
UNIT 6 DISCRETE PROBABILITY DISTRIBUTIONS

Structure

- 6.0 Introduction
- 6.1 Objectives
- 6.2 Bernoulli Distribution and its Properties
- 6.3 Binomial Probability Function
- 6.4 Poisson Distribution
- 6.5 Discrete Uniform Distribution
- 6.6 Hypergeometric Distribution
- 6.7 Summary
- 6.8 Solutions/Answers
- 6.9 References
- Appendix 1

6.0 INTRODUCTION

In Unit 5 of the Course, you have studied random variables, their probability functions and distribution functions. You have come to know how the expectations of random variables are obtained. Unit 5 also discussed the definitions and properties of general discrete and probability distributions.

The present unit is devoted to the study of some special discrete distributions, including Bernoulli, Binomial, Poisson, Uniform and Hypergeometric distributions. First, this unit defines the Bernoulli distribution and its properties.

Binomial distribution and its applications are covered next, which is applied in the cases where the probability of success and failure do not differ much from each other and the number of trials in a random experiment is finite. However, there may be practical situations where the probability of success is very small, that is, there may be situations where the event occurs rarely and the number of trials may not be known. For instance, the number of accidents occurring at a particular spot on a road everyday is a rare event. For such rare events, we cannot apply the binomial distribution. To these situations, we apply Poisson distribution. The concept of Poisson distribution was developed by a French mathematician, Simeon Denis Poisson (1781-1840) in the year 1837. Poisson distribution is also discussed in this unit.

Discrete uniform distribution is applicable to those experiments where the different values of random variable are equally likely. If the population is finite and the sampling is done without replacement i.e. if the events are random but not independent, then we use Hypergeometric distribution. This unit discusses discrete uniform distribution and hypergeometric distribution.

6.1 OBJECTIVES

Study of the present unit will enable you to:

- define the Bernoulli distribution and to establish its properties;
- define the binomial distribution and establish its properties;

- identify the situations where these distributions are applied;
- know the situations where Poisson distribution is applied;
- define and explain Poisson distribution;
- define the discrete uniform and hypergeometric distributions;
- compute mean and variance for different discrete distributions.

6.2 BERNOULLI DISTRIBUTION AND ITS PROPERTIES

There are experiments where the outcomes can be divided into two categories with reference to presence or absence of a particular attribute or characteristic. A convenient method of representing the two is to designate either of them as success and the other as failure. For example, head coming up in the toss of a fair coin may be treated as a success and tail as failure, or vice-versa. Accordingly, probabilities can be assigned to the success and failure.

Suppose a piece of a product is tested which may be defective (failure) or non-defective (a success). Let p be the probability that it found non-defective and $q = 1 - p$ be the probability that it is defective. Let X be a random variable such that it takes value 1 when success occurs and 0 if failure occurs.

Therefore,

$$P[X = 1] = p, \text{ and}$$

$$P[X = 0] = q = 1 - p.$$

The above experiment is a Bernoulli trial, the r.v. X defined in the above experiment is a Bernoulli variate and the probability distribution of X as specified above is called the Bernoulli distribution in honour of J. Bernoulli (1654-1705).

Definition

A discrete random variable X is said to follow Bernoulli distribution with parameter p if its probability mass function is given by

$$P[X = x] = \begin{cases} p^x (1-p)^{1-x} & ; x = 0, 1 \\ 0 & ; \text{elsewhere} \end{cases}$$

$$\text{i.e. } P[X = 1] = p^1 (1-p)^{1-1} = p \quad [\text{putting } x = 1]$$

$$\text{and } P[X = 0] = p^0 (1-p)^{1-0} = 1-p \quad [\text{putting } x = 0]$$

The Bernoulli probability distribution, in tabular form, is given as

X	0	1
$p(x)$	$1-p$	p

Remark 1: The Bernoulli distribution is useful whenever a random experiment has only two possible outcomes, which may be labelled as success and failure.

Moments of Bernoulli Distribution

The r^{th} moment about origin of a Bernoulli variate X is given as

$$\mu'_r = E(X^r)$$

$$\begin{aligned}
 &= \sum_{x=0}^1 x^r p(x) && [\text{See Unit 5 of this course}] \\
 &= (0)^r p(0) + (1)^r p(1) \\
 &= (0)(1-p) + (1)p \\
 &= p
 \end{aligned}$$

$$\Rightarrow \mu'_1 = p, \mu'_2 = p, \mu'_3 = p, \mu'_4 = p.$$

Hence,

$$\text{Mean} = \mu'_1 = p,$$

$$\text{Variance } (\mu_2) = \mu'_2 - (\mu'_1)^2 = p - p^2 = p(1-p),$$

Example 1: Let X be a random variable having Bernoulli distribution with paramete $p = 0.4$. Find its mean and variance.

Solution:

$$\text{Mean} = p = 0.4,$$

$$\text{Variance} = p(1 - p) = (0.4)(1 - 0.4) = (0.4)(0.6) = 0.24$$

Single trial is taken into consideration in Bernoulli distribution. But, if trials are performed repeatedly a finite number of times and we are interested in the distribution of the sum of independent Bernoulli trials with the same probability of success in each trial, then we need to study binomial distribution which has been discussed in the next section.

6.3 BINOMIAL PROBABILITY FUNCTION

Here, in this section, we will discuss binomial distribution which was discovered by J. Bernoulli (1654-1705) and was first published eight years after his death i.e. in 1713 and is also known as “Bernoulli distribution for n trials”. Binomial distribution is applicable for a random experiment comprising a finite number (n) of independent Bernoulli trials having the constant probability of success for each trial.

Before defining binomial distribution, let us consider the following example: Suppose a man fires 3 times independently to hit a target. Let p be the probability of hitting the target (success) for each trial and $q (= 1 - p)$ be the probability of his failure.

Let S denote the success and F the failure. Let X be the number of successes in 3 trials,

$$\begin{aligned}
 P[X = 0] &= \text{Probability that target is not hit at all in any trial} \\
 &= P[\text{Failure in each of the three trials}] \\
 &= P(F \cap F \cap F) \\
 &= P(F).P(F).P(F) \quad [\because \text{trials are independent}] \\
 &= q.q.q \\
 &= q^3
 \end{aligned}$$

This can be written as

$$P[X = 0] = {}^3C_0 p^0 q^{3-0}$$

$$[\because {}^3C_0 = 1, p^0 = 1, q^{3-0} = q^3. \text{ Recall } {}^nC_x = \frac{n!}{x!(n-x)!} \text{ (see Unit 4 of MST-001)}]$$

$P[X = 1]$ = Probability of hitting the target once

= [(Success in the first trial and failure in the second and third trial) or
(success in the second trial and failure in the first and third trials)
or (success in the third trial and failure in the first two trials)]

$$= P[(S \cap F \cap F) \text{ or } (F \cap S \cap F) \text{ or } (F \cap F \cap S)]$$

$$= P(S \cap F \cap F) + P(F \cap S \cap F) + P(F \cap F \cap S)$$

$$= P(S).P(F).P(F) + P(F).P(S).P(F) + P(F).P(F).P(S)$$

[$\because$ trials are independent]

$$= p.q.q + q.p.q + q.q.p$$

$$= pq^2 + pq^2 + pq^2$$

$$= 3pq^2$$

This can also be written as

$$P[X = 1] = {}^3C_1 p^1 q^{3-1} \quad [\because {}^3C_1 = 3, p^1 = p, q^{3-1} = q^2]$$

$P[X = 2]$ = Probability of hitting the target twice

= P[(Success in each of the first two trials and failure in the third trial) or (Success in first and third trial and failure in the second trial) or (Success in the last two trials and failure in the first trial)]

$$= P[(S \cap S \cap F) \cup (S \cap F \cap S) \cup (F \cap S \cap S)]$$

$$= P[S \cap S \cap F] + P[S \cap F \cap S] + P[F \cap S \cap S]$$

$$= P(S).P(S).P(F) + P(S).P(F).P(S) + P(F).P(S).P(S)$$

$$= p.p.q + p.q.p + q.p.p$$

$$= 3p^2q$$

This can also be written as

$$P[X = 2] = {}^3C_2 p^2 q^{3-2} \quad [\because {}^3C_2 = 3, q^{3-2} = q]$$

$P[X = 3]$ = Probability of hitting the target thrice

= [Success in each of the three trials]

$$= P[S \cap S \cap S]$$

$$= P(S).P(S).P(S)$$

$$= p.p.p$$

$$= p^3$$

This can also be written as

$$P[X = 3] = {}^3C_3 p^3 q^{3-3} \quad [\because {}^3C_3 = 1, q^{3-3} = 1]$$

From the above four enrectangled results, we can write

$$P[X = r] = {}^3C_r p^r q^{3-r}; r = 0, 1, 2, 3.$$

which is the probability of r successes in 3 trials. 3C_r , here, is the number of ways in which r successes can happen in 3 trials.

The result can be generalized for n trials in the similar fashion and is given as

$$P[X = r] = {}^nC_r p^r q^{n-r}; r = 0, 1, 2, \dots, n.$$

This distribution is called the binomial probability distribution. The reason behind giving the name binomial probability distribution for this probability distribution is that the probabilities for $x = 0, 1, 2, \dots, n$ are the respective probabilities ${}^nC_0 p^0 q^{n-0}, {}^nC_1 p^1 q^{n-1}, \dots, {}^nC_n p^n q^{n-n}$ which are the successive terms of the binomial expansion $(q + p)^n$.

$$[\because (q + p)^n = {}^nC_0 q^n p^0 + {}^nC_1 q^{n-1} p^1 + \dots + {}^nC_n q^0 p^n]$$

Binomial Expansion:

‘Bi’ means ‘Two’. ‘Binomial expansion’ means ‘Expansion of expression having two terms, e.g.

$$(X + Y)^2 = X^2 + 2XY + Y^2 = {}^2C_0 X^2 Y^0 + {}^2C_1 X^{2-1} Y^1 + {}^2C_2 X^{2-2} Y^2,$$

$$(X + Y)^3 = X^3 + 3X^2 Y + 3XY^2 + Y^3$$

$$= {}^3C_0 X^3 Y^0 + {}^3C_1 X^{3-1} Y^1 + {}^3C_2 X^{3-2} Y^2 + {}^3C_3 X^{3-3} Y^3$$

So, in general,

$$(X + Y)^n = {}^nC_0 X^n Y^0 + {}^nC_1 X^{n-1} Y^1 + {}^nC_2 X^{n-2} Y^2 + \dots + {}^nC_n X^{n-n} Y^n$$

The above discussion leads to the following definition.

Definition:

A discrete random variable X is said to follow binomial distribution with parameters n and p if it assumes only a finite number of non-negative integer values and its probability mass function is given by

$$P[X = x] = \begin{cases} {}^nC_x p^x q^{n-x}; & x = 0, 1, 2, \dots, n \\ 0; & \text{elsewhere} \end{cases}$$

where, n is the number of independent trials,

x is the number of successes in n trials,

p is the probability of success in each trial, and

$q = 1 - p$ is the probability of failure in each trial.

Remark 2:

- i) The binomial distribution is the probability distribution of sum of n independent Bernoulli variates.
- ii) If X is binomially distributed r.v. with parameters n and p , then we may write it as $X \sim B(n, p)$.
- iii) If X and Y are two binomially distributed independent random variables with parameters (n_1, p) and (n_2, p) respectively then their sum also follows a binomial distribution with parameters $n_1 + n_2$ and p . But, if the probability of success is not the same for the two random variables, then this property does not hold.

Example 2: An unbiased coin is tossed six times. Find the probability of obtaining

- (i) exactly 3 heads
- (ii) less than 3 heads

- (iii) more than 3 heads
- (iv) at most 3 heads
- (v) at least 3 heads
- (vi) more than 6 heads

Solution: Let p be the probability of getting head (success) in a toss of the coin and n be the number of trials.

$$\therefore n = 6, p = \frac{1}{2} \text{ and hence } q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}.$$

Let X be the number of successes in n trials,

$\therefore$ by binomial distribution, we have

$$P[X = x] = {}^nC_x p^x q^{n-x}; x = 0, 1, 2, \dots, n$$

$$= {}^6C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{6-x}; x = 0, 1, 2, \dots, 6$$

$$= {}^6C_x \left(\frac{1}{2}\right)^6; x = 0, 1, 2, \dots, 6.$$

$$= \frac{1}{64} \cdot {}^6C_x; x = 0, 1, 2, \dots, 6.$$

Therefore,

$$(i) P[\text{exactly 3 heads}] = P[X = 3]$$

$$= \frac{1}{64} ({}^6C_3) = \frac{1}{64} \left[\frac{6 \times 5 \times 4}{3 \times 2} \right] = \frac{5}{16}$$

$$[\because \text{Recall } {}^nC_x = \frac{n!}{x!(n-x)!} \text{ (see Unit 4 of MST-001)}]$$

$$(ii) P[\text{less than 3 heads}] = P[X < 3]$$

$$= P[X = 2 \text{ or } X = 1 \text{ or } X = 0]$$

$$= P[X = 2] + P[X = 1] + P[X = 0]$$

$$= \frac{1}{64} \cdot {}^6C_2 + \frac{1}{64} \cdot {}^6C_1 + \frac{1}{64} \cdot {}^6C_0$$

$$= \frac{1}{64} [{}^6C_2 + {}^6C_1 + {}^6C_0] = \frac{1}{64} \left[\frac{6 \times 5}{2} + 6 + 1 \right]$$

$$= \frac{22}{64} = \frac{11}{32}.$$

$$(iii) P[\text{more than 3 heads}] = P[X > 3]$$

$$= P[X = 4 \text{ or } X = 5 \text{ or } X = 6] \left[\because \text{in 6 trials one can have at most 6 heads} \right]$$

$$= P[X = 4] + P[X = 5] + P[X = 6]$$

$$= \frac{1}{64} \cdot {}^6C_4 + \frac{1}{64} \cdot {}^6C_5 + \frac{1}{64} \cdot {}^6C_6$$

$$= \frac{1}{64} [{}^6C_4 + {}^6C_5 + {}^6C_6]$$

$$= \frac{1}{64} \left[\frac{6 \times 5}{2} + 6 + 1 \right] = \frac{22}{64} = \frac{11}{32}.$$

(iv) $P[\text{at most 3 heads}] = P[3 \text{ or less than 3 heads}]$

$$\begin{aligned} &= P[X = 3] + P[X = 2] + P[X = 1] + P[X = 0] \\ &= \frac{1}{64} \cdot {}^6C_3 + \frac{1}{64} \cdot {}^6C_2 + \frac{1}{64} \cdot {}^6C_1 + \frac{1}{64} \cdot {}^6C_0 \\ &= \frac{1}{64} [{}^6C_3 + {}^6C_2 + {}^6C_1 + {}^6C_0] \\ &= \frac{1}{64} [20 + 15 + 6 + 1] = \frac{42}{64} = \frac{21}{32}. \end{aligned}$$

(v) $P[\text{at least 3 heads}] = P[3 \text{ or more heads}]$

$$\begin{aligned} &= P[X = 3] + P[X = 4] + P[X = 5] + P[X = 6] \\ &\quad \text{or} \\ &= 1 - (P[X = 0] + P[X = 1] + P[X = 2]) \\ &\quad \left[\because \text{sum of probabilities of all possible} \right. \\ &\quad \left. \text{values of a random variable is 1} \right] \\ &= 1 - \left(\frac{11}{32} \right) \quad \left[\text{Already obtained in} \right. \\ &\quad \left. \text{part (ii) of this example} \right] \\ &= \frac{21}{32}. \end{aligned}$$

(vi) $P[\text{more than 6 heads}] = P[7 \text{ or more heads}]$

$$\begin{aligned} &= P[\text{an impossible event}] \quad \left[\because \text{in six tosses, it} \right. \\ &\quad \left. \text{is impossible to get} \right. \\ &\quad \left. \text{more than six heads} \right] \\ &= 0 \end{aligned}$$

Example 3: The chances of catching cold by workers working in an ice factory during winter are 25%. What is the probability that out of 5 workers 4 or more will catch cold?

Solution: Let catching cold be the success and p be the probability of success for each worker.

$\therefore$ Here, $n = 5$, $p = 0.25$, $q = 0.75$ and by binomial distribution

$$\begin{aligned} P[X = x] &= {}^nC_x p^x q^{n-x}; x = 0, 1, 2, \dots, n \\ &= {}^5C_x (0.25)^x (0.75)^{5-x}; 0, 1, 2, \dots, 5 \end{aligned}$$

Therefore, the required probability $= P[X \geq 4]$

$$\begin{aligned} &= p[X = 4 \text{ or } X = 5] \\ &= P[X = 4] + P[X = 5] \\ &= {}^5C_4 (0.25)^4 (0.75)^1 + {}^5C_5 (0.25)^5 (0.75)^0 \\ &= (5)(0.002930) + (1)(0.000977) \\ &= 0.014650 + 0.000977 \\ &= 0.015627 \end{aligned}$$

Example 4: Let X and Y be two independent random variables such that $X \sim B(4, 0.7)$ and $Y \sim B(3, 0.7)$. Find $P[X + Y \leq 1]$.

Solution: We know that if X and Y are independent random variables each following binomial distribution with parameters (n_1, p) and (n_2, p) , then $X + Y \sim B(n_1 + n_2, p)$.

Therefore, here $X + Y$ follows binomial distribution with parameters $4 + 3$ and 0.7 , i.e. 7 and 0.7 . So, here, $n = 7$ and $p = 0.7$.

$$\begin{aligned}\text{Thus, the required probability} &= P[X + Y \leq 1] \\ &= P[X + Y = 1] + P[X + Y = 0] \\ &= {}^7C_1 (0.7)^1 (0.3)^6 + {}^7C_0 (0.7)^0 (0.3)^7 \\ &= 7(0.7)(0.000729) + 1(1)(0.0002187) \\ &= 0.0035721 + 0.0002187 \\ &= 0.0037908\end{aligned}$$

Moments of Binomial Distribution

The r^{th} order moment about origin of a binomial variate X is given as

$$\mu'_r = E(X^r) = \sum_{x=0}^n x^r \cdot P[X = x]$$

Using this moment function the following useful results are drawn for the binomial distribution:

$$\begin{aligned}\therefore \text{Mean} &= \text{First order moment about origin} \\ &= \mu'_1 \\ &= np.\end{aligned}$$

Mean = np

$$\begin{aligned}\therefore \text{Variance } (\mu_2) &= \mu'_2 - (\mu'_1)^2 \\ &= n^2 p^2 - n p^2 + np - (np)^2 \\ &= n^2 p^2 - np^2 + np - n^2 p^2 \\ &= np - np^2 \\ &= np(1 - p) \\ &= npq\end{aligned}$$

$$\therefore \text{Variance} = npq$$

Remark 3:

- (i) As $0 < q < 1$
 $\Rightarrow q < 1$ [Multiplying both sides by $np > 0$]
 $\Rightarrow npq < np$
 $\Rightarrow \text{Variance} < \text{Mean}$
Hence, for binomial distribution
Mean $>$ Variance
- (ii) As variance of $X \sim B(n, p)$ is npq ,
 $\therefore$ its standard deviation is $\sqrt{npq}$.

Example 5: For a binomial distribution with $p = \frac{1}{4}$ and $n = 10$, find mean and variance.

Solution: As $p = \frac{1}{4}$, $\therefore q = 1 - \frac{1}{4} = \frac{3}{4}$.

$$\text{Mean} = np = 10 \times \frac{1}{4} = \frac{5}{2},$$

$$\text{Variance} = npq = 10 \times \frac{1}{4} \times \frac{3}{4} = \frac{15}{8}.$$

Example 6: The mean and standard deviation of binomial distribution are 4 and $\frac{2}{\sqrt{3}}$ respectively. Find $P[X \geq 1]$.

Solution: Let $X \sim B(n, p)$, then

$$\text{Mean} = np = 4$$

and variance $= npq = \left(\frac{2}{\sqrt{3}}\right)^2$ [$\because$ S.D. $= \frac{2}{\sqrt{3}}$ and variance is square of S.D.]

Dividing the second equation by the first equation, we have

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

$$\Rightarrow q = \frac{1}{3}$$

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

Putting $p = \frac{2}{3}$ in the equation of mean, we have

$$n\left(\frac{2}{3}\right) = 4 \Rightarrow n = 6$$

$\therefore$ by binomial distribution,

$$\begin{aligned} P[X = x] &= {}^nC_x p^x q^{n-x} \\ &= {}^6C_x \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{6-x}; \quad x = 0, 1, 2, \dots, 6. \end{aligned}$$

Thus, the required probability

$$\begin{aligned} P[X \geq 1] &= P[X = 1] + P[X = 2] + P[X = 3] + \dots + P[X = 6] \\ &= 1 - P[X = 0] \\ &= 1 - {}^6C_0 \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^{6-0} = 1 - (1)(1) \frac{1}{729} = \frac{728}{729}. \end{aligned}$$

Check Your Progress 1

- 1) The probability of a man hitting a target is $\frac{1}{4}$. He fires 5 times. What is the probability of his hitting the target at least twice?

.....
.....

- 2) A policeman fires 6 bullets on a dacoit. The probability that the dacoit will be killed by a bullet is 0.6. What is the probability that the dacoit is still alive?

.....
.....
.....
.....

- 3) Comment on the following:

The mean of a binomial distribution is 3 and variance is 4.

.....
.....

- 4) Find the binomial distribution when sum of mean and variance of 5 trials is 4.8.

.....
.....
.....
.....

- 5) The mean of a binomial distribution is 30 and standard deviation is 5. Find the values of n , p and q .

.....
.....

6.4 POISSON DISTRIBUTION

In case of binomial distributions, as discussed in the last section, we deal with events whose occurrences and non-occurrences are almost equally important. However, there may be events which do not occur as outcomes of a definite number of trials of an experiment but occur rarely at random points of time and for such events our interest lies only in the number of occurrences and not in its non-occurrences. Examples of such events are:

- Our interest may lie in how many printing mistakes are there on each page of a book but we are not interested in counting the number of words without any printing mistake.
- In production where control of quality is the major concern, it often requires counting the number of defects (and not the non-defects) per item.
- One may intend to know the number of accidents during a particular time interval.

Under such situations, binomial distribution cannot be applied as the value of n is not definite and the probability of occurrence is very small. Other such situations can be thought of yourself. Poisson distribution discovered by S.D. Poisson (1781-1840) in 1837 can be applied to study these situations.

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

- n , the number of trials is indefinitely large, i.e. $n \rightarrow \infty$.
- p , the constant probability of success for each trial is very small, i.e. $p \rightarrow 0$.
- np is a finite quantity say ' λ '.

Definition: A random variable X is said to follow Poisson distribution if it assumes indefinite number of non-negative integer values and its probability mass function is given by:

$$p(x) = P(X = x) = \begin{cases} \frac{e^{-\lambda} \cdot \lambda^x}{x!}; & x = 0, 1, 2, 3, \dots \text{ and } \lambda > 0. \\ 0; & \text{elsewhere} \end{cases}$$

where e = base of natural logarithm, whose value is approximately equal to 2.7183 corrected to four decimal places. Value of $e^{-\lambda}$ can be written from the table given in the Appendix 1 at the end of this unit, or, can be seen from the reference.

Remark 4

- i) If X follows Poisson distribution with parameter λ then we shall use the notation $X \sim P(\lambda)$.
- ii) If X and Y are two independent Poisson variates with parameters λ_1 and λ_2 respectively, then $X + Y$ is also a Poisson variate with parameter $\lambda_1 + \lambda_2$. This is known as **additive property of Poisson distribution**.

Moments Of Poisson Distribution

r^{th} order moment about origin of Poisson variate is

$$\mu'_r = E(X^r) = \sum_{x=0}^{\infty} x^r p(x) = \sum_{x=0}^{\infty} x^r \frac{e^{-\lambda} \lambda^x}{x!}$$

$$\therefore \text{Mean} = \lambda$$

$$\begin{aligned} \therefore \text{Variance of } X \text{ is given as } V(X) &= \mu_2 = \mu'_2 - (\mu'_1)^2 \\ &= \lambda^2 + \lambda - (\lambda)^2 \\ &= \lambda \end{aligned}$$

Remark 5

- i) Mean and variance of Poisson distribution are always equal. In fact, this is the only discrete distribution for which Mean = Variance = the third central moment.
- ii) Moments of the Poisson distribution can be deduced from those of the binomial distribution also as explained below:

For a binomial distribution,

$$\text{Mean} = np$$

$$\text{Variance} = npq$$

Now as the Poisson distribution is a limiting form of binomial distribution under the conditions:

$$(i) n \rightarrow \infty, (ii) p \rightarrow 0 \text{ i.e. } q \rightarrow 1, \text{ and } (iii) np = \lambda \text{ (a finite quantity);}$$

$\therefore$ Mean, Variance and other moments of the Poisson distribution are given as:

$$\text{Mean} = \text{Limiting value of } np = \lambda$$

$$\begin{aligned}\text{Variance} &= \text{Limiting value of } npq \\ &= \text{Limiting value of } (np) (q) \\ &= (\lambda) (1) = \lambda\end{aligned}$$

Now, let's give some examples of the Poisson distribution.

Example 7: It is known that the number of heavy trucks arriving at a railway station follows the Poisson distribution. If the average number of truck arrivals during a specified period of an hour is 2, find the probabilities that during a given hour

- no heavy truck arrive,
- at least two trucks will arrive.

Solution: Here, the average number of truck arrivals is 2

i.e. mean = 2

$$\Rightarrow \lambda = 2$$

Let X be the number of trucks arrive during a given hour,

$\therefore$ by Poisson distribution, we have

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!} = \frac{e^{-2} (2)^x}{x!}; x = 0, 1, 2, \dots$$

Thus, the desired probabilities are:

$$(a) P[\text{arrival of no heavy truck}] = P[X = 0]$$

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

$$= e^{-2}$$

$$= 0.1353$$

[See the table given
in the Appendix at
the end of this unit]

$$(b) P[\text{arrival of at least two trucks}] = P[X \geq 2]$$

$$= P[X = 2] + P[X = 3] + \dots$$

$$= 1 - [P[X = 1] + P[X = 0]]$$

[$\because$ sum of all the
probabilities is 1]

$$= 1 - \left[\frac{e^{-2} 2^0}{0!} + \frac{e^{-2} 2^1}{1!} \right]$$

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

$$= 1 - (0.1353)(3) = 1 - 0.4059 = 0.5941$$

Note: In most of the cases for Poisson distribution, if we are to compute the probabilities of the type $P[X > a]$ or $P[X \geq a]$, we write them as $P[X > a] = 1 - P[X \leq a]$ and $P[X \geq a] = 1 - P[X < a]$, because n may not be definite and hence we cannot go up to the last value and hence the probability is written in terms of its complementary probability.

Example 8: If the probability that an individual suffers a bad reaction from an injection of a given serum is 0.001, determine the probability that out of 500 individuals

- i) exactly 3,
- ii) more than 2

individuals suffer from bad reaction

Solution: Let X be the Poisson variate, "Number of individuals suffering from bad reaction". Then,

$$n = 1500, p = 0.001,$$

$$\therefore \lambda = np = (1500)(0.001) = 1.5$$

$\therefore$ By Poisson distribution,

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots$$

$$= \frac{e^{-1.5} \cdot (1.5)^x}{x!}; x = 0, 1, 2, \dots$$

Thus,

- i) The desired probability = $P[X = 3]$

$$= \frac{e^{-1.5} \cdot (1.5)^3}{3!}$$

$$= \frac{(0.2231)(3.375)}{6} = 0.1255$$

$$\left[\begin{array}{l} \because e^{-0.5} = 0.6065, e^{-1} = 0.3679, \text{so} \\ e^{-1.5} = e^{-1} \times e^{-0.5} = (0.3679)(0.6065) = 0.2231 \\ \text{See the table given in the Appendix} \\ \text{at the end of this unit} \end{array} \right]$$

- ii) The desired probability = $P[X > 2]$

$$= 1 - P[X \leq 2]$$

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

$$= 1 - \left[\frac{e^{-1.5} \cdot (1.5)^2}{2!} + \frac{e^{-1.5} \cdot (1.5)^1}{1!} + \frac{e^{-1.5} \cdot (1.5)^0}{0!} \right]$$

$$= 1 - e^{-1.5} \left[\frac{2.25}{2} + 1.5 + 1 \right] = 1 - (3.625)e^{-1.5}$$

$$= 1 - (3.625)(0.2231) = 1 - 0.8087 = 0.1913$$

Example 9: If the mean of a Poisson distribution is 1.44, find the values of variance and the central moments of order 3 and 4.

Solution: Here, mean = 1.44

$$\Rightarrow \lambda = 1.44$$

Hence, Variance = $\lambda = 1.44$

$$\mu_3 = \lambda = 1.44$$

$$\mu_4 = 3\lambda^2 + \lambda = 3(1.44)^2 + 1.44 = 7.66.$$

Example 10: If a Poisson variate X is such that $P[X = 1] = 2P[X = 2]$, find the mean and variance of the distribution.

Solution: Let λ be the mean of the distribution, hence by Poisson distribution,

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}; x = 0, 1, 2, \dots$$

$$\text{Now, } P[X = 1] = 2P[X = 2]$$

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

$$\Rightarrow \lambda = \lambda^2 \Rightarrow \lambda^2 - \lambda = 0 \Rightarrow \lambda(\lambda - 1) = 0 \Rightarrow \lambda = 0, 1$$

But $\lambda = 0$ is rejected

[$\because$ if $\lambda = 0$ then either $n = 0$ or $p = 0$ which implies that Poisson distribution does not exist in this case.]

$$\therefore \lambda = 1$$

Hence mean = $\lambda = 1$, and

Variance = $\lambda = 1$.

Check Your Progress 2

- 1) Assume that the chance of an individual coal miner being killed in a mine accident during a year is $\frac{1}{1400}$. Use the Poisson distribution to calculate the probability that in a mine employing 350 miners, there will be at least one fatal accident in a year. (use $e^{-0.25} = 0.78$)

.....

.....

- 2) The mean and standard deviation of a Poisson distribution are 6 and 2 respectively. Test the validity of this statement.

.....

.....

- 3) For a Poisson distribution, it is given that $P[X = 1] = P[X = 2]$, find the value of the mean of the distribution. Hence find $P[X = 0]$ and $P[X = 4]$.

.....

.....

.....

.....

6.5 DISCRETE UNIFORM DISTRIBUTION

Discrete uniform distribution can be conceived in practice if under the given experimental conditions, the different values of the random variable are equally likely. For example, the number on an unbiased die when thrown may be 1 or 2 or 3 or 4 or 5 or 6. These values of random variable, “the number on an unbiased die when thrown” are equally likely and for such an experiment, the discrete uniform distribution is appropriate.

Definition: A random variable X is said to have a discrete uniform (rectangular) distribution if it takes any positive integer value from 1 to n , and its probability mass function is given by

$$P[X = x] = \begin{cases} \frac{1}{n} & \text{for } x = 1, 2, \dots, n \\ 0, & \text{otherwise.} \end{cases}$$

where n is called the parameter of the distribution.

For example, the random variable X , “the number on the unbiased die when thrown”, takes on the positive integer values from 1 to 6 follows discrete uniform distribution having the probability mass function.

$$P[X = x] = \begin{cases} \frac{1}{6}, & \text{for } x = 1, 2, 3, 4, 5, 6. \\ 0, & \text{otherwise.} \end{cases}$$

Mean and Variance of the Distribution

$$\begin{aligned} \text{Mean} = E(X) &= \sum_{x=1}^n x p(x) = \sum_{x=1}^n x \cdot \left(\frac{1}{n}\right) = \frac{1}{n} \sum_{x=1}^n x \\ &= \frac{1}{n} [1 + 2 + 3 + \dots + n] \\ &= \frac{1}{n} \cdot \frac{n(n+1)}{2} \left[\because \text{sum of first } n \text{ natural numbers} = \frac{n(n+1)}{2} \right] \\ &\quad \left[\text{(see Unit 3 of Course MST – 001)} \right] \\ &= \frac{n+1}{2}. \end{aligned}$$

$$\text{Variance} = E(X^2) - [E(X)]^2 \quad \left[\because \mu_2 = \mu_2' - (\mu_1')^2 \right]$$

where

$$E(X) = \frac{n+1}{2} \quad [\text{Obtained above}]$$

$$E(X^2) = \sum_{x=1}^n x^2 \cdot p(x)$$

$$\text{and } E(X^2) = \sum_{x=1}^n x^2 \cdot \frac{1}{n}$$

$$= \frac{1}{n} [1^2 + 2^2 + 3^2 + \dots + n^2]$$

$$= \frac{1}{n} \left[\frac{n(n+1)(2n+1)}{6} \right] \left[\begin{array}{l} \therefore \text{sum of squares of first } n \\ \text{natural numbers} = \frac{n(n+1)(2n+1)}{6} \\ \text{(see Unit 3 of Course MST - 001)} \end{array} \right]$$

$$= \frac{(n+1)(2n+1)}{6}$$

$$\therefore \text{Variance} = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2} \right)^2$$

$$= \frac{(n+1)}{12} [2(2n+1) - 3(n+1)]$$

$$= \frac{n+1}{12} [4n+2-3n-3] = \frac{(n+1)}{12} (n-1) = \frac{n^2-1}{12}$$

Example 11: Find the mean and variance of a number on an unbiased die when thrown.

Solution: Let X be the number on an unbiased die when thrown,

$\therefore$ X can take the values 1, 2, 3, 4, 5, 6 with

$$P[X = x] = \frac{1}{6}; x = 1, 2, 3, 4, 5, 6.$$

Hence, by uniform distribution, we have

$$\text{Mean} = \frac{n+1}{2} = \frac{6+1}{2} = \frac{7}{2}, \text{ and}$$

$$\text{Variance} = \frac{n^2-1}{12} = \frac{(6)^2-1}{12} = \frac{35}{12}.$$

Uniform Frequency Distribution

If an experiment, satisfying the requirements of discrete uniform distribution, is repeated N times, then expected frequency of a value of random variable is given by

$$f(x) = N \cdot P[X = x]; x = 1, 2, \dots, n$$

$$= N \cdot \frac{1}{n}; x = 1, 2, 3, \dots, n.$$

Example 12: If an unbiased die is thrown 120 times, find the expected frequency of appearing 1, 2, 3, 4, 5, 6 on the die.

Solution: Let X be the uniform discrete random variable, “the number on the unbiased die when thrown”.

$$\therefore P[X = x] = \frac{1}{6}; x = 1, 2, \dots, 6$$

Hence, the expected frequencies of the value of random variable are given as computed in the following table:

X	$P[X = x]$	Expected/Theoretical frequencies $f(x) = N \cdot P[X = x] = 120 \cdot P[X = x]$
1	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
2	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
3	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
4	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
5	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$
6	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$

6.6 HYPERGEOMETRIC DISTRIBUTION

In a discrete uniform probability distribution, the probability distribution is obtained for the possible outcomes in a single trial, however, if there are more than one but finite trials with only two possible outcomes in each trial, we apply some other distribution. One such distribution which is applicable in such a situation is binomial distribution. The binomial distribution deals with finite and independent trials, each of which has exactly two possible outcomes (Success or Failure) with constant probability of success in each trial. For example, if we again consider the example of drawing ticket randomly from an urn containing 10 tickets bearing numbers from 1 to 10. Then, the probability that the drawn ticket bears an odd number is $\frac{5}{10} = \frac{1}{2}$. If we replace the ticket back, then the

probability of drawing a ticket bearing an odd number is again $\frac{5}{10} = \frac{1}{2}$. So, if we draw ticket again and again with replacement, trials become independent and

probability of getting an odd number is same in each trial. Suppose, it is asked that what is the probability of getting 2 tickets bearing odd number in 3 draws then we apply binomial distribution as follows:

Let X be the number of times an odd number appears in 3 draws, then by binomial distribution,

$$P[X = 2] = {}^3C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{3-2} = (3) \left(\frac{1}{2}\right) \left(\frac{1}{2}\right) = \frac{3}{8}.$$

But, if in the example discussed above, we do not replace the ticket after any draw the probability of getting an odd number gets changed in each trial and the trials remain no more independent and hence in this case binomial distribution is not applicable. Suppose, in this case also, we are interested in finding the probability of getting ticket bearing odd number twice in 3 draws, then it is computed as follows:

Let A_i be the event that i^{th} ticket drawn bears odd number and $\bar{A}_i$ be the event that i^{th} ticket drawn does not bear odd number.

$\therefore$ Probability of getting ticket bearing odd number twice in 3 draws

$$\begin{aligned} &= P[A_1 \cap A_2 \cap \bar{A}_3] + P[A_1 \cap \bar{A}_2 \cap A_3] + P[\bar{A}_1 \cap A_2 \cap A_3] \\ &= P[A_1]P[A_2 | A_1]P[\bar{A}_3 | A_1 \cap A_2] + P[A_1]P[\bar{A}_2 | A_1]P[A_3 | A_1 \cap \bar{A}_2] \\ &\quad + P[\bar{A}_1]P[A_2 | \bar{A}_1]P[A_3 | \bar{A}_1 \cap A_2] \\ &\quad \text{[Multiplication theorem for dependent events]} \\ &= \frac{5}{10} \cdot \frac{4}{9} \cdot \frac{5}{8} + \frac{5}{10} \cdot \frac{5}{9} \cdot \frac{4}{8} + \frac{5}{10} \cdot \frac{5}{9} \cdot \frac{4}{8} \\ &= 3 \times \frac{5 \times 5 \times 4}{10 \times 9 \times 8} \end{aligned}$$

This result can be written in the following form also:

$$\begin{aligned} &= \frac{5 \times 4 \times 5 \times 3 \times 2}{2 \times 10 \times 9 \times 8} \quad \text{[Multiplying and Dividing by 2]} \\ &= \frac{5 \times 4}{2} \times 5 \times \frac{1}{10 \times 9 \times 8} = {}^5C_2 \times {}^5C_1 \times \frac{1}{{}^{10}C_3} = \frac{{}^5C_2 \times {}^5C_1}{{}^{10}C_3} \end{aligned}$$

In the above result, 5C_2 is representing the number of ways of selecting 2 out of 5 tickets bearing odd number, 5C_1 is representing the number of ways of selecting 1 out of 5 tickets bearing even number i.e. not bearing odd number, and ${}^{10}C_3$ is representing the number of ways of selecting 3 out of total 10 tickets.

Let us consider another similar example of a bag containing 20 balls out of which 5 are white and 15 are black. Suppose 10 balls are drawn at random one by one without replacement, then as discussed in the above example, the probability that in these 10 draws, there are 2 white and 8 black balls is

$$\frac{{}^5C_2 \times {}^{15}C_8}{{}^{20}C_{10}}.$$

Note: The result remains exactly same whether the items are drawn one by one without replacement or drawn at once.

Let us now generalise the above argument for N balls, of which M are white and $N - M$ are black. Of these, n balls are chosen at random without replacement. Let X be a random variable that denote the number of white balls drawn. Then, the probability of $X = x$ white balls among the n balls drawn is given by

$$P[X = x] = \frac{{}^MC_x \cdot {}^{N-M}C_{n-x}}{{}^NC_n}$$

$$[\text{For } x = 0, 1, 2, \dots, n (n \leq M) \text{ or } x = 0, 1, 2, \dots, M (n > M)]$$

The above probability function of discrete random variable X is called the Hypergeometric distribution.

Remark 6: We have a hypergeometric distribution under the following conditions:

- i) There are finite number of dependent trials
- ii) A single trial results in one of the two possible outcomes-Success or Failure
- iii) Probability of success and hence that of failure is not same in each trial i.e. sampling is done without replacement

Remark 7: If number (n) of balls drawn is greater than the number (M) of white balls in the bag, then if $n \leq M$, the number (x) of white balls drawn cannot be greater than n and if $n > M$, then number of white balls drawn cannot be greater than M . So, x can take the values upto n (if $n \leq M$) and M (if $n > M$) i.e. x can take the value upto n or M , whichever is less, i.e. $x = \min \{n, M\}$.

The discussion leads to the following definition

Definition: A random variable X is said to follow the hypergeometric distribution with parameters N , M and n if it assumes only non-negative integer values and its probability mass function is given by

$$P[X = x] = \begin{cases} \frac{{}^MC_x \cdot {}^{N-M}C_{n-x}}{{}^NC_n} & \text{for } x = 0, 1, 2, \dots, \min \{n, M\} \\ 0, & \text{otherwise} \end{cases}$$

where n , M , N are positive integers such that $n \leq N$, $M \leq N$.

Mean and Variance

$$\begin{aligned} \text{Mean} = E(X) &= \sum_{x=0}^n x \cdot p[X = x] \\ &= \sum_{x=1}^n x \cdot \frac{{}^MC_x \cdot {}^{N-M}C_{n-x}}{{}^NC_n} \\ &\dots \end{aligned}$$

$$= \frac{M}{N} \binom{N-1}{C_{n-1}}$$

[This result is obtained using properties of binomial coefficients and involves lot of calculations and hence its derivation may be skipped. It may be noticed that in this result the left upper suffix and also the right lower suffix is the sum of the corresponding suffices of the binomial coefficients involved in each product term.]

Also,

$$\begin{aligned} V(X) &= E(X^2) - [E(X)]^2 \\ &= \frac{NM(N-M)(N-n)}{N^2(N-1)} \quad \text{[Given without proof]} \end{aligned}$$

Example 13: A jury of 5 members is drawn at random from a voters' list of 100 persons, out of which 60 are non-graduates and 40 are graduates. What is the probability that the jury will consist of 3 graduates?

Solution: The computation of the actual probability is hypergeometric, which is shown as follows:

$$\begin{aligned} P[2 \text{ non-graduates and } 3 \text{ graduates}] &= \frac{{}^{60}C_2 \cdot {}^{40}C_3}{{}^{100}C_5} \\ &= \frac{60 \times 59 \times 40 \times 39 \times 38 \times 5 \times 4 \times 3 \times 2}{2 \times 6 \times 100 \times 99 \times 98 \times 97 \times 96} \\ &= 0.2323 \end{aligned}$$

Example 14: Let us suppose that in a lake there are N fish. A catch of 500 fish (all at the same time) is made, and these fish are returned alive into the lake after marking each with a red spot. After two days, assuming that during this time these 'marked' fish have been distributed themselves 'at random' in the lake and there is no change in the total number of fish, a fresh catch of 400 fish (again, all at once) is made. What is the probability that of these 400 fish, 100 will be having red spots?

Solution: The computation of the probability is hypergeometric and is shown as follows: As marked fish in the lake are 500 and other are $N - 500$,

$$\therefore P[100 \text{ marked fish and } 300 \text{ others}] = \frac{{}^{500}C_{100} \cdot {}^{N-500}C_{300}}{{}^N C_{400}}.$$

We cannot numerically evaluate this if N is not given. Though N can be estimated using method of Maximum likelihood estimation which you will read in later units. We are not going to estimate it.

You will agree that the exact computation of this probability is complicated. Such problem is normally there with the use of hypergeometric distribution, especially, if N and M are large. However, if n is small compared to N i.e. if n

is such that $\frac{n}{N} < 0.05$, say then there is not much difference between sampling

with and without replacement and hence in such cases, the probability obtained by binomial distribution comes out to be approximately equal to that obtained using hypergeometric distribution.

Check Your Progress 3

- 1) Obtain the mean, variance of the discrete uniform distribution for the random variable, “the number on a ticket drawn randomly from an urn containing 10 tickets numbered from 1 to 10”. Also obtain the expected frequencies if the experiment is repeated 150 times.

.....
.....

- 2) A lot of 25 units contains 10 defective units. An engineer inspects 2 randomly selected units from the lot. He/She accepts the lot if both the units are found in good condition, otherwise all the remaining units are inspected. Find the probability that the lot is accepted without further inspection.

.....
.....

We now conclude this unit by giving a summary of what we have covered in it.

6.7 SUMMARY

The following main points have been covered in this unit:

- 1) A discrete random variable X is said to follow **Bernoulli distribution** with parameter p if its probability mass function is given by

$$P[X = x] = \begin{cases} p^x (1-p)^{1-x} & ; x = 0, 1 \\ 0 & ; \text{elsewhere} \end{cases}$$

Its **mean** and **variance** are p and $p(1-p)$, respectively.

- 2) A discrete random variable X is said to follow **binomial distribution** if it assumes only a finite number of non-negative integer values and its probability mass function is given by

$$P[X = x] = \begin{cases} {}^n C_x p^x q^{n-x} & ; x = 0, 1, 2, \dots, n \\ 0 & ; \text{elsewhere} \end{cases}$$

where, n is the number of independent trials, x is the number of successes in n trial, p is the probability of success in each trial, and $q = 1 - p$ is the probability of failure in each trial.

- 3) The **constants of Binomial distribution** are:

Mean = np , Variance = npq ,

- 4) For a binomial distribution, **Mean > Variance**.

- 5) A random variable X is said to follow **Poisson distribution** if it assumes indefinite number of non-negative integer values and its probability mass function is given by:

$$p(x) = P(X = x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!} & ; x = 0, 1, 2, 3, \dots \text{ and } \lambda > 0. \\ 0 & ; \text{elsewhere} \end{cases}$$

- 6) For Poisson distribution, **Mean = Variance** = λ .
- 7) A random variable X is said to have a **discrete uniform (rectangular)** distribution if it takes any positive integer value from 1 to n , and its probability mass function is given by

$$P[X = x] = \begin{cases} \frac{1}{n} & \text{for } x = 1, 2, \dots, n \\ 0, & \text{otherwise.} \end{cases}$$

where n is called the parameter of the distribution.

- 8) For **discrete uniform** distribution, **mean** = $\frac{n+1}{2}$ and **variance** = $\frac{n^2-1}{12}$.
- 9) A random variable X is said to follow the **hypergeometric distribution** with parameters N , M and n if it assumes only non-negative integer values and its probability mass function is given by

$$P(X = x) = \begin{cases} \frac{{}^M C_x \cdot {}^{N-M} C_{n-x}}{{}^N C_n} & \text{for } x = 0, 1, 2, \dots, \min\{n, M\} \\ 0, & \text{otherwise} \end{cases}$$

where n , M , N are positive integers such that $n \leq N$, $M \leq N$.

- 10) For **hypergeometric** distribution, **mean** = $\frac{nM}{N}$ and
variance = $\frac{NM(N-M)(N-n)}{N^2(N-1)}$.

6.8 SOLUTIONS/ANSWERS

Check Your Progress 1

- 1) Let p be the probability of hitting the target (success) in a trial.

$$\therefore n = 5, p = \frac{1}{4}, q = 1 - \frac{1}{4} = \frac{3}{4},$$

and hence by binomial distribution, we have

$$P[X = x] = {}^n C_x p^x q^{n-x} = {}^5 C_x \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{5-x}; x = 0, 1, 2, 3, 4, 5.$$

$$\therefore \text{Required probability} = P[X \geq 2]$$

$$= P[X = 2] + P[X = 3] + P[X = 4] + P[X = 5]$$

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

$$= 1 - \left[{}^5 C_0 \left(\frac{1}{4}\right)^0 \left(\frac{3}{4}\right)^{5-0} + {}^5 C_1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^{5-1} \right]$$

$$= 1 - \left[\frac{243}{1024} + \frac{405}{1024} \right] = \frac{376}{1024} = \frac{47}{128}$$

- 2) Let p be the probability that the dacoit will be killed (success) by a bullet.

$\therefore n = 6, p = 0.6, q = 1 - p = 1 - 0.6 = 0.4$, and hence by binomial distribution, we have

$$P[X = x] = {}^nC_x p^x q^{n-x}; x = 0, 1, 2, \dots, n$$

$$= {}^6C_x (0.6)^x (0.4)^{6-x}; x = 0, 1, 2, \dots, 6.$$

$$\begin{aligned}\therefore \text{The required probability} &= P[\text{The dacoit is still alive}] \\ &= P[\text{No bullet kills the dacoit}] \\ &= P[\text{Number of successes is zero}] \\ &= P[X = 0] = {}^6C_0 (0.6)^0 (0.4)^6 \\ &= 0.0041\end{aligned}$$

$$3) \text{ Mean} = np = 3 \quad \dots (1)$$

$$\text{Variance} = npq = 4 \quad \dots (2)$$

$\therefore$ Dividing (2) by (1), we have

$$q = \frac{4}{3} > 1 \text{ and hence not possible}$$

[$\because$ q, being probability, cannot be greater than 1]

4) Let $X \sim B(n, p)$, then

$n = 5$ and

$$np + npq = 4.8 \quad [\because \text{given that Mean} + \text{Variance} = 4.8]$$

$$\Rightarrow 5p + 5pq = 4.8$$

$$\Rightarrow 5[p + p(1-p)] = 4.8$$

$$\Rightarrow 5[p + p - p^2] = 4.8$$

$$\Rightarrow 5p^2 - 10p + 4.8 = 0$$

$$\Rightarrow 25p^2 - 50p + 24 = 0 \quad [\text{Multiplying by 5}]$$

$$\Rightarrow 25p^2 - 30p - 20p + 24 = 0$$

$$\Rightarrow 5p(5p - 6) - 4(5p - 6) = 0$$

$$\Rightarrow (5p - 6)(5p - 4) = 0$$

$$\Rightarrow p = \frac{6}{5}, \frac{4}{5}$$

The first value $p = \frac{6}{5}$ is rejected [$\because$ probability can never exceed 1]

$$\therefore p = \frac{4}{5} \text{ and hence } q = 1 - p = \frac{1}{5}.$$

Thus, the binomial distribution is

$$P[X = x] = {}^nC_x p^x q^{n-x}$$

$$= {}^5C_x \left(\frac{4}{5}\right)^x \left(\frac{1}{5}\right)^{5-x}; x = 0, 1, 2, 3, 4, 5.$$

The binomial distribution in tabular form is given as

X	p(x)
0	${}^5C_0 \left(\frac{4}{5}\right)^0 \left(\frac{1}{5}\right)^5 = \frac{1}{3125}$
1	${}^5C_1 \left(\frac{4}{5}\right)^1 \left(\frac{1}{5}\right)^4 = \frac{20}{3125}$
2	${}^5C_2 \left(\frac{4}{5}\right)^2 \left(\frac{1}{5}\right)^3 = \frac{160}{3125}$
3	${}^5C_3 \left(\frac{4}{5}\right)^3 \left(\frac{1}{5}\right)^2 = \frac{640}{3125}$
4	${}^5C_4 \left(\frac{4}{5}\right)^4 \left(\frac{1}{5}\right)^1 = \frac{1280}{3125}$
5	${}^5C_5 \left(\frac{4}{5}\right)^5 \left(\frac{1}{5}\right)^0 = \frac{1024}{3125}$

- 5) Given that Mean = 30 and S.D. = 5

$$\text{Thus, } np = 30, \sqrt{npq} = 5$$

$$\Rightarrow np = 30, npq = 25$$

$$\frac{npq}{np} = \frac{25}{30} = \frac{5}{6} \Rightarrow q = \frac{5}{6}, p = 1 - q = 1 - \frac{5}{6} = \frac{1}{6}, n \left(\frac{1}{6}\right) = 30 \Rightarrow n = 180$$

Check Your Progress 2

- 1) Let X be the Poisson variable “Number of fatal accidents in a year”.

$$\text{Here } n = 350, p = \frac{1}{1400}$$

$$\Rightarrow \lambda = np = (350) \left(\frac{1}{1400}\right) = 0.25.$$

By Poisson distribution,

$$P[X = x] = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, x = 0, 1, 2, \dots$$

$$= \frac{e^{-0.25} (0.25)^x}{x!}, x = 0, 1, 2, \dots$$

Therefore, P [at least one fatal accident]

$$= P[X \geq 1] = 1 - P[X < 1] = 1 - P[X = 0]$$

$$= 1 - \frac{e^{-0.25} (0.25)^0}{0!} = 1 - e^{-0.25} = 1 - 0.78 = 0.22$$

- 2) As mean = 6, therefore, $\lambda = 6$.

As standard deviation is 2, therefore, variance = 4 $\Rightarrow \lambda = 4$.

We get two different values of λ , which is impossible. Hence, the statement is invalid.

- 3) Let λ be the mean of the distribution,
 $\therefore$ by Poisson distribution, we have

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}; x = 0, 1, 2, 3, \dots$$

Given that $P[X = 1] = P[X = 2]$,

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

$$\Rightarrow \lambda = \frac{\lambda^2}{2} \Rightarrow \lambda^2 - 2\lambda = 0 \Rightarrow \lambda(\lambda - 2) = 0$$

$$\Rightarrow \lambda = 0, 2.$$

$\lambda = 0$ is rejected,

$$\therefore \lambda = 2$$

Hence, Mean = 2.

$$\text{Now, } P[X = 0] = \frac{e^{-\lambda} \lambda^0}{0!} = e^{-\lambda} = e^{-2} = 0.1353, \text{ [See the Appendix 1]}$$

$$\begin{aligned} \text{and } P[X = 4] &= \frac{e^{-\lambda} \lambda^4}{4!} = \frac{e^{-2} (2)^4}{24} = \frac{e^{-2} (16)}{24} = \frac{2}{3} (0.1353) \\ &= 2(0.0451) \\ &= 0.0902. \end{aligned}$$

Check Your Progress 3

- 1) Let X be the number on the ticket drawn randomly from an urn containing tickets numbered from 1 to 10.

$\therefore X$ is a discrete uniform random variable having the values

1, 2, 3, 4, ..., 10 with probability of each of these values equal to $\frac{1}{10}$.

Thus, the expected frequencies for the values of X are obtained as in the following table:

X	$P(X = x)$	Expected/Theoretical frequency $f(x) = N.P[X = x]$ $= 150.P[X = x]$
1	$\frac{1}{10}$	$150 \times \frac{1}{10} = 15$
2	$\frac{1}{10}$	$150 \times \frac{1}{10} = 15$
3	$\frac{1}{10}$	$150 \times \frac{1}{10} = 15$
...(for 4, 5, 6, 7, 8, 9 and 10)	$\frac{1}{10}$ in each case	$150 \times \frac{1}{10} = 15$

- 2) Here $N = 25$, $M = 10$ and $n = 2$.

$$\text{The desired probability} = P \left[\begin{array}{l} \text{none of the 2 randomly selected} \\ \text{units is found defective} \end{array} \right]$$

$$= \frac{{}^{10}C_0 \cdot {}^{25-10}C_2}{{}^{25}C_2} = \frac{(1) \cdot {}^{15}C_2}{{}^{25}C_2} = \frac{15 \times 14}{25 \times 24} = 0.35.$$

6.9 REFERENCES

IGNOU Course Material

1. Post-Graduate Diploma in Applied Statistics (PGDAST), MST 003: Probability Theory, Block 3: Discrete Probability Distributions, UNIT 9: Binomial Distribution (for Sections 6.2, 6.3 and related portions in Sections 6.0 , 6.1, 6.7, 6.8 and Check Your Progress questions)
2. PGDAST, MST 003: Probability Theory, Block 3: Discrete Probability Distributions, UNIT 10: Poisson Distribution (for Section 6.4 and related portions in Sections 6.0 , 6.1, 6.7, 6.8 and Check Your Progress questions)
3. PGDAST, MST 003: Probability Theory, Block 3: Discrete Probability Distributions, UNIT 11: Discrete Uniform and Hypergeometric Distributions (for Sections 6.5, 6.6 and related portions in Sections 6.0 , 6.1, 6.7, 6.8 and Check Your Progress questions)

Reference Book

4. “Introduction to Probability and Statistics for Engineers and Scientists”, 3rd Edition, Sheldon M. Ross, Elsevier Academic Press, 2004

Appendix 1

Value of $e^{-\lambda}$ (For Computing Poisson Probabilities) ($0 < \lambda < 1$)

λ	0	1	2	3	4	5	6	7	8	9
0.0	1.0000	0.9900	0.9802	0.9704	0.9608	0.9512	0.9418	0.9324	0.9231	0.9139
0.1	0.9048	0.8958	0.8860	0.8781	0.8694	0.8607	0.8521	0.8437	0.8353	0.8270
0.2	0.7187	0.8106	0.8025	0.7945	0.7866	0.7788	0.7711	0.7634	0.7558	0.7483
0.3	0.7408	0.7334	0.7261	0.7189	0.7118	0.7047	0.6970	0.6907	0.6839	0.6771
0.4	0.6703	0.6636	0.6570	0.6505	0.6440	0.6376	0.6313	0.6250	0.6188	0.6125
0.5	0.6065	0.6005	0.5945	0.5886	0.5827	0.5770	0.5712	0.5655	0.5599	0.5543
0.6	0.5448	0.5434	0.5379	0.5326	0.5278	0.5220	0.5160	0.5113	0.5066	0.5016
0.7	0.4966	0.4916	0.4868	0.4810	0.4771	0.4724	0.4670	0.4630	0.4584	0.4538
0.8	0.4493	0.4449	0.4404	0.4360	0.4317	0.4274	0.4232	0.4190	0.4148	0.4107
0.9	0.4066	0.4026	0.3985	0.3946	0.3906	0.3867	0.3829	0.3791	0.3753	0.3716
($\lambda=1, 2, 3, \dots, 10$)										
λ	1	2	3	4	5	6	7	8	9	10
$e^{-\lambda}$	0.3679	0.1353	0.0498	0.0183	0.0070	0.0028	0.0009	0.0004	0.0001	0.00004

Note: To obtain values of $e^{-\lambda}$ for other values of λ , use the laws of exponents i.e.

$$e^{-(a+b)} = e^{-a} \cdot e^{-b} \text{ e. g. } e^{-2.25} = e^{-2} \cdot e^{-0.25} = (0.1353)(0.7788) = 0.1054.$$