Probability and Statistics

UNIT -2

Probability distributions: Recap of probability theory (definition, addition rule, multiplication rule, conditional probability). Random variables, Discrete and continuous probability distributions. Binomial, Poisson, exponential and normal distributions (derivation of mean and variance for all distributions).

Probability:

Probability is a numerical measure which indicates the chance of occurrence.

Exhaustive event:

An event consisting of all the various possibilities is called an Exhaustive event.

Mutually exclusive events:

Two or more events are said to be mutually exclusive if the happening of one event prevents the simultaneous happenings of the others.

Example:

In tossing a coin, getting head and tail are mutually exclusive in view of the fact that if head is the turn out, getting tail is not possible.

Independent Events:

Two or more events are said to be independent if the happening or not happening of one event does not prevent the happening or non-happening of the other.

Example:

When two coins are tossed the event of getting head is an independent event as both the coins can turn out heads.

Probability: (Mathematical definition)

If S is an outcome set (finite set) and E is an event (which is a specified subset of S) then the probability of E denoted by $P(E)$ is given by

$$P(E) = \frac{O(E)}{O(S)} \qquad (1)$$

Note:

(1) If $E = \phi$ then $P(E) = P(\phi) = 0$.

(2) If $E = S$ then $P(E) = P(S) = 1$.

(3) If E is a proper subset of S then $0 < P(E) < 1$. Thus $P(E)$ always satisfies the satisfies the inequality $0 \leq P(E) \leq 1$.

(4) If $P(E) = 0$ then we say that $P(E)$ is an impossible probability. If $P(E) = 1$ we say that $P(E)$ is a sure Probability.

(5) If $\overline{E} = S - E$, that is if $\overline{E}$ is the complement of E in S, then the formula (1) gives

$$P(\overline{E}) = \frac{O(\overline{E})}{O(S)} = \frac{O(S) - O(E)}{O(S)} = 1 - P(E)$$

$$P(\overline{E}) = 1 - P(E) \, (2)$$

Addition Theorem:

If E and F are two events then $E \cup F$ is an event consisting of outcomes that are in E or F or both and $E \cap F$ is an event consisting of outcomes that are common to E and F, so that

$$O(E \cup F) = O(E) + O(F) - O(E \cap F)$$

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

Mutually exclusive events:

Two events E and F are called mutually exclusive events if they are disjoint sub sets of the outcome set. That is $E \cap F = \phi$.

Then the formula $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

becomes $P(E \cup F) = P(E) + P(F)$. (4)

Note: In general if $E_1, E_2, E_3, ...E_n$ are n mutually exclusive (disjoint) events, then

$$P(E_1 \cup E_2 \cup E_3 \cup, ... \cup E_n) = P(E_1) + P(E_2) + P(E_3) + ... + P(E_n)$$

Conditional Probability:

Let E be an event in a finite outcome set S with $P(E) > 0$. Then the probability that an event A occurs when E has already occurred called the probability of A relative to E (or the conditional probability of A given E) and denoted by $P(A/E)$ is defined by

$$P(A/E) = \frac{O(A \cap E)}{O(E)} \qquad (1)$$

Multiplication Theorem:

Since E and $A \cap E$ are events of S, we have

$$P(E) = \frac{O(E)}{O(S)},\ P(A \cap E) = \frac{O(A \cap E)}{O(S)}$$

$$\therefore \frac{P(A \cap E)}{P(E)} = \frac{O(A \cap E)}{O(E)} = P(A/E)$$

So that $P(A \cap E) = P(E \cap A) = P(E)P(A/E)$ (2)

This result is called the multiplication theorem for conditional probability or theorem of compound probability.

Multiplication Theorem for Mutually Independent events:

Two events E and F are said to be mutually independent if the probability of E is equal to the probability of E relative to $F\{i.e..\ P(E) = P(E/F)\}$ and vice-versa. For such events, the result (2) becomes $P(E \cap F) = P(E)P(F)$

Baye's Theorem: An event A corresponds to a number of exhaustive events $B_1, B_2, B_3, ...B_n$. If $P(B_i)$ and $P(A/B_i)$ are given, then $P(B_i / A) = \dfrac{P(B_i)P(A/B_i)}{\sum P(B_i)P(A/B_i)}$.

Random Variable: (Stochastic variable)

Random Variable is a function which assigns a real number to every outcome in the sample space. The set of such real values is the range of the random variable.

(or)

In a random experiment if a real variable x is associated with every outcome then it is called a random variable or stochastic variable. This is equivalent to having a function defined on the sample space S and this function is called a random function or stochastic function.

Note: It may be noted that different random variables may be associated with the same sample space S.

Example: While tossing a coin, suppose the value 1 is associated for the outcome head and 0 for the outcome tail. We have

$S = \{H, T\}$ and the corresponding random variable $X = \{1, 0\}$.

Example: Suppose a coin is tossed twice we shall associate two different random variables X, Y defined below. We have the sample space

$S = \{HH, HT, TH, TT\}$

X – Number of heads in the outcome.

The association of the elements in S to X is as below

Outcome	HH	HT	TH	TT
Random variable X	2	1	1	0

Suppose Y - Number of tails in the outcome.

The association of the elements in S to Y is as below

Outcome	HH	HT	TH	TT
Random variable Y	0	1	1	2

Discrete Random Variable:

If a random variable takes finite or countable infinite (i.e. sequence of real numbers) number of values then it is called a discrete random variable.

Continuous Random Variable:

If a random variable takes non – countable infinite number of values then it is called a non discrete or continuous random variable.

Discrete (finite) Probability distribution:

Definition:

Let X be a discrete random variable and the possible values which it can assume be x_1, x_2, x_3, …. . If for each x_i we assign a real number $p(x_i)$ such that

(i) $p(x_i) \geq 0$

(ii) $\sum_i p(x_i) = 1$

then $P(X = x_i) = p(x_i)$ is called a probability density function or probability distribution (p. d. f.). The set of values $[x_1, p(x_i)]$ constitute a discrete probability distribution.

Cumulative distribution function (c. d. f.):

The distribution function $f(x)$ of a discrete variable X is defined to be

$$f(x) = P(X = x_i) = \sum_{i=1}^{i} p(x_i)$$

where i is any integer is called as the cumulative distribution function (c. d. f.).

Mean,Variance and Standard deviation of the discrete probability distribution:

Mean $= \mu = \sum_i x_i p(x_i)$

Variance $= V = \sum_i (x_i - \mu)^2 \ p(x_i)$

Standard Deviation $= \sigma = \sqrt{V}$

Note:

Variance can also be put in the form $= V = \sum_i x_i^2 p(x_i) - \left\{\sum_i x_i p(x_i)\right\}^2 = \sum_i x_i^2 p(x_i) - \mu^2$

Problem: A coin is tossed twice. A random variable X represents the number of heads turning up. Find the discrete probability distribution for X. Also find its mean and variance.

Solution: $S = \{HH, HT, TH, TT\}$, the association of the elements of S to the random variable X are respectively 2, 1, 1, 0.

$$P(HH) = \frac{1}{4}, P(HT) = \frac{1}{4}, P(TH) = \frac{1}{4}, P(TT) = \frac{1}{4}$$

Now $p(X = 0$ i.e. no head$) = p(TT) = \frac{1}{4}$

$p(X = 1$ i.e. one head$) = p(HT \cup TH) = p(HT) + p(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$p(X = 2$ i.e. two heads$) = p(HH) = \frac{1}{4}$

The discrete probability distribution for X is as follows

$X = x_i$	0	1	2
$p(x_i)$	1/4	1/2	1/4

We observe that $p(x_i) \geq 0$ and $\sum p(x_i) = 1$

$$\text{Mean } = \mu = \sum_i x_i p(x_i) = (0)(1/4) + (1)(1/2) + (2)(1/4) = 1$$

$$\text{Variance } = V = \sum_i (x_i - \mu)^2 \; p(x_i) = (0-1)^2(1/4) + (1-1)^2(1/2) + (2-1)^2(1/4) = 1/2$$

Thus we have mean $\mu = 1$ and Variance $V = 1/2$.

Problem: Show that the following distribution represents a discrete probability distribution. Find mean and variance.

x	10	20	30	40
$p(x)$	1/8	3/8	3/8	1/8

Solution: We observe that $p(x_i) \geq 0 \; \forall x$ and $\sum p(x_i) = 1$

Hence the given distribution represents a discrete probability distribution.

$$\text{Mean } = \mu = \sum_i x_i p(x_i) = (10)(1/8) + (20)(3/8) + (30)(3/8) + (40)(1/8) = \frac{200}{8} = 25$$

$$\text{Variance } = V = \sum_i (x_i - \mu)^2 \; p(x_i)$$

$$= (10-25)^2(1/8) + (20-25)^2(3/8) + (30-25)^2(3/8) + (40-25)^2(1/8) = \frac{600}{8} = 75$$

Thus we have mean $\mu = 25$ and Variance $V = 75$.

Problem: Find the value of the constant *k* and the mean and the variance of the following distribution. Also find *P*(*X*> -1)

x	-2	-1	0	1	2	3
$p(x)$	0.1	k	0.2	$2k$	0.3	k

Solution: Since the given distribution is a discrete probability distribution

Therefore $\sum p(x) = 1$

$$\Rightarrow \sum p(x) = 0.1 + k + 0.2 + 2k + 0.3 + k = 1$$

$$4k + 0.6 = 1$$

$$4k = 0.4$$

$$k = 0.1$$

The discrete probability distribution is as follows

x	-2	-1	0	1	2	3
$p(x)$	0.1	0.1	0.2	0.2	0.3	0.1

Mean $= \mu = \sum_i x_i p(x_i) = (-2)(0.1) + (-1)(0.1) + (0)(0.2) + (1)(0.2) + (2)(0.3) + (3)(0.1) = 0.8$

Variance $= V = \sum_i (x_i - \mu)^2 \; p(x_i) = (-2-0.8)^2(0.1) + (-1-0.8)^2(0.1) + (0-0.8)^2(0.2)$

$+ (1-0.8)^2(0.2) + (2-0.8)^2(0.3) + (3-0.8)^2(0.1) = 2.16$

P(*X*> -1) = *p*(0) + *p*(1) + *p*(2) + *p*(3) = 0.2 + 0.2 + 0.3 + 0.1 = 0.8.

Problem: The probability density function of a variable X is

x	0	1	2	3	4	5	6
$p(x)$	k	$3k$	$5k$	$7k$	$9k$	$11k$	$13k$

Find (i) $P(X< 4)$ (ii) $P(X\geq 5)$ (iii) $P(3 <X\leq 6)$

Solution:

We must have $p(x)\geq 0\ \forall x$ and $\sum p(x)=1$.

Therefore, $k + 3k + 5k + 7k + 9k + 11k + 13k = 1$

$49k = 1 \Rightarrow k = \dfrac{1}{49}$

The given table becomes

x	0	1	2	3	4	5	6
$p(x)$	$\frac{1}{49}$	$\frac{3}{49}$	$\frac{5}{49}$	$\frac{7}{49}$	$\frac{9}{49}$	$\frac{11}{49}$	$\frac{13}{49}$

(i) $P(x<4)=p(0)+p(1)+p(2)+p(3)$

$$P(x<4)=\frac{1}{49}+\frac{3}{49}+\frac{5}{49}+\frac{7}{49}=\frac{16}{49}=0.3265$$

(ii) $P(x\geq 4)=p(4)+p(5)+p(6)$

$$P(x\geq 4)=\frac{9}{49}+\frac{11}{49}+\frac{13}{49}=\frac{33}{49}=0.6734$$

(iii) $P(3<x\leq 6)=p(4)+p(5)+p(6)=\dfrac{9}{49}+\dfrac{11}{49}+\dfrac{13}{49}=\dfrac{33}{49}=0.6734$

Problem: Find the mean and variance of the probability distribution given by the following table

x_i	1	2	3	4	5
$p(x_i)$	0.2	0.35	0.25	0.15	0.05

Solution:

Mean $=\mu=\sum_i x_i p(x_i)=(1)(0.2)+(2)(0.35)+(3)(0.25)+(4)(0.15)+(5)(0.05)=2.5$

Variance $=V=\sum_i \left(x_i-\mu\right)^2 p(x_i)$

$$=(1-2.5)^2(0.2)+(2-2.5)^2(0.35)+(3-2.5)^2(0.25)$$
$$+(4-2.5)^2(0.15)+(5-2.5)^2(0.05)=1.25$$

Problem: Find the value of k such that the following distribution represents a finite probability distribution. Hence find its mean and standard deviation. Also find $P(X \leq 1)$, $P(X > 1)$ and $P(-1 < X \leq 2)$.

x	-3	-2	-1	0	1	2	3
$p(x)$	k	$2k$	$3k$	$4k$	$3k$	$2k$	k

Solution: The given distribution is a probability distribution, hence we must have

$p(x) \geq 0 \ \forall x$ and $\sum p(x) = 1$.

Therefore, $k + 2k + 3k + 4k + 3k + 2k + k = 1$

$16k = 1 \Rightarrow k = \dfrac{1}{16}$

The given table becomes

x	-3	-2	-1	0	1	2	3
$p(x)$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{1}{4}$	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

$$\text{Mean } = \mu = \sum_i x_i p(x_i) = (-3)\left(\frac{1}{16}\right) + (-2)\left(\frac{1}{8}\right) + (-1)\left(\frac{3}{16}\right) + (0)\left(\frac{1}{4}\right)$$

$$+(1)\left(\frac{3}{16}\right) + (2)\left(\frac{1}{8}\right) + (3)\left(\frac{1}{16}\right) = 0$$

$$\text{Variance } = V = \sum_i \left(x_i - \mu\right)^2 p(x_i)$$

$$= (-3-0)^2\left(\frac{1}{16}\right) + (-2-0)^2\left(\frac{1}{8}\right) + (-1-0)^2\left(\frac{3}{16}\right) + (0-0)^2\left(\frac{1}{4}\right)$$

$$+(1-0)^2\left(\frac{3}{16}\right) + (2-0)^2\left(\frac{1}{8}\right) + (3-0)^2\left(\frac{1}{16}\right)$$

$$= \frac{9}{16} + \frac{4}{8} + \frac{3}{16} + 0 + \frac{3}{16} + \frac{4}{8} + \frac{9}{16} = \frac{40}{16} = 2.5$$

$$P(X \leq 1) = \frac{1}{16} + \frac{1}{8} + \frac{3}{16} + \frac{1}{4} + \frac{3}{16} = \frac{13}{16}$$

$$P(X > 1) = \frac{1}{8} + \frac{1}{16} = \frac{3}{16}$$

P(-1 <X≤ 2). $P(-1 < X \leq 2) = \frac{1}{4} + \frac{3}{16} + \frac{1}{8} = \frac{9}{16}$

Problem: Define random variable. A random variable $X(= x)$ has the following probability distribution

x	0	1	2	3	4	5	6	7
$p(x)$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

Find (1) k (ii) $P(X< 6)$ (iii) $P(X\geq 6)$.

Solution: The given distribution is a probability distribution, hence we must have

$p(x) \geq 0 \ \forall x$ and $\sum p(x) = 1$.

Therefore, $0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$

$10k^2 + 9k - 1 = 0$

$10k^2 + 10k - k - 1 = 0$

$10k(k+1) - (k+1) = 0$

$(10k-1)(k+1) = 0$

$\Rightarrow k = \frac{1}{10}, k = -1$

Since $p(x) \geq 0$, hence $k \neq -1$

Hence $k = \frac{1}{10}$

The given table becomes

x	0	1	2	3	4	5	6	7
$p(x)$	0	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{1}{100}$	$\frac{1}{50}$	$\frac{17}{100}$

$$P(X < 6) = 0 + \frac{1}{10} + \frac{1}{5} + \frac{1}{5} + \frac{3}{10} + \frac{1}{100} = \frac{81}{100}$$

$$P(X \geq 6) = \frac{1}{50} + \frac{17}{100} = \frac{19}{100}$$

Problem: The probability density function of a variable X is

x	0	1	2	3	4	5
$p(x)$	k	$3k$	$5k$	$7k$	$9k$	$11k$

Find (i) $P(X< 3)$ (iii) $P(3 <X \leq 5)$

Solution: The given distribution is a probability distribution, hence we must have

$p(x) \geq 0\ \forall x$ and $\sum p(x) = 1$.

Therefore, $k + 3k + 5k + 7k + 9k + 11k = 1$

$36k = 1 \Rightarrow k = \frac{1}{36}$

The given table becomes

x	0	1	2	3	4	5
$p(x)$	$\frac{1}{36}$	$\frac{1}{12}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{1}{4}$	$\frac{11}{36}$

(i) $P(x < 3) = \frac{1}{36} + \frac{1}{12} + \frac{5}{36} = \frac{9}{36} = 0.25$

(ii) $P(3 < X \leq 5) = \frac{1}{4} + \frac{11}{36} = \frac{20}{36} = \frac{5}{9}$

Problem: Suppose a random variable X takes the values -3, -1, 2 and 5 with respective probabilities $\frac{2k-3}{10}, \frac{k-2}{10}, \frac{k-1}{10}, \frac{k+1}{10}$ find the value of k and

(i) $P(-3 <X< 4)$ (ii) $P(X\leq 2)$.

Solution: The given probability distribution is

x	-3	-1	2	5
$p(x)$	$\frac{2k-3}{10}$	$\frac{k-2}{10}$	$\frac{k-1}{10}$	$\frac{k+1}{10}$

The given distribution is a probability distribution, hence we must have

$p(x) \geq 0\ \forall x$ and $\sum p(x) = 1$.

Therefore, $\frac{2k-3}{10}+\frac{k-2}{10}+\frac{k-1}{10}+\frac{k+1}{10}=1$

$5k-5=10 \Rightarrow 5k=15 \Rightarrow k=3$

The given table becomes

x	-3	-1	2	5
$p(x)$	0.3	0.1	0.2	0.4

(i) $P(-3<X<5)=0.1+0.2=0.3$

(ii) $P(X\le 2)=0.3+0.1+0.2=0.6$

Problem: A coin is such that head occurs 3 times as often as tails. This coin is tossed 3 times. Let *X* be the random variable, the number of heads that appear. Find the probability mass function of *X*.

Bernoulli Process:

A random variable X that takes the values either 0 or 1 is known as a Bernoulli variable. Equivalently we can say that a random experiment with only two types of outcomes, success or failure is called a Bernoulli trial. The corresponding distribution is known as Bernoulli distribution. Thus we have

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

Bernoulli's Theorem:

The probability of x success out of n trials is ${}^nC_x p^x q^{n-x}$

Binomial Distribution:

If p is the probability of success and q is the probability of failure, the probability of x success out of n trials is given by ${}^nC_x p^x q^{n-x}$. We form the following probability distribution of $[x, p(x)]$ where $x = 0, 1, 2, 3, \ldots\ldots. n$.

No. of success x	0	1	2	3	-	-	-	-	n
Probability $p(x)$	q^n	${}^nC_1 pq^{n-1}$	${}^nC_2 p^2 q^{n-2}$	${}^nC_3 p^3 q^{n-3}$	-	-	-	-	p^n

Here $p(x) \geq 0 \quad \forall x$ and

$$\sum p(x) = q^n + {}^nC_1 pq^{n-1} + {}^nC_2 p^2 q^{n-2} + {}^nC_3 p^3 q^{n-3} + \ldots\ldots\ldots\ldots + p^n = 1$$

Hence $p(x)$ is a probability density function.

Note 1: The values of $p(x)$ for different values of $x = 0, 1, 2, 3\ldots\ldots n$ are the various terms in the expansion of $(p+q)^n$. Therefore, this distribution is called as Binomial distribution.

Note 2: The probability function $p(x) = {}^nC_x p^x q^{n-x}$ is also denoted by

$$b(n, p, x) = b(x) = {}^nC_x p^x q^{n-x} = {}^nC_x p^x (1-p)^{n-x}$$

Mean and Standard Deviation of the Binomial Distribution:

$$\text{Mean} = \mu = \sum_{x=0}^{n} xp(x) = \sum_{x=0}^{n} x \ {}^nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x \left\{ \frac{n!}{x!(n-x)!} \right\} p^x q^{n-x}$$

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x q^{n-x} = \sum_{x=1}^{n} \frac{n(n-1)!}{(x-1)!(n-x)!} pp^{x-1} q^{n-x}$$

$$= np\sum_{x=1}^{n}\frac{(n-1)!}{(x-1)!\{(n-1)-(x-1)\}!}\,p^{x-1}q^{(n-1)-(x-1)}$$

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

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

$$\mu = np$$

Thus the mean of the Binomial distribution is $\mu = np$.

$$\text{Variance } = V = \sum_{x=0}^{n} x^2 p(x) - \mu^2 \qquad (1)$$

Consider $\sum_{x=0}^{n} x^2 p(x) = \sum_{x=0}^{n}\{x(x-1)+x\}\,{}^{n}C_x p^x q^{n-x}$

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

$$= \sum_{x=0}^{n} x(x-1)\left\{\frac{n!}{x!(n-x)!}\right\} p^x q^{n-x} + \mu$$

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

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

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

$$\sum_{x=0}^{n} x^2 p(x) = n(n-1)p^2 + np$$

Substituting above value in (1) we get

$$V = n(n-1)p^2 + np - \mu^2 = n(n-1)p^2 + np - n^2p^2 \quad \because \mu = np$$

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

$$= np(1-p) = npq$$

$$V = npq$$

Standard deviation = $\sigma = \sqrt{V} = \sqrt{npq}$

Problem: Four coins are tossed 100 times and the following results were obtained. Fit a binomial distribution for the data and calculate the theoretical frequencies

No of heads	0	1	2	3	4
Frequency	5	29	36	25	5

Solution: We shall first find the mean for the given distribution using

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{(0)(5)+(1)(29)+(2)(36)+(3)(25)+(4)(5)}{100} = \frac{196}{100} = 1.96$$

But for the binomial distribution $\bar{x} = np$ where $n = 4$.

Therefore $\bar{x} = np = 1.96$

$$p = \frac{1.96}{n} = \frac{1.96}{4} = 0.49$$

Also $q = 1 - p = 1 - 0.49 = 0.51$

Now the binomial distribution of fit is $N(p+q)^n$

Where N = Total frequency = $\sum f$

Thus the required fit is $N(p+q)^4 = N(q^4 + {}^4C_1 q^3 p + {}^4C_2 q^2 p^2 + {}^4C_3 q p^3 + p^4)$

$$N(p+q)^4 = Nq^4 + N\,{}^4C_1 q^3 p + N\,{}^4C_2 q^2 p^2 + N\,{}^4C_3 q p^3 + p^4$$

The theoretical frequencies are obtained by writing the various terms in the above binomial expansion individually as shown below

$Nq^4 = 100(0.51)^4 = 7$

$N\,{}^4C_1 q^3 p = 100 \times 4 \times (0.51)^3 (0.49) = 26$

$N\,{}^4C_2 q^2 p^2 = 100 \times 6 \times (0.51)^2 (0.49)^2 = 37$

$N\,{}^4C_3 q p^3 = 100 \times 4 \times (0.51)(0.49)^3 = 24$

$Np^4 = 100(0.49)^4 = 6$

The theoretical frequencies are 7, 26, 37, 24, 6.

Problem: A die is tossed thrice. A success is getting 1 or 6. Find the mean and variance of the number of success.

Solution: If a die is tossed, the probability of getting 1 or 6 $=\frac{2}{6}=\frac{1}{3}$

Therefore $q=1-p=1-\frac{1}{3}=\frac{2}{3}$ and $n = 3$.

Mean $=\mu=np=3\times\frac{1}{3}=1$

Variance $=V=npq=3\times\frac{1}{3}\times\frac{2}{3}=\frac{2}{3}$

Problem: In 800 families with 5 children each how many families would be expected to have

(i) 3 boys and 2 girls (ii) 2 boys and 3 girls (iii) no girls (iv) at most 2 girls by assuming probability for boys and girls to be equal.

Solution: p = Probability of having a boy = $1/2$

q = Probability of having a girl = $1/2$

Also N = 800

The binomial distribution is $=N(p+q)^n=800(p+q)^5$

(i) Number of families with 3 boys and 2 girls $=800\,{}^5C_3p^3q^2=800\,{}^5C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2$

$$=800\times10\times\left(\frac{1}{2}\right)^5=8000\times\frac{1}{32}=250$$

(ii) Number of families with 2 boys and 3 girls $=800\,{}^5C_2p^2q^3=800\,{}^5C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3$

$$=800\times10\times\left(\frac{1}{2}\right)^5=8000\times\frac{1}{32}=250$$

(iii) Number of families with no girls that is 5 boys $=800\,{}^5C_5p^5q^0=800\,{}^5C_5\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^0$

$$= 800\left(\frac{1}{2}\right)^5 = 800 \times \frac{1}{32} = 25$$

(iv) Number of families at most 2 girls i.e. having 5 boys and 0 girls (or) 4 boys and 1 girl (or) 3 boys and 2 girls

$$= 800\left\{{}^5C_5 p^5 q^0 + {}^5C_4 p^4 q^1 + {}^5C_3 p^3 q^2\right\} = 800\left\{\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^0 + 5\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^1 + 10\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2\right\}$$

$$= 800 \times \frac{1}{32} \times \{1 + 5 + 10\} = 400.$$

Problem: In a sampling a large number of parts manufactured by a machine, the mean number of defectives in a sample of 20 is 2. Out of 1000 such samples, how many would be expected to contain at least 3 defective parts?

Solution: Let p be the probability that a sample of items contain defective part. It is given that the mean $\mu = 2$ for a sample size of $n = 20$. Therefore, the formula $\mu = np$ (binomial distribution) gives $p = \frac{\mu}{n} = \frac{2}{20} = 0.1$, hence $q = 1 - p = 1 - 0.1 = 0.9$

Thus the probability function for the distribution is $b(n, p, x) = {}^nC_x p^x q^{n-x}$

$$b(20, 0.1, x) = {}^{20}C_x\,(0.1)^x\,(0.9)^{20-x} \qquad (1)$$

The probability that there are at least 3 defective parts in a sample is

$$b(20, 0.1, x \geq 3) = 1 - b(20, 0.1, x < 3)$$

$$= 1 - \{b(20, 0.1, 0) + b(20, 0.1, 1) + b(20, 0.1, 2)\}$$

$$= 1 - \left\{{}^{20}C_0\,(0.1)^0\,(0.9)^{20-0} + {}^{20}C_1\,(0.1)^1\,(0.9)^{20-1} + {}^{20}C_2\,(0.1)^2\,(0.9)^{20-2}\right\}$$

$$= 1 - \left\{(0.9)^{20} + 20 \times (0.1)(0.9)^{19} + 190 \times (0.1)^2\,(0.9)^{18}\right\} = 0.323$$

Hence out of 1000 samples the expected number of samples that contain at least 3 defectives is $= 0.323 \times 1000 = 323$.

Problem: Let X be a binomial distributed random variable based on 6 repetitions of an experiment. If $p = 0.3$, evaluate (i) $P(X \leq 3)$ (ii) $P(X = 4)$ (iii) $P(X > 4)$.

Solution: Here $p = 0.3$, therefore $q = 1 - p = 1 - 0.3 = 0.7$

The binomial distribution function is $b(n, p, x) = {}^nC_x p^x q^{n-x}$

$$b(6, 0.3, x) = {}^6C_x\,(0.3)^x\,(0.7)^{6-x}$$

(i) $P(X \le 3) = b(6,0.3,0) + b(6,0.3,1) + b(6,0.3,2) + +b(6,0.3,3)$

$$= {}^6C_0(0.3)^0(0.7)^{6-0} + {}^6C_1(0.3)^1(0.7)^{6-1} + {}^6C_2(0.3)^2(0.7)^{6-2} + {}^6C_3(0.3)^3(0.7)^{6-3}$$

$$= (0.7)^6 + 6(0.3)(0.7)^5 + 15(0.3)^2(0.7)^4 + 20(0.3)^3(0.7)^3 = 0.9295$$

$P(X \le 3) = 0.9295$

(ii) $P(X = 4) = b(6,0.3,4) = {}^6C_4(0.3)^4(0.7)^{6-4} = 15 \times (0.3)^4(0.7)^2 = 0.0595$

$P(X = 4) = 0.0595$

(iii) $P(X > 4) = b(6,0.3,5) + b(6,0.3,6) = {}^6C_5(0.3)^5(0.7)^{6-5} + {}^6C_6(0.3)^6(0.7)^{6-6}$

$$= 6 \times (0.3)^5(0.7) + (0.3)^6 = 0.0109$$

$P(X > 4) = 0.0109$

Problem: In a shop there are 5 chewing gums. Two of them have photo stickers of Sachin Tendulkar hidden inside their outer cover. A boy buys 2 chewing gums by random choice. Find the probability that he gets at least one photo sticker.

Solution: The probability that there is a photo sticker in a chosen item is $= p = \frac{2}{5} = 0.4$

Thus $q = 1 - p = 1 - \frac{2}{5} = \frac{3}{5} = 0.6$

The binomial distribution function is $b(n,p,x) = {}^nC_x p^x q^{n-x}$

$b(2,0.4,x) = {}^2C_x(0.4)^x(0.6)^{2-x}$

The probability that there is at least one photo sticker in two choices is

$= b(2,0.4,1) + b(2,0.4,2) = {}^2C_1(0.4)^1(0.6)^{2-1} + {}^2C_2(0.4)^2(0.6)^{2-2}$

$= 2 \times (0.4)(0.6) + (0.4)^2 = 0.64$

Problem: Six coins are tossed; find the probability of getting (i) exactly 3 heads (ii) at most 3 heads (iii) at least 1 head.

Solution: Here n = 6 and $p = \frac{1}{2}$ and $q = 1 - p = \frac{1}{2}$

The binomial distribution function is $b(n,p,x) = {}^nC_x p^x q^{n-x}$

$$b(6,0.5,x) = {}^6C_x\,(0.5)^x\,(0.5)^{6-x} = {}^6C_x\,(0.5)^6 = \frac{1}{64}\,{}^6C_x$$

(i) The probability of getting exactly 3 heads $b(6,0.5,3) = \frac{1}{64}\,{}^6C_3 = 0.3125$

(ii) The probability of getting at most 3 heads
$b(6,0.5,x \le 3) = b(6,0.5,0) + b(6,0.5,1) + b(6,0.5,2) + b(6,0.5,3)$

$$b(6,0.5,x \le 3) = \frac{1}{64}\,{}^6C_0 + \frac{1}{64}\,{}^6C_1 + \frac{1}{64}\,{}^6C_2 + \frac{1}{64}\,{}^6C_3$$

$$= \frac{1}{64}\left({}^6C_0 + {}^6C_1 + {}^6C_2 + {}^6C_3\right)$$

$$= \frac{1}{64}(1+6+15+20) = \frac{42}{64} = 0.6562$$

(ii) The probability of getting at least 1 head
$b(6,0.5,x \ge 1) = 1 - b(6,0.5,x<1) = 1 - b(6,0.5,0)$

$$= 1 - \frac{1}{64}\,{}^6C_0 = \frac{63}{64} = 0.9843$$

Problem: If the chance that one of the ten telephone lines is busy at an instant is 0.2. (i) What is the chance that 5 of the lines are busy? (ii) What is the probability that all the lines are busy? (iii) What is the most probable number of busy lines? What is the probability of this number?

Solution: Here $p = 0.2$, therefore $q = 1 - p = 1 - 0.2 = 0.8$

The binomial distribution function is $b(n,p,x) = {}^nC_x\,p^x q^{n-x}$

$$b(10,0.2,x) = {}^{10}C_x\,(0.2)^x\,(0.8)^{10-x}$$

(i)What is the chance that 5 of the lines are busy?

$$b(10,0.2,5) = {}^{10}C_5\,(0.2)^5\,(0.8)^{10-5} = {}^{10}C_5\,(0.2)^5\,(0.8)^5 = 0.0264$$

ii) What is the probability that all the lines are busy?

$$b(10,0.2,10) = {}^{10}C_{10}\,(0.2)^{10}\,(0.8)^{10-10} = (0.2)^{10} = 0.024\times10^{-7}$$

(iii) What is the most probable number of busy lines? What is the probability of this number?

Most probable number of busy lines $= 0.2\times10 = 2$

The probability of most probable number of busy lines is

$b(10,0.2,2) = {}^{10}C_2\,(0.2)^2\,(0.8)^{10-2} = {}^{10}C_2\,(0.2)^2\,(0.8)^8 = 0.3019$

Problem: Two persons A and B play a game in which their chances of winning are in the ratio 3:2. Find A's chance of winning at least three games out of 6 games played.

Solution: From the given data

The probability that A wins the game is $= p = \dfrac{3}{5}$, thus $q = 1 - p = 1 - \dfrac{3}{5} = \dfrac{2}{5}$

Here $n = 6$

Therefore the probability function giving x success for A in 6 games

$$b(6,\frac{3}{5},x) = {}^{6}C_x \left(\frac{3}{5}\right)^x \left(\frac{2}{5}\right)^{6-x}$$

Therefore the probability of A winning at least three games

$$= b(6,\frac{3}{5},x \geq 3) = 1 - b(6,\frac{3}{5},x < 3) = 1 - \left\{b(6,\frac{3}{5},0) + b(6,\frac{3}{5},1) + b(6,\frac{3}{5},2)\right\}$$

$$= 1 - \left\{ {}^{6}C_0 \left(\frac{3}{5}\right)^0 \left(\frac{2}{5}\right)^{6-0} + {}^{6}C_1 \left(\frac{3}{5}\right)^1 \left(\frac{2}{5}\right)^{6-1} + {}^{6}C_2 \left(\frac{3}{5}\right)^2 \left(\frac{2}{5}\right)^{6-2} \right\}$$

$$= 1 - \left\{ \left(\frac{2}{5}\right)^6 + 6 \times \left(\frac{3}{5}\right)\left(\frac{2}{5}\right)^5 + 15 \times \left(\frac{3}{5}\right)^2 \left(\frac{2}{5}\right)^4 \right\} = 0.8208$$

Problem: The probability that a man aged 60 will live to be 70 is 0.65. What is the probability that out of ten men now aged 60.

(i) at least 7 will live to be 70.

(ii) exactly 9 will live up to 70.

(iii) at most 9 will live up to 70.

Solution: Here $p = 0.65$, therefore $q = 1 - p = 1 - 0.65 = 0.35$

The binomial distribution function is $b(n, p, x) = {}^{n}C_x p^x q^{n-x}$

$$b(10,0.65,x) = {}^{10}C_x\,(0.65)^x\,(0.35)^{10-x}$$

(i)Probability that at least 7 will live to be 70.

$b(10,0.65,x \geq 7) = b(10,0.65,7) + b(10,0.65,8) + b(10,0.65,9) + b(10,0.65,10)$

$$= {}^{10}C_7\,(0.65)^7\,(0.35)^{10-7} + {}^{10}C_8\,(0.65)^8\,(0.35)^{10-8} + {}^{10}C_9\,(0.65)^9\,(0.35)^{10-9} + {}^{10}C_{10}\,(0.65)^{10}\,(0.35)^{10-10}$$

$$= {}^{10}C_7\,(0.65)^7\,(0.35)^3 + {}^{10}C_8\,(0.65)^8\,(0.35)^2 + {}^{10}C_9\,(0.65)^9\,(0.35) + {}^{10}C_{10}\,(0.65)^{10}$$

$= 0.2522 + 0.1756 + 0.0724 + 0.0134 = 0.5136$

ii) Probability that exactly 9 will live to be 70.

$b(10, 0.65, 9) = {}^{10}C_9\,(0.65)^9\,(0.35)^{10-9} = {}^{10}C_9\,(0.65)^9\,(0.35) = 0.07249$

(iii) Probability that at most 9 will live to be 70.

$b(10, 0.65, x \le 9) = 1 - b(10, 0.65, x > 9) = 1 - b(10, 0.65, 10)$

$= 1 - 0.0134 = 0.9866$

Problem: A cricket team has 2/3 chance of winning a match. If it plays 4 matches, find the probability that it wins (i) 2 matches (ii) at least one match.

Solution: Probability of a cricket team's chance of winning a match $p = 2/3$, thus $q = 1 - p = 1 - (2/3) = 1/3$

Here $n = 4$

The binomial distribution function is $b(n, p, x) = {}^{n}C_x\, p^x q^{n-x}$

$b(4, 2/3, x) = {}^{4}C_x\,(2/3)^x\,(1/3)^{4-x}$

(i) Probability that the cricket team wins 2 matches

$b(4, 2/3, 2) = {}^{4}C_2\,(2/3)^2\,(1/3)^{4-2}$

$= {}^{4}C_2\,(2/3)^2\,(1/3)^2 = 0.2962$

(ii) Probability that the cricket team wins at least one match

$b(4, 2/3, x \ge 1) = 1 - b(4, 2/3, x < 1)$

$= 1 - b(4, 2/3, 0) = 1 - {}^{4}C_0\,(2/3)^0\,(1/3)^{4-0}$

$= 1 - (1/3)^4 = 0.9876$

Problem: When a coin is tossed 4 times, find the probability of getting (i) exactly one head (ii) at most 3 heads and (iii) at least two heads.

Solution: If a coin is tossed probability of getting head $= p = 1/2$

$q = 1 - p = 1 - 1/2 = 1/2$, $n = 4$

The binomial distribution function is $b(n, p, x) = {}^{n}C_x\, p^x q^{n-x}$

$b(4, 0.5, x) = {}^{4}C_x\,(0.5)^x\,(0.5)^{4-x} = {}^{4}C_x\,(0.5)^4 = {}^{4}C_x\,(0.0625)$

$b(4,0.5,x) = (0.0625)\ {}^4C_x$

(i)Probability of getting exactly one head

$b(4,0.5,1) = (0.0625)\ {}^4C_1 = 0.25$

(ii)Probability of getting at most 3 heads

$b(4,0.5,x \le 3) = 1 - b(4,0.5,x > 3) = 1 - b(4,0.5,4)$

$= 1 - (0.0625)\ {}^4C_4 = 0.9375$

(iii)Probability of getting at least two heads

$b(4,0.5,x \ge 2) = 1 - b(4,0.5,x < 2) = 1 - \{b(4,0.5,0) + b(4,0.5,1)\}$

$= 1 - \{(0.0625)\ {}^4C_0 + (0.0625)\ {}^4C_1\} = 1 - 0.0625\{{}^4C_0 + {}^4C_1\} = 0.6875$

Problem: The probability of a shooter hitting a target is $1/3$. How many times he should shoot so that the probability of hitting the target at least once is more than 3/4.

Solution: Probability of hitting the target is $p = 1/3$, thus $q = 1 - p = 1 - (1/3) = 2/3$

Here we have to find *n* such that

$P(X \ge 1) > 3/4$

i.e. $1 - P(X < 1) > 3/4$

$1 - P(0) > 3/4$

$$1 - {}^nC_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^n > 3/4$$

$$\left(\frac{2}{3}\right)^n < \frac{1}{4}$$

We can find n by inspection as we have $\frac{2}{3} = 0.67, \left(\frac{2}{3}\right)^2 = 0.44, \left(\frac{2}{3}\right)^3 = 0.3, \left(\frac{2}{3}\right)^4 = 0.2$

Therefore, $n = 4$.

Problem: A die is thrown 8 times. Find the probability that 3 falls

(i) Exactly 2 times

(ii) At least once

(iii) at the most 7 times.

Solution: If a die is tossed, the probability of getting 3 $= \frac{1}{6}$

Therefore $q = 1 - p = 1 - \frac{1}{6} = \frac{5}{6}$ and $n = 8$.

The binomial distribution function is $b(n, p, x) = {}^{n}C_{x} p^{x} q^{n-x}$

$$b\left(8, \frac{1}{6}, x\right) = {}^{8}C_{x} \left(\frac{1}{6}\right)^{x} \left(\frac{5}{6}\right)^{8-x}$$

(i)Probability that 3 falls exactly 2 times

$$= b\left(8, \frac{1}{6}, 2\right) = {}^{8}C_{2} \left(\frac{1}{6}\right)^{2} \left(\frac{5}{6}\right)^{8-2} = {}^{8}C_{2} \left(\frac{1}{6}\right)^{2} \left(\frac{5}{6}\right)^{6} = 0.2604$$

(ii)Probability that 3 falls at least once

$$= b\left(8, \frac{1}{6}, x \geq 1\right) = 1 - b\left(8, \frac{1}{6}, x < 1\right)$$

$$= 1 - b\left(8, \frac{1}{6}, 0\right) = 1 - {}^{8}C_{0} \left(\frac{1}{6}\right)^{0} \left(\frac{5}{6}\right)^{8-0} = 1 - \left(\frac{5}{6}\right)^{8} = 0.7674$$

(iii)Probability that 3 falls at the most 7 times

$$= b\left(8, \frac{1}{6}, x \leq 7\right) = 1 - b\left(8, \frac{1}{6}, x > 7\right)$$

$$= 1 - b\left(8, \frac{1}{6}, 8\right) = 1 - {}^{8}C_{8} \left(\frac{1}{6}\right)^{8} \left(\frac{5}{6}\right)^{8-8} = 1 - \left(\frac{1}{6}\right)^{8} = 0.9999$$

Problem: In a quiz contest of answering yes or no what is the probability of guessing at least 6 answers correctly out of 10 questions asked? Also find the probability of the same if there are 4 options for a correct answer.

Solution:

(i)The probability of answer a question in the quiz correctly if there are two options $= p = 1/2$

$q = 1 - p = 1 - 1/2 = 1/2$, $n = 10$

The binomial distribution function is $b(n, p, x) = {}^{n}C_x p^x q^{n-x}$

$b(10, 0.5, x) = {}^{10}C_x (0.5)^x (0.5)^{10-x} = {}^{10}C_x (0.5)^{10}$

The probability of guessing at least 6 answers correctly out of 10 questions asked is

$= b(10, 0.5, x \geq 6) = b(10, 0.5, 6) + b(10, 0.5, 7) + b(10, 0.5, 8) + b(10, 0.5, 9) + b(10, 0.5, 10)$

$= {}^{10}C_6 (0.5)^{10} + {}^{10}C_7 (0.5)^{10} + {}^{10}C_8 (0.5)^{10} + {}^{10}C_9 (0.5)^{10} + {}^{10}C_{10} (0.5)^{10}$

$= (0.5)^{10} \left({}^{10}C_6 + {}^{10}C_7 + {}^{10}C_8 + {}^{10}C_9 + {}^{10}C_{10} \right) = 0.376$

(ii)The probability of answer a question in the quiz correctly if there are four options $= p = 1/4$

$q = 1 - p = 1 - 1/4 = 3/4$, $n = 10$

The binomial distribution function is $b(n, p, x) = {}^{n}C_x p^x q^{n-x}$

$b(10, 0.25, x) = {}^{10}C_x (0.25)^x (0.75)^{10-x}$

The probability of guessing at least 6 answers correctly out of 10 questions asked is

$= b(10, 0.5, x \geq 6) = b(10, 0.5, 6) + b(10, 0.25, 7) + b(10, 0.25, 8) + b(10, 0.25, 9) + b(10, 0.25, 10)$

$= {}^{10}C_6 (0.25)^6 (0.75)^4 + {}^{10}C_7 (0.25)^7 (0.75)^3 + {}^{10}C_8 (0.25)^8 (0.75)^2 + {}^{10}C_9 (0.25)^9 (0.75) + {}^{10}C_{10} (0.25)^{10}$

$= 0.01931$

Problem: The probability that a pen manufactured by accompany will be defective is 0.1. If 12 such pens are selected, find the probability that
(i)exactly two will be defective
(ii) at least two will be defective
(iii) none will be defective.

Solution: The probability that a pen manufactured by accompany will be defective is $p = 0.1$

$q = 1 - p = 1 - 0.1 = 0.9$, $n = 12$.

Thus the probability function for the distribution is $b(n, p, x) = {}^{n}C_x p^x q^{n-x}$

$b(12, 0.1, x) = {}^{12}C_x (0.1)^x (0.9)^{12-x}$

(i)The probability that exactly two will be defective in a sample is

$b(12, 0.1, 2) = {}^{12}C_2 (0.1)^2 (0.9)^{12-2} = {}^{12}C_2 (0.1)^2 (0.9)^{10} = 0.23$

(ii)The probability that there are at least two defective parts in a sample is

$$b(12, 0.1, x \geq 2) = 1 - b(20, 0.1, x < 2)$$

$$= 1 - \{b(12, 0.1, 0) + b(12, 0.1, 1)\}$$

$$= 1 - \{{}^{12}C_0 (0.1)^0 (0.9)^{12-0} + {}^{12}C_1 (0.1)^1 (0.9)^{12-1}\}$$

$$= 1 - \{(0.9)^{12} + 12 \times (0.1)(0.9)^{11}\} = 0.3409$$

(iii)The probability that none will be defective parts in a sample is

$$= b(12, 0.1, 0) = {}^{12}C_0 (0.1)^0 (0.9)^{12} = (0.9)^{12} = 0.2824$$

Problem: Alpha particles are emitted by a radioactive source at an average of 5 emissions in 20 minutes. What is the probability that there will be (i) exactly two emissions (ii) at least two emissions in 20 minutes?

Solution: Let p be the probability that alpha particles are emitted by a radioactive source. It is given that the mean $\mu = 5$ for a sample size of $n = 20$.

Therefore, the formula $\mu = np$ (binomial distribution) gives $p = \frac{\mu}{n} = \frac{5}{20} = 0.25$, hence

$q = 1 - p = 1 - 0.25 = 0.75$

Thus the probability function for the distribution is $b(n, p, x) = {}^{n}C_x p^x q^{n-x}$

$b(20, 0.25, x) = {}^{20}C_x (0.25)^x (0.75)^{20-x}$

(i)The probability that there are exactly two emissions in 20 minutes

$$= b(20, 0.25, 2) = {}^{20}C_2\,(0.25)^2\,(0.75)^{20-2} = {}^{20}C_2\,(0.25)^2\,(0.75)^{18} = 0.0669$$

(ii)The probability that there are at least two emissions in 20 minutes

$$= b(20, 0.25, x \geq 2) = 1 - b(20, 0.25, x < 2)$$

$$= 1 - \{b(20, 0.25, 0) + b(20, 0.25, 1)\}$$

$$= 1 - \{{}^{20}C_0\,(0.25)^0\,(0.75)^{20-0} + {}^{20}C_1\,(0.25)^1\,(0.75)^{20-1}\}$$

$$= 1 - \{(0.75)^{20} + 20(0.25)(0.75)^{19}\} = 0.9756$$

Poisson distribution:

Poisson distribution is the limiting form of the binomial distribution when *n* is very large ($n\to\infty$) and *p* the probability of success is very small ($p\to 0$) so that *np* tends to a finite constant.

If *n* is very large and *p* is very small but $\mu = np$ is finite it can be proved that the binomial probability function $b(n,p,x) = {}^{n}C_{x}p^{x}q^{n-x}$ tends to

$$b(n,p,x) = \frac{e^{-\mu}\mu^{x}}{x!} \qquad (1)$$

The RHS of (1) is called the Poisson probability function. It can be written as

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

The probability distribution determined by this function is called the Poisson distribution and μ is called the parameter of the distribution. The random variable x associated with a Poisson distribution is called a Poisson variate. It is an infinite discrete probability distribution.

We easily verify that $p(\mu,x)$ is indeed a probability function in the sense that (i) $p(\mu,x)\ge 0$ and (ii) $\sum p(\mu,x)=1$.

Prove that Poisson distribution is the limiting form of Binomial distribution:

Poisson distribution is regarded as the limiting form of the binomial distribution when *n* is very large ($n\to\infty$) and *p* the probability of success is very small ($p\to 0$) so that *np* tends to a finite constant say μ.

Binomial probability distribution function is

$$b(n,p,x) = {}^{n}C_{x}p^{x}q^{n-x}$$

$$= \frac{n(n-1)(n-2)............(n-\overline{x-1})}{x!}p^{x}q^{n-x}$$

$$= \frac{nn\left(1-\frac{1}{n}\right)n\left(1-\frac{2}{n}\right)..........n\left(1-\frac{x-1}{n}\right)}{x!}p^{x}q^{n-x}$$

$$= \frac{n^{x}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)..........\left(1-\frac{x-1}{n}\right)}{x!}p^{x}q^{n-x}$$

$$= \frac{(np)^x \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\cdots\cdots\left(1-\frac{x-1}{n}\right)}{x!q^x} q^n$$

But $np = \mu$ and $q^n = (1-p)^n = \left(1-\frac{\mu}{n}\right)^n = \left[\left(1-\frac{\mu}{n}\right)^{-n/\mu}\right]^{-\mu}$

Denoting $-\frac{\mu}{n} = k$, we have $q^n = \left[(1+k)^{1/k}\right]^{-\mu} \to e^{-\mu}$ as $n \to \infty$ or $k \to 0$.

Since $\lim_{k\to 0}(1+k)^{1/k} = e$

Further $q^x = (1-p)^x \to 1$ for a fixed x as $p \to 0$.

Also the factors $\left(1-\frac{1}{n}\right)$, $\left(1-\frac{2}{n}\right)$ …….. $\left(1-\frac{x-1}{n}\right) \to 1$ as $n \to \infty$).

Thus we get $b(n, p, x) = \frac{e^{-\mu}\mu^x}{x!} = p(\mu, x)$

Mean and Standard deviation of the Poisson distribution:

Mean = $= \mu = \sum_{x=0}^{\infty} xp(\mu, x)$

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

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

$$= \mu e^{-\mu} \left\{1 + \frac{\mu}{1!} + \frac{\mu^2}{2!} + \frac{\mu^3}{3!} + \cdots\cdots\cdots\right\}$$

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

Therefore mean = μ

Variance $= V = \sum_{x=0}^{\infty} x^2 p(\mu, x) - \mu^2$

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

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

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

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

$$= \mu^2 e^{-\mu} \left\{1 + \frac{\mu}{1!} + \frac{\mu^2}{2!} + \frac{\mu^3}{3!} + ...\right\} + \mu - \mu^2$$

$$= \mu^2 e^{-\mu} e^{\mu} + \mu - \mu^2 = \mu^2 + \mu - \mu^2 = \mu$$

Variance $V = \mu$

Therefore Standard deviation $= \sigma = \sqrt{V} = \sqrt{\mu}$.

Problem: The number of accidents per day (*x*) as recorded in a textile industry over a period of 400 days is given. Fit a Poisson distribution for the data and calculate the theoretical frequencies

x	0	1	2	3	4	5
f	173	168	37	18	3	1

Solution: For the Poisson distribution we have Mean $= \mu = \dfrac{\sum fx}{\sum f}$

$$= \frac{(173)(0) + (168)(1) + (37)(2) + (18)(3) + (3)(4) + (1)(5)}{400} = \frac{313}{400} = 0.7825$$

Thus Poisson distribution function is $p(\mu, x) = \dfrac{e^{-\mu}\mu^x}{x!}$

$$p(0.7825, x) = \frac{e^{-0.7825}(0.7825)^x}{x!}$$

Let $f(x) = 400\,p(\mu, x) = 400\dfrac{e^{-0.7825}(0.7825)^x}{x!} = \dfrac{182.9 \times (0.7825)^x}{x!}$

Theoretical frequencies are got by substituting x = 0, 1, 2, 3, 4, 5 in $f(x)$ and they are as follows 183, 143, 56, 15, 3, 0.

Problem: A certain screw making machine produces on an average two defectives out of 100 and packs them in boxes of 500. Find the probability that the box contains

(i)Three defectives (ii)At least one defective (iii) Between two and four defectives (iv) 15 defective screws.

Solution: : p = Probability of a defective screw produced $= \dfrac{2}{100} = 0.02$

In a packet of 500, the mean number of defective screws is $\mu = np = 500 \times 0.02 = 10$

The Poisson probability function is $p(\mu, x) = \dfrac{e^{-\mu}\mu^{x}}{x!}$

$$p(10,x) = \frac{e^{-10}(10)^{x}}{x!} = p(x) \qquad \text{(say)}$$

(i) Probability that the box contains three defective screws

$$= p(10,3) = \frac{e^{-10}(10)^{3}}{3!} = e^{-10} \times \frac{1000}{6} = e^{-10} \times \frac{500}{3} = 0.007566$$

(ii) Probability that the box contains at least one defective screw

$$= p(10, x \geq 1) = 1 - p(10, x < 1) = 1 - \{p(10,0) + p(10,1)\}$$

$$= 1 - \left\{ \frac{e^{-10}(10)^{0}}{0!} + \frac{e^{-10}(10)^{1}}{1!} \right\}$$

$$= 1 - e^{-10}(1+10) = 1 - 11e^{-10} = 0.9995$$

(iii) Probability that the box contains between two and four defective screws

$$= p(10, 2 \leq x \leq 4) = p(10,2) + p(10,3) + p(10,4)$$

$$= \frac{e^{-10}(10)^{2}}{2!} + \frac{e^{-10}(10)^{3}}{3!} + \frac{e^{-10}(10)^{4}}{4!}$$

$$= e^{-10}\left(50 + \frac{500}{3} + 2500\right) = 0.1233$$

(iv) Probability that the box contains 15 defective screws

$$= p(10,15) = \frac{e^{-10}(10)^{15}}{15!} = 0.03471$$

Problem: Given that 2% of the fuses manufactured by a firm are found to be defective. Find the probability that a box containing 200 fuses contains (i) no defective fuse (ii) 3 or more defective fuses (iii) at least one defective fuse.

Solution: : p = Probability of a defective fuses manufactured = 2% = $\frac{2}{100} = 0.02$

In a packet of 200, the mean number of defective fuses is $\mu = np = 200 \times 0.02 = 4$

The Poisson probability function is $p(\mu, x) = \frac{e^{-\mu} \mu^x}{x!}$

$$p(4, x) = \frac{e^{-4}(4)^x}{x!} = p(x) \qquad \text{(say)}$$

(i) Probability that the box contains no defective fuse

$$= p(4,0) = \frac{e^{-4}(4)^0}{0!} = e^{-4} = 0.01831$$

(ii) Probability that the box contains 3 or more defective fuses

$$= p(4, x \geq 3) = 1 - p(4, x < 3)$$

$$= 1 - \{p(4,0) + p(4,1) + p(4,2)\}$$

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

$$= 1 - e^{-4}\{1 + 4 + 8\} = 1 - 13e^{-4} = 0.7618$$

(iii) Probability that the box contains at least one defective fuse

$$= p(4, x \geq 1) = 1 - p(4, x < 1) = 1 - \{p(4,0)\}$$

$$= 1 - \left\{\frac{e^{-4}(4)^0}{0!}\right\}$$

$$= 1 - e^{-4} = 0.9816$$

Problem: A car hire firm has 2 cars, which it hires out day by day. The demand for a car on each day is distributed as a Poisson distribution with mean 1.5. Calculate the probability that on a certain day (i) neither car is used and (ii) some demand is refused.

Solution: Here $\mu = 1.5$ and the Poisson probability function is $p(\mu, x) = \dfrac{e^{-\mu}\mu^x}{x!}$

$$p(1.5, x) = \frac{e^{-1.5}(1.5)^x}{x!} = p(x) \qquad \text{(say)}$$

The probability that neither car is used is $p(0) = \dfrac{e^{-1.5}(1.5)^0}{0!} = e^{-1.5} = 0.2231$

The probability that some demand is refused

= Probability that there are more than 2 demands

$= p(x > 2)$

$= 1 - p(x \le 2)$

$= 1 - \{p(0) + p(1) + p(2)\}$

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

$= 1 - e^{-1.5}\{1 + 1.5 + 1.125\} = 0.1912$

Problem: In a certain factory turning out blades, there is a small chance of 0.002 for any blade to be defective. The blades are supplied in packets of 10. Using Poisson distribution find the approximate number of packets containing (i) no defective blade (ii) one defective blade in a consignment of 10,000 packets.

Solution: p = Probability of a defective blade = 0.002

In a packet of 10, the mean number of defective blades is $\mu = np = 10 \times 0.002 = 0.02$

Therefore the Poisson distribution function is $p(\mu, x) = \dfrac{e^{-\mu}\mu^x}{x!}$

$$p(0.02, x) = \frac{e^{-0.02}(0.02)^x}{x!} = p(x) \qquad \text{(say)}$$

(i) Probability of no defective blade $= p(0) = \dfrac{e^{-0.02}(0.02)^0}{0!} = e^{-0.02} = 0.9802$

Therefore approximate number of packets containing no defective blade in a consignment of 10,000 packets = 10,000 × 0.9802 = 9802

(ii) Probability of one defective blade

$$= p(1) = \frac{e^{-0.02}\,(0.02)^1}{1!} = e^{-0.02} \times 0.02 = 0.9802 \times 0.02 = 0.019604$$

Therefore approximate number of packets containing one defective blade in a consignment of 10,000 packets = 10,000 × 0.019604 = 196.04 ≈ 196.

Problem: The number of accidents in a year to taxi drivers in a city follows a Poisson distribution with mean 3 out of 1000 taxi drivers find approximately the number of drivers with (i) no accident in a year (ii) more than 3 accidents in a year.

Solution: Here $\mu = 3$ and the Poisson probability function is $p(\mu, x) = \dfrac{e^{-\mu}\mu^x}{x!}$

$$p(3, x) = \frac{e^{-3}\,(3)^x}{x!} = p(x) \qquad \text{(say)}$$

(i) Probability of drivers with no accident in a year $= p(3,0) = \dfrac{e^{-3}\,(3)^0}{0!} = e^{-3} = 0.04978$

Therefore, approximate number of drivers with no accident in a year

$$= 0.04978 \times 1000 = 49.78 \approx 49 \text{ drivers}$$

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

$$= p(3, x \geq 3) = 1 - p(3, x < 3)$$

$$= 1 - \{p(3,0) + p(3,1) + p(3,2)\}$$

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

$$= 1 - e^{-3}\left\{1 + 3 + \frac{9}{2}\right\} = 0.5768$$

Therefore, approximate number of drivers with more than 3 accidents in a year

$$= 0.5768 \times 1000 = 576.8 \approx 576 \text{ drivers}$$

Problem: If the number of accidents occurring on a highway each day is a Poisson distribution with mean equal to 3. What is the probability that no accident occur today?

Solution: Here $\mu = 3$ and the Poisson probability function is $p(\mu, x) = \dfrac{e^{-\mu}\mu^x}{x!}$

$$p(3, x) = \frac{e^{-3}(3)^x}{x!} = p(x) \qquad \text{(say)}$$

Probability of drivers with no accident occur today $= p(3,0) = \dfrac{e^{-3}(3)^0}{0!} = e^{-3} = 0.04978$

Problem: The probability that an individual suffers a bad reaction from an injection is 0.001. Find the probability that out of 2000 individuals, more than 2 will get bad reaction.

Solution: p = Probability of an individual suffers a bad reaction from an injection = 0.001

In a group of 2000, the mean number of individual suffers a bad reaction from an injection $\mu = np = 2000 \times 0.001 = 2$

Therefore the Poisson distribution function is $p(\mu, x) = \dfrac{e^{-\mu}\mu^x}{x!}$

$$p(2, x) = \frac{e^{-2}(2)^x}{x!} = p(x) \qquad \text{(say)}$$

Probability of more than 2 will get bad reaction

$$= p(2, x \geq 2) = 1 - p(2, x < 2)$$

$$= 1 - \{p(2,0) + p(2,1)\}$$

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

$$= 1 - e^{-2}\{1 + 2\} = 1 - 3e^{-2} = 0.5939$$