(X = x)$$
$$= 1 - \sum_{x=0}^{3} \frac{e^{-0.4} 0.4^x}{x!}$$
$$= 0.00077$$

Example 16

If two cards are drawn from a pack of 52 cards which are diamonds, using Poisson distribution, find the probability of getting two diamonds at least 3 times in 51 consecutive trials of two cards drawing each time.

Solution

Let p be the probability of getting two diamonds from a pack of 52 cards.

$$p = \frac{{}^{13}C_2}{{}^{52}C_2} = \frac{3}{51}, \quad n = 51$$

Since p is very small and n is large, Poisson distribution is used.

$$\lambda = np = 51\left(\frac{3}{51}\right) = 3$$

Let X be the random variable which denotes the drawing of two diamond cards.

Probability of x trials of drawing two diamond cards in 51 trials

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

Probability of getting two diamond cards at least 3 times in 51 trials

$$P(X \geq 3) = 1 - P(X < 3)$$
$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$
$$= 1 - \sum_{x=0}^{2} \frac{e^{-3} 3^x}{x!}$$
$$= 0.5768$$

Example 17

Suppose a book of 585 pages contains 43 typographical errors. If these errors are randomly distributed throughout the book, what is the probability that 10 pages, selected at random, will be free from errors?

Solution

Let p be the probability of errors in a page.

$$p = \frac{43}{585} = 0.0735, \quad n = 10$$

Since p is very small and n is large, Poisson distribution is used.

$\lambda = np = 10(0.0735) = 0.735$

Let X be the random variable which denotes the errors in the pages.

Probability of x errors in a page in a book of 585 pages

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-0.735}0.735^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Probability that a random sample of 10 pages will contain no error.

$$P(X = 0) = \frac{e^{-0.735}0.735^0}{0!} = 0.4795$$

Example 18

A hospital switchboard receives an average of 4 emergency calls in a 10-minute interval. What is the probability that (i) there are at most 2 emergency calls? (ii) there are exactly 3 emergency calls in an interval of 10 minutes?

Solution

Let p be the probability of receiving emergency calls per minute.

$$p = \frac{4}{10} = 0.4, \quad n = 10$$

$\lambda = np = 10(0.4) = 4$

Let X be the random variable which denotes the number of emergency calls per minute.

Probability of x emergency calls per minute

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-4}4^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Probability that there are at most 2 emergency calls

$$\begin{aligned} P(X \le 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= \sum_{x=0}^{2} P(X = x) \\ &= \sum_{x=0}^{2} \frac{e^{-4}4^x}{x!} \\ &= 0.238 \end{aligned}$$

Probability that there are exactly 3 emergency calls

$$P(X = 3) = \frac{e^{-4}4^3}{3!} = 0.1954$$

Example 19

A manufacturer, who produces medicine bottles, finds that 0.1% of the bottles are defective. The bottles are packed in boxes containing 500 bottles. A drug manufacturer buys 100 boxes from the producer of bottles. Using Poisson distribution, find how many boxes will contain (i) no defective bottles and (ii) at least 2 defective bottles.

Solution

Let p be the probability of deflective bottles.

$p = 0.1\% = 0.001$

$n = 500$

$\lambda = np = 500(0.001) = 0.5$

Let X be the random variable which denotes the number of defective bottles in a box of 500.

Probability of x defective bottles in a box of 500

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-0.5}0.5^x}{x!}, \quad x = 0, 1, 2, \ldots$$

(i) Probability of no defective bottles in a box

$$P(X = 0) = \frac{e^{-0.5}0.5^0}{0!} = 0.6065$$

Number of boxes containing no defective bottles

$f(x) = N\,P(x = 0) = 100(0.6065) = 61$

(ii) Probability of at least 2 defective bottles

$$\begin{aligned} P(X \ge 2) &= 1 - P(X < 2) \\ &= 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - \sum_{x=0}^{1} P(X = x) \\ &= 1 - \sum_{x=0}^{1} \frac{e^{-0.5}0.5^x}{x!} \\ &= 0.0902 \end{aligned}$$

Number of boxes containing at least 2 defective bottles

$f(x) = N\,P(X \ge 2) = 100\,(0.0902) = 9$

Example 20

In a certain factory turning out blades, there is a small chance of $\frac{1}{500}$ for any blade to be defective. The blades are supplied in packets of 10. Use the Poisson distribution to calculate the approximate number of packets containing no defective, one defective, and two defective blades in a consignment of 10000 packets.

Solution

Let p be the probability of defective blades in a packet.

$$p = \frac{1}{500}, \quad n = 10, \quad N = 10000$$

$$\lambda = np = 10\left(\frac{1}{500}\right) = 0.02$$

Let X be the random variable which denotes the number of defective blades in a packet.

Probability of x defective blades in a packet

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-0.02}\, 0.02^x}{x!}, \quad x = 0, 1, 2, \ldots$$

(i) Probability of no defective blades in a packet

$$P(X = 0) = \frac{e^{-0.02}\, 0.02^0}{0!} = 0.9802$$

Number of packets with no defective blades

$f(x) = N\,P(X = 0) = 10000(0.9802) = 9802$

(ii) Probability of one defective blade in a packet

$$P(X = 1) = \frac{e^{-0.02}\, 0.02^1}{1!} = 0.0196$$

Number of packets with one defective blade

$f(x) = N\,P(X = 1) = 10000\,(0.0196) = 196$

(iii) Probability of two defective blades in a packet

$$P(X = 2) = \frac{e^{-0.02}\, 0.02^2}{2!} = 1.96 \times 10^{-4}$$

Number of packets with 2 defective blades

$f(x) = N\,P(X = 2) = 10000\,(1.96 \times 10^{-4}) = 1.96 \approx 2$

Example 21

The number of accidents in a year attributed to taxi drivers in a city follows Poisson distribution with a mean of 3. Out of 1000 taxi drivers, find approximately the number of drivers with (i) no accidents in a year, and (ii) more than 3 accidents in a year.

Solution

For a Poisson distribution,

$$\lambda = 3, \ N = 1000$$

Probably of x accidents in year

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

(i) Probability of no accidents in a year

$$P(X = 0) = \frac{e^{-3}3^0}{0!} = 0.0498$$

Number of drivers with no accidents

$f(x) = N\,P(X = 0) = 1000(0.0498) = 49.8 \approx 50$

(ii) Probability of more than 3 accidents in a year

$$\begin{aligned} P(X > 3) &= 1 - P(X \le 3) \\ &= 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)] \\ &= 1 - \sum_{x=0}^{3} P(X = x) \\ &= 1 - \sum_{x=0}^{3} \frac{e^{-3}3^x}{x!} \\ &= 0.3528 \end{aligned}$$

Number of drivers with more than 3 accidents

$f(x) = N\,P(X > 3) = 1000\,(0.3528) = 3528 \approx 353$

Example 22

Fit a Poisson distribution to the following data:

Number of deaths (x)	0	1	2	3	4
Frequency (f)	122	60	15	2	1

Solution

$$\begin{aligned} \text{Mean} &= \frac{\sum fx}{\sum f} \\ &= \frac{122(0) + 60(1) + 15(2) + 2(3) + 1(4)}{122 + 60 + 15 + 2 + 1} \\ &= \frac{100}{200} \\ &= 0.5 \end{aligned}$$

For a Poisson distribution,

$$\lambda = 0.5$$

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-0.5}0.5^x}{x!}, \quad x = 0, 1, 2, 3, 4$$

$$N = \sum f = 100$$

Theoretical or expected frequency $f(x) = N\,P(X = x)$

$$f(x) = \frac{200\,e^{-0.5}\,0.5^x}{x!}$$

$$f(0) = \frac{200e^{-0.5}\,0.5^0}{0!} = 121.31 \approx 121$$

$$f(1) = \frac{200\,e^{-0.5}\,0.5^1}{1!} = 60.65 \approx 61$$

$$f(2) = \frac{200e^{-0.5}\,0.5^2}{2!} = 15.16 \approx 15$$

$$f(3) = \frac{200e^{-0.5}\,0.5^3}{3!} = 2.53 \approx 3$$

$$f(4) = \frac{200e^{-0.5}\,0.5^4}{4!} = 0.32 \approx 0$$

Poisson Distribution

Number of deaths (x)	0	1	2	3	4
Expected Poisson frequency $f(x)$	121	61	15	3	0

Example 23

Assuming that the typing mistakes per page committed by a typist follows a Poisson distribution, find the expected frequencies for the following distribution of typing mistakes:

Number of mistakes per page	0	1	2	3	4	5
Number of pages	40	30	20	15	10	5

Solution

$$\text{Mean} = \frac{\sum fx}{\sum f}$$

$$= \frac{40(0)+30(1)+20(2)+15(3)+10(4)+5(5)}{40+30+20+15+10+5}$$

$$= \frac{180}{120}$$

$$= 1.5$$

For a Poisson distribution,

$$\lambda = 1.5$$

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-1.5}1.5^x}{x!}, \quad x = 0, 1, 2, 3, 4, 5$$

$$N = \sum f = 120$$

Expected frequency $f(x) = N\,P(X = x)$

$$f(x) = \frac{120\,e^{-1.5}\,1.5^x}{x!}$$

$$f(0) = \frac{120e^{-1.5}\,1.5^0}{0!} = 26.78 \approx 27$$

$$f(1) = \frac{120e^{-1.5}\,1.5^1}{1!} = 40.16 \approx 40$$

$$f(2) = \frac{120e^{-1.5}\,1.5^2}{2!} = 30.12 \approx 30$$

$$f(3) = \frac{120e^{-1.5}\,1.5^3}{3!} = 15.06 \approx 15$$

$$f(4) = \frac{120e^{-1.5}\,1.5^4}{4!} = 5.65 \approx 6$$

$$f(5) = \frac{120e^{-1.5}\,1.5^5}{5!} = 1.69 \approx 2$$

EXERCISE 5.2

1. The mean and variance of a probability distribution is 2. Write down the distribution. $\left[\textbf{Ans.: } P(X = x) = \frac{e^{-2}2^x}{x!}, \ x = 0, 1, 2, \ldots\right]$

2. In a Poisson distribution, the probability $P(X = 0)$ is 20 per cent. Find the mean of the distribution. [**Ans.:** 2.9957]

3. If X is a Poisson variate and $P(X = 0) = 6\,P(X = 3)$, find $P(X = 2)$. [**Ans.:** 0.1839]

4. The standard deviation of a Poisson distribution is 3. Find the probability of getting 3 successes.

[Ans.: 0.0149]

5. The probability that a Poisson variable X takes a positive value is $1 - e^{-1.5}$. Find the variance and the probability that X lies between -1.5 and 1.5.

[Ans.: 1.5, 0.5578]

6. If 2 per cent bulbs are known to be defective bulbs, find the probability that in a lot of 300 bulbs, there will be 2 or 3 defective bulbs using Poisson distribution.

[Ans.: 0.1338]

7. In a certain manufacturing process, 5% of the tools produced turn out to be defective. Find the probability that in a sample of 40 tools, at most 2 will be defective.

[Ans.: 0.675]

8. If the probability that an individual suffers a bad reaction from a particular injection is 0.001, determine the probability that out of 2000 individuals (i) exactly three, and (ii) more than two individuals suffer a bad reaction.

[Ans.: (i) 0.1804 (ii) 0.3233]

9. It is known from past experience that in a certain plant, there are on the average 4 industrial accidents per year. Find the probability that in a given year, there will be less than 4 accidents. Assume Poisson distribution.

[Ans.: 0.43]

10. Find the probability that at most 5 defective fuses will be found in a box of 200 fuses, if experience shows that 2% of such fuses are defective.

[Ans.: 0.7851]

11. Assume that the probability of an individual coal minor being killed in a mine accident during a year is $\frac{1}{2400}$. Use appropriate statistical distribution to calculate the probability that in a mine employing 200 miners, there will be at least one fatal accident every year.

[Ans.: 0.07]

12. Between the hours of 2 and 4 p.m., the average number of phone calls per minute coming into the switchboard of a company is 2.5. Find the probability that during a particular minute, there will be (i) no phone call at all, (ii) 4 or less calls, and (iii) more than 6 calls.

[Ans.: (i) 0.0821 (ii) 0.8909 (iii) 0.0145]

13. Suppose that a local appliances shop has found from experience that the demand for tubelights is roughly distributed as Poisson with a mean of 4 tubelights per week. If the shop keeps 6 tubelights during a particular week, what is the probability that the demand will exceed the supply during that week?

[Ans.: 0.1106]

14. The distribution of the number of road accidents per day in a city is Poisson with a mean of 4. Find the number of days out of 100 days when there will be (i) no accident, (i) at least 2 accidents, and (iii) at most 3 accidents.

[Ans.: (i) 2 (ii) 91 (iii) 44]

15. A manufacturer of electric bulbs sends out 500 lots each consisting of 100 bulbs. If 5% bulbs are defective, in how many lot can we expect (i) 97 or more good bulbs? (ii) less than 96 good bulbs?

[Ans.: (i) 62 (ii) 132]

16. A firm produces articles, 0.1 per cent of which are defective. It packs them in cases containing 500 articles. If a wholesaler purchases 100 such cases, how many cases can be expectecd (i) to be free from defects? (ii) to have one defective article?

[Ans.: (i) 16 (ii) 30]

17. In a certain factory producing certain articles, the probability that an article is defective is $\frac{1}{500}$. The articles are supplied in packets of 20. Find approximately the number of packets containing no defective, one defective, two defectives in a consignment of 20000 packets.

[Ans.: 19200, 768, 15]

18. In a certain factory manufacturing razor blades, there is a small chance, $\frac{1}{50}$ for any blade to be defective. The blades are placed in packets, each containing 10 blades. Using the Poisson distribution, calculate the approximate number of packets containing not more than 2 defective blades in a consignment of 10000 packets.

[Ans.: 9988]

19. It is known that 0.5% of ballpen refills produced by a factory are defective. These refills are dispatched in packaging of equal numbers. Using a Poisson distribution, determine the number of refills in a packing to be sure that at least 95% of them contain no defective refills.

[Ans.: 10]

20. A manufacturer finds that the average demand per day for the mechanics to repair his new product is 1.5 over a period of one year and the demand per day is distributed as a Poisson variate. He employs two mechanics. On how many days in one year (i) would both mechanics would be free? (ii) some demand is refused?

[Ans.: (i) 81.4 days (ii) 69.8 days]

21. Fit a Poisson distribution to the following data:

X	0	1	2	3	4
f	211	90	19	5	0

[Ans.: $\lambda = 0.44$, Frequencies : 209, 92, 20, 3, 1]

22. Fit a Poisson distribution to the following data:

No. of defects per piece	0	1	2	3	4
No. of pieces	43	40	25	10	2

[Ans.: Frequencies: 42, 44, 24, 8, 2]

23. Fit a Poisson distribution to the following data:

X	0	1	2	3	4	5
f	142	156	69	27	5	1

[Ans.: Frequencies: 147, 147, 74, 24, 6, 2]

24. Fit a Poisson distribution to the following data:

X	0	1	2	3	4	5	6	7	8
f	56	156	132	92	37	22	4	0	1

[Ans.: Frequency : 70, 137, 135, 89, 44, 17, 6, 2, 0]

5.4 NORMAL DISTRIBUTION

A continuous random variable X is said to follow normal distribution with mean μ and variance σ^2, if its probability function is given by

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

$$-\infty < X < \infty, -\infty < \mu < \infty, \sigma > 0$$

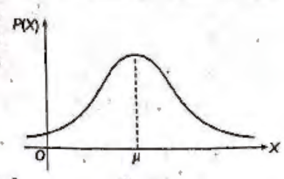


Fig. 5.1

where μ and σ are called parameters of the normal distribution. The curve representing the normal distribution is called the normal curve (Fig. 5.1).

5.4.1 Properties of the Normal Distribution

A normal probability curve, or normal curve, has the following properties:

(i) It is a bell-shaped symmetrical curve about the ordinate $X = \mu$. The ordinate is maximum at $X = \mu$.

(ii) It is a unimodal curve and its tails extend infinitely in both the directions, i.e., the curve is asymptotic to X-axis in both the directions.

(iii) All the three measures of central tendency coincide, i.e.,

mean = median = mode

(iv) The total area under the curve gives the total probability of the random variable X taking values between $-\infty$ to ∞. Mathematically,

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

(v) The ordinate at $X = \mu$ divides the area under the normal curve into two equal parts, i.e.,

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

(vi) The value of $f(x)$ is always nonnegative for all values of X, i.e., the whole curve lies above the X-axis.

(vii) The points of inflexion (the point at which curvature changes) of the curve are at $X = \mu + \sigma$ and the curve changes from concave to convex at $X = \mu + \sigma$ to $X = \mu - \sigma$.

(viii) The area under the normal curve (Fig. 5.2) is distributed as follows:

(a) The area between the ordinates at $\mu - \sigma$ and $\mu + \sigma$ is 68.27%

(b) The area between the ordinates at $\mu - 2\sigma$ and $\mu + 2\sigma$ is 95.45%

(c) The area between the ordinates at $\mu - 3\sigma$ and $\mu + 3\sigma$ is 99.74%

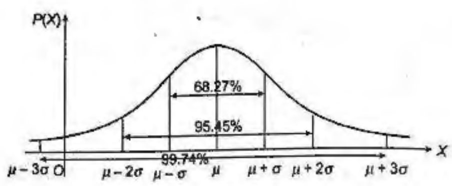


Fig. 5.2

5.4.2 Constants of the Normal Distribution

1. Mean of the Normal Distribution

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

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

Putting $\frac{x-\mu}{\sigma} = t$, $dx = \sigma\,dt$

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

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

Putting $t^2 = u$ in the second integral,

$2t\,dt = du$

When $t \to \infty$, $u \to \infty$

When $t \to -\infty$, $u \to \infty$

$$E(X) = \mu\frac{1}{\sqrt{2\pi}}\cdot\sqrt{2\pi} + \int_{\infty}^{\infty} \sigma\cdot\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u}\frac{du}{2} \quad \left[\because \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2} dt = \sqrt{2\pi}\right]$$

$= \mu + 0$ [∵ the limits of integration are same]

$= \mu$

2. Variance of the Normal Distribution

$$Var(X) = E(X-\mu)^2$$

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

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

Putting $\frac{x-\mu}{\sigma} = t$, $dx = \sigma\,dt$

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

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

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} t^2 e^{-\frac{1}{2}t^2} dt \quad [\because \text{ integral is an even function}]$$

Putting $\frac{t^2}{2} = u$,

$t = \sqrt{2u}$

$dt = \sqrt{2}\frac{1}{2\sqrt{u}}du = \frac{1}{\sqrt{2u}}du$

When $t = 0$, $u = 0$

When $t = \infty$, $u = \infty$

$$Var(X) = \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} 2u e^{-u} \frac{1}{\sqrt{2u}} du$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \int_{0}^{\infty} e^{-u} u^{\frac{1}{2}} du$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) \quad \left[\because \int_{0}^{\infty} e^{-x} x^{n-1} dx = \Gamma n\right]$$

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

$$= \frac{2\sigma^2}{\sqrt{\pi}} \frac{1}{2} \sqrt{\pi}$$

$$= \sigma^2$$

3. Standard Deviation of the Normal Distribution

$SD = \sigma$

4. Mode of the Normal Distribution

Mode is the value of x for which $f(x)$ is maximum. Mode is given by

$$f'(x)=0 \text{ and } f''(x)<0$$

For normal distribution,

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

Differentiating w.r.t. x,

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

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

When $f'(x)=0$, $x-\mu=0$

$$x=\mu$$

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

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

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

At $x=\mu$,

$$f''(x)=-\frac{f(x)}{\sigma^2}=-\frac{1}{\sigma^3\sqrt{2\pi}}<0$$

Hence, $x=\mu$ is the mode of the normal distribution.

5. Median of the Normal Distribution

If M is median of the normal distribution,

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

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

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\mu}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx+\frac{1}{\sigma\sqrt{2\pi}}\int_{\mu}^{M}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx=\frac{1}{2} \quad \text{...(5.3)}$$

Putting $\frac{x-\mu}{\sigma}=t$ in the first integral,

$$dx=\sigma\,dt$$

When $x=-\infty$, $t=-\infty$

When $x=\mu$, $t=0$

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\mu}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{0}e^{-\frac{1}{2}t^2}\sigma\,dt$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}e^{-\frac{1}{2}t^2}dt$$

$$=\frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}e^{-\frac{1}{2}t^2}dt \quad \text{[By symmetry]}$$

$$=\frac{1}{\sqrt{2\pi}}\sqrt{\frac{\pi}{2}}$$

$$=\frac{1}{2} \quad \text{...(5.4)}$$

From Eqs (5.3) and (5.4),

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

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

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

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

Hence, mean = median for the normal distribution.

Note For normal distribution,

mean = median = mode = μ

Hence, the normal distribution is symmetrical.

5.4.3 Probability of a Normal Random Variable in an Interval

Let X be a normal random variable with mean μ and standard deviation σ. The probability of X lying in the interval (x_1, x_2) (Fig. 5.3) is given by

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

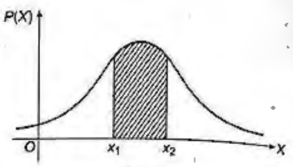


Fig. 5.3

Hence, the probability is equal to the area under the normal curve between the ordinates $X = x_1$ and $X = x_2$ respectively. $P(x_1 < X < x_2)$ can be evaluated easily by converting a normal random variable into another random variable.

Let $\quad Z = \dfrac{X-\mu}{\sigma}$ be a new random variable.

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\left[E(X) - \mu\right] = 0$$

$$\text{Var}(Z) = \text{Var}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}\text{Var}(X-\mu) = \frac{1}{\sigma^2}\text{Var}(X) = 1$$

The distribution of Z is also normal. Thus, if X is a normal random variable with mean μ and standard deviation σ then $Z = \dfrac{X-\mu}{\sigma}$ is a normal random variable with mean 0 and standard deviation 1. Since the parameters of the distribution of Z are fixed, it is a known distribution and is termed *standard normal distribution*. Further, Z is termed as a *standard normal variate*. Thus, the distribution of any normal variate X can always be transformed into the distribution of the standard normal variate Z.

$$P(x_1 \le X \le x_2) = P\left[\left(\frac{x_1-\mu}{\sigma}\right) \le \left(\frac{X-\mu}{\sigma}\right) \le \left(\frac{x_2-\mu}{\sigma}\right)\right]$$
$$= P(z_1 \le Z \le z_2)$$

where $\quad z_1 = \dfrac{x_1-\mu}{\sigma}$ and $z_2 = \dfrac{x_2-\mu}{\sigma}$

This probability is equal to the area under the standard normal curve between the ordinates at $Z = z_1$ and $Z = z_2$.

Case I If both z_1 and z_2 are positive (or both negative) (Fig. 5.4),

$P(x_1 \le X \le x_2)$
$= P(z_1 \le Z \le z_2)$
$= P(0 \le Z \le z_2) - P(0 \le Z \le z_1)$
= (Area under the normal curve from 0 to z_2)
− (Area under the normal curve from 0 to z_1)

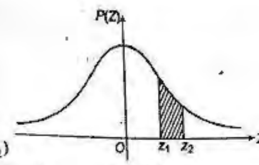


Fig. 5.4

Case II If $z_1 < 0$ and $z_2 > 0$ (Fig. 5.5),

$P(x_1 \le X \le x_2)$
$= P(-z_1 \le Z \le z_2)$
$= P(-z_1 \le Z \le 0) + P(0 \le Z \le z_2)$
$= P(0 \le Z \le z_1) + P(0 \le Z \le z_2)$
[By symmetry]
= (Area under the normal curve from 0 to z_1)
+ (Area under the normal curve from 0 to z_2)

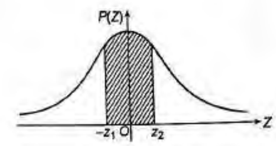


Fig. 5.5

When $X > x_1$, $Z > z_1$, the probability $P(Z > z_1)$ can be found for two cases as follows:

Case I If $z_1 > 0$ (Fig. 5.6),

$P(X > x_1) = P(Z > z_1)$
$= 0.5 - P(0 \le Z \le z_1)$
= 0.5 − (Area under the curve from 0 to z_1)

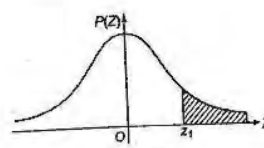


Fig. 5.6

Case II If $z_1 < 0$ (Fig. 5.7),

$P(X > x_1) = P(Z > -z_1)$
$= 0.5 + P(-z_1 < Z < 0)$
$= 0.5 + P(0 < Z < z_1)$
[By symmetry]
= 0.5 + (Area under the curve from 0 to z_1)

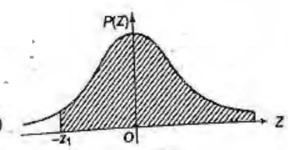


Fig. 5.7

When $X < x_1$, $Z < z_1$, the probability $P(Z < z_1)$ can be found for two cases as follows:

Case I If $z_1 > 0$ (Fig. 5.8),

$$
\begin{aligned}
P(X < x_1) &= P(Z < z_1) \\
&= 1 - P(Z \geq z_1) \\
&= 1 - [0.5 - P(0 < Z < z_1)] \\
&= 0.5 + P(0 < Z < z_1) \\
&= 0.5 + (\text{Area under the curve from 0 to } z_1)
\end{aligned}
$$

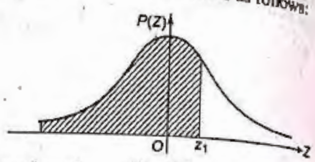


Fig. 5.8

Case II If $z_1 < 0$ (Fig. 5.9),

$$
\begin{aligned}
P(X < x_1) &= P(Z < -z_1) \\
&= 1 - P(Z \geq -z_1) \\
&= 1 - [0.5 + P(-z_1 \leq Z \leq 0)] \\
&= 1 - [0.5 + P(0 \leq Z \leq z_1)] \quad \text{[By symmetry]} \\
&= 0.5 - P(0 \leq Z \leq z_1) \\
&= 0.5 - (\text{Area under the curve from 0 to } z_1)
\end{aligned}
$$

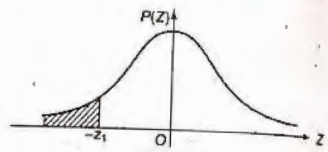


Fig. 5.9

Note

(i) $P(X < x_1) = F(x_1) = \int_{-\infty}^{x_1} f(x)\,dx$

Hence, $P(X < x_1)$ represents the area under the curve from $X = -\infty$ to $X = x_1$.

(ii) If $P(X < x_1) < 0.5$, the point x_1 lies to the left of $X = \mu$ and the corresponding value of standard normal variate will be negative (Fig. 5.10).

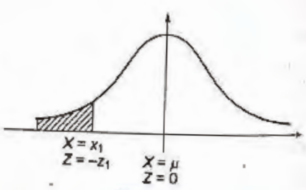


Fig. 5.10

(iii) If $P(X < x_1) > 0.5$, the point x_1 lies to the right of $x = \mu$ and the corresponding value of standard normal variate will be positive (Fig. 5.11).

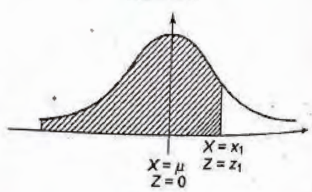
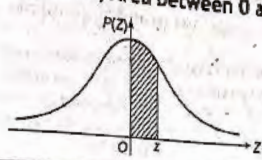


Fig. 5.11

Standard Normal (Z) Table, Area between 0 and z



Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3990	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4115	0.4131	0.4147	0.4162
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

5.4.4 Uses of Normal Distribution

(i) The normal distribution can be used to approximate binomial and Poisson distributions.

(ii) It is used extensively in sampling theory. It helps to estimate parameters from statistics and to find confidence limits of the parameter.

(iii) It is widely used in testing statistical hypothesis and tests of significance in which it is always assumed that the population from which the samples have been drawn should have normal distribution.

(iv) It serves as a guiding instrument in the analysis and interpretation of statistical data.

(v) It can be used for smoothing and graduating a distribution which is not normal simply by contracting a normal curve.

Example 1

What is the probability that a standard normal variate Z will be (i) greater than 1.09? (ii) less than −1.65? (iii) lying between −1 and 1.96? (iv) lying between 1.25 and 2.75?

Solution

(i) $Z > 1.09$ (Fig. 5.12)

$P(Z > 1.09) = 0.5 - P(0 \le Z \le 1.09)$
$= 0.5 - 0.3621$
$= 0.1379$

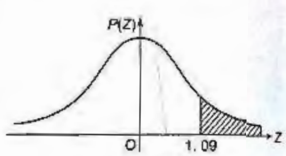


Fig. 5.12

(ii) $Z \le -1.65$ (Fig. 5.13)

$P(Z \le -1.65) = 1 - P(Z > -1.65)$
$= 1 - [0.5 + P(-1.65 < Z < 0)]$
$= 1 - [0.5 + P(0 < Z < 1.65)]$ [By symmetry]
$= 0.5 - P(0 < Z < 1.65)$
$= 0.5 - 0.4505$
$= 0.0495$

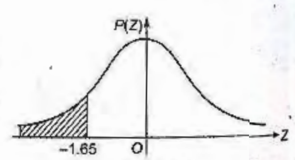


Fig. 5.13

(iii) $-1 < Z < 1.96$ (Fig. 5.14)

$P(-1 < Z < 1.96)$
$= P(-1 < Z < 0) + P(0 < Z < 1.96)$
$= P(0 < Z < 1) + P(0 < Z < 1.96)$ [By symmetry]
$= 0.3413 + 0.4750$
$= 0.8163$

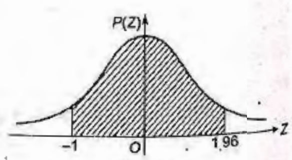


Fig. 5.14

(iv) $1.25 < Z < 2.75$ (Fig. 5.15)

$P(1.25 < Z < 275)$
$= P(0 < Z < 2.75) - P(0 < Z < 1.25)$
$= 0.4970 - 0.3944$
$= 0.1026$

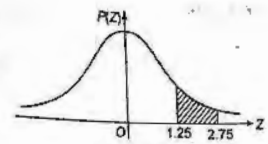


Fig. 5.15

Example 2

If X is a normal variate with a mean of 30 and an SD of 5, find the probabilities that (i) $26 \le X \le 40$, and (ii) $X \ge 45$.

Solution

$\mu = 30, \quad \sigma = 5$

$Z = \frac{X - \mu}{\sigma}$

(i) When $X = 26$, $Z = \frac{26-30}{5} = -0.8$

When $X = 40$, $Z = \frac{40-30}{5} = 2$

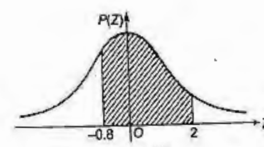


Fig. 5.16

$P(26 \le X \le 40) = P(-0.8 \le Z \le 2)$ (Fig. 5.16)
$= P(-0.8 \le Z \le 0) + P(0 \le Z \le 2)$
$= P(0 \le Z \le 0.8) + P(0 \le Z \le 2)$ [By symmetry]
$= 0.2881 + 0.4772$
$= 0.7653$

(ii) When $X = 45$, $Z = \frac{45-30}{5} = 3$

$P(X \ge 45) = P(Z \ge 3)$ (Fig. 5.17)
$= 0.5 - P(0 < Z < 3)$
$= 0.5 - 0.4987$
$= 0.0013$

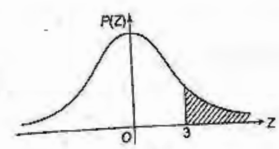


Fig. 5.17

Example 3

X is normally distributed and the mean of X is 12 and the SD is 4. Find out the probability of the following:
(i) $X \ge 20$ (ii) $X \le 20$ (iii) $0 \le X \le 12$.

Solution

$\mu = 12, \quad \sigma = 4$

$$Z = \frac{X-\mu}{\sigma}$$

(i) When $X = 20, Z = \frac{20-12}{4} = 2$

$$\begin{aligned} P(X \ge 20) &= P(Z \ge 2) \text{ (Fig. 5.18)} \\ &= 0.5 - P(0 < Z < 2) \\ &= 0.5 - 0.4772 \\ &= 0.0228 \end{aligned}$$

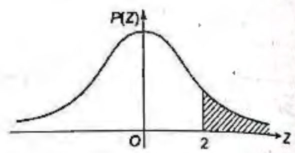


Fig. 5.18

(ii) $$\begin{aligned} P(X \le 20) &= 1 - P(X > 20) \\ &= 1 - 0.0228 \\ &= 0.9772 \end{aligned}$$

(iii) When $X = 0, Z = \frac{0-12}{4} = -3$

When $X = 12, Z = \frac{12-12}{4} = 0$

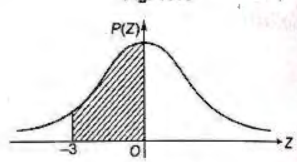


Fig. 5.19

$$\begin{aligned} P(0 \le X \le 12) &= P(-3 \le Z \le 0) \text{ (Fig. 5.19)} \\ &= P(0 \le Z \le 3) \quad \text{[By symmetry]} \\ &= 0.4987 \end{aligned}$$

Example 4

If X is normally distributed with a mean of 2 and an SD of 0.1, find $P(|X-2|) \ge 0.01)$?

Solution:

$\mu = 2, \quad \sigma = 0.1$

$$Z = \frac{X-\mu}{\sigma}$$

When $X = 1.99, Z = \frac{1.99-2}{0.1} = -0.1$

When $X = 2.01, Z = \frac{2.01-2}{0.1} = 0.1$

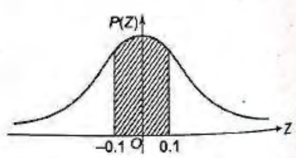


Fig. 5.20

$$\begin{aligned} P(|X-2| \le 0.01) &= P(1.99 \le X \le 2.01) \text{ (Fig. 5.20)} \\ &= P(-0.1 \le Z \le 0.1) \\ &= P(-0.1 \le Z \le 0) + P(0 \le Z \le 0.1) \\ &= P(0 \le Z \le 0.1) + P(0 \le Z \le 0.1) \quad \text{[By symmetry]} \\ &= 2P(0 < Z \le 0.1) \\ &= 2(0.0398) \\ &= 0.0796 \end{aligned}$$

$$\begin{aligned} P(|X-2| \ge 0.01) &= 1 - P(|X-2| < 0.01) \\ &= 1 - 0.0796 \\ &= 0.9204 \end{aligned}$$

Example 5

If X is a normal variate with a mean of 120 and a standard deviation of 10, find c such that (i) $P(X > c) = 0.02$, and (ii) $P(X < c) = 0.05$.

Solution

For normal variate X,

$\mu = 120, \quad \sigma = 10$

$$Z = \frac{X-\mu}{\sigma}$$

(i) $$\begin{aligned} P(X > c) &= 0.02 \\ P(X < c) &= 1 - P(X \ge c) \\ &= 1 - 0.02 \\ &= 0.98 \end{aligned}$$

Since $P(X < c) > 0.5$, the corresponding value of Z will be positive.

$$\begin{aligned} P(X > c) &= P(Z > z_1) \text{ (Fig. 5.21)} \\ 0.02 &= 0.5 - P(0 \le Z \le z_1) \\ P(0 \le Z \le z_1) &= 0.48 \end{aligned}$$

$\therefore\ z_1 = 2.05$ [From normal table]

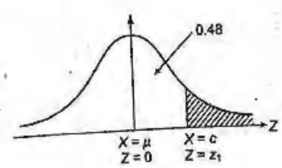


Fig. 5.21

$$Z = \frac{c-120}{10} = z_1 = 2.05$$

$$c = 2.05(10) + 120 = 140.05$$

(ii) Since $P(X < c) < 0.5$, the corresponding value of Z will be negative.

$$\begin{aligned} P(X < c) &= P(Z < -z_1) \text{ (Fig. 5.22)} \\ 0.05 &= 1 - P(Z \ge -z_1) \\ 0.05 &= 1 - [0.5 + P(-z_1 \le Z \le 0)] \end{aligned}$$

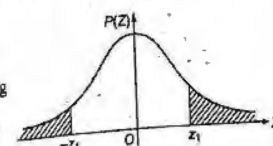


Fig. 5.22

$$0.05 = 1 - [0.5 + P(0 \le Z \le z_1)] \quad \text{[By symmetry]}$$

$$0.05 = 0.5 - P(0 \le Z \le z_1)$$

$$P(0 \le Z \le z_1) = 0.5 - 0.05 = 0.45$$

$$\therefore \quad z_1 = -1.64 \quad \text{[From normal table]}$$

$$Z = \frac{c - 120}{10} = z_1 = -1.64$$

$$c = 10(-1.64) + 120 = 103.6$$

Example 6

A manufacturer knows from his experience that the resistances of resistors he produces is normal with $\mu = 100$ ohms and $SD = \sigma = 2$ ohms. What percentage of resistors will have resistances between 98 ohms and 102 ohms?

Solution

Let X be the random variable which denotes the resistances of the resistors.

$$\mu = 100, \quad \sigma = 2$$

$$Z = \frac{X - \mu}{\sigma}$$

When $X = 98$, $Z = \frac{98 - 100}{2} = -1$

When $X = 102$, $Z = \frac{102 - 100}{2} = 1$

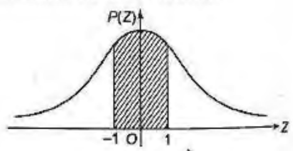


Fig. 5.23

$$P(98 \le X \le 102) = P(-1 \le Z \le 1) \text{ (Fig. 5.23)}$$

$$= P(-1 \le Z \le 0) + P(0 \le Z \le 1)$$

$$= P(0 \le Z \le 1) + P(0 \le Z \le 1) \quad \text{[By symmetry]}$$

$$= 2P(0 \le Z \le 1)$$

$$= 2(0.3413)$$

$$= 0.6826$$

Hence, the percentage of resistors have resistances between 98 ohms and 102 ohms = 68.26%.

Example 7

The average seasonal rainfall in a place is 16 inches with an SD of 4 inches. What is the probability that the rainfall in that place will be between 20 and 24 inches in a year?

Solution

Let X be the random variable which denotes the seasonal rainfall in a year.

$$\mu = 16, \quad \sigma = 4$$

$$Z = \frac{X - \mu}{\sigma}$$

When $X = 20$, $Z = \frac{20 - 16}{4} = 1$

When $X = 24$, $Z = \frac{24 - 16}{4} = 2$

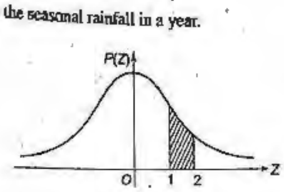


Fig. 5.24

$$P(20 < X < 24) = P(1 < Z < 2) \text{ (Fig. 5.24)}$$

$$= P(0 < Z < 2) - P(0 < Z < 1)$$

$$= 0.4772 - 0.3413$$

$$= 0.1359$$

Example 8

The lifetime of a certain kind of batteries has a mean life of 400 hours and the standard deviation as 45 hours. Assuming the distribution of lifetime to be normal, find (i) the percentage of batteries with a lifetime of at least 470 hours, (ii) the proportion of batteries with a lifetime between 385 and 415 hours, and (iii) the minimum life of the best 5% of batteries.

Solution

Let X be the random variable which denotes the lifetime of a certain kind of batteries.

$$\mu = 400, \quad \sigma = 45$$

$$Z = \frac{X - \mu}{\sigma}$$

(i) When $X = 470$,

$$Z = \frac{470 - 400}{45} = 1.56$$

Fig. 5.25

$$P(X \ge 470) = P(Z \ge 1.56) \text{ (Fig. 5.25)}$$

$$= 0.5 - P(0 < Z < 1.56)$$

$$= 0.5 - 0.4406$$

$$= 0.0594$$

Hence, the percentage of batteries with a lifetime of at least 470 hours = 5.94%.

(ii) When $X = 385$,

$$Z = \frac{385-400}{45} = -0.33$$

When $X = 415$,

$$Z = \frac{415-400}{45} = 0.33$$

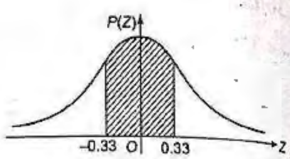


Fig. 5.26

$$\begin{aligned}P(385 < X < 415) &= P(-0.33 < Z < 0.33) \quad \text{(Fig. 5.26)}\\ &= P(-0.33 < Z < 0) + P(0 < Z < 0.33)\\ &= P(0 < Z < 0.33) + P(0 < Z < 0.33) \quad \text{[By symmetry]}\\ &= 2P(0 < Z < 0.33)\\ &= 2(0.1293)\\ &= 0.2586\end{aligned}$$

Hence, the proportion of batteries with a lifetime between 385 and 415 hours = 25.86%.

(iii) $P(X > x_1) = 0.05$ (Fig. 5.27)

$P(X > x_1) = P(Z > z_1)$

$0.05 = 0.5 - P(0 \le Z \le z_1)$

$P(0 \le Z \le z_1) = 0.5 - 0.05 = 0.45$

$\therefore$ $z_1 = 1.65$ [From normal table]

$$Z = \frac{x_1 - 400}{45} = z_1 = 1.65$$

$\therefore$ $x_1 = 1.65(45) + 400 = 474.25$ hours

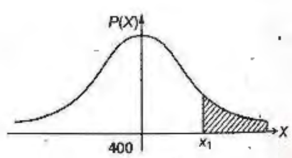


Fig. 5.27

Example 9

If the weights of 300 students are normally distributed with a mean of 68 kg and a standard deviation of 3 kg, how many students have weights (i) greater than 72 kg? (ii) less than or equal to 64 kg? (iii) between 65 kg and 71 kg inclusive?

Solution

Let X be the random variable which denotes the weight of a student.

$\mu = 68, \quad \sigma = 3, \quad N = 300$

$$Z = \frac{X-\mu}{\sigma}$$

(i) When $X = 72$, $Z = \frac{72-68}{3} = 1.33$

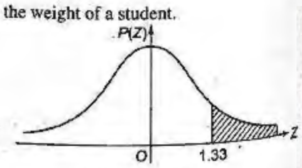


Fig. 5.28

$$\begin{aligned}P(X > 72) &= P(Z > 1.33) \quad \text{(Fig. 5.28)}\\ &= 0.5 - P(0 \le Z \le 1.33)\\ &= 0.5 - 0.4082\\ &= 0.0918\end{aligned}$$

$$\begin{aligned}\text{Number of students with weights more than 72 kg} &= N\,P(X > 72)\\ &= 300(0.0918)\\ &= 27.54\\ &\approx 28\end{aligned}$$

(ii) When $X = 64$, $Z = \frac{64-68}{3} = -1.33$

$$\begin{aligned}P(X \le 64) &= P(Z \le -1.33) \quad \text{(Fig. 5.29)}\\ &= P(Z \ge 1.33) \quad \text{[By symmetry]}\\ &= 0.5 - P(0 < Z < 1.33)\\ &= 0.5 - 0.4082\\ &= 0.0918\end{aligned}$$

Number of students with weights less than or equal to 64 kg

$= N\,P(X \le 64)$

$= 300(0.0918)$

$= 27.54$

≈ 28

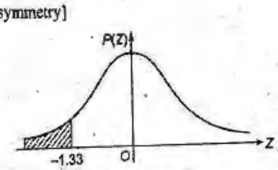


Fig. 5.29

(iii) When $X = 65$, $Z = \frac{65-68}{3} = -1$

When $X = 71$, $Z = \frac{71-68}{3} = 1$

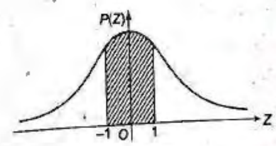


Fig. 5.30

$$\begin{aligned}P(65 \le X \le 71) &= P(-1 \le Z \le 1) \quad \text{(Fig. 5.30)}\\ &= P(-1 \le Z \le 0) + P(0 \le Z \le 1)\\ &= P(0 \le Z \le 1) + P(0 \le Z \le 1) \quad \text{[By symmetry]}\\ &= 2P(0 \le Z \le 1)\\ &= 2(0.3413)\\ &= 0.6826\end{aligned}$$

$$\begin{aligned}\text{Number of students with weights between 65 and 71 kg} &= N\,P(65 \le X \le 71)\\ &= 300(0.6826)\\ &= 204.78\\ &\approx 205\end{aligned}$$

Example 10

The mean yield far a one-acre plot is 662 kg with an SD of 32 kg. Assuming normal distribution, how many one-acre plots in a batch of 1000 plots would you expect to have yields (i) over 700 kg? (ii) below 650 kg? (iii) What is the lowest yield of the best 100 plots?

Solution

Let X be the random variable which denotes the yield for the one-acre plot.

$$\mu = 662, \quad \sigma = 32, \quad N = 1000$$

$$Z = \frac{X-\mu}{\sigma}$$

(i) When $X = 700$, $Z = \frac{700-662}{32} = 1.19$

$$P(X > 700) = P(Z > 1.19) \text{ (Fig. 5.31)}$$
$$= 0.5 - P(0 \le Z \le 1.19)$$
$$= 0.5 - 0.3830$$
$$= 0.1170$$

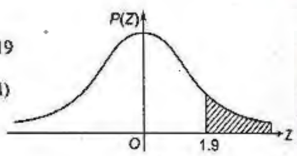


Fig. 5.31

Expected number of plots with yields over 700 kg $= N\,P(X > 700)$

$$= 1000(0.1170)$$
$$= 117$$

(ii) When $X = 650$,

$$Z = \frac{650-662}{32} = -0.38$$

$$P(X < 650) = P(Z < -0.38) \text{ (Fig. 5.32)}$$
$$= P(Z > 0.38) \quad \text{[By symmetry]}$$
$$= 0.5 - P(0 \le Z \le 0.38)$$
$$= 0.5 - 0.1480$$
$$= 0.352$$

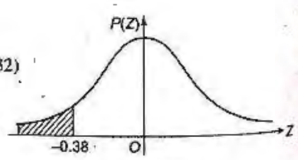


Fig. 5.32

Expected number of plots with yields below 650 kg $= N\,P(X < 650)$

$$= 1000(0.352)$$
$$= 352$$

(iii) The lowest yield, say, x_1 of the best 100 plots is given by

$$P(X > x_1) = \frac{100}{1000} = 0.1$$

When $X = x_1$, $Z = \frac{x_1 - 662}{32} = z_1$

$$P(X > x_1) = P(Z > z_1)$$
$$0.1 = 0.5 - P(0 \le Z \le z_1)$$
$$P(0 \le Z \le z_1) = 0.4$$
$$\therefore \quad z_1 = 1.2 \text{ (approx.) [From normal table]}$$
$$\frac{x_1 - 662}{32} = 1.28$$
$$x_1 = 702.96$$

Hence, the best 100 plots have yields over 702.96 kg.

Example 11

Assume that the mean height of Indian soldiers is 68.22 inches with a variance of 10.8 inches. How many soldiers in a regiment of 1000 would you expect to be over 6 feet tall?

Solution

Let X be the continuous random variable which denotes the heights of Indian soldiers.

$$\mu = 68.22, \quad \sigma^2 = 10.8, \quad N = 1000$$
$$\sigma = 3.29$$
$$Z = \frac{X-\mu}{\sigma}$$

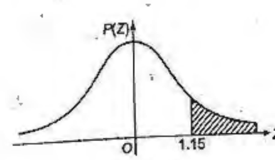


Fig. 5.33

When $X = 6$ feet $= 72$ inches,

$$Z = \frac{72-68.22}{3.29} = 1.15$$

$$P(X > 72) = P(Z > 1.15) \text{ (Fig. 5.33)}$$
$$= 0.5 - P(0 \le Z \le 1.15)$$
$$= 0.5 - 0.3749$$
$$= 0.1251$$

Expected number of Indian soldiers having heights over 6 feet (72 inches)

$$= N\,P(X > 72)$$
$$= 1000(0.1251)$$
$$= 125.1$$
$$\approx 125$$

Example 12

The marks obtained by students in a college are normally distributed with a mean of 65 and a variance of 25. If 3 students are selected at random from this college, what is the probability that at least one of them would have scored more than 75 marks?

Solution

Let X be the continuous random variable which denotes the marks of a student.

$$\mu = 65, \quad \sigma^2 = 25$$
$$\sigma = 5$$
$$Z = \frac{X-\mu}{\sigma}$$

When $X = 75$, $Z = \frac{75-65}{5} = 2$

$$\begin{aligned} P(X > 75) &= P(Z > 2) \quad \text{(Fig. 5.34)} \\ &= 0.5 - P(0 \le Z \le 2) \\ &= 0.5 - 0.4772 \\ &= 0.0228 \end{aligned}$$

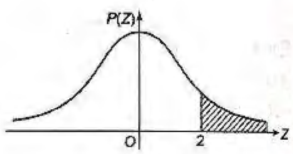


Fig. 5.34

If p is the probability of scoring more than 75 marks,

$p = 0.0228$, $q = 1 - p = 1 - 0.0228 = 0.9772$

P(at least one student would have scored more than 75 marks)

$$\begin{aligned} &= \sum_{x=1}^{3} {}^3C_x\, p^x q^{n-x} \\ &= \sum_{x=1}^{3} {}^3C_x (0.0228)^x (0.9772)^{3-x} \\ &= 0.0668 \end{aligned}$$

Example 13

Find the mean and standard deviation in which 7% of items are under 35 and 89% are under 63.

Solution

Let μ be the mean and σ be standard deviation of the normal curve.

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

$$P(X > 63) = 1 - P(X < 63) = 1 - 0.89 = 0.11$$
$$Z = \frac{X-\mu}{\sigma}$$

Since $P(X < 35) < 0.5$, the corresponding value of Z will be negative.

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

Since $P(X < 63) > 0.5$, the corresponding value of Z will be positive.

When $X = 63$, $Z = \frac{63-\mu}{\sigma} = z_2$ (say)

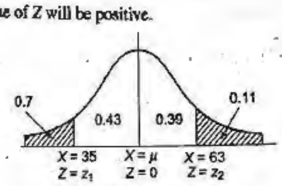


Fig. 5.35

From Fig. 5.35,

$$P(Z < -z_1) = 0.07$$
$$P(Z > z_2) = 0.11$$

$$\begin{aligned} P(0 < Z < z_1) &= P(-z_1 < Z < 0) \\ &= 0.5 - P(Z \le -z_1) \\ &= 0.5 - 0.07 \\ &= 0.43 \\ z_1 &= 1.48 \quad \text{[From normal table]} \end{aligned}$$

$$\begin{aligned} P(0 < Z < z_2) &= 0.5 - P(Z \ge z_2) \\ &= 0.5 - 0.11 \\ &= 0.39 \\ z_2 &= 1.23 \quad \text{[From normal table]} \end{aligned}$$

Hence, $\frac{35-\mu}{\sigma} = -1.48$

$$-1.48\,\sigma + \mu = 35 \qquad \text{...(1)}$$

and $\frac{63-\mu}{\sigma} = 1.23$

$$1.23\,\sigma + \mu = 63 \qquad \text{...(2)}$$

Solving Eqs (1) and (2),

$$\mu = 50.29, \qquad \sigma = 10.33$$

Example 14

In an examination, it is laid down that a student passes if he secures 40 % or more. He is placed in the first, second, and third division according to whether he secures 60% or more marks, between 50% and 60% marks and between 40% and 50% marks respectively. He gets a distinction in case he secures 75% or more. It is noticed from the result that 10% of *the students failed in the examination, whereas 5% of them obtained distinction. Calculate the percentage of students placed in the second division. (Assume normal distribution of marks.)*

Solution

Let X be the random variable which denotes the marks of students in the examination. Let μ be the mean and σ be the standard deviation of the normal distribution of marks.

$$P(X < 40) = 0.10$$

$$P(X \geq 75) = 0.05$$

$$P(X < 75) = 1 - P(X \geq 75) = 1 - 0.05 = 0.95$$

$$Z = \frac{X-\mu}{\sigma}$$

Since $P(X < 40) < 0.5$, the corresponding value of Z will be negative.

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

Since $P(X < 75) < 0.5$, the corresponding value of Z will be positive.

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

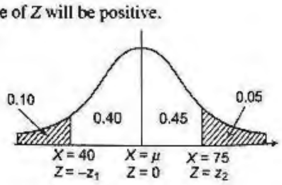


Fig. 5.36

From Fig. 5.36,

$$P(Z < -z_1) = 0.10$$

$$P(Z > z_2) = 0.05$$

$$P(0 < Z < z_1) = P(-z_1 < Z < 0)$$
$$= 0.5 - P(Z \leq -z_1)$$
$$= 0.5 - 0.10$$
$$= 0.40$$

$$z_1 = 1.28 \quad \text{[From normal table]}$$

$$P(0 < Z < z_2) = 0.5 - P(Z \geq z_2)$$
$$= 0.5 - 0.05$$
$$= 0.45$$

$$z_2 = 1.64 \quad \text{[From normal table]}$$

Hence, $\frac{40-\mu}{\sigma} = -1.28$

$$\mu - 1.28\,\sigma = 40 \qquad \text{...(1)}$$

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

$$\mu + 1.64\,\sigma = 75 \qquad \text{...(2)}$$

Solving Eqs (1) and (2),

$$\mu = 55.34 \approx 55$$

$$\sigma = 11.98 \approx 12$$

Probability that a student is placed in the second division is equal to the probability that his score lies between 50 and 60

When $X = 50$, $Z = \frac{50-55}{12} = -0.42$

When $X = 60$, $Z = \frac{60-55}{12} = 0.42$

$$P(50 < X < 60) = P(-0.42 < Z < 0.42)$$
$$= P(-0.42 < Z < 0) + P(0 < Z < 0.42)$$
$$= P(0 < Z < 0.42) + P(0 < Z < 0.42) \quad \text{[By symmetry]}$$
$$= 2\,P(0 < Z < 0.42)$$
$$= 2(0.1628)$$
$$= 0.3256$$
$$\approx 0.32$$

Hence, the percentage of students placed in the second division = 32%.

5.4.5 Fitting a Normal Distribution

Fitting a normal distribution or a normal curve to the data means to find the equation of the curve in the form $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ which will be as close as possible to the points given. There are two purposes of fitting a normal curve:

(i) To judge the whether the normal curve is the best fit to the sample data.

(ii) To use the normal curve to estimate the characteristics of a population.

The area method for fitting a normal curve is given by the following steps:

(i) Find the mean μ and standard deviation σ for the given data if not given.

(ii) Write the class intervals and lower limits X of class intervals in two columns.

(iii) Find $Z = \frac{X-\mu}{\sigma}$ for each class interval.

(iv) Find the area corresponding to each Z from the normal table.

(v) Find the area under the normal curve between the successive values of Z. These are obtained by subtracting the successive areas when the corresponding Z's have the same sign and adding them when the corresponding Z's have the same sign and adding them when the corresponding Z's opposite sign.

(vi) Find the expected frequencies by multiplying the relative frequencies by the number of observations.

Example 1

Fit a normal curve from the following distribution. It is given that the mean of the distribution is 43.7 and its standard distribution is 14.8.

Class interval	11–20	21–30	31–40	41–50	51–60	61–70	71–80
Frequency	20	28	40	60	32	20	8

Solution

$\mu = 43.7, \quad \sigma = 14.8 \quad N = \Sigma f = 200$

The series is converted into an inclusive series.

Class Interval	Lower class	$Z = \frac{X-\mu}{\sigma}$	Area from 0 to Z	Area in class Interval	Expected Frequencies
10.5–20.5	10.5	–2.24	0.4875	0.0457	9.14 ≈ 9
20.5–30.5	20.5	–1.57	0.4418	0.1285	25.7 ≈ 26
30.5–40.5	30.5	–0.89	0.3133	0.2262	45.24 ≈ 45
40.5–50.5	40.5	–0.22	0.0871	0.2643	52.86 ≈ 53
50.5–60.5	50.5	0.46	0.1772	0.1957	39.14 ≈ 39
60.5–70.5	60.5	1.14	0.3729	0.092	18.4 ≈ 18
70.5–80.5	70.5	1.81	0.4649	0.0287	5.74 = 65
	80.5	2.49	0.4936		

Example 2

Fit a normal distribution to the following data:

X	125	135	145	155	165	175	185	195	205
Y	1	1	14	22	25	19	13	3	2

It is given that $\mu = 165.5$ *and* $\sigma = 15.26$.

Solution

$\mu = 165.5, \quad \sigma = 15.26 \quad N = \Sigma f = 100$

The data is first converted into class intervals with inclusive series.

Class Interval	Lower class	$Z = \frac{X-\mu}{\sigma}$	Area from 0 to Z	Area in class Interval	Expected Frequencies
120–130	120	–2.98	0.4986	0.0085	0.85 ≈ 1
130–140	130	–2.33	0.4901	0.0376	3.74 ≈ 4
140–150	140	–1.67	0.4525	0.1064	10.64 ≈ 11
150–160	150	–1.02	0.3461	0.2055	20.55 ≈ 21
160–170	160	–0.36	0.1406	0.2547	25.47 = 25
170–180	170	0.29	0.1141	0.2148	21.48 ≈ 21
180–190	180	0.95	0.3289	0.1174	11.74 = 12
190–200	190	1.61	0.4463	0.0418	4.18 ≈ 4
200–210	200	2.26	0.4881	0.0101	1.01 ≈ 1
210–220	210	2.92	0.4982		

EXERCISE 5.3

1. If X is normally distributed with a mean and standard deviation of 4, find (i) $P(5 \le X \le 10)$, (ii) $P(X \ge 15)$, (iii) $P(10 \le X \le 15)$, and (iv) $P(X \le 5)$.
   [Ans.: (i) 0.3345 (ii) 0.003 (iii) 0.0638 (iv) 0.4013]

2. A normal distribution has a mean of 5 and a standard deviation of 3. What is the probability that the deviation from the mean of an item taken at random will be negative?
   [Ans.: 0.0575]

3. If X is a normal variate with a mean of 30 and an SD of 6, find the value of $X = x_1$ such that $P(X \ge x_1) = 0.05$.
   [Ans.: 39.84]

4. If X is a normal variate with a mean of 25 and SD of 5, find the value of $X = x_1$ such that $P(X \le x_1) = 0.01$.
   [Ans.: 11.02]

5. The weights of 4000 students are found to be normally distributed with a mean of 50 kg and an SD of 5 kg. Find the probability that a student selected at random will have weight (i) less than 45 kg, and (ii) between 45 and 60 kg.
   [Ans.: (i) 0.1587 (ii) 0.8185]

6. The daily sales of a firm are normally distributed with a mean of ₹ 8000 and a variance of ₹ 10000. (i) What is the probability that on a certain

day the sales will be less than ₹ 8210? (ii) What is the percentage of days on which the sales will be between ₹ 8100 and ₹ 8200?

[Ans.: (i) 0.482 (ii) 14%]

7. The mean height of Indian soldiers is 68.22″ with a variance of 10.8″. Find the expected number of soldiers in a regiment of 1000 whose height will be more than 6 feet.

[Ans.: 125]

8. The life of army shoes is normally distributed with a mean of 8 months and a standard deviation of 2 months. If 5000 pairs are issued, how many pairs would be expected to need replacement after 12 months?

[Ans.: 2386]

9. In an intelligence test administered to 1000 students, the average was 42 and the standard deviation was 24. Find the number of students (i) exceeding 50, (ii) between 30 and 54, and (iii) the least score of top 1000 students.

[Ans.: (i) 129 (ii) 383 (iii) 72.72]

10. In a test of 2000 electric bulbs, it was found that the life of a particular make was normally distributed with an average of life of 2040 hours and a standard deviation of 60 hours. Estimate the number of bulbs likely to burn for (i) more than 2150 hours, and (ii) less than 1950 hours.

[Ans.: (i) 67 (ii) 184]

11. The marks of 1000 students of a university are found to be normally distributed with a mean of 70 and a standard of deviation 5. Estimate the number of students whose marks will be (i) between 60 and 75, (ii) more than 75, and (iii) less than 68.

[Ans.: (i) 910 (ii) 23 (iii) 37]

12. In a normal distribution, 31% items are under 45 and 8% are over 64. Find the mean and standard deviation. Find also, the percentage of items lying between 30 and 75.

[Ans.: 50, 10, 0.957]

13. Of a large group of men, 5% are under 60 inches in height and 40% are between 60 and 65 inches. Assuming a normal distribution, find the mean and standard deviation of distribution.

[Ans.: 65.42, 3.27]

14. The marks obtained by students in an examination follow a normal distribution. If 30% of the students got marks below 35 and 10% got marks above 60, find the mean and percentage of students who got marks between 40 and 50.

[Ans.: 42.23, 13.88, 28%]

15. Fit a normal distribution to the following data:

Class	60–65	65–70	70–75	75–80	80–85	85–90	90–95	95–100
Frequency	3	21	150	335	326	135	26	4

[Ans.: Expected frequency: 3, 31, 148, 322, 319, 144, 30, 3]

5.5 EXPONENTIAL DISTRIBUTION

A continuous random variable X is said to follow exponential distribution if its probability function is given by

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0$$
$$= 0 \quad , \quad x \le 0$$

where $\lambda > 0$ is called the rate of the distribution.

5.5.1 Memoryless Property of the Exponential Distribution

The exponential distribution has the memoryless (forgetfulness) property. This property indicates that the distribution is independent of its part, that means future happening of an event has no relation to whether or not this even has happened in the past. This property is as follows:

If X is exponentially distributed, and s, t are two positive real numbers then

$$P[(X > s+t)/(X > s)] = P(X > t)$$

Proof: $P[(X > s+t)/(X > s)] = \dfrac{P[(X > s+t) \cap (X > s)]}{P(X > s)}$

[using conditional probability]

$$= \frac{P(X > s+t)}{P(X > s)}$$

$$= \frac{\int_{s+t}^{\infty} \lambda e^{-\lambda x} dx}{\int_{s}^{\infty} \lambda e^{-\lambda x} dx}$$

$$= \frac{\lambda\left|\frac{e^{-\lambda x}}{-\lambda}\right|_{s+t}^{\infty}}{\lambda\left|\frac{e^{-\lambda x}}{-\lambda}\right|_{s}^{\infty}}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}$$

$$= e^{-\lambda t} \qquad \text{...(5.5)}$$

$$P(X > t) = \int_t^{\infty} \lambda e^{-\lambda x}\,dx$$

$$= \lambda\left|\frac{e^{-\lambda x}}{-\lambda}\right|_t^{\infty}$$

$$= e^{-\lambda t} \qquad \text{...(5.6)}$$

From Eq. (5.5) and Eq. (5.6),

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

5.5.2 Constants of the Exponential Distribution

1. Mean of the Exponential Distribution

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

$$= \int_0^{\infty} x\,\lambda e^{-\lambda x}\,dx$$

$$= \lambda\left|x\cdot\frac{e^{-\lambda x}}{-\lambda} - 1\cdot\frac{e^{-\lambda x}}{\lambda^2}\right|_0^{\infty}$$

$$= \lambda\cdot\frac{1}{\lambda^2}$$

$$= \frac{1}{\lambda}$$

2. Variance of the Exponential Distribution

$$\text{Var}(X) = E(X^2) - [E(X)]^2 \qquad \text{...(5.7)}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$= \int_0^{\infty} x^2 \lambda e^{-\lambda x}\,dx$$

$$= \lambda\left|x^2\,\frac{e^{-\lambda x}}{-\lambda} - 2x\,\frac{e^{-\lambda x}}{\lambda^2} + 2\,\frac{e^{-\lambda x}}{-\lambda^3}\right|_0^{\infty}$$

$$= \lambda\left(\frac{2}{\lambda^3}\right)$$

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

Substituting in Eq (5.7),

$$\text{Var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \qquad \left[\because \mu = \frac{1}{\lambda}\right]$$

3. Standard Deviation of the Exponential Distribution

$$\text{SD} = \sqrt{\text{Var}(X)} = \sqrt{\frac{1}{\lambda^2}} = \frac{1}{\lambda}$$

4. Mode of the Exponential Distribution

Mode is the value of x for which $f(x)$ is maximum.

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0$$
$$= 0 \quad , \quad x \le 0$$

$f(x)$ will be maximum when $e^{-\lambda x}$ is maximum.

Maximum value of $e^{-\lambda x} = 1$, which is at $x = 0$.

Hence, $x = 0$ is the mode of the exponential distribution

5. Median of the Exponential Distribution

If M is the median of the exponential distribution,

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

$$\int_0^{M} \lambda e^{-\lambda x}\,dx = \frac{1}{2}$$

$$\lambda\left|\frac{e^{-\lambda x}}{-\lambda}\right|_0^{M} = \frac{1}{2}$$

$$-(e^{-\lambda M} - 1) = \frac{1}{2}$$

$$-e^{-\lambda M} = \frac{1}{2} - 1 = -\frac{1}{2}$$

$$e^{-\lambda M} = \frac{1}{2}$$

$$-\lambda M \log e = \log\frac{1}{2} = -\log 2$$

$$\lambda M = \log 2$$

$$M = \frac{1}{\lambda}\log 2$$

Example 1

Let X be a random variable with pdf

$$f(x) = \begin{cases} \frac{1}{5}e^{-\frac{x}{5}} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find (i) $P(X > 5)$ (ii) $P(3 \le X \le 6)$ (iii) mean (iv) variance.

Solution

$$\lambda = \frac{1}{5}$$

(i) $P(X > 5) = \int_5^\infty f(x)\,dx$

$$= \int_5^\infty \frac{1}{5}e^{-\frac{x}{5}}\,dx$$

$$= \frac{1}{5}\left|\frac{e^{-\frac{x}{5}}}{-\frac{1}{5}}\right|_5^\infty$$

$$= -\left|e^{-\frac{x}{5}}\right|_5^\infty$$

$$= -(e^{-\infty} - e^{-1})$$

$$= e^{-1}$$

$$= 0.3679$$

(ii) $P(3 \le X \le 6) = \int_3^6 f(x)\,dx$

$$= \int_3^6 \frac{1}{5}e^{-\frac{x}{5}}\,dx$$

$$= \frac{1}{5}\left|\frac{e^{-\frac{x}{5}}}{-\frac{1}{5}}\right|_3^6$$

$$= -\left|e^{-\frac{x}{5}}\right|_3^6$$

$$= -\left(e^{-\frac{6}{5}} - e^{-\frac{3}{5}}\right)$$

$$= e^{-\frac{3}{5}} - e^{-\frac{6}{5}}$$

$$= 0.2476$$

(iii) Mean $\mu = \frac{1}{\lambda} = \frac{1}{\left(\frac{1}{5}\right)} = 5$

(iv) Variance $= \text{Var}(X) = \frac{1}{\lambda^2} = \frac{1}{\left(\frac{1}{5}\right)^2} = 25$

Example 2

A random variable has pdf $f(x) = ce^{-2x}$ for $x > 0$. Find (i) $P(X > 2)$ (ii) $P\left(X < \frac{1}{c}\right)$.

Solution

Since $f(x)$ is a probability density function,

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

$$\int_0^\infty ce^{-2x}\,dx = 1$$

$$\left|\frac{ce^{-2x}}{-2}\right|_0^\infty = 1$$

$$-\frac{c}{2}\left|e^{-2x}\right|_0^{\infty}=1$$

$$-\frac{c}{2}(e^{-\infty}-e^{0})=1$$

$$\frac{c}{2}=1$$

$$c=2$$

$$\therefore f(x)=2e^{-2x}, \quad x>0$$

(i) $P(X>2)=\int_2^{\infty} f(x)\,dx$

$$=\int_2^{\infty} 2e^{-2x}\,dx$$

$$=2\left|\frac{e^{-2x}}{-2}\right|_2^{\infty}$$

$$=-\left|e^{-2x}\right|_2^{\infty}$$

$$=-(e^{-\infty}-e^{-4})$$

$$=e^{-4}$$

$$=0.0183$$

(ii) $P\left(X<\frac{1}{c}\right)=P\left(X<\frac{1}{2}\right)$

$$=\int_0^{\frac{1}{2}} f(x)\,dx$$

$$=\int_0^{\frac{1}{2}} 2e^{-2x}\,dx$$

$$=2\left|\frac{e^{-2x}}{-2}\right|_0^{\frac{1}{2}}$$

$$=-\left|e^{-2x}\right|_0^{\frac{1}{2}}$$

$$=-(e^{-1}-e^{0})$$

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

$$=0.6321$$

Example 3

If X is random variable which follows an exponential distribution with parameter λ with $P(X\le 1)=P(X>1)$, find Var(X).

Solution

Since X is random variable which follows an exponential distribution,

$$f(x)=\lambda e^{-\lambda x}, \quad x\ge 0$$

$$P(X\le 1)=P(X>1)$$

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

$$2P(X>1)=1$$

$$P(X>1)=\frac{1}{2}$$

$$\int_1^{\infty} f(x)\,dx=\frac{1}{2}$$

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

$$\lambda\left|\frac{e^{-\lambda x}}{-\lambda}\right|_1^{\infty}=\frac{1}{2}$$

$$-\left|e^{-\lambda x}\right|_1^{\infty}=\frac{1}{2}$$

$$-(e^{-\infty}-e^{-\lambda})=\frac{1}{2}$$

$$e^{-\lambda}=\frac{1}{2}$$

$$\frac{1}{e^{\lambda}}=\frac{1}{2}$$

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

$$\lambda=\log_e 2$$

$$\text{Var}(X)=\frac{1}{\lambda^2}=\frac{1}{(\log_e 2)^2}$$

Example 4

If X is a exponentially distributed random variable with parameter λ, find the value of k such that $\dfrac{P(X>k)}{P(X\le k)}=a$.

Solution

$$\frac{P(X>k)}{P(X\le k)}=a$$

$$\frac{P(X>k)}{1-P(X>k)}=a$$

$$P(X>k)=a\left[1-P(X>k)\right]$$

$$P(X>k)(1+a)=a$$

$$P(X>k)=\frac{a}{1+a}$$

$$\int_k^{\infty} f(x)\,\mathrm{d}x=\frac{a}{1+a}$$

$$\int_k^{\infty} \lambda e^{-\lambda x}\,\mathrm{d}x=\frac{a}{1+a}$$

$$\lambda\left|\frac{e^{-\lambda x}}{-\lambda}\right|_k^{\infty}=\frac{a}{1+a}$$

$$-\left|e^{-\lambda x}\right|_k^{\infty}=\frac{a}{1+a}$$

$$-(e^{-\infty}-e^{-\lambda k})=\frac{a}{1+a}$$

$$e^{-\lambda k}=\frac{a}{1+a}$$

$$\frac{1}{e^{\lambda k}}=\frac{a}{1+a}$$

$$e^{\lambda k}=\frac{1+a}{a}$$

$$\lambda k=\log\left(\frac{1+a}{a}\right)$$

$$k=\frac{1}{\lambda}\log\left(\frac{1+a}{a}\right)$$

Example 5

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

Solution

Since $f(x)$ is a density function,

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

$$\int_a^{\infty} ce^{-b(x-a)}\,\mathrm{d}x=1$$

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

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

$$-\frac{c}{b}(e^{-\infty}-e^{0})=1$$

$$\frac{c}{b}=1$$

$$b=c \qquad \text{...(1)}$$

$$\mu=E(X)=\int_a^{\infty} bxe^{-b(x-a)}\,\mathrm{d}x$$

$$=be^{ab}\left|x\left(\frac{e^{-bx}}{-b}\right)-\frac{e^{-bx}}{b^2}\right|_a^{\infty}$$

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

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

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

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

$$=b\left(\frac{a^2}{b}+\frac{2a}{b^2}+\frac{2}{b^3}\right)$$

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

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

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

$$= \frac{1}{b^2}$$

$$\sigma = \frac{1}{b} \qquad \text{...(3)}$$

From Eq. (1) and (3),

$$b = c = \frac{1}{\sigma}$$

Subtracting Eq. (3) from Eq. (2),

$$\mu - \sigma = a$$

$$\therefore \quad a = \mu - \sigma$$

Example 6

The mileage which car owners get with a certain kind of radial tire is a random variable having an exponential distribution with mean 4000 km. Find the probabilities that one of these tires will last (i) at least 2000 km (ii) at most 3000 km.

Solution

Let X be the random variable which denotes the mileage obtained with the tire.

$$\text{Mean } \mu = \frac{1}{\lambda} = 4000 \text{ km}$$

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

$$= \frac{1}{4000} e^{-\frac{1}{4000}x}, \qquad x > 0$$

(i) $P(X \ge 2000) = \int_{2000}^{\infty} f(x)\,dx$

$$= \int_{2000}^{\infty} \frac{1}{4000} e^{-\frac{1}{4000}x}\,dx$$

$$= \frac{1}{4000} \left| \frac{e^{-\frac{1}{4000}x}}{-\frac{1}{4000}} \right|_{2000}^{\infty}$$

$$= -\left| e^{-\frac{1}{4000}x} \right|_{2000}^{\infty}$$

$$= -(e^{-\infty} - e^{-0.5})$$

$$= e^{-0.5}$$

$$= 0.6065$$

(ii) $P(X \le 3000) = \int_0^{3000} f(x)\,dx$

$$= \int_0^{3000} \frac{1}{4000} e^{-\frac{1}{4000}x}\,dx$$

$$= \frac{1}{4000} \left| \frac{e^{-\frac{1}{4000}x}}{-\frac{1}{4000}} \right|_0^{3000}$$

$$= -\left| e^{-\frac{1}{4000}x} \right|_0^{3000}$$

$$= -(e^{-0.75} - e^{0})$$

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

$$= 0.5270$$

Example 7

If the number of kilometers that a car can run before its battery wears out is exponentially distributed with an average value of 10000 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 same time.

Solution

Let X be the random variable which denotes the number of kilometers that a car can run before its battery wears out.

$$\text{Mean } \mu = \frac{1}{\lambda} = 10000$$

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

$$= \frac{1}{1000} e^{-\frac{1}{10000}x}, x > 0$$

$$P(X > 5000) = \int_{5000}^{\infty} f(x)\,dx$$
$$= \int_{5000}^{\infty} \frac{1}{10000} e^{-\frac{1}{10000}x}\,dx$$
$$= \frac{1}{10000} \left| \frac{e^{-\frac{1}{10000}x}}{-\frac{1}{10000}} \right|_{5000}^{\infty}$$
$$= -\left| e^{-\frac{1}{10000}x} \right|_{5000}^{\infty}$$
$$= -(e^{-\infty} - e^{-0.5})$$
$$= e^{-0.5}$$
$$= 0.6065$$

Example 8

The average time it takes to serve a customer at a petrol pump is 6 minutes. The service time follows exponential distribution. Calculate the probability that

(i) A customer will take less than 2 minutes to complete the service.
(ii) A customer will take between 4 and 5 minutes to get the service.
(iii) A customer will take more than 10 minutes for his service.

Solution

Let X be the random variable which denotes the service time.

$$\text{Mean } \mu = \frac{1}{\lambda} = 6$$
$$f(x) = \lambda e^{-\lambda x}, x > 0$$
$$= \frac{1}{6} e^{-\frac{1}{6}x}, x > 0$$

(i) $P(X < 2) = \int_0^2 f(x)\,dx$

$$= \int_0^2 \frac{1}{6} e^{-\frac{1}{6}x}\,dx$$

$$= \frac{1}{6} \left| \frac{e^{-\frac{1}{6}x}}{-\frac{1}{6}} \right|_0^2$$
$$= -\left| e^{-\frac{1}{6}x} \right|_0^2$$
$$= -(e^{-\frac{1}{3}} - e^0)$$
$$= -e^{-\frac{1}{3}} + 1$$
$$= 0.2835$$

(ii) $P(4 < X < 5) = \int_4^5 f(x)\,dx$

$$= \int_4^5 \frac{1}{6} e^{-\frac{1}{6}x}\,dx$$
$$= \frac{1}{6} \left| \frac{e^{-\frac{1}{6}x}}{-\frac{1}{6}} \right|_4^5$$
$$= -\left| e^{-\frac{1}{6}x} \right|_4^5$$
$$= -\left(e^{-\frac{5}{6}} - e^{-\frac{2}{3}}\right)$$
$$= 0.0788$$

(iii) $P(X > 10) = \int_{10}^{\infty} f(x)\,dx$

$$= \int_{10}^{\infty} \frac{1}{6} e^{-\frac{1}{6}x}\,dx$$
$$= \frac{1}{6} \left| \frac{e^{-\frac{1}{6}x}}{-\frac{1}{6}} \right|_{10}^{\infty}$$
$$= -\left| e^{-\frac{1}{6}x} \right|_{10}^{\infty}$$

$$= -\left(e^{-\infty} - e^{-\frac{10}{6}}\right)$$
$$= e^{-\frac{10}{6}}$$
$$= 0.1889$$

Example 9

The length of time X to complete a job is exponentially distributed with $E(X) = \mu = \frac{1}{\lambda} = 10$ hours. (i) Compute the probability of job completion between two consecutive jobs exceeding 20 hours. (ii) The cost of job completion is given by $C = 4 + 2X + 2X^2$. Find the expected value of C.

Solution

Let X be a random variable which denotes the length of time to complete a job.

$$E(X) = \mu = \frac{1}{\lambda} = 10$$
$$f(x) = \lambda e^{-\lambda x}$$
$$= \frac{1}{10} e^{-\frac{1}{10}x}$$

(i) $P(X > 20) = \int_{20}^{\infty} f(x)\,dx$

$$= \int_{20}^{\infty} \frac{1}{10} e^{-\frac{1}{10}x}\,dx$$
$$= \frac{1}{10}\left|\frac{e^{-\frac{1}{10}x}}{-\frac{1}{10}}\right|_{20}^{\infty}$$
$$= -\left|e^{-\frac{1}{10}x}\right|_{20}^{\infty}$$
$$= -(e^{-\infty} - e^{-2})$$
$$= e^{-2}$$
$$= 0.1353$$

(ii) For an exponential random variable,

$$E(X) = \mu = \frac{1}{\lambda} = 10$$
$$\text{Var}(X) = \frac{1}{\lambda^2}$$
$$\text{Var}(X) = E(X^2) - \mu^2$$
$$E(X^2) = \text{Var}(X) + \mu^2$$
$$= \frac{1}{\lambda^2} + \frac{1}{\lambda^2}$$
$$= \frac{2}{\lambda^2}$$
$$= 200$$
$$E(C) = E(4 + 2X + 2X^2)$$
$$= E(4) + 2E(X) + 2E(X^2)$$
$$= 4 + 2(10) + 2(200)$$
$$= 424$$

Example 10

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

(i) What is the probability that the repair time exceeds 2 hours?

(ii) What is the conditional probability that a repair takes at least 11 hours given that its direction exceeds 8 hours?

Solution

Let X be the random variable which denotes the time to repair the machine.

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

Scanned with CamScanner

(i) $P(X>2)=\int_2^{\infty} f(x)\,dx$

$$=\int_2^{\infty}\frac{1}{2}e^{-\frac{1}{2}x}\,dx$$

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

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

$$=-(e^{-\infty}-e^{-1})$$

$$=e^{-1}$$

$$=0.3679$$

(ii) $P(X\ge 11/X>9)=P(X>3)$ (By the memoryless property)

$$=\int_3^{\infty} f(x)\,dx$$

$$=\int_3^{\infty}\frac{1}{2}e^{-\frac{1}{2}x}\,dx$$

$$=\frac{1}{2}\left|\frac{e^{-\frac{1}{2}x}}{-\frac{1}{2}}\right|_3^{\infty}$$

$$=-\left|e^{-\frac{1}{2}x}\right|_3^{\infty}$$

$$=-(e^{-\infty}-e^{-1.5})$$

$$=e^{-1.5}$$

$$=0.2231$$

Example 11

The daily consumption of milk in excess of 20000 gallons is approximately exponentially distributed with $\lambda=\frac{1}{3000}$. The city has a daily stock of 35000 gallons. What is the probability that of 2 days selected at random, the stock is insufficient for both the days.

Solution

Let Y be a random variable which denotes the daily consumption of milk consumed in a day. The random variable $X=Y-20000$ has an exponential distribution.

$$\lambda=\frac{1}{3000}$$

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

$$=\frac{1}{3000}e^{-\frac{1}{3000}x},\quad x>0$$

Probability that the stock is insufficient for both days

$$P(Y>35000)=P(X>15000)$$

$$=\int_{15000}^{\infty} f(x)\,dx$$

$$=\int_{15000}^{\infty}\frac{1}{3000}e^{-\frac{1}{3000}x}\,dx$$

$$=\frac{1}{3000}\left|\frac{e^{-\frac{1}{3000}x}}{-\frac{1}{3000}}\right|_{15000}^{\infty}$$

$$=-\left|e^{-\frac{1}{3000}x}\right|_{15000}^{\infty}$$

$$=-(e^{-\infty}-e^{-5})$$

$$=e^{-5}$$

$$=0.0067$$

EXERCISE 5.4

1. If X is exponentially distributed, prove that probability that X exceeds its expected value is less than 0.5.
2. The amount of time that a watch will run without having to be reset is a random variable having an exponential distribution with mean 120 days. Find the probability that such a watch will
   (a) have to be set in less than 24 days.
   (b) not have to be reset in at least 180 days.
   [Ans.: (a) 0.1813, (b) 0.2231]

3. 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 minutes? If a shower has already lasted for 2 minutes, what is the probability that it will last for at least one more minute?

[Ans.: (a) 0.0025, (b) 0.1353]

4. If X is exponentially distributed with parameter λ, find the value of k such that $P(X > k)/P(X \leq k) = a$.

$$\left[\text{Ans.: } \lambda^{-1}\log\left(1+\frac{1}{a}\right)\right]$$

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

[Ans.: Process I is advantageous to the manufacturer]

6. The life of an electronic component follows exponential distribution with a mean of 4 years. The manufacturer of this component gives a replacement warranty of 3 years.
   (a) What proportion of components will be replaced in the period of warranty?
   (b) What is the probability that a randomly selected component will have life within two standard deviations of the mean life?

[Ans.: (a) 0.5276, (b) 0.9502]

5.6 GAMMA DISTRIBUTION

A continuous random variable X is said to follow exponential distribution if its probability function is given by

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

5.6.1 Constants of the Gamma Distribution

1. Mean of the Gamma Distribution

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$$
$$= \int_0^{\infty} x \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}\,dx$$
$$= \frac{\lambda^r}{\Gamma(r)} \int_0^{\infty} e^{-\lambda x} x^r\,dx$$
$$= \frac{\lambda^r}{\Gamma(r)} \frac{\Gamma(r+1)}{\lambda^{r+1}} \qquad \left[\because \int_0^{\infty} e^{-kx} x^{n-1}\,dx = \frac{\Gamma(n)}{k^n}\right]$$
$$= \frac{\lambda^r r\Gamma(r)}{\Gamma(r)\cdot\lambda^{r+1}}$$
$$= \frac{r}{\lambda}$$

2. Variance of the Gamma Distribution

$$\text{Var}(X) = E(X^2) - [E(X)]^2 \qquad \text{...(5.8)}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$
$$= \int_0^{\infty} x^2 \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}\,dx$$
$$= \frac{\lambda^r}{\Gamma(r)} \int_0^{\infty} e^{-\lambda x} x^{r+1}\,dx$$
$$= \frac{\lambda^r}{\Gamma(r)} \frac{\Gamma(r+2)}{\lambda^{r+2}} \qquad \left[\because \int_0^{\infty} e^{-kx} x^{n-1}\,dx = \frac{\Gamma(n)}{k^n}\right]$$
$$= \frac{(r+1)r\Gamma(r)}{\Gamma(r)\cdot\lambda^2}$$
$$= \frac{r^2+r}{\lambda^2}$$

Substituting in Eq. (5.8),

$$\text{Var}(X) = \frac{r^2+r}{\lambda^2} - \frac{r^2}{\lambda^2}$$
$$= \frac{r}{\lambda^2}$$

3. Standard Deviation of the Gamma Distribution

$$\text{SD} = \sqrt{\text{Var}(X)} = \sqrt{\frac{r}{\lambda^2}} = \frac{\sqrt{r}}{\lambda}$$

4. Mode of the Gamma Distribution

Mode is the value of x for which $f(x)$ is maximum.

$$f(x) = \frac{\lambda^r}{\overline{|r}} x^{r-1} e^{-\lambda x}, \quad x > 0$$
$$= 0 \quad , \quad x \le 0$$

Differentiating w.r.t. x,

$$f'(x) = \frac{\lambda^r}{\overline{|r}}\left[(r-1)x^{r-2}e^{-\lambda x} + x^{r-1}e^{-\lambda x}(-\lambda)\right]$$
$$= \frac{\lambda^r}{\overline{|r}} x^{r-2} e^{-\lambda x}\left[(r-1) - \lambda x\right]$$

For maximum value of $f(x)$,

$$f'(x) = 0$$
$$(r-1) - \lambda x = 0$$
$$x = \frac{r-1}{\lambda}$$

Differentiating $f'(x)$ w.r.t. x,

$$f''(x) = \frac{\lambda^r}{\overline{|r}}\Big[(r-2)x^{r-3}e^{-\lambda x}(r-1-\lambda x)$$
$$+ x^{r-2}e^{-\lambda x}(-\lambda)(r-1-\lambda x) + x^{r-2}e^{-\lambda x}(-\lambda)\Big]$$
$$= \frac{\lambda^r}{\overline{|r}} x^{r-3} e^{-\lambda x}\left[(r-2)(r-1-\lambda x) - \lambda x(r-1-\lambda x) - \lambda x\right]$$

Putting $x = \frac{r-1}{\lambda}$,

$$f''(x) = \frac{\lambda^r}{\overline{|r}} x^{r-3} e^{-\lambda x}\left[(r-2)(r-1-r+1) - \lambda x(r-1-r+1) - (r-1)\right]$$
$$= \frac{\lambda^r}{\overline{|r}} x^{r-3} e^{-\lambda x}(1-r)$$

$f(x)$ is maximum when $x = \frac{r-1}{\lambda}$, if $f''(x) < 0$,

$$f''(x) < 0 \text{ if } 1 - r < 0$$
$$1 < r$$
$$\text{or} \quad r > 1$$

Hence, $x = \frac{r-1}{\lambda}$ is the mode of the gamma distribution for $r > 1$.

Example 1

Given a Gamma random variable X with $r = 3$ and $\lambda = 2$. Compute $E(X)$, $\text{Var}(X)$ and $P(X \le 1.5 \text{ years})$.

Solution

$$\lambda = 2, \quad r = 3$$
$$f(x) = \frac{\lambda^r}{\overline{|r}} x^{r-1} e^{-\lambda x}, \quad x > 0$$

(a) $E(X) = \frac{r}{\lambda} = \frac{3}{2} = 1.5$ years

(b) $\text{Var}(X) = \frac{r}{\lambda^2} = \frac{3}{(2)^2} = 0.75$

(c) $P(X \le 1.5 \text{ years}) = \int_0^{1.5} f(x)\,dx$

$$= \int_0^{1.5} \frac{2^3}{\overline{|3}} x^2 e^{-2x}\,dx$$
$$= 4\left|x^2\left(\frac{e^{-2x}}{-2}\right) - 2x\left(\frac{e^{-2x}}{4}\right) + 2\left(\frac{e^{-2x}}{-8}\right)\right|_0^{1.5}$$
$$= 4\left[(1.5)^2\left(\frac{e^{-3}}{-2}\right) - 2(1.5)\left(\frac{e^{-3}}{4}\right) + 2\left(\frac{e^{-3}}{-8}\right) + \frac{1}{4}\right]$$
$$= 0.5768$$

Example 2

The daily consumption of milk in a city, in excess 20000 litres, is approximately distributed as a Gamma variate with parameters $\lambda = \frac{1}{10000}$

and $r = 2$. The city has a daily stock of 30000 litres. What is the probability that the stock is insufficient on a particular day?

Solution

Let Y be the random variable which denotes the daily consumption of milk (in litres) in a city. The random variable $X = Y - 20000$ has a gamma distribution.

$$\lambda = \frac{1}{10000}, r = 2$$

$$f(x) = \frac{\lambda^r}{\overline{|r}} x^{r-1} e^{-\lambda x}, \quad x > 0$$

$$= \frac{\left(\frac{1}{10000}\right)^2}{\overline{|2}} x^{2-1} e^{-\frac{1}{10000}x}$$

$$= \frac{xe^{-\frac{1}{1000}x}}{(10000)^2}$$

Probability that the stock is insufficient on a particular day

$$P(Y > 30000) = P(X > 10000)$$

$$= \int_{10000}^{\infty} f(x)dx$$

$$= \int_{1000}^{\infty} \frac{xe^{-\frac{1}{1000}x}}{(10000)^2} dx$$

$$= \frac{1}{10^8} \int_{10^4}^{\infty} xe^{-10^{-4}x} dx$$

$$= \frac{1}{10^8} \left| \frac{x \cdot e^{-10^{-4}x}}{-10^{-4}} - \frac{1 \cdot e^{-10^{-4}x}}{(-10^{-4})^2} \right|_{10^4}^{\infty}$$

$$= \frac{1}{10^8} \left(\frac{e^{-1}}{10^{-8}} + \frac{e^{-1}}{10^{-8}} \right)$$

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

$$= 2e^{-1}$$

$$= 0.7358$$

Example 3

In a certain city, the daily consumption of electric power in millions of kilowatt hours can be treated as a random variable having gamma distribution with parameters $\lambda = \frac{1}{2}$ and $r = 3$. 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.

Solution

Let X be a random variable which denotes the daily consumption of electric power in millions killowatt-hours.

$$f(x) = \frac{\lambda^r}{\overline{|r}} x^{r-1} e^{-\lambda x}, \quad x > 0$$

$$= \frac{\left(\frac{1}{2}\right)^3}{\overline{|3}} x^2 e^{-\frac{1}{2}x}$$

$$P(\text{power supply is inadequate}) = P(X > 12)$$

$$= \int_{12}^{\infty} f(x)dx$$

$$= \int_{12}^{\infty} \frac{1}{\overline{|3}} \frac{1}{2^3} x^2 e^{-\frac{1}{2}x} dx$$

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

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

$$= 25e^{-6}$$

$$= 0.062$$

Example 4

If a company employs n sales persons, its gross sales in thousands of rupees may be regarded as a random variable having a gamma distribution with $\lambda = \frac{1}{2}$ and $r = 80\sqrt{n}$. If the sales cost is ₹8000 per *salesperson, how many salespersons should the company employ to maximise the expected profit?*

Solution

Let X be the random variable which denotes the gross sales in rupees by n salespersons.

$$\lambda = \frac{1}{2}, \quad r = 80000\sqrt{n}$$

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

$$E(X) = \frac{r}{\lambda} = \frac{80000\sqrt{n}}{\frac{1}{2}} = 160000\sqrt{n}$$

If y denotes the total expected profit of the company,

$$\begin{aligned} y &= \text{total expected sales} - \text{total sales cost} \\ &= 160000\sqrt{n} - 8000n \end{aligned}$$

$$\frac{dy}{dn} = \frac{80000}{\sqrt{n}} - 8000$$

For maximum profits,

$$\frac{dy}{dn} = 0$$

$$\frac{80000}{\sqrt{n}} - 8000 = 0$$

$$\frac{80000}{\sqrt{n}} = 8000$$

$$\sqrt{n} = 10$$

$$n = 100$$

$$\frac{d^2y}{dn^2} = -\frac{40000}{n^{\frac{3}{2}}}$$

When $n = 100$, $\frac{d^2y}{dn^2} = -40 < 0$

$\therefore$ y is maximum when $n = 100$.

Hence, the company should employ 100 salespersons to maximise the expected profit.

Example 5

Consumer demand for milk in a certain locality, per month, is known to be a general gamma random variable. If the average demand is 'a' litres and the most likely demand is 'b' litres (b < 0), what is the variance of the demand?

Solution

Let X be the random variable which denotes 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 probability density function is maximum.

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

$$f'(x) = \frac{\lambda^r}{\lceil r}\left[(r-1)x^{r-2}e^{-\lambda x} - \lambda x^{r-1}e^{-\lambda x}\right]$$

$$= \frac{\lambda^r}{\lceil r} x^{r-2} e^{-\lambda x}\left[(r-1) - \lambda x\right]$$

For maximum value of $f(x)$,

$$f'(x) = 0$$

$$(r-1) - \lambda x = 0$$

$$x = \frac{r-1}{\lambda}$$

Differentiating $f'(x)$ w.r.t. x,

$$f''(x) = \frac{\lambda^r}{\lceil r}\left[-\lambda x^{r-2}e^{-\lambda x} + \{(r-1) - \lambda x\}\frac{d}{dx}\left[x^{r-2}e^{-\lambda x}\right]\right]$$

$$= \frac{\lambda^r}{\lceil r} x^{r-3} e^{-\lambda x}(1-r)$$

$f''(x) < 0$ when $x = \frac{r-1}{\lambda}$

$f(x)$ is maximum when $x = \frac{r-1}{\lambda}$ if $f''(x) < 0$

$$f''(x) < 0 \text{ if } 1 - r < 0$$

$$1 < r$$

or $$r > 1$$

Most likely demand $= \dfrac{r-1}{\lambda} = b, r > 1$

$$\frac{r-1}{\lambda} = b$$

$$\frac{r}{\lambda} = b + \frac{1}{\lambda} \quad \text{...(1)}$$

Average demand $= E(X) = \dfrac{r}{\lambda} = a$ ...(2)

Putting in Eq. (1),

$$a = b + \frac{1}{\lambda}$$

$$\frac{1}{\lambda} = a - b \quad \text{...(3)}$$

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

$$= a(a-b) \qquad \text{[from Eq. (2) and (3)]}$$

EXERCISE 5.5

1. Find the probabilities that the value of a random variable will exceed 4, if it has gamma distribution with

   (a) $\lambda = \dfrac{1}{3}, r = 2$ (b) $\lambda = \dfrac{1}{4}, r = 3$

   [Ans.: (a) 0.5551 (b) 4]

2. If X follows the gamma distribution with parameter λ and r, prove that the expected value of the positive square root of X is $\dfrac{\Gamma\left(r+\frac{1}{2}\right)}{\sqrt{\lambda}\,\Gamma(r)}$.

3. 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.