UNIT 9 BINOMIAL DISTRIBUTION

Structure

9.1 INTRODUCTION

In Unit 5 of the Course, you have studied random variables, their probability functions and distribution functions. In Unit 8 of the Course, you have come to know as to how the expectations and moments of random variables are obtained. In those units, the definitions and properties of general discrete and probability distributions have been discussed.

The present block is devoted to the study of some special discrete distributions and in this list, Bernoulli and Binomial distributions are also included which are being discussed in the present unit of the course.

Sec. 9.2 of this unit defines Bernoulli distribution and its properties. Binomial distribution and its applications are covered in Secs. 9.3 and 9.4 of the unit.

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 as to how binomial distribution is fitted to the given data; and
- solve various practical problems related to these distributions.

9.2 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 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,

$P[X=1]=p,$ 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

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

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

and $P[X=0]=p^0(1-p)^{1-0}=1-p$ [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 rth moment about origin of a Bernoulli variate X is given as

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

$$= \sum_{x=0}^{1} x^r p(x) \quad \text{[See Unit 8 of this course]}$$

$$=(0)^r p(0)+(1)^r p(1)$$

$$=(0)(1-p)+(1)p$$

$$=p$$

$\Rightarrow \mu_1'=p,\ \mu_2'=p,\ \mu_3'=p,\ \mu_4'=p$.

Hence,

Mean = $\mu_1'=p$,

Variance $(\mu_2)=\mu_2'-(\mu_1')^2=p-p^2=p(1-p)$,

Third order central moment $(\mu_3)=\mu_3'-3\mu_2'(\mu_1')+2(\mu_1')^3$

$$=p-3pp+2(p)^3$$

$$= p - 3p^2 + 2p^3$$
$$= p\left(2p^2 - 3p + 1\right) = p\left(2p - 1\right)\left(p - 1\right)$$
$$= p\left(1 - p\right)\left(1 - 2p\right)$$

Fourth order central moment $(\mu_4) = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\left(\mu_1'\right)^2 - 3\left(\mu_1'\right)^4$

$$= p - 4p.p + 6p\left(p\right)^2 - 3\left(p\right)^4$$
$$= p - 4p^2 + 6p^3 - 3p^4$$
$$= p\left[1 - 4p + 6p^2 - 3p^3\right]$$
$$= p\left(1 - p\right)\left(1 - 3p + 3p^2\right)$$

[**Note:** For relations of central moments in terms of moments about origin, see Unit 3 of MST-002.]

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

Solution:

Mean = p = 0.4,
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.

9.3 BINOMIAL PROBABILITY FUNCTION

Here, is 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 $q\left(=1-p\right)$ 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,

P[X = 0] = Probability that target is not hit at all in any trial
= P [Failure in each of the three trials]

$$= P\left(F \cap F \cap F\right)$$
$$= P\left(F\right).P\left(F\right).P\left(F\right) \quad [\because \text{ trials are independent}]$$
$$= q.q.q$$
$$= q^3$$

This can be written as

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

[$\because\ {}^3C_0 = 1,\ p^0 = 1,\ q^{3-0} = q^3$. Recall ${}^nC_x = \frac{\lfloor n}{\lfloor x \lfloor n-x}$ (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)]

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

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

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

[$\because$ 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

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

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)]

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

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

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

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

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

This can also be written as

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

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

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

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

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, …, n are the respective probabilities ${}^nC_0 p^0 q^{n-0}$, ${}^nC_1 p^1 q^{n-1}$, ..., ${}^nC_n p^n q^{n-n}$ which are the successive terms of the binomial expansion $(q + p)^n$.

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

Binomial Expansion:

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

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

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

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

So, in general,

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

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

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

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

$$P\left[X=x\right]={}^{n}C_{x}p^{x}q^{n-x}; x=0, 1, 2, ..., n$$

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

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

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

Therefore,

(i) P[exactly 3 heads] = P [X = 3]

$$=\frac{1}{64}\left({}^{6}C_{3}\right)=\frac{1}{64}\left[\frac{6\times5\times4}{3\times2}\right]=\frac{5}{16}$$

$$\left[\because\ \text{Recall}\ {}^{n}C_{x}=\frac{\lfloor n}{\lfloor x\lfloor n-x}\ \text{(see Unit 4 of MST- 001)}\right]$$

(ii) P[less than 3 heads] = P[X < 3]

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

$$=P\left[X=2\right]+P\left[X=1\right]+P\left[X=0\right]$$

$$=\frac{1}{64}.{}^{6}C_{2}+\frac{1}{64}.{}^{6}C_{1}+\frac{1}{64}.{}^{6}C_{0}$$

$$= \frac{1}{64}\left[{}^{6}C_{2} + {}^{6}C_{1} + {}^{6}C_{0}\right] = \frac{1}{64}\left[\frac{6\times 5}{2} + 6 + 1\right]$$

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

(iii) P[more than 3 heads] = P[X > 3]

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

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

$$= \frac{1}{64}.\,{}^{6}C_{4} + \frac{1}{64}.\,{}^{6}C_{5} + \frac{1}{64}.\,{}^{6}C_{6}$$

$$= \frac{1}{64}\left[{}^{6}C_{4} + {}^{6}C_{5} + {}^{6}C_{6}\right]$$

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

(iv) P[at most 3 heads] = P [3 or less than 3 heads]

$$= P[X = 3] + P[X = 2] + P[X = 1] + P[X = 0]$$

$$= \frac{1}{64}.\,{}^{6}C_{3} + \frac{1}{64}.\,{}^{6}C_{2} + \frac{1}{64}.\,{}^{6}C_{1} + \frac{1}{64}.\,{}^{6}C_{0}$$

$$= \frac{1}{64}\left[{}^{6}C_{3} + {}^{6}C_{2} + {}^{6}C_{1} + {}^{6}C_{0}\right]$$

$$= \frac{1}{64}\left[20 + 15 + 6 + 1\right] = \frac{42}{64} = \frac{21}{32}.$$

(v) P[at least 3 heads] = P[3 or more heads]

$$= P[X = 3] + P[X = 4] + P[X = 5] + P[X = 6]$$

or

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

$$\left[\begin{array}{l}\because \text{ sum of probabilities of all possible} \\ \text{ values of a random variable is 1}\end{array}\right]$$

$$= 1 - \left(\frac{11}{32}\right) \quad \left[\begin{array}{l}\text{Already obtained in} \\ \text{part (ii) of this example}\end{array}\right]$$

$$= \frac{21}{32}.$$

(vi) P [more than 6 heads] = P [7 or more heads]

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

$$= 0$$

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

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

$$= {}^{5}C_{x}\,(0.25)^{x}(0.75)^{5-x} \;;\; 0, 1, 2, ..., 5$$

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

$$= p[X = 4 \text{ or } X = 5]$$

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

$$= {}^{5}C_{4}(0.25)^{4}(0.75)^{1} + {}^{5}C_{5}(0.25)^{5}(0.75)^{0}$$

$$= (5)(0.002930) + (1)(0.000977)$$

$$= 0.014650 + 0.000977$$

$$= 0.015627$$

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.

Thus, the required probability = $[X + Y \leq 1]$

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

$$= {}^{7}C_{1}(0.7)^{1}(0.3)^{6} + {}^{7}C_{0}(0.7)^{0}(0.3)^{7}$$

$$= 7(0.7)(0.000729) + 1(1)(0.0002187)$$

$$= 0.0035721 + 0.0002187$$

$$= 0.0037908$$

Now, we are sure that you can try the following exercises:

E1) 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?

E2) 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?

9.4 MOMENTS OF BINOMIAL DISTRIBUTION

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

$$\mu_r' = E\left(X^r\right) = \sum_{x=0}^{n} x^r . P\left[X = x\right]$$

$$\therefore \quad \mu_1' = E\left(X\right) = \sum_{x=0}^{n} x \; .P\left[X = x\right]$$

$$= \sum_{x=0}^{n} x.\,{}^nC_x\, p^x q^{n-x} \quad \left[\because P\left[X = x\right] = {}^nC_x p^x q^{n-x} \,;\, x = 0, 1, 2, ..., n\right]$$

$$= \sum_{x=1}^{n} x.\,{}^nC_x\, p^x q^{n-x} \quad \left[\begin{array}{l}\because \text{first term with } x = 0 \text{ will be zero} \\ \text{and hence we may start from } x = 1\end{array}\right]$$

$$= \sum_{x=1}^{n} x.\frac{n}{x}.\;{}^{n-1}C_{x-1}\, p^x q^{n-x}$$

$$\left[\begin{array}{l}\because \; {}^nC_x = \dfrac{\underline{|n}}{\underline{|x}\,\underline{|n-x}} = \dfrac{n\underline{|n-1}}{x\underline{|x-1}\,\underline{|\left(n-1\right)-\left(x-1\right)}} = \dfrac{n}{x}\;{}^{n-1}C_{x-1}, \\ \left(\text{see Unit 4 of MST} - 001\right)\end{array}\right]$$

$$= \sum_{x=1}^{n} n.\;{}^{n-1}C_{x-1}\, p^{x-1}.p^1.q^{(n-1)-(x-1)} \quad [n - x = (n - 1) - (x - 1)]$$

$$= np \sum_{x=1}^{n} \;{}^{n-1}C_{x-1}\, p^{x-1}.q^{(n-1)-(x-1)}$$

$$= np\left[\begin{array}{r}{}^{n-1}C_0\, p^0\, q^{(n-1)-0} + {}^{n-1}C_1\, p^1\, q^{(n-1)-1} + {}^{n-1}C_2\, p^2 q^{(n-1)-2} + ... \\ + \;{}^{n-1}C_{n-1}.\, p^{n-1}\, q^{(n-1)-(n-1)}\end{array}\right]$$

$$= np \times \left[\begin{array}{c}\text{Sum of probabilities of all possible values of a} \\ \text{binomial variate with parameters } n-1 \text{ and } p\end{array}\right]$$

$$= np \times 1 \qquad \left[\begin{array}{l}\because \text{sum of probabilities of all possible} \\ \text{values of a random variable is 1}\end{array}\right]$$

$$= np.$$

$\therefore$ Mean = First order moment about origin

$= \mu_1'$

$= np.$

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

$$\mu_2' = E\left(X^2\right) = \sum_{x=0}^{n} x^2.P\left[X = x\right] = \sum_{x=0}^{n} x^2.\,{}^nC_x p^x q^{n-x}$$

Here, we will write x^2 as $x\left(x-1\right) + x$ $\quad [\because x\left(x-1\right) + x = x^2 - x + x = x^2]$

This is done because in the following expression, we get $x\left(x-1\right)$ in the denominator:

$$\left[\because {}^{n}C_x = \frac{\lfloor n}{\lfloor x \times \lfloor n-x} = \frac{n(n-1)\lfloor n-2}{x(x-1)\lfloor x-2 \times \lfloor (n-2)-(x-2)} = \frac{n(n-1)}{x(x-1)}.{}^{n-2}C_{x-2}\right]$$

$$\therefore \mu_2' = \sum_{x=0}^{n}\left[x(x-1)+x\right]{}^{n}C_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)\,{}^{n}C_x p^x q^{n-x} + \sum_{x=0}^{n} x.\,{}^{n}C_x p^x q^{n-x}$$

$$= \left[\sum_{x=2}^{n} x(x-1)\,{}^{n}C_x p^x q^{(n-2)-(x-2)}\right] + \left(\mu_1'\right)$$

$$= \left[\sum_{x=2}^{n} x(x-1).\frac{n(n-1)}{x(x-1)}\;{}^{n-2}C_{x-2} p^x q^{n-x}\right] + \mu_1'$$

$$= \left[\sum_{x=2}^{n} n(n-1)\,{}^{n-2}C_{x-2} p^{x-2}.p^2 q^{(n-2)-(x-2)}\right] + \mu_1'$$

$$= \left[n(n-1)p^2 \sum_{x=2}^{n} {}^{n-2}C_{x-2} p^2 q^{(n-2)-(x-2)}\right] + \mu_1'$$

$$= n(n-1)p^2 \times \left[\begin{array}{c}\text{Sum of probabilities of all possible values of a} \\ \text{binomial variate with parameters } n-2 \text{ and } p\end{array}\right] + \mu_1'$$

$$= n(n-1)\,p^2\,(1) + np \qquad \left[\because \mu_1' = np\right]$$

$$= n^2 p^2 - np^2 + np$$

$\therefore$ Variance $(\mu_2) = \mu_2' - \left(\mu_1'\right)^2$ [See Unit 3 of MST-002]

$$= n^2p^2 - n p^2 + np - (np)^2$$

$$= n^2p^2 - np^2 + np - n^2p^2$$

$$= np - np^2$$

$$= np(1-p)$$

$$= npq$$

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

$$\mu_3' = \sum_{x=0}^{n} x^3.P\left[X = x\right]$$

Here, we will write x^3 as $x(x-1)(x-2)+3x(x-1)+x$

Let $x^3 = x(x-1)(x-2)+Bx(x-1)+Cx$
Comparing coefficients of x^2, we have
$0 = -3 + B \Rightarrow B = 3$
Comparing coeffs of x, we have
$0 = 2 - B + C \Rightarrow C = B - 2 = 3 - 2 \Rightarrow C = 1$

$$\therefore \quad \mu_3' = \sum_{x=0}^{n}\left[x(x-1)(x-2)+3x(x-1)+x\right].{}^nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)(x-2)\,{}^nC_x p^x q^{n-x} + 3\sum_{x=0}^{n} x(x-1)\,{}^nC_x p^x q^{n-x} + \sum_{x=0}^{n} x.\,{}^nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)(x-2)\frac{n}{x}.\frac{n-1}{x-1}.\frac{n-2}{x-2}\,{}^{n-3}C_{x-3}p^x q^{n-x} + 3\left[n(n-1)p^2\right]+\left[np\right]$$

[The expression within brackets in the second term is the first term of R.H.S. in the derivation of μ_2' and the expression in the third term is μ_1' as already obtained.]

$$\left[\because\ {}^nC_x = \frac{\underline{|n}}{\underline{|x}\,\underline{|n-x}} = \frac{n(n-1)(n-2)\underline{|n-3}}{x(x-1)(x-2)\underline{|x-3}\,\underline{|(n-3)-(x-3)}} = \frac{n(n-1)(n-2)}{x(x-1)(x-2)}.\,{}^{n-3}C_{x-3}\right]$$

$$= \sum_{x=3}^{n} n(n-1)(n-2).\,{}^{n-3}C_{x-3}p^3p^{x-3}q^{(n-3)-(x-3)} + 3n(n-1)p^2 + np$$

$$= n(n-1)(n-2)p^3\sum_{x=3}^{n} {}^{n-3}C_{x-3}p^{x-3}q^{(n-3)-(x-3)} + 3n(n-1)p^2 + np$$

$$= n(n-1)(n-2)p^3(1) + 3n(n-1)p^2 + np$$

$\therefore$ Third order central moment is given by

$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\left(\mu_1'\right)^3$ [See Unit 4 of MST-002]

$= npq(q-p)$ [On simplification]

$$\boxed{\mu_3 = npq(q-p)}$$

$$\mu_4' = \sum_{x=0}^{n} x^4\,P(X=x)$$

Writing

$$x^4 = x(x-1)(x-2)(x-3)+6x(x-1)(x-2)+7x(x-1)+x$$

and proceeding in the similar fashion as for $\mu_1', \mu_2', \mu_3',$ we have

$$\mu_4' = n(n-1)(n-2)(n-3)p^4 + 6n(n-1)(n-2)p^3 + 7n(n-1)p^2 + np$$

and hence

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\left(\mu_1'\right)^2 - \left(\mu_1'\right)^4$$

$\mu_4 = npq\left[1+3(n-2)pq\right]$ [On simplification]

Now, recall the measures of skewness and kurtosis which you have studied in Unit 4 of MST-002

These measures are given as follows:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\left[npq(q-p)\right]^2}{[npq]^3} = \frac{(q-p)^2}{npq},$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{npq\left[1+3(n-2)pq\right]}{[npq]^2} = 3 + \frac{1-6pq}{npq},$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{npq}}, \text{ and}$$

$$\gamma_2 = \beta_2 - 3 = \frac{1-6pq}{npq}$$

Remark 3:

(i) As $0 < q < 1$

$\Rightarrow q < 1$

$\Rightarrow npq < np$ [Multiplying both sides by $np > 0$]

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

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

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

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

Example 5: 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

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 second equation by the first equation, we have

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

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

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

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,

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

$$= {}^{6}C_x \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{6-x} \;; \; x = 0, 1, 2, \ldots, 6.$$

Thus, the required probability

$$P[X \geq 1] = P[X = 1] + P[X = 2] + P[X = 3] + \ldots + P[X = 6]$$

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

$$= 1 - {}^{6}C_0 \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^{6-0} = 1 - (1)(1)\frac{1}{729} = \frac{728}{729}.$$

Example 6: If X ~ B (n, p). Find p if n = 6 and 9P[X = 4] = P[X = 2].

Solution: As X ~ B(n, p) and n = 6,

$$\therefore \; P[X = x] = {}^{6}C_x p^x (1-p)^{6-x} \; ; x = 0, 1, 2, \ldots, 6.$$

Now, $9P[X = 4] = P[X = 2]$

$$\Rightarrow 9 \times {}^{6}C_4 \times p^4 (1-p)^{6-4} = {}^{6}C_2 \times p^2 (1-p)^4$$

$$\Rightarrow 9 \times \frac{6 \times 5}{2} \times p^4 \times (1-p)^2 = \frac{6 \times 5}{2} p^2 (1-p)^4$$

$$\Rightarrow 9p^2 = (1-p)^2$$

$$\Rightarrow 9p^2 = 1 + p^2 - 2p$$

$$\Rightarrow 8p^2 + 2p - 1 = 0$$

$$\Rightarrow 8p^2 + 4p - 2p - 1 = 0$$

$$\Rightarrow 4p(2p+1) - 1(2p+1) = 0$$

$$\Rightarrow (2p+1)(4p-1) = 0$$

$$\Rightarrow (2p+1) = 0 \text{ or } (4p-1) = 0$$

$$\Rightarrow p = -\frac{1}{2} \text{ or } \frac{1}{4}$$

But $p = -\frac{1}{2}$ rejected [$\because$ probability can never be negative]

Hence, p = $\frac{1}{4}$

Now, you can try the following exercises:

E3) Comment on the following:

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

E4) Find the binomial distribution when sum of mean and variance of 5 trails is 4.8.

E5) The mean of a binomial distribution is 30 and standard deviation is 5. Find the values of

i) n, p and q,

ii) Moment coefficient of skewness, and

iii) Kurtosis.

9.5 FITTING OF BINOMIAL DISTRIBUTION

To fit a binomial distribution, we need the observed data which is obtained from repeated trials of a given experiment. On the basis of the observed data, we find the theoretical (or expected) frequencies corresponding to each value of the binomial variable. Process of finding the probabilities corresponding to each value of the binomial variable becomes easy if we use the recurrence relation for the probabilities of Binomial distribution. So, in this section, we will first establish the recurrence relation for probabilities and then define the binomial frequency distribution followed by process of fitting a binomial distribution.

Recurrence Relation for the Probabilities of Binomial Distribution

You have studied that binomial probability function is

$$p(x) = P[X = x] = {}^nC_x p^x q^{n-x} \qquad \dots (1)$$

If we replace x by x + 1, we have

$$p(x+1) = {}^nC_{x+1} p^{x+1} q^{n-(x+1)} \qquad \dots (2)$$

Dividing (2) by (1), we have

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

$$= \frac{\lfloor n \, \lfloor x \, \lfloor n-x}{\lfloor x+1 \, \lfloor n-x-1 \, \lfloor n} \times \frac{p}{q} \qquad \left[\because \; {}^nC_{x+1} = \frac{\lfloor n}{\lfloor x+1 \, \lfloor n-x-1} \text{ and } {}^nC_x = \frac{\lfloor n}{\lfloor x \, \lfloor n-x}\right]$$

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

$$\Rightarrow p(x+1) = \frac{n-x}{x+1} \frac{p}{q} p(x) \qquad \dots (3)$$

Putting $x = 0$, 1, 2, 3,…in this equation, we get p(1) in terms of p(0), p(2) in terms of p(1), p(3) in terms of p(2), and so on. Thus, if p(0) is known, we can find p(1) then p(2), p(3) and so on.

So, eqn. (3) is the recurrence relation for finding the probabilities of binomial distribution. The initial probability i.e. p(0) is obtained from the following formula:

$$p(0) = q^n$$

[$\because$ $p(x) = {}^nC_x p^x q^{n-x}$ putting x = 0, we have p(0) = ${}^nC_0 p^0 q^n = q^n$]

Binomial Frequency Distribution

We have studied that in a random experiment with n trials and having p as the probability of success in each trial,

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

where x is the number of successes. Now, if such a random experiment of n trials is repeated say N times, then the expected (or theoretical) frequency of getting x successes is given by

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

i.e. probability is multiplied by N to get the corresponding expected frequency.

Process of Fitting a Binomial Distribution

Suppose we are given the observed frequency distribution. We first find the mean from the given frequency distribution and equate it to np. From this, we can find the value of p. After having obtained the value of p, we obtain $p(0) = q^n$, where q = 1 – p.

Then the recurrence relation i.e. $p(x+1) = \frac{n-x}{x+1} p(x)$ is applied to find the values of p(1), p(2),…. After that, the expected (theoretical) frequencies f(0), (1), f(2), … are obtained on multiplying each of the corresponding probabilities i.e. p(0), p(1), p(2), … by N.

In this way, the binomial distribution is fitted to the given data. Thus, fitting of a binomial distribution involves comparing the observed frequencies with the expected frequencies to see how best the observed results fit with the theoretical (expected) results.

Example 7: Four coins were tossed and number of heads noted. The experiment is repeated 200 times.

The number of tosses showing 0, 1, 2, 3 and 4 heads were found distributed as under. Fit a binomial distribution to these observed results assuming that the nature of the coins is not known.

Number of heads:	0	1	2	3	4
Number of tosses	15	35	90	40	20

Solution: Here n = 4, N = 200.

First, we obtain the mean of the given frequency distribution as follows:

Number of head X	Number of tosses f	fX
0	15	0
1	35	35
2	90	180
3	40	120
4	20	80
Total	200	415

$\therefore$ Mean = $\frac{\sum f(x)}{\sum f}$ [See Unit 1of MST-002]

$= \frac{415}{200}$

$= 2.075$

As mean for binomial distribution is np,

$\therefore$ np = 2.075

$\Rightarrow p = \frac{2.075}{4}$

$= 0.5188$

$\Rightarrow q = 1 - p$

$= 1 - 0.5188$

$= 0.4812$

$\therefore$ p(0) = q^n

$= (0.4812)^4$

$= 0.0536$

Now, using the recurrence relation

$$p(x+1) = \frac{n-x}{x+1}.\frac{p}{q}p(x); x = 0, 1, 2, 3, 4;$$

we obtain the probabilities for different values of the random variable X i.e. p(1) is obtained on multiplying p(0) with $\frac{4-0}{0+1}$, p(2) is obtained on multiplying p(1) with $\frac{4-1}{1+1}$, and so on; i.e. the values as shown in col. 3 of the following table are obtained on multiplying the preceding values of col. 2 and col 3, except the first value which has been obtained using p(0) = q^n as above.

Number of Heads (X) (1)	$\frac{n-x}{x+1}\cdot\frac{p}{q}=\frac{4-x}{x+1}\left(\frac{0.5188}{0.4812}\right)$ $=\frac{4-x}{x+1}(1.07814)$ (2)	$p(x)$ (3)	Expected or theoretical frequency $f(x)$ (4)
0	$\frac{4-0}{0+1}(1.07814)=4.31256$	p(0) = 0.0536	10.72 ≈ 11
1	$\frac{4-1}{1+1}(1.07814)=1.61721$	p(1) = 4.31256 × 0.0536 = 0.23115	46.23 ≈ 46
2	$\frac{4-2}{2+1}(1.07814)=.071876$	p(2) = 1.61721 × 0.23115 = 0.37382	74.76 ≈ 75
3	$\frac{4-3}{3+1}(1.07814)=0.26954$	p(3) = 0.71876 × 0.37382 = 0.26869	53.73 ≈ 54
4	$\frac{4-4}{4+1}(1.07814)=0$	p(4) = 0.26954 × .26869 = 0.0724	14.48 ≈ 14

Remark 3: In the above example, if the nature of the coins had been known e.g. if it had been given that "the coins are unbiased" then we would have taken

$p = \frac{1}{2}$ and then the observed data would not have been used to find p. Such a situation can be seen in the problem **E6).**

Here are two exercises for you:

E6) Seven coins are tossed and number of heads noted. The experiment is repeated 128 times and the following distribution is obtained:

Number of heads	0	1	2	3	4	5	6	7
Frequencies	7	6	19	35	30	23	7	1

Fit a binomial distribution assuming the coin is unbiased.

E7) Out of 800 families with 4 children each, how many families would you expect to have 3 boys and 1 girl, assuming equal probability of boys and girls?

Now before ending this unit, let's summarize what we have covered in it.

9.6 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. **Third** and **fourth central moments** of this distribution are $p(1-p)(1-2p)$ and $p(1-p)(1-3p+3p^2)$ 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

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

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,

$\mu_3 = npq(q-p)$, $\mu_4 = npq[1 + 3(n-2)pq]$

$\beta_1 = \dfrac{(q-p)^2}{npq}$, $\beta_2 = 3 + \dfrac{1-6pq}{npq}$,

$\gamma_1 = \dfrac{1-2p}{\sqrt{npq}}$, and $\gamma_2 = \dfrac{1-6pq}{npq}$

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

5) **Recurrence relation for the probabilities of binomial distribution** is

$$p(x+1) = \frac{n-x}{x+1} \cdot \frac{p}{q} \cdot p(x),\ x = 0, 1, 2, \ldots, n-1$$

6) The **expected frequencies of the binomial distribution** are given by

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

9.7 SOLUTIONS/ANSWERS

E1) 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

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

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

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

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

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

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

E2) 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

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

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

$\therefore$ The required probability = P[The dacoit is still alive]

= P[No bullet kills the dacoit]

= P[Number of successes is zero]

= P[X = 0] = ${}^6C_0 (0.6)^0 (0.4)^6$

= 0.0041

E3) Mean = np = 3 ... (1)

Variance = npq = 4 ... (2)

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

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

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

E4) Let $X \sim B(n, p)$, then

n = 5 and

$np + npq = 4.8$ [$\because$ given that Mean + Variance = 4.8]

$\Rightarrow 5p + 5pq = 4.8$

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

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

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

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

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

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

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

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

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

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

Thus, the binomial distribution is

$P\left[X = x\right] = {}^{n}C_{x}p^{x}q^{n-x}$

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

The binomial distribution in tabular form is given as

X	p(x)
0	${}^{5}C_{0}\left(\frac{4}{5}\right)^{0}\left(\frac{1}{5}\right)^{5} = \frac{1}{3125}$
1	${}^{5}C_{1}\left(\frac{4}{5}\right)^{1}\left(\frac{1}{5}\right)^{4} = \frac{20}{3125}$
2	${}^{5}C_{2}\left(\frac{4}{5}\right)^{2}\left(\frac{1}{5}\right)^{3} = \frac{160}{3125}$
3	${}^{5}C_{3}\left(\frac{4}{5}\right)^{3}\left(\frac{1}{5}\right)^{2} = \frac{640}{3125}$
4	${}^{5}C_{4}\left(\frac{4}{5}\right)^{4}\left(\frac{1}{5}\right)^{1} = \frac{1280}{3125}$
5	${}^{5}C_{5}\left(\frac{4}{5}\right)^{5}\left(\frac{1}{5}\right)^{0} = \frac{1024}{3125}$

E5) Given that Mean = 30 and S.D. = 5

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

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

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

ii) $\mu_2=npq=180\times\frac{1}{6}\times\frac{5}{6}=25$

$\mu_3=npq\left(q-p\right)=25\left(\frac{5}{6}-\frac{1}{6}\right)=\frac{50}{3}$

$\Rightarrow\beta_1=\frac{\mu_3^2}{\mu_2^3}=\frac{4}{225}$

$\therefore$ Moment coefficient of skewness is given by

$\gamma_1=\sqrt{\beta_1}=\frac{2}{15}$

iii) $\beta_2=3+\frac{1-6pq}{npq}=3+\frac{1-6\times\frac{1}{6}\times\frac{5}{6}}{25}=3+\frac{1}{150}$

$\Rightarrow\ \gamma_2=\beta_2-3=\frac{1}{150}>0$

So, the curve of the binomial distribution is leptokurtic.

E6) As the coin is unbiased, $\therefore\ p=\frac{1}{2}$.

Here, n = 7, N = 128, $p=\frac{1}{2}, q=1-p=\frac{1}{2}$.

$\Rightarrow$ p(0) = $q^n=\left(\frac{1}{2}\right)^7=\frac{1}{128}$.

Expected frequencies are, therefore, obtained as follows:

Number of heads (X)	$\frac{n-x}{x+1}.\frac{p}{q}=\frac{7-x}{x+1}.\frac{\frac{1}{2}}{\frac{1}{2}}$ $=\frac{7-x}{x+1}$	$p\left(x\right)$	Expected or theoretical Frequency $f\left(x\right)=N.p\left(x\right)$ $=128.p\left(x\right)$
0	$\frac{7-0}{0+1}=7$	$\frac{1}{128}$	1
1	$\frac{7-1}{1+1}=3$	$7\times\frac{1}{128}=\frac{7}{128}$	7
2	$\frac{7-2}{2+1}=\frac{5}{3}$	$3\times\frac{7}{128}=\frac{21}{128}$	21
3	$\frac{7-3}{3+1}=1$	$\frac{5}{3}\times\frac{21}{128}=\frac{35}{128}$	35

4	$\frac{7-4}{4+1}=\frac{3}{5}$	$1\times\frac{35}{128}=\frac{35}{128}$	35
5	$\frac{7-5}{5+1}=\frac{1}{3}$	$\frac{3}{5}\times\frac{35}{128}=\frac{21}{128}$	21
6	$\frac{7-6}{6+1}=\frac{1}{7}$	$\frac{1}{3}\times\frac{21}{128}=\frac{7}{128}$	7
7	$\frac{7-7}{7+1}=0$	$\frac{1}{7}\times\frac{7}{128}=\frac{1}{128}$	1

E7) Here, probability (p) to have a boy is $\frac{1}{2}$ and the probability (q) to have a girl is $\frac{1}{2}$, n = 4, N = 800.

Let X be the number of boys in a family.

$\therefore$ by binomial distribution, the probability of having 3 boys in a family of 4 children

= P[X = 3] $\quad\quad$ $\left[\because\ P\left[X=x\right]={}^{n}C_{x}p^{x}q^{n-x}\right]$

$$={}^{4}C_{3}\left(\frac{1}{2}\right)^{3}\left(\frac{1}{2}\right)^{4-3}=\ 4\left(\frac{1}{2}\right)^{4}$$

Hence, the expected number of families having 3 boys and 1 girl

$$=\ N.p\left(3\right)\ =128\left(\frac{1}{4}\right)\ =32$$

UNIT 10 POISSON DISTRIBUTION

Structure

10.1 INTRODUCTION

In Unit 9, you have studied binomial distribution which is applied in the cases where the probability of success and that of 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.

In this unit, we define and explain Poisson distribution in Sec. 10.2. Moments of Poisson distribution are described in Sec. 10.3 and the process of fitting a Poisson distribution is explained in Sec. 10.4.

Objectives

After studing this unit, you would be able to:

- know the situations where Poisson distribution is applied;
- define and explain Poisson distribution;
- know the conditions under which binomial distribution tends to Poisson distribution;
- compute the mean, variance and other central moments of Poisson distribution;
- obtain recurrence relation for finding probabilities of this distribution; and
- know as to how a Poisson distribution is fitted to the observed data.

10.2 POISSON DISTRIBUTION

In case of binomial distributions, as discussed in the last unit, 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:

i) 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.

ii) 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.

iii) 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:

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

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

iii) 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:

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

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 at the end of this unit, or, can be seen from any book of log tables.

Remark 1

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 repectively, then X + Y is also a Poisson variate with parameter $\lambda_1+\lambda_2$. This is known as **additive property of Poisson distribution**.

10.3 MOMENTS OF POISSON DISTRIBUTION

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

$$\mu_r' = E\left(X^r\right) = \sum_{x=0}^{\infty} x^r p\left(x\right) = \sum_{x=0}^{\infty} x^r \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$\mu_1' = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{\lfloor x} = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x\lfloor x-1} = \sum_{x=1}^{\infty} \frac{e^{-\lambda}.\lambda^x}{\lfloor x-1}$$

$$= e^{-\lambda}\left[\frac{\lambda^1}{\lfloor 0} + \frac{\lambda^2}{\lfloor 1} + \frac{\lambda^3}{\lfloor 2} + \ldots\right]$$

$$= \lambda e^{-\lambda}\left[1 + \frac{\lambda^1}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \ldots\right]$$

$$= e^{-\lambda}\lambda e^{\lambda} \quad \left[\because e^{\lambda} = 1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \frac{\lambda^3}{\lfloor 3} + \ldots\left(\text{see Unit 2 of MST-001}\right)\right]$$

$$= \lambda$$

$\therefore$ Mean = λ

$$\mu_2' = \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= \sum_{x=0}^{\infty}\left[x\left(x-1\right) + x\right]\frac{e^{-\lambda}.\lambda^x}{\lfloor x} \qquad \text{[As done in Unit 9 of this Course]}$$

$$= \sum_{x=0}^{\infty}\left(x\left(x-1\right)\frac{e^{-\lambda}\lambda^x}{\lfloor x} + x\frac{e^{-\lambda}\lambda^x}{\lfloor x}\right)$$

$$= \sum_{x=2}^{\infty} x\left(x-1\right)\frac{e^{-\lambda}\lambda^x}{x\left(x-1\right)\lfloor x-2} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda^x}{\lfloor x-2} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda}\left[\frac{\lambda^2}{\lfloor 0} + \frac{\lambda^3}{\lfloor 1} + \frac{\lambda^4}{\lfloor 2} + \ldots\right] + \mu_1'$$

$$= e^{-\lambda}\lambda^2\left[1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \ldots\right] + \mu_1'$$

$$= e^{-\lambda}\lambda^2 e^{\lambda} + \mu_1'$$

$$= \lambda^2 + \lambda$$

$\therefore$ Variance of X is given as V(X) = $\mu_2 = \mu_2' - \left(\mu_1'\right)^2$

$$= \lambda^2 + \lambda - \left(\lambda\right)^2$$

$$= \lambda$$

$$\mu_3' = \sum_{x=0}^{3} x^3 p(x)$$

Writing x^3 as $x(x-1)(x-2)+3x(x-1)+x,$ we have

$\left[\begin{array}{l}\text{See Unit 9 of this course where} \\ \text{the expression of } \mu_3' \text{ is obtained}\end{array}\right]$

$$= \sum_{x=0}^{\infty} \left[x(x-1)(x-2)+3x(x-1)+x\right]\frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= \sum_{x=3}^{\infty} x(x-1)(x-2)\frac{e^{-\lambda}\lambda^x}{\lfloor x} + 3\sum_{x=2}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{\lfloor x} + \sum_{x=1}^{\infty} x\frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda}\sum_{x=3}^{\infty} x(x-1)(x-2)\frac{\lambda^x}{x(x-1)(x-2)\lfloor x-3} + 3(\lambda^2) + (\lambda)$$

$$= e^{-\lambda}\sum_{x=3}^{\infty} \frac{\lambda^x}{\lfloor x-3} + 3\lambda^2 + \lambda$$

$$= e^{-\lambda}\left(\frac{\lambda^3}{\lfloor 0} + \frac{\lambda^4}{\lfloor 1} + \frac{\lambda^5}{\lfloor 2} + \ldots\right) + 3\lambda^2 + \lambda$$

$$= e^{-\lambda}\lambda^3\left(1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \ldots\right) + 3\lambda^2 + \lambda$$

$$= e^{-\lambda}\lambda^3 e^{\lambda} + 3\lambda^2 + \lambda$$

$$= \lambda^3 + 3\lambda^2 + \lambda$$

Third order central moment is

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\left(\mu_1'\right)^3$$

$$= \lambda \qquad \text{[On simplification]}$$

$$\mu_4' = \sum_{x=3}^{\infty} x^4 . \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

Now writing $x^4 = x(x-1)(x-2)(x-3)+6x(x-1)(x-2)+7x(x-1)+x,$

and proceeding in the similar fashion as done in case of μ_3', we have

$$\mu_4' = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$$

$\therefore$ Fourth order central moment is

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\left(\mu_1'\right)^2 - 3\left(\mu_1'\right)^4$$

$$= 3\lambda^2 + \lambda \qquad \text{[On simplification]}$$

Therefore, measures of skewness and kurtosis are given by

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda}, \qquad \gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}}; \text{ and}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\lambda^2 + \lambda}{\lambda^2} = 3 + \frac{1}{\lambda}, \qquad \gamma_2 = \beta_2 - 3 = \frac{1}{\lambda}.$$

Now as γ_1 is positive, therefore the Poisson distribution is always positively skewed distribution. Also as $\gamma_2 >$ ($\because$ $\lambda > 0$), the curve of the distribution is Leptokurtic.

Remark 2

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,

Mean = np

Variance = npq

$$\mu_3 = npq(q-p)$$

$$\mu_4 = npq\left[1+3pq(n-2)\right] = npq\left[1+3npq-6pq\right]$$

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

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

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

Mean = Limiting value of np = λ

Variance = Limiting value of npq

= Limiting value of (np) (q)

= $(\lambda)(1) = \lambda$

μ_3 = Limiting value of npq (q – p)

= Limiting value of (npq) (q – p)

= $(\lambda)(1 - 0)$

= λ

μ_4 = Limiting value of npq [1 + 3npq – 6pq]

= Limiting value of (npq) [1 + 3(npq) –6 (p)(q)]

= $(\lambda)[1 + 3(\lambda) - 6(0)(1)]$

= $\lambda[1 + 3\lambda] = 3\lambda^2 + \lambda$

Now let's give some examples of Poisson distribution.

Example 1: 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

a) no heavy truck arrive,

b) 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

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

Thus, the desired probabilities are:

(a) P[arrival of no heavy truck] = P[X = 0]

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

$$= e^{-2}$$

$$= 0.1353 \quad \begin{bmatrix} \text{See the table given} \\ \text{in the Appendix at} \\ \text{the end of this unit} \end{bmatrix}$$

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

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

$$= 1 - \left[P[X = 1] + P[X = 0]\right]$$

$$\begin{bmatrix} \because \text{ sum of all the} \\ \text{probabilities is 1} \end{bmatrix}$$

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

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

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

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 2: 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,

$$P[X = x] = \frac{e^{-\lambda}\lambda^x}{\underline{|x}}, x = 0, 1, 2, ...$$

$$= \frac{e^{-1.5}.(1.5)^x}{\underline{|x}}; x = 0, 1, 2, ...$$

Thus,

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

$$= \frac{e^{-1.5}.(1.5)^3}{\underline{|3}}$$

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

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

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

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

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

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

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

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

Example 3: 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 = λ = 1.44

$\mu_3 = \lambda = 1.44$

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

Example 4: 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,

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

Now, $P[X = 1] = 2P[X = 2]$

$$\Rightarrow \frac{e^{-\lambda}.\lambda^1}{\lfloor 1} = 2\frac{e^{-\lambda}.\lambda^2}{\lfloor 2}$$

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

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 = λ = 1, and

Variance = λ = 1.

Example 5: If X and Y be two independent Poisson variates having means 1 and 2 respectively, find P[X + Y < 2].

Solution: As X ~ P(1), Y ~ P(2), therefore,

X + Y follows Poisson distribution with mean = 1 + 2 = 3.

Let X + Y = W. Hence, probability function of W is

$$P[W = w] = \frac{e^{-3}.3^w}{\lfloor w}; w = 0, 1, 2, \ldots.$$

Thus, the required probability= P[X + Y < 2]

= P[W < 2]

= P[W = 0] + P[W = 1]

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

$= (0.0498)(1 + 3)$ [From Table, $e^{-3} = 0.0498$]

$= 0.1992.$

You may now try these exercises.

E1) 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 $e^{-0.25} = 0.78$)

E2) The mean and standard deviation of a Poisson distribution are 6 and 2 respectively. Test the validity of this statement.

E3) For a Poisson distribution, it is given that P[X = 1] = P[X = 2], find the value of mean of distribution. Hence find P[X = 0] and P[X = 4].

We now explain as to how the Poisson distribution is fitted to the observed data.

10.4 FITTING OF POISSON DISTRIBUTION

To fit a Poisson distribution to the observed data, we find the theoretical (or expected) frequencies corresponding to each value of the Poisson variate. Process of finding the probabilities corresponding to each value of the Poisson variate becomes easy if we use the recurrence relation for the probabilities of Poisson distribution. So, in this section, we will first establish the recurrence relation for probabilities and then define the Poisson frequency distribution followed by the process of fitting a Poisson distribution.

Recurrence Formula for the Probabilities of Poisson Distribution

For a Poisson distribution with parameter λ, we have

$$p(x) = \frac{e^{-\lambda}\lambda^x}{\lfloor x} \qquad \ldots (1)$$

Changing x to x + 1, we have

$$p(x+1) = \frac{e^{-\lambda}\lambda^{x+1}}{\lfloor x+1} \qquad \ldots (2)$$

Dividing (2) by (1), we have

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

$$\Rightarrow p(x+1) = \frac{\lambda}{x+1} p(x) \qquad \ldots (3)$$

This is the recurrence relation for probabilities of Poisson distribution. After obtaining the value of p(0) using Poisson probability function i.e.

$p(0) = \frac{e^{-\lambda}\lambda^0}{(0)} = e^{-\lambda}$, we can obtain p(1), p(2), p(3),…, on putting x = 0, 1, 2, …. successively in (3).

Poisson Frequency Distribution

If an experiment, satisfying the requirements of Poisson distribution, is repeated N times, then the expected frequency of getting x successes is given by

$$f(x) = N.P[X = x] = N.\frac{e^{-\lambda}\lambda^x}{\lfloor x};\ x = 0, 1, 2, \ldots$$

Example 5: 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 at least two defective bottles.

Solution: Let X be the Poisson variate, "the number of defective bottles in a box". Here, number of bottles in a box (n) = 500, therefore, the probability (p) of a bottle being defective is

$$p = 0.1\% = \frac{0.1}{100} = 0.001$$

Number of boxes (N) = 100

$$\lambda = np = 500 \times .001 = 0.5$$

Using Poisson distribution, we have

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

$$= \frac{e^{-0.5}(0.5)^x}{\lfloor x};\ x = 0, 1, 2, \ldots$$

$\therefore$ Probability that a box contain at least two defective bottles

$$= P[X \ge 2]$$

$$= 1 - P[X < 2]$$

$$= 1 - \left[P[X = 0] + P[X = 1]\right]$$

$$= 1 - \left[\frac{e^{-0.5}(0.5)^0}{\lfloor 0} + \frac{e^{-0.5}(0.5)^1}{\lfloor 1}\right] = 1 - e^{-0.5}[1 + 0.5]$$

$$= 1 - (0.6065)(1.5) = 1 - 0.90975 = 0.09025.$$

Hence, the expected number of boxes containing at least two defective bottles

$= N.P[X \geq 2]$

$= (100)\ (0.09025)$

$= 9.025$

Process of Fitting a Poison Distribution

For fitting a Poisson distribution to the observed data, you are to proceed as described in the following steps.

- First we obtain mean of the given distribution i.e. $\frac{\sum fx}{\sum f}$, being mean, take this as the value of λ.
- Next we obtain $p(0) = e^{-\lambda}$ [Use table given in Appendix at the end of this unit.]
- The recurrence relation $p(x+1) = \frac{\lambda}{x+1} p(x)$ is then used to compute the values of p(1), p(2), p(3), …
- The probabilities obtained in the preceding two steps are then multiplied with N to get expected/theoretical frequencies i.e. $f(x) = N.P[X = x]; x = 0, 1, 2, ...$

Example 6: The following data give frequencies of aircraft accidents experienced by 2480 pilots during a certain period:

Number of Accidents	0	1	2	3	4	5
Frequencies	1970	422	71	13	3	1

Fit a Poisson distribution and calculate the theoretical frequencies.

Solution: Let X be the number of accidents of the pilots. Let us first obtain the mean number of accidents as follows:

Number of Accidents (X)	Frequency (f)	f X
0	1970	0
1	422	422
2	71	142
3	13	39
4	3	12
5	1	5
Total	2480	620

$\therefore$ Mean = $\lambda = \dfrac{\sum fx}{\sum f} = \dfrac{620}{2480}$

$\Rightarrow \lambda = 0.25$

$\therefore$ by Poisson distribution,

$p(0) = e^{-\lambda} = e^{-0.25}$

$= 0.7788 \quad \begin{bmatrix} \text{See table given in the Appendix} \\ \text{at the end of this unit} \end{bmatrix}$

Now, using the recurrence relation for probabilities of Poisson distribution i.e. $p(x+1) = \dfrac{\lambda}{x+1} p(x)$ and then multiplying each probability with N, we get the expected frequencies as shown in the following table

Number of Accidents (X) (1)	$\dfrac{\lambda}{x+1} = \dfrac{0.25}{x+1}$ (2)	$p(x) = P[X = x]$ (3)	Expected/ Theoretical frequency f(x) = 2480p(x) (4)
0	$\dfrac{0.25}{0+1} = 0.25$	p(0) = 0.7788	1931. 4 $\approx$ 1931
1	$\dfrac{0.25}{1+1} = 0.125$	p(1) = 0.25 × 0.7788 = 0.1947	482. 9 $\approx$ 483
2	$\dfrac{0.25}{2+1} = 0.0833$	p(2) = 0.125 × 0.1947 = 0.0243	60.3 $\approx$ 60
3	$\dfrac{0.25}{3+1} = 0.0625$	p(3)= 0.0833 × 0.0243 = 0.0020	4.96 $\approx$ 5
4	$\dfrac{0.25}{4+1} = 0.05$	p(4)= 0.0625 × 0.0020 = 0.0001	0.248 $\approx$ 0
5	$\dfrac{0.25}{5+1} = 0.0417$	p(5)= 0.05 × 0.0001 = 0.000005	0

You can now try the following exercises

E4) In a certain factory turning out fountain pens, there is a small chance, $\dfrac{1}{500}$, for any pen to be defective. The pens are supplied in packets of 10. Calculate the approximate number of packets containing (i) one defective (ii) two defective pens in a consignment of 20000 packets.

E5) A typist commits the following mistakes per page in typing 100 pages. Fit a Poisson distribution and calculate the theoretical frequencies.

Mistakes per page(X)	0	1	2	3	4	5
Frequency (f)	42	33	14	6	4	1

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

10.5 SUMMARY

The following main points have been covered in this unit:

1. 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:

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

2. For Poisson distribution, **Mean = Variance = μ_3** $= \lambda$, $\mu_4 = 3\lambda^2 + \lambda$

3. $\beta_1 = \dfrac{1}{\lambda}$, $\gamma_1 = \dfrac{1}{\sqrt{\lambda}}$, $\beta_2 = 3 + \dfrac{1}{\lambda}$, $\gamma_2 = \dfrac{1}{\lambda}$ for this distribution.

4. **Recurrence relation for probabilities of Poisson distribution** is

$$p(x+1) = \frac{\lambda}{x+1}.p(x),\ x = 0, 1, 2, 3, \ldots$$

5. **Expected frequencies for a Poisson distribution** are given by

$$f(x) = N.P[X = x] = N.\frac{e^{-\lambda}\lambda^{x}}{\underline{|x}}; x = 0, 1, 2, \ldots$$

If you want to see what our solutions/answers to the exercises in the unit are, we have given them in the following section.

10.6 SOLUTIONS/ANSWERS

E1) Let X be the Poisson variable "Number of fatal accidents in a year".

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

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

By Poisson distribution,

$$P[X = x] = \frac{e^{-\lambda}.\lambda^x}{\lfloor x}, x = 0, 1, 2, \ldots$$

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

Therefore, P [at least one fatal accident]

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

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

E2) 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.

E3) Let λ be the mean of the distribution,

$\therefore$ by Poisson distribution, we have

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

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

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

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

$\Rightarrow \lambda = 0, 2.$

$\lambda = 0$ is rejected,

$\therefore \lambda = 2$

Hence, Mean = 2.

Now, $P[X = 0] = \frac{e^{-\lambda}\lambda^0}{\lfloor 0} = e^{-\lambda} = e^{-2} = 0.1353,$

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

and $$P[X = 4] = \frac{e^{-\lambda}\lambda^4}{\lfloor 4} = \frac{e^{-2}(2)^4}{\lfloor 4} = \frac{e^{-2}(16)}{24} = \frac{2}{3}(0.1353)$$

$$= 2(0.0451)$$

$$= 0.0902.$$

E4) Here $p = \frac{1}{500}$, $n = 10$, $N = 20000$,

$\therefore \lambda = np = 10 \times \frac{1}{500} = 0.02$

By Poisson frequency distribution

$f(x) = N.P[X = x]$

$= (20000)\frac{e^{-\lambda}\lambda^{x}}{\underline{|x}}; x = 0, 1, 2, \ldots$

Now,

i) The number of packets containing one defective

$= f(1)$

$= (20000)\frac{e^{-0.02}.(0.02)^{1}}{\underline{|1}}$

$= (20000)\ (0.9802)\ (0.02)$ $\left[\begin{array}{l}\text{See the table given}\\ \text{in the Appendix}\end{array}\right]$

$= 392.08 \sqcup 392$; and

ii) The number of packets containing two defectives

$= f(2) = 20000\frac{e^{-0.02}(0.02)^{2}}{\underline{|2}}$

$= (20000)\frac{(0.9802)(0.0004)}{2} = 3.9208 \sqcup 4$

E5) The mean of the given distribution is computed as follows

X	f	fX
0	42	0
1	33	33
2	14	28
3	6	18
4	4	16
5	1	5
Total	100	100

$\therefore$ Mean $\lambda = \frac{\sum fx}{\sum f} = \frac{100}{100} = 1$

$\Rightarrow p(0) = e^{-\lambda} = e^{-1} = 0.3679.$

Now, we obtain p(1), p(2), p(3), p(4), p(5) using the recurrence relation for probabilities of Poisson distribution i.e.

$p(x+1)=\frac{\lambda}{x+1}p(x); x=0, 1, 2, 3, 4$ and then obtain the expected frequencies as shown in the following table:

X	$\frac{\lambda}{x+1}=\frac{1}{x+1}$	$p(x)$	Expected/Theoretical frequency $f(x)=N.P(X=x)$ $=100.P(X=x)$
0	$\frac{1}{0+1}=1$	$p(0)=0.3679$	36.79 □ 37
1	$\frac{1}{1+1}=0.5$	$p(1)=1\times 0.3679=0.3679$	36.79 □ 37
2	$\frac{1}{2+1}=0.3333$	$p(2)=0.5\times 0.3679=0.184$	18.4 □ 18
3	$\frac{1}{3+1}=0.25$	$p(3) = 0.3333\times 0.184 = 0.0613$	6.13 □ 6
4	$\frac{1}{4+1}=0.2$	$p(4)=0.25 \times 0.0613 = 0.0153$	1.53 □ 2
5	$\frac{1}{5+1}=0.1667$	$p(5)=0.2 \times 0.0153 = 0.0031$	0.3 □ 0

Appendix

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

(0 < λ < I)

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	06703	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
(λ=1, 2, 3, ...,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.

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

UNIT 11 DISCRETE UNIFORM AND HYPERGEOMETRIC DISTRIBUTIONS

Structure

11.1 INTRODUCTION

In the previous two units, we have discussed binomial distribution and its limiting form i.e. Poisson distribution. Continuing the study of discrete distributions, in the present unit, two more discrete distributions – Discrete uniform and Hypergeometric distributions are discussed.

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 Hypergemetric distribution.

In this unit, discrete uniform distribution and hypergeometric distribution are discussed in Secs. 11.2 and 11.3, respectively. We shall be discussing their properties and applications also in these sections.

Objectives

After studing this unit, you should be able to:

- define the discrete uniform and hypergeometric distributions;
- compute their means and variances;
- compute probabilities of events associated with these distributions; and
- know the situations where these distributions are applicable.

11.2 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

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

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

$$\text{Mean} = E(X) = \sum_{x=1}^{n} x\, p(x) = \sum_{x=1}^{n} x.\left(\frac{1}{n}\right) = \frac{1}{n}\sum_{x=1}^{n} x$$

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

$$= \frac{1}{n}.\frac{n(n+1)}{2} \quad \begin{bmatrix} \because \text{sum of first n natural numbers} = \frac{n(n+1)}{2} \\ \text{(see Unit 3 of Course MST} - 001) \end{bmatrix}$$

$$= \frac{n+1}{2}.$$

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

where

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

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

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

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

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

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

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

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

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

Example 1: 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

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

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

Example 2: 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".

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

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 $f(x) = N.P[X = x] = 120.P[X = x]$
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	$\frac{1}{6}$	$120 \times \frac{1}{6} = 20$

Now, you can try the following exercise:

E1) 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.

11.3 HYPERGEOMETRIC DISTRIBUTION

In the last section of this unit, we have studied discrete uniform probability distribution wherein the probability distribution is obtained for the possible outcomes in a single trial like drawing a ticket from an urn containing 10 tickets as mentioned in exercise **E1).** But, 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 which you have studied in Unit 9. 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 on odd number appears in 3 draws, then by binomial distribution,

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

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

$$= P\left[A_1 \cap A_2 \cap \bar{A}_3\right] + P\left[A_1 \cap \bar{A}_2 \cap A_3\right] + P\left[\bar{A}_1 \cap A_2 \cap A_3\right]$$

[As done in Unit 3 of this Course]

$$= P[A_1]P[A_2 \mid A_1]P\left[\bar{A}_3 \mid A_1 \cap A_2\right] + P[A_1]P\left[\bar{A}_2 \mid A_1\right]P\left[A_3 \mid A_1 \cap \bar{A}_2\right]$$
$$+ P\left[\bar{A}_1\right]P\left[A_2 \mid \bar{A}_1\right]P\left[A_3 \mid \bar{A}_1 \cap A_2\right]$$

[Multiplication theorem for dependent events (See Unit 3 of this Course)]

$$= \frac{5}{10}.\frac{4}{9}.\frac{5}{8} + \frac{5}{10}.\frac{5}{9}.\frac{4}{8} + \frac{5}{10}.\frac{5}{9}.\frac{4}{8}$$

$$= 3 \times \frac{5\times5\times4}{10\times9\times8}$$

This result can be written in the following form also:

$$= \frac{5\times4\times5\times3\times2}{2\times10\times9\times8} \qquad \text{[Multiplying and Dividing by 2]}$$

$$= \frac{5\times4}{2}\times5\times\frac{1}{\dfrac{10\times9\times8}{3\times2}} = {}^5C_2 \times {}^5C_1 \times \frac{1}{{}^{10}C_3} = \frac{{}^5C_2 \times {}^5C_1}{{}^{10}C_3}$$

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

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

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

Let us now generalize 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

$$P[X = x] = \frac{{}^{M}C_x . {}^{N-M}C_{n-x}}{{}^{N}C_n}$$

[For $x = 0, 1, 2, ..., n(n \le M)$ or $x = 0, 1, 2, ..., M(n > M)$]

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

Remark 1: 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 2: If number (n) of balls drawn is greater than the number (M) of white balls in the bag, then if $n \le M$, 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 $(\text{if } n \le M)$ 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

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

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

Mean and Variance

$$\text{Mean} = E(X) = \sum_{x=0}^{n} x.p[X = x]$$

$$= \sum_{x=1}^{n} x.\frac{{}^{M}C_x . {}^{N-M}C_{n-x}}{{}^{N}C_n}$$

$$= \sum_{x=1}^{n} x.\frac{M}{x}.\frac{{}^{M-1}C_{x-1}.{}^{N-M}C_{n-x}}{{}^{N}C_{n}}$$

$$= \frac{M}{{}^{N}C_{n}}\sum_{x=1}^{n}\left({}^{M-1}C_{x-1}.{}^{N-M}C_{n-x}\right)$$

$$= \frac{M}{{}^{N}C_{n}}\left[{}^{M-1}C_{0}.{}^{N-M}C_{n-1} + {}^{M-1}C_{1}.{}^{N-M}C_{n-2} + \ldots + {}^{M-1}C_{n-1}.{}^{N-M}C_{0}\right]$$

$$= \frac{M}{{}^{N}C_{n}}\left({}^{N-1}C_{n-1}\right)$$

[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. However, the result used in the above expression is enrectangled below for the interesting learners.]

> We know that
>
> $$(1+x)^{m+n} = (1+x)^{m}.(1+x)^{n} \qquad \text{[By the method of indices]}$$
>
> Expanding using binomial theorem as explained in Unit 9 of this course, we have
>
> $${}^{m+n}C_{0}.\,x^{m+n} + {}^{m+n}C_{1}.x^{m+n-1} + {}^{m+n}C_{2}.x^{m+n-2} + \ldots + {}^{m+n}C_{m+n}$$
>
> $$= \left({}^{m}C_{0}x^{m} + {}^{m}C_{1}\,x^{m-1} + {}^{m}C_{2}\,x^{m-2} + \ldots + {}^{m}C_{m}\right)$$
>
> $$.\left({}^{n}C_{0}x^{n} + {}^{n}C_{1}\,x^{n-1} + {}^{n}C_{2}\,x^{n-2} + \ldots + {}^{n}C_{n}\right)$$
>
> Comparing coefficients of X^{m+n-r}, we have
>
> $${}^{m+n}C_{r} = \left({}^{m}C_{0}.{}^{n}C_{r} + {}^{m}C_{1}.{}^{n}C_{r-1} + \ldots + {}^{m}C_{r}.{}^{n}C_{0}\right)$$

$$= \frac{M\underline{|n}\,\underline{|N-n}}{\underline{|N}}.\frac{\underline{|N-1}}{\underline{|N-n}\,\underline{|n-1}}$$

$$= \frac{M.n\underline{|n-1}}{N.\underline{|N-1}}.\frac{\underline{|N-1}}{\underline{|n-1}} = \frac{nM}{N}.$$

$$E\left(X^{2}\right) = E\left[X\left(X-1\right)+X\right]$$

$$= E\left[X\left(X-1\right)\right] + E\left(X\right)$$

$$= \left[\sum_{x=0}^{n} x\left(x-1\right).\frac{{}^{M}C_{x}.{}^{N-M}C_{n-x}}{{}^{N}C_{n}}\right] + \left(\frac{nM}{N}\right)$$

$$= \sum_{x=0}^{n}\left[x\left(x-1\right).\frac{M}{x}.\frac{M-1}{x-1}.\frac{{}^{M-2}C_{x-2}.{}^{N-M}C_{n-x}}{{}^{N}C_{n}}\right] + \left(\frac{nM}{N}\right)$$

$$= \frac{M(M-1)}{^N C_n}\left[\sum_{x=0}^{n}\left({}^{M-2}C_{x-2}.{}^{N-M}C_{n-x}\right)\right]+\left(\frac{nM}{N}\right)$$

$$= \frac{M(M-1)}{^N C_n}\left({}^{N-2}C_{n-2}\right)+\left(\frac{nM}{N}\right)$$

[The result in the first term has been obtained using a property of binomial coefficients as done above for finding E(X).]

$$= \frac{M(M-1)\underline{|N-n}\,\underline{|n}}{\underline{|N}}.\frac{\underline{|N-2}}{\underline{|n-2}\,\underline{|N-n}}+\frac{nM}{N}$$

$$= \frac{M(M-1)n(n-1)}{N(N-1)}+\frac{nM}{N}$$

Thus,

$$V(X)=E(X^2)-[E(X)]^2=\frac{M(M-1)n(n-1)}{N(N-1)}+\frac{nM}{N}-\left(\frac{nM}{N}\right)^2$$

$$= \frac{NM(N-M)(N-n)}{N^2(N-1)} \qquad \text{[On simplification]}$$

Example 2: 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:

P **[**2 non-graduates and 3 graduates**]** $= \frac{^{60}C_2.\,^{40}C_3}{^{100}C_5}$

$$= \frac{60\times59\times40\times39\times38\times5\times4\times3\times2}{2\times6\times100\times99\times98\times97\times96}$$

$$= 0.2323$$

Example 3: 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 making 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 marked fish and 300 others] = $\frac{^{500}C_{100}.\,^{N-500}C_{300}}{^N C_{400}}$.

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 Unit 2 of MST-004 We are not going to estimate it. You may try it as an exercise after reading Unit 2 of MST-004.

Here, let us take an assumed value of N say 5000.

Then,

$$P[X=100]=\frac{{}^{500}C_{100}\cdot{}^{4500}C_{300}}{{}^{5000}C_{400}}$$

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.

You may now try the following exercise.

E2) 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.

11.4 SUMMARY

The following main points have been covered in this unit:

1) 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

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

where n is called the parameter of the distribution.

2) For **discrete uniform** distribution, **mean** = $\frac{n+1}{2}$ and **variance** = $\frac{n^2-1}{12}$.

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

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

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

4) For **hypergeometric** distribution, **mean** $=\frac{nM}{N}$ and

variance $=\frac{NM(N-M)(N-n)}{N^2(N-1)}$.

11.5 SOLUTIONS/ANSWERS

E1) 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	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
4	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
5	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
6	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
7	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
8	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
9	$\frac{1}{10}$	$150\times\frac{1}{10}=15$
10	$\frac{1}{10}$	$150\times\frac{1}{10}=15$

E2) Here N = 25, M = 10 and n = 2.

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

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

UNIT 12 GEOMETRIC AND NEGATIVE BINOMIAL DISTRIBUTIONS

Structure

12.1 INTRODUCTION

In Units 9 and 11, we have studied the discrete distributions – Bernoulli, Binomial, Discrete Uniform and Hypergeometric. In each of these distributions, the random variable takes finite number of values. There may also be situations where the discrete random variable assumes countably infinite values. Poisson distribution, wherein discrete random variable takes an indefinite number of values with very low probability of occurrence of event, has already been discussed in Unit 10. Dealing with some more situations where discrete random variable assumes countably infinite values, we, in the present unit, discuss geometric and negative binomial distributions. It is pertinent to mention here that negative binomial distribution is a generalization of geometric distribution. Some instances where these distributions can be applied are "deaths of insects", "number of insect bites".

Like binomial distribution, geometric and negative binomial distributions also have independent trials with constant probability of success in each trial. But, in binomial distribution, the number of trials (n) is fixed whereas in geometric distribution, trials are performed till first success and in negative binomial distribution trials are performed till a certain number of successes.

Secs. 12.2 and 12.3 of this unit discuss geometric and negative binomial distribution, respectively along with their properties.

Objectives

After studing this unit, you would be able to:

- define the geometric and negative binomial distributions;
- calculate the mean and variance of these distributions;
- compute probabilities of events associated with these distributions;
- identify the situations where these distributions can be applied; and
- know about distinguishing features of these distributions like memoryless property of geometric distribution.

12.2 GEOMETRIC DISTRIBUTION

Let us consider Bernoulli trials i.e. independent trials having the constant probability 'p' of success in each trial. Each trial has two possible outcomes – success or failure. Now, suppose the trial is performed repeatedly till we get the success. Let X be the number of failures preceding the first success. Example of such a situation is "tossing a coin until head turns up". X defined above may take the values 0, 1, 2, …. Letting q be the probability of failure in each trial, we have

$$P[X=0] = P[\text{Zero failure preceding the first success}]$$

$$= P(S)$$

$$= p,$$

$$P[X=1] = P[\text{One failure preceding the first success}]$$

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

$$= P(F)\,P(S) \;\; [\because \text{trials are independent}]$$

$$= qp$$

$$P[X = 2] = P[\text{Two failures preceding the first success}]$$

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

$$= P(F)\,P(F)\,P(S)$$

$$= qqp$$

$$= q^2 p$$

and so on.

Therefore, in general, probability of x failures preceding the first success is

$$P[X = x] = q^x p;\;\; x = 0, 1, 2, 3, \ldots$$

Notice that for $x = 0, 1, 2, 3, \ldots$ the respective probabilities $p, qp, q^2p, q^3p, \ldots$ are the terms of geometric progression series with common ratio q. That is why, the above probability distribution is known as geometric distribution [see Unit 3 of MST-001].

Hence, the above discussion leads to the following definition:

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

$$P[X=x] = \begin{cases} q^x p & \text{for } x = 0, 1, 2, \ldots \\ 0, & \text{otherwise} \end{cases}$$

Notice that

$$\sum_{x=0}^{\infty} q^x p = p + q\,p + q^2 p + q^3 p + \ldots$$

$$= p[1 + q + q^2 + q^3 + \ldots]$$

$$= p\left(\frac{1}{1-q}\right) = \frac{p}{p} = 1$$

[$\because$ sum of infinite terms of G.P. $= \frac{a}{1-r} = \frac{1}{1-q}$ (see Unit 3 of MST-001)]

Now, let us take up some examples of this distribution.

Example 1: An unbiased die is cast until 6 appear. What is the probability that it must be cast more than five times?

Solution: Let p be the probability of a success i.e. getting 6 in a throw of the die

$\therefore\ p = \frac{1}{6}$ and $q = 1 - p = \frac{5}{6}$

Let X be the number of failures preceding the first success.

$\therefore$ by geometric distribution,

$$P[X = x] = q^x p;\ x = 0, 1, 2, 3, \ldots$$

$$= \left(\frac{5}{6}\right)^x \left(\frac{1}{6}\right) \text{ for } x = 0, 1, 2, 3, \ldots$$

Thus, the desired probability = P[The die is to be cast more than five times]

= P [The number of throws is at least 6]

$$= P\left[\begin{array}{l}\text{The number of failures preceding}\\ \text{the first success is at least 5}\end{array}\right]$$

$= P[X \geq 5]$

$= P[X = 5] + P[X = 6] + P[X = 7] + \ldots$

$$= \left(\frac{5}{6}\right)^5\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)^6\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)^7\left(\frac{1}{6}\right) + \ldots$$

$$= \left(\frac{5}{6}\right)^5\left(\frac{1}{6}\right)\left[1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \ldots\right]$$

$$= \left(\frac{5}{6}\right)^5\left(\frac{1}{6}\right)\left[\frac{1}{1-\frac{5}{6}}\right] = \left(\frac{5}{6}\right)^5$$

Let us now discuss some properties of geometric distribution.

Mean and Variance

Mean of the geometric distribution is given as

$$\text{Mean} = E(X) = \sum_{x=0}^{\infty} x\,q^x p = p\sum_{x=1}^{\infty} x\,q^x = p\sum_{x=1}^{\infty} x\,q^{x-1}.q$$

$$= pq\sum_{x=1}^{\infty} x\, q^{x-1} = p\,q\sum_{x=1}^{\infty}\frac{d}{dq}\left(q^{x}\right)$$

$\frac{d}{dq}\left(q^{x}\right)$ is the differentiation of q^x w.r.t. q where x is kept as constant

[$\because \frac{d}{dx}\left(x^{m}\right) = mx^{m-1}$, where m is constant (see Unit 6 of MST-001)]

$$= pq\frac{d}{dq}\left[\sum_{x=1}^{\infty} q^{x}\right] \qquad \left[\begin{array}{l}\because \text{sum of the derivatives is} \\ \text{the derivatives of the sums}\end{array}\right]$$

$$= pq\frac{d}{dq}\left[\sum_{x=1}^{\infty} q^{x}\right]$$

$$= pq\frac{d}{dq}\left[q + q^{2} + q^{3} + \ldots\right]$$

$$= pq\frac{d}{dq}\left[\frac{q}{1-q}\right]$$

$$= pq\left[\frac{(1-q) - q(-1)}{(1-q)^{2}}\right] \qquad \left[\begin{array}{l}\text{Applying quotient rule} \\ \text{of differentiation}\end{array}\right]$$

$$= pq\left[\frac{1-q+q}{p^{2}}\right]$$

$$= \frac{q}{p}. \qquad \ldots (1)$$

Variance of the geometric distribution is

$V(X) = E(X^2) – [E(X)]^2$,

where

$$E(X^2) = \sum_{x=\infty}^{\infty} x^{2} p(x)$$

$$= \sum_{x=0}^{\infty}\left[x(x-1) + x\right]p(x)$$

[$\because x^{2} = x(x-1) + x$ (it has already been discussed in Unit 9)]

$$= \sum_{x=0}^{\infty} x(x-1)p(x) + \sum_{x=0}^{\infty} x\,p(x)$$

$$= \sum_{x=2}^{\infty} x(x-1)q^{x}p + \left(\frac{q}{p}\right) \qquad \text{[Using (1) in second term]}$$

$$= pq^{2}\sum_{x=2}^{\infty} x(x-1)q^{x-2} + \left(\frac{q}{p}\right) \qquad [\because q^{x} = q^{x-2}.q^{2}]$$

$$= pq^2 \sum_{x=2}^{\infty} \frac{d^2}{dq^2}\left(q^x\right) + \frac{q}{p} \quad \left[\begin{array}{l} \because \dfrac{d^2}{dq^2}\left(q^x\right) = \dfrac{d}{dq}\left(\dfrac{d}{dq}q^x\right) \\ = \dfrac{d}{dq}\left(xq^{x-1}\right) = x\left(x-1\right)q^{x-2} \\ \text{treating x as constant} \end{array}\right]$$

$$= pq^2 \frac{d^2}{dq^2}\left(\sum_{x=2}^{\infty} q^x\right) + \frac{q}{p}$$

$$= pq^2 \frac{d^2}{dq^2}\left[q^2 + q^3 + q^4 + \ldots\right] + \frac{q}{p}$$

$$= pq^2 \frac{d^2}{dq^2}\left[\frac{q^2}{1-q}\right] + \frac{q}{p}$$

$$= pq^2 \frac{d}{dq}\left[\frac{(1-q)2q - q^2(-1)}{(1-q)^2}\right] + \frac{q}{p}$$

$$= pq^2 \frac{d}{dq}\left[\frac{2q - 2q^2 + q^2}{(1-q)^2}\right] + \frac{q}{p}$$

$$= pq^2 \frac{d}{dq}\left[\frac{2q - q^2}{(1-q)^2}\right] + \frac{q}{p}$$

$$= pq^2\left[\frac{(1-q)^2(2-2q) - \left(2q - q^2\right).2(1-q)^1(-1)}{(1-q)^4}\right] + \frac{q}{p}$$

$$= pq^2\left[\frac{(1-q)\left\{2(1-q)^2 + 2\left(2q - q^2\right)\right\}}{(1-q)^4}\right] + \frac{q}{p}$$

$$= pq^2.\frac{p\left[2p^2 + 2q(2-q)\right]}{p^4} + \frac{q}{p} \qquad \text{as } p = 1 - q$$

$$= 2\left(\frac{q}{p}\right)^2\left[p^2 + q(2-q)\right] + \frac{q}{p}$$

$$= 2\left(\frac{q}{p}\right)^2\left[(1-q)^2 + q(2-q)\right] + \frac{q}{p}$$

$$= 2\left(\frac{q}{p}\right)^2\left[1 + q^2 - 2q + 2q - q^2\right] + \frac{q}{p}$$

$$= 2\frac{q^2}{p^2}(1) + \frac{q}{p} = \frac{2q^2}{p^2} + \frac{q}{p}$$

$$\therefore \quad V(X) = E(X^2) - [E(X)]^2$$

$$= \frac{2q^2}{p^2} + \frac{q}{p} - \left(\frac{q}{p}\right)^2$$

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

$$= \frac{q}{p}\cdot\frac{1}{p} = \frac{q}{p^2}$$

Remark 1: Variance = $\frac{q}{p^2} = \frac{q}{p.p} = \frac{\text{Mean}}{p}$

$\Rightarrow$ Variance > Mean $\qquad [\because p < 1 \Rightarrow \frac{\text{Mean}}{p} > \text{Mean}]$

Hence, unlike binomial distribution, variance of the geometric distribution is greater than mean.

Example 2: Comment on the following:

The mean and variance of geometric distribution are 4 and 3 respectively.

Solution: If the given geometric distribution has parameter p(probability of success in each trial).

Then,

Mean = $\frac{q}{p} = 4$ and Variance = $\frac{q}{p^2} = 3$

$$\Rightarrow \frac{1}{p} = \frac{3}{4}$$

$\Rightarrow\ p = \frac{4}{3}$, which is impossible, since probability can never exceed unity.

Hence, the given statement is wrong.

Now, you can try the following exercises.

E1) Probability of hitting a target in any attempt is 0.6, what is the probability that it would be hit on fifth attempt?

E2) Determine the geometric distribution for which the mean is 3 and variance is 4.

Lack of Memory Property

Now, let us discuss the distinguishing property of the geometric distribution i.e. the 'lack of memory' property or 'forgetfulness property'. For example, in a random experiment satisfying geometric distribution the wait up to 3 trials (say) for the first success does not affect the probability that one will have to wait for a further 5 trials if it is given that the first two trials are failures. The geometric distribution is the only discrete distribution which has the forgetfulness (memoryless) property. However, there is one continuous distribution which also has the memoryless property and that is the exponential distribution which we will study in Unit 15 of MST-003. The exponential distribution is also the only continuous distribution having this property. It is pertinent to mention here that in several aspects, the geometric distribution is discrete analogs of the exponential distribution.

Let us now give mathematical/statistical discussion on 'memoryless property' of geometric distribution.

Suppose an event occurs at one of the trials 1, 2, 3, 4,… and the occurrence time X has a geometric distribution with probability p. Let X be the number of trials preceding to which one has to wait for successful attempt.

Thus, $P[X \geq j] = P[X = j] + P[X = j+1] + \ldots$

$$= q^j p + q^{j+1} p + q^{j+2} p + \ldots$$

$$= q^j p\left[1 + q + q^2 + \ldots\right]$$

$$= q^j p\left[\frac{1}{1-q}\right] = q^j p\left(\frac{1}{p}\right) = q^j$$

Now, let us consider the event $[X > j+k]$

Now, $P[X \geq j+k \mid X \geq j]$ means the conditional probability of waiting for at least j + k unsuccessful trials given that we waited for at least j unsuccessful attempts; and is given by

$$P[X \geq j+k \mid X \geq j] = \frac{P[X \geq j+k \mid X \geq j]}{P[X \geq j]}$$

$$= \frac{P\left[(X \geq j+k) \cap (X \geq j)\right]}{P[X \geq j]}$$

$$= \frac{P[X \geq j+k]}{P[X \geq j]} \qquad [\because X \geq j+k \text{ implies that} \geq j]$$

$$= \frac{q^{j+k}}{q^j} = q^k$$

$$= P[X \geq k] \qquad \left[\begin{array}{l}\because P[X \geq j] = q^j \text{ already} \\ \text{obtained in this section}\end{array}\right]$$

So, $P[X \geq j+k \mid X \geq j] = P[X \geq k]$

The above result reveals that the conditional probability of at least first j+k trials are unsuccessful before the first success given that at least first j trial were unsuccessful, is the same as the probability that the first k trials were unsuccessful. So, the probability to get first success remains same if we start counting of k unsuccessful trials from anywhere provided all the trials preceding to it are unsuccessful i.e. the future does not depend on past, it depends only on the present. So, the geometric distribution forgets the preceding trials and hence this property is given the name "forgetfulness property" or "Memoryless property" or "lack of memory" property.

12.3 NEGATIVE BINOMIAL DISTRIBUTION

Negative binomial distribution is a generalisation of geometric distribution. Like geometric distribution, variance of this distribution is also greater than its mean. There are many instances including 'deaths of insects' and 'number of insect bites' where negative binomial distribution is employed.

Negative binomial distribution is a generalisation of geometric distribution in the sense that geometric distribution is the distribution of 'number of failures preceding the first success' whereas the negative binomial distribution is the distribution of 'number of failures preceding the r^{th} success'.

Let X be the random variable which denote the number of failures preceding the r^{th} success. Let p be the probability of a success and let x failures are there preceding the r^{th} success and hence for this the number of trials is $x + r$.

Now, $(x+r)^{th}$ trial is success, but the remaining (r – 1) successes in the $x+r-1$ trials can happen in any r – 1 trials out of the $(x+r-1)$ trials. Thus, happening of first (r – 1) successes in $(x+r-1)$ trials follow binomial distribution with 'p' as the probability of success in each trial and thus is given by

$^{x+r-1}C_{r-1}\, p^{r-1}\, q^{(x+r-1)-(r-1)} = {}^{x+r-1}C_{r-1}\, p^{r-1} q^{x}$, where q = 1 – p

[$\because$ by binomial distribution, the probability of x successes in n trials with p as the probability of success is $^{n}C_{x}\, p^{x}\, q^{n-x}$.]

Therefore,

P[x failures preceding the r^{th} success]

=P[{First (r – 1) successes in $(x+r-1)$ trials} $\cap$ {success in $(x+r)^{th}$ trial}]

=P[First(r – 1) successes in $(x+r-1)$ trails]. P[success in $(x+r)^{th}$ trial]

$= \left({}^{x+r-1}C_{r-1}\, p^{r-1}\, q^{x}\right) p$

$= {}^{x+r-1}C_{r-1}\, p^{r}\, q^{x}$

The above discussion leads to the following definition:

Definition: A random variable X is said to follow a negative binomial distribution with parameters r (a positive integer) and p (0 < p < 1) if its probability mass function is given by:

$$P[X = x] = \begin{cases} {}^{x+r-1}C_{r-1}\, p^r q^x & \text{for } x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

Now, as we know that

${}^nC_r = {}^nC_{n-r}$, [See 'combination' in Unit 4 of MST-001]

$\therefore\ {}^{x+r-1}C_{r-1}$ can be written as

$${}^{x+r-1}C_{(x+r-1)-(r-1)} = {}^{x+r-1}C_x$$

$$= \frac{\lfloor x+r-1}{\lfloor x \lfloor x+r-1-x} = \frac{\lfloor r+x-1}{\lfloor x \lfloor r-1}$$

$$= \frac{(r+(x-1))(r+(x-2))\ldots(r+1)(r)\lfloor r-1}{\lfloor x \lfloor r-1}$$

$$= \frac{(r+x-1)(r+x-2)\ldots(r+1)r}{\lfloor x}$$

$$= \frac{\{-(-r-x+1)\}\{-(-r-x+2)\}\ldots\{-(-r-1)\}\{-(-r)\}}{\lfloor x}$$

[$\because$ Numerator is product of x terms from r + 0 to $r+(x-1)$ and we have taken common (-1) from each of these x terms in the product.]

$$= (-1)^x \frac{(-r-x+1)(-r-x+2)\ldots(-r-1)(-r)}{\lfloor x}$$

$$= \frac{(-1)^x(-r)(-r-1)(-r-2)\ldots\{-r-(x-1)\}}{\lfloor x}$$

[Writing the terms in the numerator in reverse order]

$$= (-1)^x \binom{-r}{x}$$

Note: The symbol $\binom{n}{x}$ stands for ${}^nC_x$ if n is positive integer and is equal to $\frac{n(n-1)(n-2)\ldots\{n-(x-1)\}}{\lfloor x}$. We may also use the symbol $\binom{n}{x}$ if n is any real but in this case though it does not stand for ${}^nC_x$, yet it is equal to $\frac{n(n-1)(n-2)\ldots\{n-(n-x)\}}{\lfloor x}$.

Hence, the probability distribution of negative binomnial distribution can be expressed in the following form:

$$P\left[X=x\right]=\left(-1\right)^{x}\binom{-r}{x}p^{r}q^{x}$$

$$=\binom{-r}{x}\left(-q\right)^{x}p^{r}\ \text{for}\ x=0,1,2,3,\ldots$$

$$=\binom{-r}{x}\left(-q\right)^{x}\left(1\right)^{-r-x}p^{r}\ \text{for}\ x=0,1,2,\ldots$$

Here, the expression $\binom{-r}{x}\left(-q\right)^{x}\left(1\right)^{-r-x}$ is similar to the binomial distribution $\binom{n}{x}p^{x}q^{n-x}$

$\therefore\ \binom{-r}{x}\left(-q\right)^{x}\left(1\right)^{-r-x}$ is the general term of $\left[1+\left(-q\right)\right]^{-r}=\left(1-q\right)^{-r}$

You have already studied in Unit 9 of this Course that $\binom{n}{x}p^{x}q^{n-x}$ is the general term of $\left[q+p\right]^{n}$.

and hence

$P\left[X=x\right]=\binom{-r}{x}\left(-q\right)^{x}\left(1\right)^{-r-x}.p^{r}$ is the general term of $\left(1-q\right)^{-r}p^{r}$

$\therefore$ P[X = 0], P[X = 1], P[X = 2],… are the successive terms of the binomial expansion $\left(1-q\right)^{-r}p^{r}$ and hence the sum of these probabilities

$$=\left(1-q\right)^{-r}p^{r}$$

$$=p^{-r}p^{r}\ \left[\because\ 1-q=p\right]$$

$$=1,$$

which must be, being a probability distribution.

Also, as the probabilities of the negative binomial distribution for $X=0,1,2,\ldots$ are the successive terms of

$\left(1-q\right)^{-r}p^{r}=\left(1-q\right)^{-r}\left(\frac{1}{p}\right)^{-r}=\left[\left(1-q\right)\frac{1}{p}\right]^{-r}=\left[\frac{1}{p}+\left(-\frac{q}{p}\right)\right]^{-r}$, which is a binomial expansion with negative index $(-r)$, it is for this reason the probability distribution given above is called the negative binomial distribution.

Mean and Variance

Mean and variance of the negative binomial distribution can be obtained on observing the form of this distribution and comparing it with the binomial distribution as follows:

The probabilities of binomial distribution for X = 0, 1, 2, … are the successive terms of the binomial expansion of $(q + p)^n$ and the mean and variance obtained for the distribution are

Mean = $np = (n)(p)$ i.e. Product of index and second term in (q + p)

Variance = npq = (n) (p) (q) i.e. Product of index, second term in (q + p) and first term in (q + p)

Similarly, the probabilities of negative binomial distribution for X = 0, 1, 2, ... are the successive term of the expansion of $\left[\frac{1}{p}+\left(-\frac{q}{p}\right)\right]^{-r}$ and thus, its mean and variance are:

Mean = (index) [second term in $\left\{\frac{1}{p}+\left(\frac{-q}{p}\right)\right\}$] = $(-r)\left[\left(-\frac{q}{p}\right)\right]=\frac{rq}{p}$, and

Variance = (index) [second term in $\left\{\frac{1}{p}+\left(-\frac{q}{p}\right)\right\}$] [First term in $\left\{\frac{1}{p}+\left(-\frac{q}{p}\right)\right\}$]

$$= (-r)\left(-\frac{q}{p}\right)\left(\frac{1}{p}\right)$$

$$=\frac{rq}{p^2}.$$

Remark 2

i) If we take r = 1, we have $P[X = x] = pq^x$ for $x = 0, 1, 2, \ldots$ which is geometric probability distribution.

Hence, geometric distribution is a particular case of negative binomial distribution and the latter may be regarded as the generalisation of the former.

ii) Putting r = 1 in the formulas of mean and variance of negative binomial distribution, we have

Mean = $\frac{(1)q}{p}=\frac{q}{p}$, and

Variance = $\frac{(1)(q)}{p^2}=\frac{q}{p^2}$,

which are the mean and variance of geometric distribution.

Example 3: Find the probability that third head turns up in 5 tosses of an unbiased coin.

Solution: It is a negative binomial situation with $p=\frac{1}{2}$, r = 3,

$x + r = 5 \Rightarrow x = 2.$

$\therefore$ by negative binomial distribution, we have

$$P[X=2] = {}^{x+r-1}C_{r-1}\, p^r q^x$$

$$= {}^{2+3-1}C_{3-1}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = {}^4C_2\left(\frac{1}{2}\right)^5 = \frac{4\times 3}{2}\times\frac{1}{32} = \frac{3}{16}$$

Example 4: Find the probability that a third child in a family is the family's second daughter, assuming the male and female are equally probable.

Solution: It is a negative binomial situation with

$p = \frac{1}{2}$ [$\because$ male and female are equally probable]

$r = 2,\ x + r = 3$

$\Rightarrow\ x = 1$

$\therefore$ by negative binomial distribution,

$$P[X=1] = {}^{x+r-1}C_{r-1}\, p^r\, q^x = {}^{1+2-1}C_{2-1}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1 = {}^2C_1\left(\frac{1}{2}\right)^3 = 2\times\frac{1}{8} = \frac{1}{4}.$$

Example 5: A proof-reader catches a misprint in a document with probability 0.8. Find the expected number of misprints in the document in which the proof-reader stops after catching the 20^{th} misprint.

Solution: Let X be the number of misprints not caught by the proof-reader and r be the number of misprints caught by him/her. It is a negative binomial situation where we are to obtain the expected (mean) number of misprints in the document i.e. E(X + r). We will first obtain mean number of misprints which could not be caught by the proof-reader i.e. E(X).

Here, p = 0.8 and hence q = 0.2, r = 20.

Now, by negative binomial distribution,

$$E(X) = \frac{rq}{p} = \frac{(20)(0.2)}{(0.8)} = 5$$

Therefore, $E(X+r) = E(X) + r$ = 5 + 20 = 25.

Hence, the expected number of misprints in the document till he catches the 20^{th} misprint is 25.

Now, we are sure that you will be able to solve the following exercises:

E3) Find the probability that fourth five is obtained on the tenth throw of an unbiased die.

E4) An item is produced by a machine in large numbers. The machine is known to produce 10 per cent defectives. A quality control engineer is testing the item randomly. What is the probability that at least 3 items are examined in order to get 2 defectives?

E5) Find the expected number of children in a family which stops producing children after having the second daughter. Assume, the male and female births are equally probable.

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

12.4 SUMMARY

The following main points have been covered in this unit:

1) A random variable X is said to follow **geometric distribution** if it assumes non-negative integer values and its probability mass function is given by

$$P[X = x] = \begin{cases} q^x p & \text{for } x = 0, 1, 2, \ldots \\ 0, & \text{otherwise} \end{cases}$$

2) For **geometric** distribution, **mean** $= \frac{q}{p}$ and **variance** $= \frac{q}{p^2}$.

3) A random variable X is said to follow a **negative binomial distribution** with parameters r (a positive integer) and p (0 < p < 1) if its probability mass function is given by:

$$P(X = x) = \begin{cases} {}^{x+r-1}C_{r-1}\, p^r q^x & \text{for } x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

4) For **negative binomial** distribution, **mean** $= \frac{rq}{p}$ and **variance** $= \frac{rq}{p^2}$.

5) For both these distributions, **variance > mean.**

12.5 SOLUTIONS/ANSWERS

E1) Let p be the probability of success i.e. hitting the target in an attempt.

$\therefore\ p = 0.6,\ q = 1 - p = 0.4$.

Let X be the number of unsuccessful attempts preceding the first successful attempt.
$\therefore$ by geometric distribution,

$$P[X = x] = q^x p \ \text{ for } x = 0, 1, 2, \ldots$$

$$= (0.4)^x (0.6) \ \text{ for } x = 0, 1, 2, \ldots$$

Thus, the desired probability = P[hitting the target in fifth attempt]

= P[The number of unsuccessful attempts before the first success is 4]

$= P[X = 4]$

$= (0.4)^4 (0.6) = (0.0256)(0.6) = 0.01536.$

E2) Let p be the probability of success in an attempt, and q = 1 – p

Now, mean = $\frac{q}{p} = 3$ and Variance = $\frac{q}{p^2} = 4$

$$\Rightarrow \frac{1}{p} = \frac{4}{3} \Rightarrow p = \frac{3}{4} \text{ and hence } q = \frac{1}{4}.$$

Now, let X be the number of failures preceding the first success,

$$\therefore P[X = x] = q^x p$$

$$= \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right) \text{ for } x = 0, 1, 2, \ldots$$

This is the desired probability distribution.

E3) It is a negative binomial situation with

$$r = 4,\ x + r = 10 \Rightarrow x = 6,\ p = \frac{1}{6} \text{ and hence } q = \frac{5}{6}$$

$$\therefore P[X = 6] = {}^{x+r-1}C_{r-1} p^r q^x$$

$$= {}^{6+4-1}C_{4-1} \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^6$$

$$= {}^9C_3 . \frac{625 \times 25}{36 \times 36 \times 36 \times 36 \times 36}$$

$$= \frac{9 \times 8 \times 7 \times 625 \times 25}{6 \times 36 \times 36 \times 36 \times 36 \times 36} = 0.0217$$

E4) It is a negative binomial situation with $r = 2,\ x + r = 3 \Rightarrow x = 1,\ p = 0.1$ and hence q = 0.9.

Now, the required probability $= P[X + r \geq 3]$

$$= P[X \geq 1]$$

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

$$= 1 - \left[{}^{0+r-1}C_{r-1}\, p^r q^0\right]$$

$$= 1 - \left[{}^{0+2-1}C_{2-1}\, (0.1)^2 (0.9)^0\right]$$

$$= 1 - (1)\,(0.01) = 0.99.$$

E5) It is a negative binomial situation with $p = \frac{1}{2},\ q = \frac{1}{2},\ r = 2.$

Let X be the number of boys in the family

$$\therefore E(X) = \frac{rq}{p} = \frac{(2)\left(\frac{1}{2}\right)}{\frac{1}{2}} = 2.$$

$$\Rightarrow E(X + r) = E(X) + r = 2 + 2 = 4$$

$\therefore$ the required expected value = 4.