Discrete Distributions · Code: 18K 2S04

Unit: 1

Moment generating function – Definition, properties, Characteristics function – Definition and properties. Inversion Theorem (statement only). Cumulants – Definition and properties. Moments – Raw moments, Central moments and their relationships.

Unit – 2

Bernoulli distribution – Definition. Binomial distribution – Definition, Derivation of Binomial probability distribution. Derivation of moments, β_1 β_2 co-efficients, Cumulants. Recurrence relation for moments, mode, ~~Additiv~~ Additive property moment generating function. Characteristics function and simple problems.

Unit – 3

Poisson distribution – Limiting form of binomial distribution – Definition, properties, Derivation of moments, β_1, β_2 Recurrence relation for moments, Cumulants, mode, Additive property, M.G.F, Characteristics function and Simple problems.

Unit – 4

Discrete Uniform distribution – Definition, derivation of mean and Variance. Negative Binomial distribution – Definition, properties, derivation of mean and Variance, moment generating function. Cumulants. Poisson distribution as a Limiting Case of Negative ~~binomial di~~ binomial distribution

Unit – 5

Geometric distribution – Definition, properties moments, moment Generating function, Hyper – geometric distribution – Definition, mean and Variance.

UNIT - I

Moment Generation Function

The moment generation function (MGF) of a random variable x (about origin) having the Probability function $f(x)$ is given by

$$M_x(t) = E(e^{tx}) = \begin{cases} \int e^{tx} f(x)\, dx, & \text{(for continuous prob distribution)} \\ \sum_x e^{tx} f(x), & \text{(for discrete prob distribution)} \end{cases}$$

the integration or summation being extend to the entire range of x, t being the real parameter and it is being assumed that the right hand side is absolutely convergent for some positive number is such that

$-h < t < h$

since, $M_x(t) = E(e^{tx})$

$= E\left(1 + tx + \frac{t^2 x^2}{2!} + \dots + \frac{t^r x^r}{r!} + \dots\right)$

$= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \dots + \frac{t^r}{r!}\mu_r'$

Properties of M.G.F

Property - 1

$M_{cx}(t) = M_x(ct)$. c being a constant

Proof By definition

LHS $= M_{cx}(t) = E(e^{tcx})$

RHS $= M_x(ct) = E(e^{ctx}) =$ LHS

Property - 2

The MGF of the sum of a number of independent random variables is equal to the product of their respective M.G.F s.

Symbolically if $x_1, x_2, \ldots, x_n$ are Independent random variables then the moment generation function of their sum $x_1 + x_2 + x_3 \ldots + x_n$ is given by

$M_{x_1 + x_2 + \ldots x_n}(t) = M_{x_1}(t)\, M_{x_2}(t) \ldots$

$M_{x_1 + x_2 + x_3 + \ldots x_n}(t) = M_{x_1}(t) . M_{x_2}(t) . . M_{x_n}(t)$

Proof

by definition.

$$M_{x_1 + x_2 + x_3 + \ldots x_n}(t) = E\left[e^{t(x_1 + x_2 + \ldots x_n)}\right]$$

$$= E\left[e^{tx_1 + tx_2 \ldots + tx_n}\right]$$

$$= E(e^{tx_1}) . E(e^{tx_2}) \ldots E(e^{tx_n})$$

$\therefore x_1, x_n \ldots x_n$ are independent.

$$= M_{x_1}(t)\, M_{x_2}(t) \ldots M_{x_n}(t)$$

Property - 3

Effect of change of origin and scal of M.G.F. Let us transform x to the new variable u by changing both the origin and Scal in x as follows

$$U = \frac{x-a}{h}$$, where, a and h are constants.

M.G.F of U (about origin) is given by:

$$M_U(t) = E(e^{tU}) = E[\exp\{t(x-a)/h\}]$$

$$= E[e^{tx/h} e^{-at/h}]$$

$$= e^{-at/h} E . e^{(tx/h)}$$

$$= e^{-at/h} E(e^{xt/h})$$

$$= e^{-at/h} M_x(t/h)$$

Where, $M_x(t)$ is the M.G.F of x about origin.

In Particular it we take

$a = E(X) = \mu$, say and $h = \sigma_x = \sigma$

say then $U = \frac{X-E(X)}{\sigma_x} = \frac{X-\mu}{\sigma} = Z$

is known as a standard variate thus the M.G.F of a standard variate z is given by

$$M_Z(t) = e^{-\mu t/\sigma} M_x(t/\sigma)$$

* Uniqueness theorem Of M.G.F

The moment generationg function of a distribution, if it exists, uniquely determines the distribution.

This implis that corresponding to a given probability distribution, there is only a one M.G.F (Provided its exists) and corresponding to a given M.G.F there is only one a Probability distribution.

Hence

$M_x(t) = M_y(t) \Rightarrow$ x and y are identically distributions.

CUMULANTS

Cumulants generating function $K(t)$ is defined as : $K_x(t) = \log_e M_x(t)$.

Provided the right-hand side can be expended as a convergent series in powers of t. Thus.

$$K_x(t) = K_1 t + K_2 \frac{t^2}{2!} + \dots + K_r \frac{t^r}{r!} + \dots = \log M_x(t)$$

$$= \log \left(1 + \mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \dots + \mu_r' \frac{t^r}{r!} + \dots\right)$$

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}$$

Where K_r = co-efficient of $\frac{t^r}{r!}$ in $K_x(t)$ is called the r^{th} cumulant.

$$K_1 t + K_2 \frac{t^2}{2!} + K_3 \frac{t^3}{3!} + K_4 \frac{t^4}{4!} + \dots$$

$$= \left(\mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \cdots\right) - \frac{1}{2}\left(\mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \cdots\right)^2 + \frac{1}{3}\left(\mu_1' t + \mu_2' \frac{t^2}{2!} + \cdots\right)^3 - \frac{1}{4}\left(\mu_1' t + \mu_2' \frac{t^2}{2!} + \cdots\right)^4 + \cdots$$

Comparing the co-efficients of like powers of 't' on both sides, we get the relationship between the moments and cumulants. Hence, we have

$$K_1 = \mu_1' = \text{Mean}, \quad \frac{K_2}{2!} = \frac{\mu_2'}{2!} - \frac{\mu_1'^2}{2!} \Rightarrow K_2 = \mu_2' - \mu_1'^2 = \mu_2$$

$(a+b+c)^2 = a^2+b^2+c^2+2ab+2bc+2ca$

$$\frac{K_3}{3!} = \frac{\mu_3'}{3!} - \frac{1}{2}\cdot\frac{2\mu_1'\mu_2'}{2!} + \frac{\mu_1'^3}{3} \Rightarrow K_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 = \mu_3$$

Also,

$(a+b)^3 = a^3+b^3-3a^2b+3ab^2$

$$\frac{K_4}{4!} = \frac{\mu_4'}{4!} - \frac{1}{2}\left(\frac{\mu_2'^2}{\cancel{2}\,4} + \frac{2\mu_1'\mu_3'}{3!}\right) + \frac{1}{3}\cdot\frac{3\mu_1'^2\mu_2'}{2} - \frac{\mu_1'^4}{4}$$

$$\Rightarrow K_4 = \mu_4' - 3\mu_2'^2 - 4\mu_1'\mu_3' + 12\mu_1'^2\mu_2' - 6\mu_1'^4$$

$(a-b)^2$

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

$$= \mu_4 - 3(\mu_2' - \mu_1'^2)^2 = \mu_4 - 3\mu_2^2 = \mu_4 - 3K_2^2$$

$$\mu_4 = K_4 + 3K_2^2$$

Hence we have obtained

$$\text{Mean} = K_1$$

$$\mu_2 = K_2 = \text{Variance}$$

$$\mu_3 = K_3$$

$$\mu_4 = K_4 + 3K_2^2$$

Properties of Cumulants.

Property 1

Additive property of cumulants

The rth cumulant of the sum of the independent random variables is equal to the sum of the rth cumulants of the individual variables. Symbolically,

$$K_r(X_1 + X_2 + \dots + X_n) = K_r(X_1) + K_r(X_2) + \dots + K_r(X_n),$$

Where X_i; $i = 1, 2, \dots n$ are independent random variables.

Proof :

Since X_i's are Independent

$$M_{X_1+X_2+\dots+X_n}(t) = M_{X_1}(t) M_{X_2}(t) \dots M_{X_n}(t)$$

Taking logarithm of each side,

$$K_{X_1+X_2+\dots+X_n}(t) = K_{X_1}(t) + K_{X_2}(t) + \dots + K_{X_n}(t)$$

Differentiating both side with respect

to t 'r' times and putting $t=0$. we get,

$$\left[\frac{d^r}{dt^r} K_{x_1+x_2+\cdots x_n}(t)\right]_{t=0} = \left[\frac{d^r}{dt^r} K_{x_1}(t)\right]_{t=0} + \cdots$$

$$+ \cdots + \left[\frac{d^r}{dt^r} K_{x_n}(t)\right]_{t=0}.$$

$$K_r(x_1+x_2+\cdots+x_n) = K_r(x_1)+K_r(x_2)+\cdots+K_r(x_n)$$

which establishes the result.

Property - 2

Effect of change of origin and scale on cumulants.

If we take

$$U = \frac{x-a}{h}, \text{ then } M_U(t) = \exp(-at/h)\, M_x(t/h)$$

$$K_U(t) = \log M_U(t) = -\frac{at}{h} + K_x(t/h)$$

$$K_1' t + K_2' \frac{t^2}{2!} + \cdots + K_r' \frac{t^r}{r!} + \cdots = -\frac{at}{h} + K_1(t/h) + K_2 \frac{(t/h)^2}{2!} + \cdots + K_r \frac{(t/h)^r}{r!} + \cdots$$

Where K_r' and K_r are the rth cumulants of U and x respectively

comparing co-efficients, we get $K_1' = \frac{K_1 - a}{h}$

and $K_r' = \frac{K_r}{h^r}$; $r = 2, 3.$

Thus we see the except the first cumulants all cumulants are independent of change of origin. But the cumulants.

CHARACTERISTIC FUNCTION - DEFINITION.

The characteristic function is defined

as

$$\phi_x(t) = E(e^{itx}) = \begin{cases} \int e^{itx} f(x)\,dx & \text{(for continuous Prob} \\ \sum_x e^{itx} p(x) & \text{(for discrete Prob dist)} \end{cases}$$

Remark:

If $F_x(x)$ is the distribution function of a continuous random variable X, then

$$\phi_x(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x)$$

Obviously $\phi(t)$ is a complex valued function of real variable t.

Properties of characteristic function.

Property - 1 Back Page

For all real t, we have

i) $\phi(0) = \int_{-\infty}^{\infty} dF(x) = 1$

ii) $|\phi(t)| \leq 1 = \phi(0)$

Property - 2

$\phi(t)$ is continuous every where, i.e., $\phi(t)$ is continuous function of t in $(-\infty, \infty)$ rather $\phi(t)$ is uniformly continuous in t.

Property - 3.

$\phi_x(-t)$ and $\phi_x(t)$ are conjucate function

i.e., $\phi_x(-t) = \overline{\phi_x(t)}$ where $\bar{a}$ is the conjucate of a.

Property - 4.

If the distribution function of a random variable x is symmetrical about zero i.e., if

$$1 - f(x) = f(-x) \Rightarrow f(-x) = f(x)$$

Then $\phi_x(t)$ is real valued and even function of t.

Property 5. Back page

If x is some random variable with characteristic function $\phi_x(t)$ and if $\mu_r' = E(x^r)$ exists then.

Property - 6.

$\phi_{cx}(t) = \phi_x(ct)$, c being a constant.

Property - 7

If x_1 and x_2 are independent random variable, then,

$$\phi_{x_1 + x_2}(t) = \phi_{x_1}(t)\, \phi_{x_2}(t)$$

More generally, for independent random variablebles $X_i = 1, 2, \ldots n$. we have

$$\phi_{X_1 + X_2 + \ldots + X_n}(t) = \phi_{X_1}(t)\, \phi_{X_2}(t) \ldots \phi_{X_n}(t)$$

Property 8:

Effect of change of origin and scale on characteristic function.

If $U = \dfrac{X - a}{h}$, a and h being constants, then

$$\phi_U(t) = e^{-iat/h}\, \phi_X(t/h)$$

Property - 9

If $|\phi_X(s)| = 1$ for some $s \neq 0$. then for some real a. $X - a$ is a lattice variable with must, $h = 2\pi/|s|$

Inversion Theorem (Levy theorem)

Let $F(x)$ and $\phi(t)$ denote respectively the distribution function and the characteristic function of a random variable X.

If $(a-h, a+h)$ is the continuity interval of the distribution function $F(x)$ then,

$$F(a+h) - F(a-h) = \lim_{T \to \infty} \frac{1}{\pi} \int_{-T}^{T} \frac{\sin hT}{t} \cdot e^{ita} \phi t \, dt$$

Moments

Raw moments

The rth moment of the variables x about any point $x = A$ usually denoted by μ_r' is given by,

$$\mu_r' = \frac{1}{N} \sum f_i (x_i - A)^r$$

where,

$$\sum f_i = N$$

$$= \frac{1}{N} \sum_i f_i d_i^r \quad \text{where } d_i = (x_i - A)$$

This is called a raw moment.

Central moments

The rth moments of a variable about the mean $\bar{x}$, usually denoted by μ_r given by.

$$\mu_r = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^r$$

$$= \frac{1}{N} \sum_i f_i z_i^r$$

where $z_i = (x_i - \bar{x})$

$\mu_2 = \sigma^2 \quad \mu_1 = 0, \quad \mu_0 = 1$

In particular,

$\mu_0 = \frac{1}{N} \sum f_i (x_i - \bar{x})^0 = \frac{1}{N} \sum f_i = 1 \quad [\sum f_i = N]$

and $\mu_0 = 1 \quad \mu_1 = \frac{1}{N} \sum_i f_i (x_i - \bar{x}) = 0$

$[\because \sum (x_i - \bar{x}) = 0]$

$\mu_1 = 0$ always

Being the algebraic sum of deviations from the mean.

also, $\mu_2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 = \sigma^2$

These results, viz, $\mu_0 = 1$, $\mu_1 = 0$ and $\mu_2 = \sigma^2$ are of fundamental important and should be committed to memory. we know that

$d_i = (x_i - A)$ then

$\bar{x} = A + \frac{1}{N} \sum f_i d_i$

$= A + \mu_1'$

<u>Relation between moments about mean</u>

Interms of moments about at any point (Relationship between central moments and raw moments)

$$\mu_r = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^r$$

$$= \frac{1}{N} \sum_i f_i (x_i - A + A - \bar{x})^r$$

$$= \frac{1}{N} \sum_i f_i (d_i + A - \bar{x})^r \quad \text{where } d_i = (x_i - A)$$

we know that

$$\bar{x} = A + \frac{1}{N} \sum f_i d_i = A + \mu_1' \rightarrow (1)$$

using (1) we get,

$$\mu_r = \frac{1}{N} \sum_i f_i (d_i - \mu_1')^r$$

$$= \frac{1}{N} \sum_i f_i d_i^r - {}^rC_1 d_i^{r-1} \mu_1' + {}^rC_2 d_i^{r-2} \mu_1'^2 - {}^rC_3 d_i^{r-3} \mu_1'^3 \dots + (-1)^r \mu_1'^r \rightarrow (2)$$

$$\mu_r = \mu_r' - {}^rC_1 \mu_{r-1}' \mu_1' + {}^rC_2 \mu_{r-2}' \mu_1'^2 - {}^rC_3 \mu_{r-3}' \mu_1'^3 + \dots + (-1)^r \mu_1'^r \rightarrow (3)$$

In particular on putting $r = 2, 3, 4 \dots$ in equating 3 and simplying.

$r = 1$

$$\mu_1 = \mu_1' - \mu_1' = 0$$

$r = 2$

$$\mu_2 = \mu_2' - 2C_1 \mu_{2-1}' \mu_1' + 2C_2 \mu_{2-2}' (\mu_1')^2$$

$$= \mu_2' - 2\mu_1'\mu_1' + \mu_1'^2$$

$$= \mu_2' - 2\mu_1'^2 + \mu_1'^2$$

$$= \mu_2' - \mu_1'^2$$

$r=3$

$$\mu_3 = \mu_3' - 3C_1 \mu_{3-1}' \mu_1' + 3C_2 \mu_{3-2}' \mu_1'^2 - 3C_3 \mu_{3-3}(\mu_1')^3$$

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

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

$$= \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$$

$r=4$

$$\mu_4 = \mu_4' - 4C_1 \mu_{4-1}\mu_1' + 4C_2 \mu_{4-2}'(\mu_1')^2 - 4C_3 \mu_{4-3}'(\mu_1')^3 + 4C_4 \mu_{4-4}(\mu_1')^4$$

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

$$= \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$$

Similarly we can know moments about any point once the means and moments about are mean.

i) Given central moment

ii) we can find out raw moments.

$$\mu_r' = \frac{1}{N}\sum_i f_i (x_i - A)^r \qquad [\because \bar{x} = A + \mu_1']$$

$$= \frac{1}{N}\sum_i f_i (x_i - \bar{x} + \bar{x} - A)^r$$

$$= \frac{1}{N}\sum_i f_i (z_i + \mu_1')^r \qquad [z_i = x_i - \bar{x}]$$

Thus

$$\mu_r' = \frac{1}{N} \Sigma f_i \left(z_i^r + {}^rC_1 z_i^{r-1} \mu_1' + {}^rC_2 z_i^{r-2} \mu_1'^2 + \dots + \mu_1'^r \right)$$

$$= \mu_r + {}^rC_1 \mu_{r-1}' + {}^rC_2 \mu_{r-2} \mu_1'^2 + \dots + \mu_1'^r$$

$r = 2, 3, 4 \dots$ in (1) we get

$$\mu_1 = 0 , \quad \mu_2 = \mu_2' + \mu_1'^2$$

$$\mu_3 = \mu_3' - 3\mu_2' \mu_1' + 2\mu_1'^3$$

$$\mu_4 = \mu_4' - 4\mu_3' \mu_1' + 6\mu_2' \mu_1' - 3\mu_1'^4$$

UNIT - II

Bernoulli Distribution

Definition

A random variable x is said to have a Bernoulli distribution with parameter 'p' if its probability mass function is given by.

$$P(X=x) = \begin{cases} p^{x}(1-p)^{1-x} & \text{for } x = 0, 1 \\ 0 & \text{otherwise} \end{cases}$$

The parameter p satisfies $0 \le p \le 1$. often $(1-p)$ is denoted as q.

A random experiment whose outcomes are of two types, success S and failure F, occurring with probabilities p and q respectively, is called a Bernoulli distribution

Binomial distribution

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,

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

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

Moments of Binomial distribution or Constants of Binomial distribution.

The first four moments about origin of Binomial distribution are obtained as follows,

$$\mu_1' = E(X) = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} \qquad \frac{n}{x} \sum_{x=1}^{n} x \binom{n-1}{x-1} p^x$$

$$= np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} q^{n-x} \qquad = np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1}$$

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

$$= np$$

$$\left[\because \binom{n}{x} = \frac{n}{x}\binom{n-1}{x-1} = \frac{n}{x}\cdot\frac{n-1}{x-1}\binom{n-2}{x-2}\right.$$

$$\left. = \frac{n}{x}\cdot\frac{n-1}{x-1}\cdot\frac{n-2}{x-2}\binom{n-3}{x-3} \text{ and so on}\right]$$

Thus the mean of the binomial distribution is np.

$$\mu_2' = E(X^2) = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \left\{x(x-1)+x\right\} \frac{n(n-1)}{x(x-1)} \cdot \binom{n-2}{x-2} p^x q^{n-x}.$$

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

$$\mu_2' = n(n-1)p^2(q+p)^{n-2} + np$$

$$\mu_2' = n(n-1)p^2 + np$$

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

$$= \sum_{x=0}^{n} \left\{x(x-1)(x-2) + 3x(x-1) + x\right\} \binom{n}{x} p^x q^{n-x}$$

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

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

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

Similarly,

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

let

$x^4 = Ax(x-1)(x-2)(x-3) + Bx(x-1)(x-2) + Cx(x-1) + x$

[By giving to x the values 1, 2 and 3 we find the values of arbitary constants A, B and C]

$$\mu_4' = E(X^4) = \sum_{x=0}^{n} x^4 \binom{n}{x} p^x q^{n-x}$$

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

(on simplification)

Central moments of Binomial distributions

$\mu_2 = \mu_2' - \mu_1'^2$

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

$= np\{1-p\} = npq$

$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$

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

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

$= np\{3np(1-p) + 2p^2 - 3p + 1 - 3npq\}$

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

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

$= npq(1-2p)$

$$= npq\{q+p-2p\}$$

$$= npq(q-p)$$

$$\mu_4 = \mu_4' - 4\mu_3' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$$

$$= npq\{1+3(n-2)pq\} \quad \text{(on simplification)}$$

Hence,

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

$$= \frac{n^2p^2q^2(q-p)^2}{n^3p^3q^3} = \frac{(q-p)^2}{npq}$$

$$= \frac{(1-2p)^2}{npq}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$= \frac{npq\{1+3(n-2)pq\}}{n^2p^2q^2} = \frac{1+3(n-2)pq}{npq}$$

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

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

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

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

$$\gamma_1 = \sqrt{\beta_1}$$

$$= \sqrt{\frac{(q-p)^2}{npq}} = \frac{q-p}{\sqrt{npq}}$$

$$\gamma_2 = \beta_2 - 3$$

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

$$= \frac{1-6pq}{npq}$$

Recurrence relation for the moments of Binomial distributions.

$$\mu_r = E\{X-E(X)\}^r$$

$$= \sum_{x=0}^{r} (x-np)^r \binom{n}{x} p^x q^{n-x}.$$

Differentiating with respect to p, we get,

$$\frac{d\mu_r}{dp} = \sum_{x=0}^{n} \binom{n}{x} -nr(x-np)^{r-1} p^x q^{n-x} + (x-np)^r \{x p^{x-1} q^{n-x} - (n-x) p^x q^{n-x-1}\}]$$

$$= -nr \sum_{x=0}^{n} \binom{n}{x} (x-np)^{r-1} p^x q^{n-x} + \sum_{x=0}^{n} \binom{n}{x} (x-np)^r p^x q^{n-x} \left(\frac{x}{p} - \frac{n-x}{q}\right)$$

$$= -nr \sum_{x=0}^{n} (x-np)^{r-1} P(x) + \sum_{x=0}^{n} (x-np)^r P(x) \frac{(x-np)}{pq}$$

$$= -nr \sum_{x=0}^{n} (x-np)^{r-1} P(x) + \frac{1}{pq} \sum_{x=0}^{n} (x-np)^{r+1} P(x)$$

$$= -nr\mu_{r-1} + \frac{1}{pq} \mu_{r+1}$$

$$\mu_{r+1} = pq\left(nr\mu_{r-1} + \frac{d\mu_r}{dp}\right) \rightarrow (1)$$

(Renovsky formula)

Putting $r = 1, 2$ and 3 we get,

$r = 1$

$\mu_2 = pq\left(n(1)\mu_0 + \frac{d\mu_1}{dp}\right)$ $\quad (\because \mu_0 = 1 \;\&\; \mu_1 = 0)$

$= npq$

$\mu_3 = pq\left[2n\mu_1 + \frac{d\mu_2}{dp}\right]$

$= pq\left[2n(0) + \frac{d\mu_2}{dp}\right]$

$= pq\left[2n(0) + \frac{d(npq)}{dp}\right]$

$= pq\left[0 + \frac{d(npq)}{dp}\right]$

$= pq \cdot \frac{d(npq)}{dp}$

$= npq \cdot d/dp\,\{p(1-p)\}$

$= npq \; d/dp\,\{p - p^2\}$

$= npq\,(1-2p)$

$= npq\,(1-p) - p$

$= npq\,(q-p)$

$\mu_4 = pq\left[3n\mu_2 + \frac{d\mu_3}{dp}\right]$ $\quad p(1-p$

$\quad p - p^2($

$= pq\; 3n(npq) + d/dp\,[npq\,(q-p)\}]$ $\quad p - 2p^2 +$

$= p - 3p^2$

$= \left[3n^2pq + n\, d/dp\,\{p(1-p)(1-2p)\}\right]$

$= pq\left[3n^2pq + n\, d/dp\; (p - 3p^2 + 2p^3)\right]$

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

$6p + 6p^2$

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

$= -6p(1+p)$

$= npq\left[3npq + 1 - 6pq\right]$

$= -6p(1+q-1)$

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

$= -6pq.$

Mode of Binomial distribution

We have,

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

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

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

$$= \frac{(n-x+1)\,p}{xq} \quad ||.$$

$$= \frac{xq + (n-x+1)p - xq}{xq}$$

$$= 1 + \frac{(n+1)p - x(p+q)}{xq}$$

$$= 1 + \frac{(n+1)p - x}{xq} \rightarrow (\ )$$

mode is the value of x for which $P(x)$

is maximum.

We discuss the following two cases.

Case - I

When $(n+1)p$ is not an integer. Let

$(n+1)p = m+f$, where m is an integer and f

is fractional such that $0 < f < 1$ substituting

in ① we get, $1 + \frac{(n+1)p - x}{xq}$

$$\frac{P(x)}{P(x-1)} = 1 + \frac{(m+f) - x}{xq} \longrightarrow (2)$$

From (2) it is obvious that,

$$\frac{P(x)}{P(x-1)} > 1 \text{ for } x = 0, 1, 2 \ldots m \text{ and}$$

$$\frac{P(x)}{P(x-1)} < 1 \text{ for } x = m+1, m+2, \ldots, n$$

Put

$x = 1$ $\quad$ $x = 2$ $\quad$ $x = m$

$\frac{P(1)}{P(0)}$ $\quad$ $\frac{P(2)}{P(1)}$ $\quad$ $\frac{P(m)}{P(m-1)}$

$$\Rightarrow \frac{P(1)}{P(0)} > 1, \frac{P(2)}{P(1)} > 1, \ldots \frac{P(m)}{P(m-1)} > 1 \text{ and}$$

Put

$x = m+1$

$$\frac{P(m+1)}{P(m+1-1)} = \frac{P(m+1)}{P(m)}$$

$x = m+2$

$$\frac{P(m+2)}{P(m+2-1)} = \frac{P(m+2)}{P(m+1)}$$

$x = n$

$$\frac{P(n)}{P(n-1)}$$

$$\frac{P(m+1)}{P(m)} < 1, \quad \frac{P(m+2)}{P(m+1)} < 1, \quad \dots, \quad \frac{P(n)}{P(n-1)} < 1$$

$$P(0) < P(1) < P(2) < \dots < P(m-1) < P(m) > P(m+1) >$$
$$P(m+2) > \dots > P(n)$$

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

Case II

When $(n+1)p$ is an integer

Let, $(n+1)p = m$ (an integer)

Substituting in (1) we get,

$$\frac{P(x)}{P(x-1)} = 1 + \frac{m-x}{xq} \longrightarrow (3)$$

From ③ it is obvious that,

$$\frac{P(x)}{P(x-1)} \begin{cases} > 1 \text{ for } x = 1, 2 \ldots m-1 \\ = 1 \text{ for } x = m \\ < 1 \text{ for } x = m+1, m+2 \ldots n \end{cases}$$

Now, proceeding as in case I, we have,

$$P(0) < P(1) < \ldots < P(m-1) = P(m) > P(m+1) > P(m+2) > \ldots > P(n)$$

Thus in this case the binomial distribution is bimodal and the two modal values are m and $m-1$.

Moment generating function of Binomial distribution

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

$$M_X(t) = E(e^{tx})$$

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

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

$$= (q + pe^t)^n$$

Moment generating function about mean of Binomial distribution

$$E\{e^{t(X-np)}\} = e^{-tnp} \cdot E(e^{tX}) \qquad (e^{-tp})^n (q+pe^t)^n$$

$$= e^{-tnp} \cdot M_X(t) \qquad = (qe^{-tp} + pe^{-tp+t})^n$$

$$= e^{-tnp}(q+pe^t)^n \qquad = qe^{-tp} + pe^{t(-p+1)}$$

$$= (qe^{-pt} + pe^{tq})^n \qquad = (qe^{-tp} + pe^{tq})^n$$

$$= \left[q\left\{1 - pt + \frac{p^2t^2}{2!} - \frac{p^3t^3}{3!} + \frac{p^4t^4}{4!} \cdots\right\} + p\left\{1 + tq + \frac{t^2q^2}{2!} + \frac{t^3q^3}{3!} + \cdots\right\}\right]^n$$

$q+p=1$

$$= \left[(q+p) + \frac{t^2}{2!}pq(q+p) + \frac{t^3}{3!}pq(q^2-p^2) + \frac{t^4}{4!}pq(q^3+p^3) + \cdots\right]^n$$

$$= \left[1 + \binom{n}{1}\left\{\frac{t^2}{2!}pq + \frac{t^3}{3!}pq(q-p) + \frac{t^4}{4!}qp(1-3pq) + \cdots\right\}\right]^n$$

$$= 1 + \binom{n}{1}\left\{\frac{t^2}{2!}pq + \frac{t^3}{3!}pq(q-p) + \frac{t^4}{4!}pq(1-3pq) + \cdots\right\}$$

$$+ \binom{n}{2}\left\{\frac{t^2}{2!}pq + \frac{t^2}{3!}pq(q-p) + \cdots\right\}^2 + \cdots\Bigg]$$

Now,

$$\mu_2 = \text{co-efficient of } \frac{t^2}{2!} = npq$$

$$\mu_3 = \text{co-efficient of } \frac{t^3}{3!} = npq(q-p)$$

$$\mu_4 = \text{co-efficient of } \frac{t^4}{4!} = npq(1-3pq) + 3n(n-1)p^2q^2$$

$$= 3n^2p^2q^2 + npq(1-6pq)$$

Additive property of Binomial distribution

Let $X \sim B(n_1, P_1)$ and $Y \sim B(n_2, P_2)$ be independent random variables then,

$$\left.\begin{aligned} M_X(t) &= (q_1 + P_1 e^t)^{n_1} \\ M_Y(t) &= (q_2 + P_2 e^t)^{n_2} \end{aligned}\right\} \longrightarrow (1)$$

We have,

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t) \quad [\because X \text{ and } Y \text{ are independent}]$$

$$= (q_1 + P_1 e^t)^{n_1} \cdot (q_2 + P_2 e^t)^{n_2} \longrightarrow (2)$$

Since (2) cannot be expressed in the form $(q + pe^t)^n$ from uniqueness theorem of MGF's it follows that $X+Y$ is not binomial variable. Hence in general the sum of two independent binomial variate. In other words, binomial distribution does not possess the additive or reproductive property.

However, if we take $P_1 = P_2 = P$ (say) then from (2).

$$M_{X+Y}(t) = (q + Pe^t)^{n_1 + n_2}$$

which is the mgf of a binomial variate with parameters $(n_1 + n_2, P)$. Hence by uniqueness theorem of m.g.f.s $X + Y \sim B(n_1 + n_2, P)$, Thus the binomial distribution prosses the additive or reproductive property if $P_1 = P_2$

Characteristic function of Binomial distribution

$$\varphi_x(t) = E(e^{itx})$$

$$= \sum_{x=0}^{n} e^{itx} p(x)$$

$$= \sum_{x=0}^{n} e^{itx} \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (pe^{it})^x q^{n-x}$$

$$= (q + pe^{it})^n$$

Cumulants of the Binomial distribution

Cumulant generating function is given by

$$K_x(t) = \log M_x(t)$$

$$= \log (q + pe^t)^n$$

$$= n \log (q + pe^t)$$

$$= n \log \left[q + p\left(1 + t + \tfrac{t^2}{2!} + \tfrac{t^3}{3!} + \tfrac{t^4}{4!} + \cdots\right)\right]$$

$$= n \log \left[1 + p\left(t + \tfrac{t^2}{2!} + \tfrac{t^3}{3!} + \tfrac{t^4}{4!} + \cdots\right)\right]$$

$$= n\left[P\left(t + t^2/2! + t^3/3! + t^4/4! + \cdots\right) - \frac{P^2}{2}\left(t + t^2/2! + t^3/3! + \cdots\right)^2\right.$$

$$\left. + \frac{P^3}{3}\left(t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\right)^3 - \frac{P^4}{4}\left(t + \frac{t^2}{2!} + \frac{t^3}{3} + \cdots\right)^4 + \cdots\right]$$

Mean $= K_1 =$ co-efficient of t in $K_X(t) = np$

$\mu_2 = K_2 =$ co-efficient of $t^2/2!$ in $K_X(t) = n(P - P^2)$

$$= nP(1-P) = npq$$

The co-efficient of t^3 in $K_X(t)$

$$= n\left[P/3! - \frac{P^2}{2!} \cdot 2 \cdot 1/2! + P^3/3\right]$$

$$= \frac{nP}{3!}(1 - 3P + 2P^2) \Rightarrow K_3 = \text{co-efficient of } \frac{t^3}{3!} \text{ in } K_X(t)$$

$$= np(1 - 3P + 2P^2)$$

$$= np(1-P)(1-2P)$$

$$= npq(1 - P - P)$$

$$= npq(q - P)$$

$\mu_3 = K_3 = npq(q - P)$

The co-efficient of t^4 in $K_X(t)$

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

$$= \frac{nP}{4!}(1 - 7P + 12P^2 - 6P^3)$$

K_4 = co-efficient of $\frac{t^4}{4!}$ in $K_X(t)$

$= np(1-p)(1-6p+6p^2)$

$= npq\,[1-6p(1-p)]$

$= npq\,(1-6pq)$

$\mu_4 = K_4 + 3K_2^2$

$= npq\,(1-6pq) + 3n^2p^2q^2$

$= npq\,(1-6pq+3npq)$

$= npq\,[1+3pq\,(n-2)]$.

Probability generating function.

The p.g.f of a random variable is defined as,

$P(s) = E(s^x)$ For binomial distribution

$$P(s) = E(s^x) = \sum_{x=0}^{n} s^x\, P(x)$$

$$= \sum_{n=0}^{n} s^x \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{n=0}^{n} \binom{n}{x} (ps)^x q^{n-x}$$

$$= (q+ps)^n$$