Binomial and Poisson Distribution

7·1. Bernoulli Distribution. A random variable X which takes two values 0 and 1, with probabilities q and p respectively, *i.e.*, $P(X = 1) = p$, $P(X = 0) = q$, $q = 1 - p$ is called a *Bernoulli variate* and is said to have a Bernoulli distribution.

7·2. Binomial Distribution. Binomial distribution was discovered by James Bernoulli (1654-1705) in the year 1700 and was first published posthumously in 1713, eight years after his death). Let a random experiment be performed repeatedly and let the occurrence of an event in a trial be called a success and its non-occurrence a failure. Consider a set of n independent Bernoullian trials (n being finite), in which the probability 'p' of success in any trial is constant for each trial. Then $q = 1 - p$, is the probability of failure in any trial.

The probability of x successes and consequently $(n - x)$ failures in n independent trials, in a specified order (say) $SSFSFFFS...FSF$ (where S represents success and F failure) is given by the compound probability theorem by the expression :

$$P(SSFSFFFS...FSF) = P(S)P(S)P(F)P(S)P(F)P(F)P(F)P(S) \times \ldots \times P(F)P(S)P(F)$$

$$= p \cdot p \cdot q \cdot p \cdot q \cdot q \cdot q \cdot p \ldots q \cdot p \cdot q$$

$$= \underbrace{p \cdot p \ldots p}_{\{x \text{ factors}\}} \quad \underbrace{q \cdot q \cdot q \ldots q}_{\{(n-x) \text{ factors}\}} = p^x q^{n-x}$$

But x successes in n trials can occur in $\binom{n}{x}$ ways and the probability for each of these ways is $p^x q^{n-x}$. Hence the probability of x successes in n trials in *any order* whatsoever is given by the addition theorem of probability by the expression:

$$\binom{n}{x} p^x q^{n-x}$$

The probability distribution of the number of successes, so obtained is called the *Binomial probability distribution*, for the obvious reason that the probabilities of 0, 1, 2, ..., n successes, *viz.*,

$q^n, \binom{n}{1} q^{n-1} p, \binom{n}{2} q^{n-2} p^2, \ldots, p^n$, are the successive terms of the binomial expansion $(q + p)^n$.

Definition. *A random variable X is said to follow binomial distribution if it assumes only non-negative values and its probability mass function is given by*

$$P(X = x) = p(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} ; x = 0,1,2,\ldots, n ; q = 1 - p \\ 0, \text{ otherwise} \end{cases} \quad \ldots(7{\cdot}2)$$

The two independent constants n and p in the distribution are known as the *parameters* of the distribution. 'n' is also, sometimes, known as the degree of the binomial distribution.

Binomial distribution is a discrete distribution as X can take only the integral values, *viz.*, 0, 1, 2,..., n. Any variable which follows binomial distribution is known as *binomial variate*.

We shall use the notation $X \sim B(n, p)$ to denote that the random variable X follows binomial distribution with parameters n and p.

The probability $p(x)$ in (7·2) is also sometimes denoted by $b(x, n, p)$.

Remarks 1. This assignment of probabilities is permissible because

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

2. Let us suppose that n trials constitute an experiment. Then if this experiment is repeated N times, the *frequency function* of the binomial distribution is given by

$$f(x) = Np(x) = N \binom{n}{x} p^x q^{n-x} \, ; x = 0, 1, 2, \ldots, n \qquad \ldots(7.3)$$

and the expected frequencies of 0, 1, 2, ..., n successes are the successive terms of the binomial expansion, $N(q+p)^n$, $q+p=1$.

3. Binomial distribution is important not only because of its wide applicability, but because it gives rise to many other probability distributions. Tables for $p(x)$ are available for various values of n and p.

4. **Physical conditions for Binomial Distribution.** We get the binomial distribution under the following experimental conditions.

(i) Each trial results in two mutually disjoint outcomes, termed as success and failure.

(ii) The number of trials 'n' is finite.

(iii) The trials are independent of each other.

(iv) The probability of success 'p' is constant for each trial.

Example 7·1.. *Ten coins are thrown simultaneously. Find the probability of getting at least seven heads.*

Solution. p = Probability of getting a head = $\frac{1}{2}$

q = Probability of not getting a head = $\frac{1}{2}$

The probability of getting x heads in a random throw of 10 coins is

$$p(x) = \binom{10}{x}\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{10-x} = \binom{10}{x}\left(\frac{1}{2}\right)^{10} ; x = 0, , 2, \ldots, 10$$

$\therefore$ Probability of getting at least seven heads is given by

$$P(X \geq 7) = p(7) + p(8) + p(9) + p(10)$$

$$= \left(\frac{1}{2}\right)^{10}\left\{\binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right\}$$

$$= \frac{120 + 45 + 10 + 1}{1024} = \frac{176}{1024}.$$

3125

Example 8·3. *A coffee connoisseur claims that he can distinguish between a cup of instant coffee and a cup of percolator coffee 75% of the time. It is agreed that his claim will be accepted if he correctly identifies at least 5 of the 6 cups. Find his chances of having the claim (i) accepted, (ii) rejected, when he does have the ability he claims.*

Solution. If p denotes the probability of a correct distinction between a cup of instant coffee and a cup of percolator coffee, then we are given :

$$p = \frac{75}{100} = \frac{3}{4} \quad \Rightarrow \quad q = 1 - p = \frac{1}{4}, \quad \text{and} \quad n = 6$$

If the random variable X denotes the number of correct distinctions, then by the Binomial probability law, the probability of x correct identifications out of 6 cups is given by :

$$P(X = x) = p(x) = \binom{6}{x}\left(\frac{3}{4}\right)^x \left(\frac{1}{4}\right)^{6-x} \quad ; x = 0, 1, 2, \ldots, 6$$

(*i*) The probability of the claim being accepted is :

$$P(X \geq 5) = p(5) + p(6) = \binom{6}{5}\left(\frac{3}{4}\right)^5 \left(\frac{1}{4}\right)^{6-5} + \binom{6}{6}\left(\frac{3}{4}\right)^6 = \frac{1458}{4096} + \frac{729}{4096} = 0{\cdot}534.$$

(*ii*) The probability of the claim being rejected is :

$$P(X \leq 4) = 1 - P(X \geq 5) = 1 - 0{\cdot}534 = 0{\cdot}466.$$

Example 8·7. *The probability of a man hitting a target is* $\frac{1}{4}$:

(i) *If he fires 7 times what is the probability of his hitting the target at least twice ?*

(ii) *How many times must he fire so that the probability of his hitting the target at least once is greater than* $\frac{2}{3}$?

Solution. p = Probability of the man hitting the target $= \frac{1}{4} \quad \Rightarrow \quad q = 1 - p = \frac{3}{4}$.

$p(x)$ = Probability of getting x hits in 7 shots $= \binom{7}{x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{7-x}$; $x = 0, 1, \ldots, 7$

(*i*) Probability of at least two hits

$$= 1 - \{p(0) + p(1)\} = 1 - \left\{ \binom{7}{0} \left(\frac{1}{4}\right)^0 \left(\frac{3}{4}\right)^{7-0} + \binom{7}{1} \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^{7-1} \right\} = \frac{4547}{8192}.$$

(*ii*) Probability of at least one hit in n shots $= 1 - p(0) = 1 - \left(\frac{3}{4}\right)^x$.

It is required to find n, so that $\quad 1 - \left(\frac{3}{4}\right)^n > \frac{2}{3} \quad \Rightarrow \quad \frac{1}{3} > \left(\frac{3}{4}\right)^n$

Taking logarithms of each side, $\log \frac{1}{3} > n \log \frac{3}{4} \quad \Rightarrow \quad \log 1 - \log 3 > n\,(\log 3 - \log 4)$

$\Rightarrow \quad 0 - 0{\cdot}4771 > n\,(0{\cdot}4771 - 0{\cdot}6021) \quad \Rightarrow \quad 0{\cdot}4771 < 0{\cdot}1250\, n$

$$\therefore \qquad n > \frac{0{\cdot}4771}{0{\cdot}1250} = 3{\cdot}8$$

Since n cannot be fractional, the required number of shots is 4.

Example 7·5 *In a precision bombing attack there is a 50% chance that any one bomb will strike the target. Two direct hits are required to destroy the target completely. How many bombs must be dropped to give a 99% chance or better of completely destroying the target ?*

Solution. We have :

p = Probability that the bomb strikes the target = 50% = $\frac{1}{2}$. Let n be the number of bombs which should be dropped to ensure 99% chance or better of completely destroying the target. This implies that "probability that out of n bombs, at least two strike the target, is greater than 0·99".

Let X be a r.v. representing the number of bombs striking the target. Then $X \sim B(n, p = \frac{1}{2})$ with

$$p(x) = P(X = x) = \binom{n}{x}\left(\frac{1}{2}\right)^x \cdot \left(\frac{1}{2}\right)^{n-x} = \binom{n}{x}\left(\frac{1}{2}\right)^n ; x = 0, 1, \ldots, n$$

We should have :

$$P(X \geq 2) \geq 0{\cdot}99$$

$$\Rightarrow \quad [1 - P(X \leq 1)] \geq 0{\cdot}99$$

$$\Rightarrow \quad [1 - \{p(0) + p(1)\}] \geq 0{\cdot}99$$

$$\Rightarrow \quad 1 - \left\{\binom{n}{0} + \binom{n}{1}\right\}\left(\frac{1}{2}\right)^n \geq 0{\cdot}99$$

$$\Rightarrow \quad 0{\cdot}01 \geq \frac{1+n}{2^n} \quad \Rightarrow \quad 2^n \times (0{\cdot}01) \geq 1 + n$$

$$\Rightarrow \quad 2^n \geq 100 + 100\,n \qquad \ldots(*)$$

By trial method, we find that the inequality (*) is satisfied by $n = 11$. Hence the minimum number of bombs needed to destroy the target completely is 11.

(*b*) A multiple choice test consists of 8 questions and 3 answers to each question, of which only one is correct. If a student answers each question by rolling a balanced die and checking the first answer if he gets 1 or 2, the second answer if he gets 3 or 4, and the third answer if he gets 5 or 6, find the probability of getting:

(*i*) exactly 3 correct answers,

(*ii*) no correct answer,

(*iii*) at least 6 correct answers. [Gauhati Univ. M.A. (Econ.), 1993]

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

Ans. 52/3125

(*b*). (*a*) In a binomial distribution consisting of 5 independent trails, probabilities of 1 and 2 successes are 0·4096 and 0·2048 respectively. Find the parameter p of the distribution. (**Ans.** 0·2)

7·3·0. Poisson Distribution (as a limiting case of Binomial Distribution). Poisson distribution was discovered by the French mathematician and physicist Simeon Denis Poisson (1781—1840) who published it in 1837. Poisson distribution is a limiting case of the 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 indefinitely small, *i.e.*, $p \to 0$.

(*iii*) $np = \lambda$, (say), is finite. Thus $p = \lambda/n$, $q = 1 - \lambda/n$, where λ is a positive real number.

The probability of x successes in a series of n independent trials is

$$b(x\,;n,p) = \binom{n}{x} p^x q^{n-x}\,;\, x = 0, 1, 2, \ldots, n \qquad \ldots(*)$$

Aliter. Poisson distribution can also be derived without using Stirling's approximation as follows :

$$b(x\,;n,p) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{n}{x}\left[\frac{p}{1-p}\right]^x (1-p)^n$$

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

$$= \frac{\left[1-\frac{1}{n}\right]\left[1-\frac{2}{n}\right]\ldots\left[1-\frac{x-1}{n}\right]}{x!\left[1-\frac{\lambda}{n}\right]^x}\,\lambda^x\left[1-\frac{\lambda}{n}\right]^n$$

$$\therefore \lim_{n\to\infty} b(x\,;n,p) = \frac{e^{-\lambda}\lambda^x}{x!}\,;\ x = 0, 1, 2, \ldots \qquad \text{[From (**)]}$$

Definition. *A random variable X is said to follow a Poisson distribution if it assumes only non-negative values and its probability mass function is given by*

$$p(x, \lambda) = P(X = x) = \frac{e^{-\lambda} \lambda^x}{x\,!}\,;\; x = 0, 1, 2, \ldots\,;\lambda > 0$$
$$= 0, \text{ otherwise} \qquad \ldots(7\cdot 14)$$

Here λ is known as the parameter of the distribution.

We shall use the notation $X \sim P(\lambda)$ to denote that X is a Poisson variate with paramter λ.

Example 7· 24 . *A car hire firm has two cars which it fires out day by day. The number of demands for a car on each day is distributed as Poisson variate with mean 1· 5., Calculate the proportion of days on which (i) neither car is used, and (ii) some demand is refused.*

Solution. The proportion of days on which there are x demands for a car

$$= P\{\text{of } x \text{ demands in a day}\}$$

$$= \frac{e^{-1\cdot 5}(1\cdot 5)^x}{x!},$$

since the number of demands for a car on any day is a Poisson variate with mean $1\cdot 5$. Thus

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

(*i*) Proportion of days on which neither car is used is given by

$$P(X = 0) = e^{-1\cdot 5}$$

$$= \left[1 - 1\cdot 5 + \frac{(1\cdot 5)^2}{2!} - \frac{(1\cdot 5)^3}{3!} + \frac{(1\cdot 5)^4}{4!} - \ldots\right]$$

$$= 0\cdot 2231$$

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

$$P(X > 2) = 1 - P(X \le 2)$$

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

$$= 1 - e^{-1\cdot 5}\left[1 + 1\cdot 5 + \frac{(1\cdot 5)^2}{2!}\right]$$

$$= 1 - 0\cdot 2231 \times 3\cdot 625 = 0\cdot 19126$$

Example 7·25. *A manufacturer of cotter pins knows that 5% of his product is defective. If he sells cotter pins in boxes of 100 and guarantees that not more than 10 pins will be defective, what is the approximate probability that a box will fail to meet the guaranteed quality ?*

Solution. We are given $n = 100$.

Let $p =$ Probability of a defective pin $= 5\% = 0 \cdot 05$

$\therefore$ $\lambda =$ Mean number of defective pins in a box of 100

$= np = 100 \times 0 \cdot 05 = 5$

Since 'p' is small, we may use Poisson distribution.

Probability of x defective pins in a box of 100 is

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

Probability that a box will fail to meet the guaranteed quality is

$$P(X > 10) = 1 - P(X \le 10) = 1 - \sum_{x=0}^{10} \frac{e^{-5}5^x}{x\,!} = 1 - e^{-5} \sum_{x=0}^{10} \frac{5^x}{x\,!}$$

Example 7·27. *In a book of 520 pages, 390 typo-graphical errors occur. Assuming Poisson law for the number of errors per page, find the probability that a random sample of 5 pages will contain no error.*

Solution. The average number of typographical errors per page in the book is given by $\lambda = (390/520) = 0 \cdot 75$

Hence using Poisson probability law, the probability of x errors per page is given by : $P(X = x) = \dfrac{e^{-\lambda} \lambda^x}{x!} = \dfrac{e^{-0 \cdot 75} (0 \cdot 75)^x}{x!}$; $x = 0, 1, 2, \ldots$

The required probability that a random sample of 5 pages will contain no error is given by : $[P(X = 0)]^5 = \left(e^{-0 \cdot 75}\right)^5 = e^{-3 \cdot 75}$

Example 8-34. *An insurance company insures 4,000 people against loss of both eyes in a car accident. Based on previous data, the rates were computed on the assumption that on the average 10 persons in 1,00,000 will have car accident each year that result in this type of injury. What is the probability that more than 3 of the insured will collect on their policy in a given year ?*

Solution. In usual notations, we are given : $n = 4,000$, and

$$p = \text{Probability of loss of both eyes in a car accident} = \frac{10}{1,00,000} = 0{\cdot}0001.$$

Since p is very small and n is large, we may approximate the given distribution by Poisson distribution. Thus the parameter λ of the Poisson distribution is :

$$\lambda = np = 4,000 \times 0{\cdot}0001 = 0{\cdot}4$$

Let the random variable X denote number of car accidents in the batch of 4,000 people. Then by Poisson probability law :

$$P(X = x) = \frac{e^{-\lambda}.\lambda^x}{x\,!} = \frac{e^{-0{\cdot}4}\,(0{\cdot}4)^x}{x\,!}\,;\; x = 0, 1, 2, \ldots \qquad \ldots (*)$$

Hence the required probability that more than 3 of the insured will collect on their policy is given by :

$$P\,(X > 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)]$$

$$= 1 - e^{-0{\cdot}4}\left\{(0{\cdot}4)^0 + (0{\cdot}4) + \frac{(0{\cdot}4)^2}{2\,!} + \frac{(0{\cdot}4)^3}{3\,!}\right\}$$

$$= 1 - 0{\cdot}6703\,(1 + 0{\cdot}4 + 0{\cdot}08 + 0{\cdot}0107) = 1 - 0{\cdot}6703 \times\; 1{\cdot}4907 = 0{\cdot}0008.$$