Poisson Distribution:

Poisson distribution, also known as the law of small numbers, is one of the important families of probability distributions named after Simeon Denis Poisson, a French mathematician, who discovered it in 1837.

A random variable following this rule is referred to as a Poisson variate, and the process of generating values of such a random variable is known as a Poisson process

The distribution has been found applicable to many processes that involve an observation falling in a given time interval or in a specified region or space. Experiments yielding numerical values of a random variable X within the given interval is often called "Poisson Experiment". The given interval may be of any length, such as a minute, a day, a week, a month or even a year. Here are some typical examples of a Poisson varite.

(i) The number of telephone calls received at a switchboard per minute.

(ii) The number of customers arriving at a bank counter per 5-minute period.

(iii) The number of machines breaking down during any one day.

(iv) The number of drivers stopped at a road block not properly licensed in a day.

(v) The number of days the city schools remain closed due to heavy cold during winter.

(vi) The number of postponed cricket matches due to heavy shower in summer.

(vii) The number of typographical errors per page in an official document.

(viii) The number of leakages in 100 kilometers of gas pipelines.

(ix) The number of defective parts per batch shipped from the factory.

(x) The number of rats per acre of cultivated land.

The above processes all follow approximately the Poisson pattern and are characterized by the expected number of successes per unit of time or space just as the binomial distribution is characterized by the number of successes in n trials. The per unit of time or space in a Poisson distribution is thus equivalent to the sample size of n in a binomial distribution.

A Poisson variable is a random variable and can assume only integral values from 0 to an infinitely large value. This random variable thus can assume any of a countably infinite set of values and is therefore a discrete variable. Notice that while the time between failures (or telephone calls) is a continuous random variable, the number of failures (or telephone calls) in a fixed time interval takes only the integer values 0, 1, 2, ... and so on and therefore is a discrete variable.

Although the Poisson variable can be given many useful interpretations, perhaps the simplest approach to the study of the Poisson process is to regard it as a special case of the binomial, where n is thought to be very large and p is very small.

Probability Distribution of Poisson Variable:

The probability distribution of the Poisson variable X is called the Poisson distribution and will be denoted by $f(x, \mu)$ since its value depends only on μ, the average number of successes occuring in the given time interval or specified region.

Defn:

Let μ be the mean or expected number of successes in a specified time or space and the random variable X designates the number of successes in a given time interval or specified region. Then the probability distribution of X is a Poisson distribution and is given by the formula

$$f(x, \mu) = \frac{e^{-\mu} \mu^{x}}{x!} \; ; \quad x = 0, 1, 2, \dots, \infty$$

where e is a constant and equal to 2.71828 and x is any positive value that X can assume.

The Poisson distribution, as compared to binomial distribution, is rather easy to apply once the value of the parameter μ (the mean of the distribution) is given.

Scanned with CamScanner

Example:

Suppose that at any time of the working day, the average number of telephone calls per minute is three ($\mu = 3$). You can obtain the value of $e^{-\mu}$ by using calculator or get it directly from the Appendix table. This table shows that $e^{-\mu} = 0.0798$. Thus, in a given minute, say from 9:00 to 9:01, the probabilities of specified number of calls or $f(x, \mu)$ and the cumulative probabilities $(P(X \le x))$ are computed as follows:

No. of calls X	Probability $p(X=x)$	Cumulative Probability $P(X \le x)$
0	$\frac{e^{-3} 3^{0}}{0!} = 0.0498$	0.0498
1	$\frac{e^{-3} 3^{1}}{1!} = 0.1494$	0.1992
2	$\frac{e^{-3} 3^{2}}{2!} = 0.2240$	0.4232
3	$\frac{e^{-3} 3^{3}}{3!} = 0.2240$	0.6472
4	$\frac{e^{-3} 3^{4}}{4!} = 0.1681$	0.8153

Characteristics of a Poisson Distribution:

A Poisson distribution is assumed to possess the following characteristics:

(1) The number of successess occuring in one time interval is independent of those occuring in any other disjoint time intervals.

(2) The probability of single success occuring during a very short time interval is proportional to the length of the time interval and does not depend on the number of successes occuring outside the interval.

(3) The probability of more than one success in such a short time interval is negligible.

Th^m :

Verify that $f(x,\mu) = \dfrac{e^{-\mu}\mu^x}{x!}$, $x = 0, 1, 2, \dots \infty$

is a probability mass function.

Proof :

It is clear that $f(x,\mu) \geq 0$ for each value of x.
In order to verify that the function $f(x,\mu)$ satisfies the requirements of every probability function, it must be shown that

$$\sum_{x=0}^{\infty} f(x,\mu) = 1$$

It is known from the elementary algebra that for any real number μ,

$$e^{\mu} = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

Therefore,

$$\sum_{x=0}^{\infty} f(x,\mu) = e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

$$= e^{-\mu} \cdot e^{\mu}$$

$$= 1$$

(Proved)

Problem:

Telephone calls arrive at a switchboard at a mean rate of 0.5 calls per minute. Calculate the probability that two calls will arrive in a particular five-minute period.

Sol^n:

Here one time unit is five minutes. This is a problem that fits the Poisson distribution with average number of calls per unit of time $\mu = 5 * 0.5 = 2.5$.
If X is the random variable, we want to compute $f(2, 2.5)$. Thus

$$f(x, \mu) = f(2, 2.5) = \frac{e^{-2.5} (2.5)^2}{2!}$$

$$\Rightarrow f(2, 2.5) = \frac{0.0821 * 6.25}{2}$$

$$= 0.257$$

which is the required probability.

Problem :

The average number of calls received by a telephone operator during a time interval of 10 minutes during 5:00 P.M to 5:10 PM daily is 3. What is the probability that the operator will receive

(i) no call

(ii) exactly one call &

(iii) at least two calls tomorrow during the same interval ?

Sol^n :

Let, X be the random variable designating the number of calls made during the interval. Assuming that X follows Poisson distribution, we will evaluate the probabilities $P(X=0)$, $P(X=1)$ & $P(X\geq 2)$

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

(ii) $P(X=1) = \frac{e^{-3}3^{1}}{1!} = 3e^{-3} = 0.1494$

(iii) $P(X\geq 2) = 1 - \sum_{x=0}^{1} P(X=x)$

Scanned with CamScanner

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

$$= 1 - (0.0498 + 0.1494)$$

$$= 0.8008$$

Properties of Poisson distribution:

Mean:

The mean of Poisson distribution $f(x,\mu)$ is μ. The value of this mean is usually a small positive quantity.

By definition, $E[X] = \sum_{x=0}^{\infty} x f(x,\mu)$

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

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

$$\Rightarrow E[X] = \mu \sum_{x=0}^{\infty} \frac{e^{-\mu}\mu^{x-1}}{(x-1)!}$$

Let, $y = x-1$ which gives

$$E[X] = \mu . \sum_{y=0}^{\infty} \frac{e^{-\mu}\mu^y}{y!} = \sum_{y=0}^{\infty} \mu . \frac{e^{-\mu}\mu^y}{y!}$$

$$= \sum_{y=0}^{\infty} \mu . f(y,\mu)$$

$$= \mu . \sum_{y=0}^{\infty} f(y,\mu)$$

$$\Rightarrow E[X] = \mu . 1 \qquad \left[\because \sum_{y=0}^{\infty} f(y,\mu) = 1\right]$$

$$\therefore E[X] = \mu$$

Hence, mean of the Poisson distribution is equal to its parameter μ.

Variance:

The variance of the Poisson distribution is also μ.

By definition,

$$V[X] = E[X - E[X]]^2$$

$$= E[(X-\mu)^2]$$

$$\Rightarrow V[X] = E[X^2] - \mu^2 \quad \text{——— (1)}$$

But

$$E[X^2] = \sum_{x=0}^{\infty} x^2 f(x,\mu)$$

$$= \sum_{x=0}^{\infty} [x(x-1)+x] f(x,\mu)$$

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

$$= \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\mu}\mu^x}{x!} + \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\mu}\mu^x}{x(x-1)(x-2)!} + \mu$$

$$= \mu^2 \sum_{x=2}^{\infty} \frac{e^{-\mu}\mu^{x-2}}{(x-2)!} + \mu$$

$$\Rightarrow E[X^2] = \mu^2 \sum_{x=2}^{\infty} \frac{e^{-\mu} \mu^{x-2}}{(x-2)!} + \mu$$

Setting $y = x-2$, we have

$$E[X^2] = \mu^2 \sum_{y=0}^{\infty} \frac{e^{-\mu} \mu^{y}}{y!} + \mu$$

$$= \mu^2 \sum_{y=0}^{\infty} f(y,\mu) + \mu$$

$$= \mu^2 . 1 + \mu \qquad \left[\because \sum_{y=0}^{\infty} f(y,\mu) = 1\right]$$

$$\therefore E[X^2] = \mu^2 + \mu \qquad \text{——— } (2)$$

From equ^n (1) we can write

$$V[X] = \mu^2 + \mu - \mu^2 \quad [\text{using } (2)]$$

$$\therefore V[X] = \mu$$

Binomial Distribution:

Binomial experiment:

Many experiments are composed of repetitions of independent trials, each with two possible outcomes: success and failure. When the probabilities of these two outcomes remain unchanged from one trial to another, we name these trials as Bernoulli trials. Thus, a binomial experiment consists of a "fixed number "n"" of Bernoulli trials.

Def^n:

When an experiment has two possible outcomes, success and failure and the experiment is repeated n times independently, and the probability p of success of any given trial remains constant from trial to trial, the experiment is known as binomial experiment.

Properties of binomial experiment:

① The overall experiment can be described in terms of a sequence of n identical experiments, which are called trials.

② Each trial results in an outcome that may be classified as a success or a failure.

③ The probability of success on a single trial is equal to some value p which remains constant from one trial to the next. The probability of failure is equal to $q = 1 - p$.

④ The repeated trials are independent.

⑤ The random variable of interest is X, the number of successes observed in n trials.

Examples of Binomial experiment:

① An insurance salesman contacts ten different families. The outcome associated with visiting each family can be referred to as a success if the family purchases an insurance policy and a failure, if not. If the probability of selling a policy is assumed to be same for each family, and the decision to purchase or not a policy by one family is not influenced by the decision of any other family, then we have a situation analogous to the binomial experiment.

② A customer enters the Aarong outlet and makes a purchase. For convenience, let us restrict our attention to the next ten customers who enter the shop. The event of interest here is to ascertain whether or not any one of the customers makes a purchase. The process can be viewed as a binomial experiment because of the reason that (i) the experiment can be described as a sequence of ten identical trials, one trial for each of the ten customers that will enter the store, (ii) the probability of purchase (p) may be assumed to be the same for all customers (iii) the purchase decision of each customer is independent of the decision of each customer is independent of the decision of the other customers.

③ An experiment consists in selecting five radios at random from a lot and inspecting and then classifying them as defective and non-defective. A defective radio is labeled a success, while a non-defective radio is labeled a failure. Suppose 20% of the radios are defective. If this probability is assumed to remain constant from trial to trial, then X, the number of successes is a binomial random variable assuming the values 0, 1, 2, 3, 4, and 5. Clearly, the experiment is a binomial experiment with $n=5$ and $p=0.20$.

④ If five cards are drawn in succession from an ordinary deck and each trial is labeled a success or failure depending on whether the card is a red or black. If each card is replaced and the deck is well shuffled before the next drawing is made, then the repeated drawing (trials) are independent and the probability of success (drawing a red ~~card~~ here) remain ~~trial to~~ constant from trial to trial. This is a binomial experiment.

In each of the above experiments, there exists a variable that denotes the number of successes in n- Bernoulli trials. This variable is known as the binomial random variable. A binomial random variable may thus be viewed as a set of values representing the number of successes generated through some binomial experiment.

Def^{n}:

If X is a random variable designating the number of successes in n Bernoulli trials taking on a set of values $\{0, 1, 2, \ldots, n\}$, then X is called the binomial random variable.

The distribution associated with a binomial random variable X defined above results in a binomial probability distribution with two parameters n, the number of trials and p, the probability of success.

Binomial probability distribution:

If a Bernoulli trial can result in a sucess with probability p and a failure with probability $q=1-p$, then the probability distribution of the binomial random variable X, the number of successes in n independent trials is

$$b(x;n,p) = \begin{cases} {}^{n}C_{x}\, p^{x} (1-p)^{n-x} ; & x = 0,1,2,\dots,n \\ 0 ; & \text{otherwise} \end{cases}$$

This is the so-called binomial probability distribution with parameters n and p.

Mean:

By definition, Mean $= E[X] = \sum_{x=0}^{n} x \; b(x;n,p)$

$$\Rightarrow E[X] = \sum_{x=0}^{n} x \; {}^{n}C_{x}\, p^{x} q^{n-x}$$

$$= 0 \cdot q^{n} + 1 \cdot {}^{n}C_{1} \cdot p \cdot q^{n-1} + 2 \cdot {}^{n}C_{2}\, p^{2} q^{n-2} + \dots + n p^{n}$$

$$= n q^{n-1} p + 2 \cdot \frac{n(n-1)}{2!} q^{n-2} p^{2} + 3 \cdot \frac{n(n-1)(n-2)}{3!} q^{n-3} p^{3} + \dots + n p^{n}$$

$$\Rightarrow E[X] = np\left\{ q^{n-1} + (n-1)q^{n-1}p + \frac{(n-1)(n-2)}{2!}q^{n-2}p^{2} + \dots + p^{n-1} \right\}$$

$$= np\,(q+p)^{n-1}$$

$$= np\,(1)^{n-1}$$

$$\therefore E[X] = np$$

$\therefore$ Mean of the binomial distribution, $E[X] = np$.

Variance :

By def^n, variance $= V[X] = E[X^2] - \{E[X]\}^2$

$$\Rightarrow V[X] = E[X^2] - (np)^2 \quad \text{——— (1)}$$

Now, $$E[X^2] = \sum_{x=0}^{n} x^2\, {}^nC_x\, p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \left[\{x + x(x-1)\}\right] {}^nC_x\, p^x q^{n-x}$$

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

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

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

$$= np + n(n-1)p^2 (q+p)^{n-2}$$

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

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

$$= np\left[1 + (n-1)p\right]$$

$$= np\left[1 + np - p\right] = np\left[1 - p + np\right]$$

$$\Rightarrow E[X^2] = np[q + np]$$

1 $\Rightarrow$

$$V[X] = np[q + np] - n^2p^2$$

$$= npq + \cancel{n^2p^2} - \cancel{n^2p^2}$$

$$\therefore V[X] = npq$$

$\therefore$ Variance of the binomial distribution, $V[X] = npq$

Moment generating function:

The moment generating function of a binomial variate X is defined by

$$M_X(t) = E[e^{tX}]$$

$$= \sum_{x=0}^{n} e^{tx}\, b(x; n, p)$$

$$= \sum_{x=0}^{n} e^{tx}\, {}^nC_x\, p^x q^{n-x}$$

$$= \sum_{x=0}^{n} {}^nC_x\, (pe^t)^x q^{n-x}$$

$$\therefore M_X(t) = (pe^t + q)^n$$

Binomial Recursion relation:

It is usually a laborious job in calculating binomial probabilities directly from the binomial function $f(x; n, p)$. A much simpler approach is to use a recursion formula which can be expressed as

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

Scanned with CamScanner

Proof:

$$b(x;n,p) = {}^nC_x\, p^x q^{n-x} = \frac{n!}{x!\,(n-x)!}\, p^x q^{n-x} \quad \text{—} ①$$

& $$b(x+1;n,p) = {}^nC_{x+1}\, p^{x+1} q^{n-x-1} = \frac{n!}{(x+1)!\,(n-x-1)!}\, p^{x+1} q^{n-x-1}$$

— ②

② ÷ ①

$$\frac{b(x+1;n,p)}{b(x;n,p)} = \frac{\dfrac{n!}{(x+1)!\,(n-x-1)!}\, p^{x+1} q^{n-x-1}}{\dfrac{n!}{x!\,(n-x)!}\, p^x q^{n-x}}$$

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

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

$$\Rightarrow \frac{b(x+1;n,p)}{b(x;n,p)} = \frac{n-x}{x+1}\, \frac{p}{q}$$

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

Scanned with CamScanner

Thus $b(x+1; n, p)$ can be obtained from $b(x; n, p)$ knowing n and p.

Problem:

Five coins are tossed and the experiment is repeated 200 times. The following table gives the frequency distribution of the number of heads obtained.

No. of heads	0	1	2	3	4	5	Total
Frequency	12	56	74	39	18	1	200

Assuming that the distribution of heads confirms to a binomial experiment, we estimate the probability of obtaining head in a single toss and hence the expected number of heads.

Sol^n:

The mean value of the distribution is

$$\bar{x} = \frac{\sum_i f_i x_i}{\sum_i f_i} = \frac{(0*12)+(1*56)+(2*74)+(3*39)+(4*18)+(5*1)}{200}$$

$$= \frac{0+56+148+117+72+5}{200}$$

$$= 1.99$$

Scanned with CamScanner

Equating this mean to the binomial mean np, we have $np = 1.99$. Hence, $p = \frac{1.99}{n} = \frac{1.99}{5} = 0.398$

Thus, the estimated function of the binomial distribution is

$$b(x; n, p) = b(x; 5, 0.398) = {}^5C_x (0.398)^x (1-0.398)^{5-x}$$

$$= {}^5C_x (0.398)^x (0.602)^{5-x};$$

for $x = 0, 1, 2, 3, 4, 5$.

The expected frequencies can be estimated from

$$N\,b(x; n, p) = 200\ {}^5C_x (0.398)^x (0.602)^{5-x}; \quad x = 0,1,2,3,4,5$$

$$b(0; 5, 0.398) = (0.602)^5 = 0.07907$$

Now, using the recursion relation we get,

$$b(x+1; n, p) = \left(\frac{n-x}{x+1}\right)\left(\frac{p}{q}\right) b(x; n, p) \quad \text{——— (1)}$$

For $x = 0, 1, 2, 3, 4$ we have from (1) successively

$$b(1; 5, 0.398) = \left(\frac{5-0}{0+1}\right)\left(\frac{0.398}{0.602}\right)(0.07907) = 0.26136$$

$$b(2; 5, 0.398) = \left(\frac{5-1}{1+1}\right)\left(\frac{0.398}{0.602}\right)(0.26136) = 0.34559$$

$$b(3;5,0.398) = \left(\frac{5-2}{2+1}\right)\left(\frac{0.398}{0.602}\right)(0.34559) = 0.22847$$

$$b(4;5,0.398) = \left(\frac{5-3}{3+1}\right)\left(\frac{0.398}{0.602}\right)(0.22847) = 0.07553$$

$$b(5;5,0.398) = \left(\frac{5-4}{4+1}\right)\left(\frac{0.398}{0.602}\right)(0.07553) = 0.00998$$

Multiplying each of the above probabilities by $N=200$, the expected frequencies for $x=0,1,2,3,4,5$ are obtained. These are shown in the following table:

$X=x$	0	1	2	3	4	5	Total
Observed f:	12	56	74	39	18	1	200
Expected f:	16	52	69	46	15	2	200

The number of successes x, together with the expected frequencies constitutes what we call the binomial frequency distribution.

Use of Binomial Probability Table:

$$P(X \le r) = \sum_{x=0}^{r} b(x; n, p) = \sum_{x=0}^{r} {}^{n}C_{x}\, p^{x} (1-p)^{n-x}$$

Also, using the Binomial table, we can compute the probability of the type $P(X \ge r)$ or $P(a \le X \le b)$

Problem:

If X is a binomial variable with $n=5$ and $p=0.3$, evaluate the following probabilities.

(i) $P(X \le 4)$

(ii) $P(X=2)$

(iii) $P(X<3)$

(iv) $P(X>1)$

(v) $P(X \ge 3)$

Solⁿ:

(i) $P(X \le 4) = \sum_{x=0}^{4} b(x; 5, 0.3)$

$$= \sum_{x=0}^{4} {}^{5}C_{x} (0.3)^{x} (0.7)^{5-x}$$

$\cancel{= 0.8370}$ (From table)

$= 0.9977$

(ii) $P(X=2) = b(2; 5, 0.3) = P(X \le 2) - P(X \le 1)$

$\Rightarrow P(X=2) = 0.8370 - 0.5283 = 0.3087$

(iii) $P(X<3) = P(X \leq 2)$

$$= \sum_{x=0}^{2} b(x; 5, 0.3)$$

$$= 0.8370$$

(iv) $P(X>1) = 1 - P(X \leq 1)$

$$= 1 - \sum_{x=0}^{1} b(x; 5, 0.3)$$

$$= 1 - 0.5283$$

$$= 0.4717$$

(v) $P(X \geq 3) = 1 - P(X<3)$

$$= 1 - 0.8370$$

$$= 0.1630$$

Binomial frequency distribution:

If n independent trials constitute one experiment and if this experiment is repeated N times, then we obtain what is known as the binomial frequency distribution. Thus we expect x successes to occur $N\,{}^nC_x\,p^x q^{n-x}$ times. This is known as the expected frequency of x successes in N experiments and the possible number of successes together with the expected (theoretical) frequencies will be said to constitute the binomial frequency distribution. It is in fact a theoretical distribution and in practice, the observed frequencies will not always coincide with the expected frequencies of this theoretical distribution because of sampling variations. Nevertheless, this is a useful device that permits a comparison of the experimental results with the theoretical frequencies of $0, 1, 2, \ldots, n$ successes are given by the successive terms in the binomial expansion of $N(q+p)^n$, where $p+q=1$.

Problem:

A biased coin is tossed 4 times and the number of heads observed is given below. The experiment is repeated 500 times in all. The results of the experiment is as follows:

No. of heads	0	1	2	3	4
Observed frequency	12	50	151	200	87

(i) find the probability of obtaining a head when the coin is tossed.

(ii) Calculate the expected frequency fitting a theoretical frequency distribution.

Sol^n:

For the observed distribution

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{(0*12)+(1*50)+(2*151)+(3*200)+(4*87)}{500}$$

$$= 2.6$$

(i) Let X be the random variable "the number of heads obtained in 4 tosses". Then X is a binomial variate with $n=4$ so that the mean, $E[X] = np$.

Therefore, the probability that the coin will show head is 0.65.

Therefore, $np = 2.6 \Rightarrow p = 2.6/4 = 0.65$.

Thus, Binomial distribution of X is $b(x; 4, 0.65)$

(ii) To evaluate the expected frequencies, we need to compute $P(X=x)$ for $x = 0, 1, 2, 3, 4$.

Since, $P(X=x) = b(x; 4, 0.65)$

$$\Rightarrow P(X=x) = {}^4C_x (0.65)^x (0.35)^{4-x} \quad \text{—— (1)}$$

Putting successively $x = 0, 1, 2, 3$ & 4, we obtain from (1)

$$P(X=0) = {}^4C_0 (0.65)^0 (0.35)^4 = 0.01500625$$

$$P(X=1) = {}^4C_1 (0.65)^1 (0.35)^3 = 0.111475$$

$$P(X=2) = {}^4C_2 (0.65)^2 (0.35)^2 = 0.3105375$$

$$P(X=3) = {}^4C_3 (0.65)^3 (0.35)^1 = 0.384475$$

$$P(X=4) = {}^4C_4 (0.65)^4 (0.35)^0 = 0.1785062$$

The theoretical distribution is now obtained by multiplying each of the above probabilities by the total frequency 500. The resulting distribution is thus

No. of heads	0	1	2	3	4
Expected frequency	8	56	155	192	89

Problem:

A fair coin is tossed 5 times. Find the probability of obtaining (i) exactly 4 heads (ii) fewer than 3 heads.

Soln:

Here, the given experiment is a binomial one, where $p = 0.5$ & $n = 5$.

(i) $P(X=4) = b(x; n, p) = b(x; 5, 0.5) = {}^{n}C_x \, p^x q^{n-x}$

$\Rightarrow P(X=4) = {}^5C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^{5-4} = \frac{5}{32}$.

(ii) $P(X<3) = \sum_{x=0}^{2} b(x; 5, 0.5) = 0.5$ [Directly from chart]

or,

$P(X<3) = \sum_{x=0}^{2} b(x; 5, 0.5)$

$= b(0; 5, 0.5) + b(1; 5, 0.5) + b(2; 5, 0.5)$

$= {}^5C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{5-0} + {}^5C_1 \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{5-1} + {}^5C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{5-2}$

$= \frac{1}{32} + \frac{5}{32} + \frac{10}{32}$

$= \frac{1}{2}$

$= 0.5$

Ans

Problem:

A traffic control officer reports that 75% of the trucks passing through a check post are from within Dhaka city. What is the probability that at least three of the next five trucks are from out of the Dhaka city?

Sol^n:

Let, the random variable X defines the number of trucks that pass through are from out of the Dhaka city.

The probability of such an event is then given by $P = 1 - 0.75 = 0.25 = \frac{1}{4}$

Hence, the required probability can be estimated by binomial distribution.

i.e $P(X \geq 3) = \sum_{x=3}^{5} b(x; 5, \frac{1}{4})$

$$= b(3; 5, \tfrac{1}{4}) + b(4; 5, \tfrac{1}{4}) + b(5; 5, \tfrac{1}{4})$$

$$= {}^5C_3 \left(\tfrac{1}{4}\right)^3 \left(\tfrac{3}{4}\right)^{5-3} + {}^5C_4 \left(\tfrac{1}{4}\right)^4 \left(\tfrac{3}{4}\right)^{5-4} + {}^5C_5 \left(\tfrac{1}{4}\right)^5 \left(\tfrac{3}{4}\right)^{5-5}$$

$$= {}^5C_3 \left(\tfrac{1}{4}\right)^3 \left(\tfrac{3}{4}\right)^2 + {}^5C_4 \left(\tfrac{1}{4}\right)^4 \left(\tfrac{3}{4}\right)^1 + {}^5C_5 \left(\tfrac{1}{4}\right)^5 \left(\tfrac{3}{4}\right)^0$$

$$= \left(10 * \tfrac{1}{64} * \tfrac{9}{16}\right) + \left(5 * \tfrac{1}{256} * \tfrac{3}{4}\right) + \left(1 * \tfrac{1}{1024} * 1\right)$$

$$= \frac{90}{1024} + \frac{15}{1024} + \frac{1}{1024}$$

$$\Rightarrow P(X \geq 3) = \frac{106}{1024} = 0.1035$$

Ans

Alternative Method:

$$P(X \geq 3) = 1 - P(X \leq 2)$$

$$= 1 - \sum_{x=0}^{2} b(x; 5, 0.25)$$

$$= \cdot$$

Problem:

The probability that a patient recovers from a delicate heart operation is 0.9. What is the probability that exactly five of the next seven patients undergoing this operation will survive?

Sol^n:

We assume that the operations are made independently and $p = 0.9$ for each of the seven patients and that $n = 7$ and $x = 5$ so that

$$b(5; 7, 0.9) = {}^{7}C_{5} (0.9)^{5} (0.1)^{7-5}$$

$$= 21 * 0.5905 * 0.01$$

$$= 0.1240$$ Ans

Problem:

It is known that 75% of the mice inoculated with a serum are protected from a certain disease. If three mice are inoculated, what is the probability that at most two of the mice will be protected from the disease? Exactly 2 will contact the disease?

Sol^n:

Let X be the number of mice inoculated. Assuming that the inoculation of one mouse is independent of inoculation of the other mice, $p = 3/4$; $n = 3$

$$P(X \le 2) = \sum_{x=0}^{2} b(x; 3, \tfrac{3}{4})$$

$$= {}^3C_0 \left(\tfrac{3}{4}\right)^0 \left(\tfrac{1}{4}\right)^{3-0} + {}^3C_1 \left(\tfrac{3}{4}\right)^1 \left(\tfrac{1}{4}\right)^{3-1} + {}^3C_2 \left(\tfrac{3}{4}\right)^2 \left(\tfrac{1}{4}\right)^{3-2}$$

$$= \left(1 * 1 * \tfrac{1}{64}\right) + \left(3 * \tfrac{3}{4} * \tfrac{1}{16}\right) + \left(3 * \tfrac{9}{16} * \tfrac{1}{4}\right)$$

$$= \frac{1}{64} + \frac{9}{64} + \frac{27}{64}$$

$$= \frac{37}{64}$$

For the second part of the given problem, let Y be the random variable denoting the number of mice will contact the disease.

Here, $p = 1/4$; $n = 3$;

$$\therefore P(Y=2) = {}^3C_2 \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^{3-2}$$

$$= \frac{9}{64}$$

Ans

Problem :

(i) In a binomial distribution, the mean and the standard deviation are 36 and 4.8. Find n and p.

(ii) Is it possible to have a binomial distribution with mean 5 and standard deviation 3 ?

(iii) For a binomial distribution $n=100$. and $p=0.4$, find the mean and variance of the distribution.

Soln :

(i) Given,

$$np = 36 \quad \& \quad \sqrt{npq} = 4.8$$

Since, $npq = (4.8)^2$

$$\Rightarrow 36q = (4.8)^2$$

$$\Rightarrow q = 0.64$$

and $p = 1 - q = 1 - 0.64 = 0.36$

$$\boxed{\therefore p = 0.36}$$

Also, we are given $np = 36$

$$\Rightarrow n = \frac{36}{0.36} = 100$$

$$\boxed{\therefore n = 100}$$

(ii) Given,

Mean = 5

standard deviation = 3

$\therefore np = 5$ & $\sqrt{npq} = 3$

$$\Rightarrow npq = 9$$

$$\Rightarrow 5q = 9$$

$$\Rightarrow q = 9/5 = 1.8$$

$$\therefore q = 1.8$$

which is impossible, since $p + q = 1$

Hence, it is not possible to have a binomial distribution with mean 5 and standard deviation 3. It is important to note that for a binomial distribution, the variance cannot exceed the mean, since $p + q = 1$.

(iii) Given,

$n = 100$

$p = 0.4$

$\therefore$ Mean $= np = 100 \times 0.4 = 40$

and the variance is $npq = np(1-p)$

$= 100 \times 0.4 \times 0.6$

$= 24$

(Ans)

Problem:

The probability that a person recovers from a rare blood disease is 0.4. If 15 people are known to have contacted this disease, what is the probability that (i) at least 10 people survive (ii) from 3 to 8 survive, (iii) exactly 5 survive?

Solⁿ:

Let X be the random variable denoting the number of people surviving. Then

Here, $n = 15$

$p = 0.4$

(i) $P(X \geq 10) = 1 - P(X < 10)$

$= 1 - \sum_{x=0}^{9} b(x; 15, 0.4)$

$= 1 - 0.9662$

$= 0.0338$

(ii) $P(3 \le X \le 8) = \sum_{x=3}^{8} b(x; 15; 0.4)$

$= \sum_{x=0}^{8} b(x; 15, 0.4) - \sum_{x=0}^{2} b(x; 15, 0.4)$

$= 0.9050 - 0.0271$

$= 0.8779$

(iii) $P(X=15) = b(5; 15, 0.4)$

$= \sum_{x=0}^{5} b(x; 15, 0.4) - \sum_{x=0}^{4} b(x; 15, 0.4)$

$= 0.4032 - 0.2173$

$= 0.1859$

Problem:

Twenty percent of the TVs produced in an industry are defective. If 4 TVs are placed in a box for marketing, in how many boxes do you expect to have (i) one defective TV, (ii) two defective TVs, (iii) at most 2 defective TVs in a consignment of 2000 such boxes?

Solⁿ:

Here, $P = 20\% = 0.2$ & $q = 1 - p = 0.8$

If X is the random variable defines the number of defective TVs, then X can assume values 0, 1, 2, 3, 4.

Hence,

(i) For $X = 1$,

$$P(X=1) = {}^4C_1 \left(\frac{1}{5}\right)^1 \left(\frac{4}{5}\right)^{4-1}$$

$$= {}^4C_1 \left(\frac{1}{5}\right) \left(\frac{4}{5}\right)^3$$

$$= \frac{256}{625}$$

$$= 0.4096$$

Hence, the number of boxes having one defective TV is $N * P(X=1) = 2000 * 0.4096 = 819$

(ii) For $X = 2$, $P(X=2) = {}^4C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^{4-2}$

$$= {}^4C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^2$$

$$= \frac{96}{625}$$

$$= 0.1536$$

Hence, the number of boxes having two defective TVs is $N * P(X=2) = 2000 * 0.1536 = 307$.

(iii) For $X \le 2$, $P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$

$$\Rightarrow P(X \le 2) = {}^4C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^4 + {}^4C_1 \left(\frac{1}{5}\right)^1 \left(\frac{4}{5}\right)^3 + {}^4C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^2$$

$$\Rightarrow P(X \le 2) = \frac{256}{625} + \frac{256}{625} + \frac{96}{625}$$

$$= \frac{608}{625}$$

$$= 0.9728.$$

Hence, the number of boxes having at most two defective TVs is $N * P(X \le 2) = 2000 * 0.9728$

$$= 1946$$

(Ans)

Problem :

Seventy percent of the passengers who travel on Mohanagar Provati to Chittagong from Dhaka buy "Daily Star" at the bookstall before they board the train. The train is full and each compartment holds eight passengers.

(a) What is the probability that all the passengers in a compartment have bought the "Daily Star"?

(b) What is the probability that none of the passengers in a compartment has bought the "Daily Star"?

(c) What is the probability that exactly three passengers in a compartment have bought the "Daily Star"?