1. Introduction

Welcome to the series of e-learning modules on Binomial Distribution. In this module we are going to cover the definition of binomial distribution, moments, mgf, cgf, skewness, kurtosis, recurrence relation for probabilities, mode, fitting binomial distribution to the given data and its applications.

By the end of this session, you will be able to explain:

- Binomial distribution
- Moments
- Nature of the distribution
- Moment Generating Function
- Cumulants Generating function
- Additive property of the distribution
- Recurrence relation for probabilities
- Mode
- Fitting binomial distribution

Binomial distribution was discovered by James Bernoulli (1654 – 1705) in the year 1700 and was first published in 1713.
This is an extension of Bernoulli distribution. If we conduct a series of n Bernoulli trials with a constant probability p of success then the total number of success in these n trials follow Binomial distribution.

Definition
A random variable X is said to follow Binomial distribution with parameters n and p if its probability mass function is given by:
P of x is equal to n c x, p power x into q power n-x, where x takes values 0, 1, 2, etc., n , p lies between 0 and 1, and p+q=1
And we write x follows binomial n p, read as x follows binomial distribution with parameters n and p.

Note that this assignment of probabilities is permissible because,
Summation over x from 0 to n, p of x is equal to summation over x, n c x, p power x into q power n-x
Is equal to q plus p power n
Which is equal to 1 as p + q=1.

Physical conditions for Binomial Distribution
We get binomial distribution under the following experimental conditions:

1. Each trail results in two mutually disjoint outcomes, termed as success and failure.
2. The number of trails ‘n’ is finite.
3. The trails are independent of each other.
4. The probability of success ‘p’ is constant for each trail.

Examples for binomial distribution:

1. Number of heads obtained in n tosses of a coin.
2. Number of defectives in a sample of size n, drawn from a large lot.
3. Number of individuals with a specified characteristic (such as blindness, deafness etc.) in a sample of size n drawn from a large population.
4. Number of boys (or girls) in a family having n children.
5. Number of seeds germinating among n seeds which were sown.
6. Number of bombs hitting a bridge among n bombs dropped on it.

2. Moments of Distribution

Let us discuss about the moments of the distribution.

First raw moment, which is nothing but mean of the distribution is given by,
Mew 1 dash is equal to Expectation of x
Is equal to summation over x, x into p of x.
Is equal to summation over x from 0 to n, x into n c x, p power x into q power n-x.
Substituting for n c x, we get,
Summation over x, x into n factorial divided by x factorial into n-x factorial, into p power x into q power n-x
We can expand n factorial as n into n-1 factorial, and can take n outside. Similarly x factorial in denominator can be written as x into x-1 factorial and cancel x in the numerator and in p power x add and subtract 1, and take p power 1 outside
Hence we get,
n into p summation over x from 1 to n, n-1 factorial divided by x-1 factorial into n-x factorial, into p power x-1 into q power n-x
is equal to n p into q+p power n-1
is equal to np.
Therefore mean of the distribution is np.

Second raw moment is given by, mew 2 dash equals to expectation of x square
Is equal to summation, x into x-1 plus x, into p of x.

Which is equal to summation over x from 0 to n, x into x-1 plus x, into n c x, p power x into q power n-x.

If we split the summation to two terms, x into x-1 and x separately, we get summation over x, x into x-1 into n factorial divided by x factorial into n-x factorial into p power x into q power n-x, plus np (as the second term is np which is nothing but mew 1 dash).

Doing similar simplification as we have done for mew 1 dash, we get
n into n-1, p square into summation over x from 2 to n, n minus 2 factorial divided by x-2 factorial into n-x factorial, into p power x-2 into q power n-x, plus n p
same as n into n-1 into p square into q+p power n-2 plus np
is equal to n into n-1 into p square plus np.

By observing expectation of x into x-1, we can write,
Expectation of x into x-1 into x-2 is equal to n into n-1 into n-2 into p cube and
Expectation of x into x-1 into x-2 into x-3 is equal to n into n-1 into n-2 into n-3 into p power 4

Also we can write, x cube is equal to x into x-1 into x-2, plus 3 x into x-1, plus x
and
X power 4 is equal to x into x-1 into x-2 into x-3, plus 6 x into x-1 into x-2, plus 7 x into x-1, plus x.

Third raw moment is given by,

Mew 3 dash is equal to expectation of x cube
Is equal to n into n-1 into n-2 into p cube, plus 3 n into n-1 into p square, plus n p

and 4th raw moment is given by,
Mew 4 dash is equal to expectation of x power 4
Is equal to n into n-1 into n-2 into n-3 into p power 4, plus 6 n into n-1 into n-2 into p cube, plus 7 n into n-1 into p square plus n p.

From raw moments let us obtain the central moments.
Mew 2 is equal to mew two dash minus mew one dash square
Is equal to n into n-1 into p square plus n p minus n p whole square
Which is equal to npq

Observe that for binomial distribution mean is greater than that of variance.

Mew 3 is equal to mew 3 dash, minus 3 mew 2 dash into mew 1 dash, plus 2 mew 1 dash cube.
Is equal to n into n-1 into n-2 into p cube plus 3 n into n-1 into p square, plus n p
minus 3 into n into n-1 into p square plus np into np, plus 2 into np whole cube.
on simplification, we get
n p q into q-p

Similarly, we can obtain
mew 4 which is equal to mew 4 dash, minus 4 mew 3 dash into mew 1 dash, plus 6 mew 2 dash into mew 1 dash square, minus 3 mew 1 dash power four.

Equal to n into n-1 into n-2 into n-3 into p power 4 plus, 6 n into n-1 into n-2 into p cube, plus 7 n into n-1 into p square plus np

minus 4 into n into n-1 into n-2 into p cube plus 3 n into n-1 into p square plus n p into n p

plus 6 into n into n-1 into p square plus n p into np whole square
minus 3 into np whole power 4.

By simplifying the above expression we get,
Mew 4 is equal to n p q into 1 plus 3 into n-2 into pq.

3. Nature of Distribution

Now let us find the nature of the distribution
Skewness of the distribution is given by
Beta 1 is equal to mew 3 square by mew 2 cube
Which is equal to q-p whole square divided by n p q

Observe that the skewness depends on the value of the parameters p.
Hence Binomial distribution is negatively skewed if p is greater than 0.5, Positively skewed if p is less than 0.5 and if p=0.5, the distribution is symmetric.

To find kurtosis we find beta 2, given by mew 4 divided by mew 2 square which is equal to 3 plus, 1 minus 6 pq divided by n p q
Hence the kurtosis goes to infinity for high and low values of p.

Moment generating function of the binomial distribution is given by,
M x of t is equal to expectation of e power tx
Is equal to summation over x from 0 to n, e power tx into n c x, p power x into q power n-x
Substituting for n c x we get
Summation over x, e power tx into n factorial by, x factorial into n-x factorial, into p power x into q power n-x
By combining the terms having power x, we get
Summation over x, n factorial by x factorial into n-x factorial, into e power t into p whole power x into q power n-x
Is equal to q+p e power t whole power n.

4. Sum of Independent Binomial Random Variables & Mode of Distribution

Let us find the distribution of sum of independent binomial random variables, known as additive property or reproductive property of binomial distribution.

Let X follows binomial distribution with parameters n one and p one, and Y follows binomial distribution with parameters n two and p two. Further X and Y are independent random variables.

Let us find the distribution of X plus Y.

First we consider the moment generating functions of x and y given by,
M x of t is equal to q one plus p one into e power t whole power n one and
M y of t is equal to q two plus p two into e power t whole power n two.

Now we consider the mgf of sum of the random variables x + y.

Since x and y are independent , we can write,
M x plus y of t is equal to M x of t into M y of t
Is equal to q one plus p one into e power t whole power n one, into q two plus p two into e power t whole power n two.

Observe that M x plus y of t cannot be written in the form q+p into e power t whole power n. Hence by uniqueness theorem of mgf it follows that X + Y is not a binomial variate. Hence, in general the sum of two independent binomial variates is not a binomial variate. In other words, binomial distribution does not possess the additive or reproductive property.

However, if we take p one is equal to p two is equal to p (say), then
M x plus y of t is equal to q+p into e power t whole power n one plus n two, which is the moment generating function of binomial distribution with parameters n one plus n two and p. Hence by uniqueness theorem of mgf, x plus y follows binomial distribution with parameters n one plus n two and p.

Using above result we can generalise the addition theorem for n independent random variables.
Suppose x i for i is equal to 1, 2, etc., k. K are independent binomial variates with parameters n i and p where i is equal to 1, 2, etc., k, then their sum, summation x i is also a binomial variable with parameters summation n i and p.

Having mgf, let us find Cumulants generating function.
It is found by taking logarithm of mgf and is denoted by k x of t.
t by 1 factorial, plus t square by 2 factorial, plus etc.

We know that mgf of the binomial distribution is given by,
M x of t is equal to q plus p e power t whole power n
Hence Cumulant generating function is given by
K x of t is equal to log m x of t
Is equal to log q plus p e power t whole power n
Is equal of n into log q plus p e power t
Now using exponential expansion for e power t we get,
N log q plus p into 1 plus t by 1 factorial, plus t square by 2 factorial, plus etc.

which is same as n log, 1 plus p into t by 1 factorial, plus t square by 2 factorial, plus etc.

By substituting the expansion for log 1 plus x, we get,
K x of t is equal to n into
P into t plus t square by 2 factorial Plus t cube by 3 factorial Plus etc.
Minus p square by two, into t plus t square by 2 factorial Plus t cube by 3 factorial Plus etc.
Plus p cube by 3 into, t plus t square by 2 factorial Plus t cube by 3 factorial Plus etc. whole power square.
Minus p power 4 by 4 into, t plus t square by 2 factorial Plus t cube by 3 factorial Plus etc. whole power 4, plus etc.

r^{th} Cumulants Kr is given by equating the coefficient of t power r by r factorial from the above equation.

Now let us find the mode of the distribution.
Mode is that value of x for which p of x is maximum. To find this maximum value, first we find the ratio of probabilities when x takes value x and x-1.

That is p of x divided by p of x-1 is equal to n c x, p power x into q power n-x divided by n c x-1, p power x-1 into q power n-x+1

By substituting for n c x as n factorial divided by x factorial into n-x factorial, and n c x-1 as n factorial divided by x-1 factorial divided by n-x+1 factorial ,
And simplifying the expression we get,
n-x+1 into p divided by x q.
Now we add and subtract xq in the numerator and split the denominator for each term, we can write the above expression as:
1 plus, n+1 into p minus x whole divided by xq.

5. Examples & Summary

Now to find maximum, we discuss the following 2 cases:

Case(i): When n+1 into p is not an integer

Let n+1 into p can be written as m + f, where m is an integer and f is fractional such that f lies between 0 and 1. Substituting for n+1 into p in the above term, we get,

P of x divided by p of x-1 is equal to 1 plus, m plus f minus x whole divided by xq

By substituting x values in the above equation we get,
P of x by p of x-1 is greater than 1 for x equal to 1, 2, etc., m and
P of x by p of x-1 is less than 1 for x equal to m plus 1, m+2 etc., n.

Which implies, p of 1 by p of 0 greater than 1, p of 2 by p of 1 greater than 1, etc., p of m by p of m-1 is greater than 1 and

P of m+1 by p of m less than 1, p of m+2 by p of m+1 is less than 1, etc., p of n by p of n-1 is less than 1.

From above 2 expressions we can write, p of zero less than p of 1 less than p of 2 less than etc., less than p of m-1 less than p of m greater than p of m+1 greater than p of m+2 greater than etc., greater than p of n.

Thus, in this case there exists unique modal value for binomial distribution and it is m, the integral part of n+1 into p.

Now let us consider Case(ii): When n+1 into p is an integer

Let n+1 into p can be written as m, where m is an integer.
Substituting for n+1 into p in the above term, we get

P of x divided by p of x-1 is equal to 1 plus, m minus x whole divided by xq

By substituting x values in the above equation we get,
P of x by p of x-1 is greater than 1 for x equal to 1, 2, etc., m-1
P of x by p of x-1 is equal 1, for x is equal to m
and P of x by p of x-1 is less than 1 for x equal to m plus 1, m+2 etc., n.

Which implies, p of 1 by p of 0 greater than 1, p of 2 by p of 1 greater than 1, etc., p of m-1 by p of m-2 is greater than 1
P of m by p of m-1 is equal to 1 and
P of m+1 by p of m less than 1, p of m+2 by p of m+1 is less than 1, etc., p of n by p of n-1 is less than 1.

From above 3 expressions we can write, p of zero less than p of 1 less than p of 2 less than etc., less than p of m-1 is equal to p of m greater than p of m+1 greater than p of m+2 greater than etc., greater than p of n.

Thus in this case the binomial distribution has two modal values namely m and m-1, that is n+1 into p and n+1 into p minus1.

Let us fit the binomial distribution to the given data

First we consider the ratio of probabilities when x takes value x and x minus 1, which is already found in mode of the distribution.

P of x divided by p of x minus 1 is equal to n minus x plus 1 into p divided by x q

Which implies,
P of x is equal to n minus x plus 1 divided by x, into p by q, into p of x minus 1.

Given the sample frequency distribution we know the value of n. To find p we use the formula p is equal to x bar by n.

The theoretical or expected frequency corresponding to the value x is given by
N into p of x.

Expected frequency for x is equal to zero, e not is found by using the formula
N into q power n where capital n is the total frequency.

And successive frequencies are found using the relation
E x is equal to n minus x plus one divided by x, into p by q, into e x minus 1
For x is equal to 1, 2, etc., n.

Here’s a summary of our learning in this session:

- Definition of binomial distribution
- Moments
- Nature of the distribution
- Moment and Cumulants Generating functions
- Additive property of the distribution
- Recurrence relation for probabilities
- Mode
- Fitting the distribution